Introduction to Target Trial Emulation

David Hajage david.hajage@aphp.fr

3 Juin 2025

SANTÉ SORBONNE UNIVERSITÉ

Sommaire

SANTÉ
SORBONNE
UNIVERSITÉ

1

Introduction

Evaluation of Health Interventions and RCTs

Randomized controlled trials : gold standard gold standard for evaluating the comparative effectiveness of treatments or interventions

However...

Evaluation of Health Interventions and RCTs

Randomized controlled trials : gold standard gold standard for evaluating the comparative effectiveness of treatments or interventions

However...

- Strict inclusion criteria can reduce generalizability
- Relatively short follow-up limits the range of outcomes studied
- Not immune to attrition bias
- Difficult to conduct when rapid decisions are needed or for evaluating complex interventions
- Often costly

Evaluation of Health Interventions and RWD

Real world data : medical records, hospital admissions, disease registries, cohorts, etc.

- Wider range of participants and follow-up.
- Generally less expensive than RCTs.
- Useful for studying treatments received under 'real' conditions (i.e., pragmatic).

However...

Evaluation of Health Interventions and RWD

Real world data : medical records, hospital admissions, disease registries, cohorts, etc.

- Wider range of participants and follow-up.
- Generally less expensive than RCTs.
- Useful for studying treatments received under 'real' conditions (i.e., pragmatic).

However...

- Potentially affected by various biases

Biases in Observational Studies

Biases are usually classified into three distinct groups:

- Confounding bias
- Selection bias
- Classification bias

What type(s) of bias have the most impact in an observational cohort study?

Classic Example: HRT and Coronary Artery Disease

- Hormone replacement therapy (HRT) in menopausal women and coronary artery disease (CAD).
- Randomized trial from the Women's Health Initiative (WHI): *20% increase in the risk of CAD* in women initiating HRT compared to non-initiators (Manson et al., NEJM 2003).
- Observational study on the health of nurses: *30% reduction in the risk of CAD* in prevalent users compared to never-users (Grodstein et al., J Women's Health 2006).

Classic Example: HRT and Coronary Artery Disease

What went wrong in the previous example?

- Confounding bias?

Classic Example: HRT and Coronary Artery Disease

What went wrong in the previous example?

- Confounding bias?
- Different questions or poor design?

Classic Example: HRT and Coronary Artery Disease

What went wrong in the previous example?

- Confounding bias?
- Different questions or poor design?

A reanalysis of the same data on nurses' health, this time comparing incident users to non-users, leads to estimates similar to the results of the WHI (Hernán et al., Epidemiology 2008).

Hernán's Two Steps Algorithm

Hernán & Robins, American Journal of Epidemiology 2016

Hernán's Two Steps Algorithm

Hernán & Robins, American Journal of Epidemiology 2016

- Define the question

Hernán's Two Steps Algorithm

Hernán & Robins, American Journal of Epidemiology 2016

- Define the question
- Answer the question

Hernán's Two Steps Algorithm

Hernán & Robins, American Journal of Epidemiology 2016

- Define the question
  - Writing a detailed randomized clinical trial protocol

Hernán's Two Steps Algorithm

Hernán & Robins, American Journal of Epidemiology 2016

- Define the question
  - Writing a detailed randomized clinical trial protocol
- Answer the question
  - Option A: Conduct the clinical trial (funding, authorization, inclusion, etc.)
  - Option B: Use already available observational data to emulate the target trial point by point.

Hernán's Two Steps Algorithm

Step 1 Define the target trial protocol

- Eligibility criteria
- Treatment strategies
- Randomization
- Start/end of follow-up
- Outcomes
- Causal contrast
- Statistical analysis plan

Hernán's Two Steps Algorithm

Step 1 Define the target trial protocol

- Eligibility criteria
- Treatment strategies
- Randomization
- Start/end of follow-up
- Outcomes
- Causal contrast
- Statistical analysis plan

Step 2 Emulate the target trial

- Eligibility criteria
- Treatment strategies
- Randomization
- Start/end of follow-up
- Outcomes
- Causal contrast
- Statistical analysis plan

Hernán's Two Steps Algorithm

(Good) RCTs suffer less from :

- *Confounding bias* because of randomization
- *Selection bias* and measurement error because their protocols plan a priori the selection criteria, the intervention, the endpoint, and the start/end of follow-up

Hernán and Robins (2016) proposed a formal approach to adopt the same design principles in research based on observational data

SANTÉ SORBONNE UNIVERSITÉ

2

Time Zero Related Biases

What Happens in a Randomized Trial?

Some Misalignment Related Biases

SANTÉ SORBONNE UNIVERSITÉ

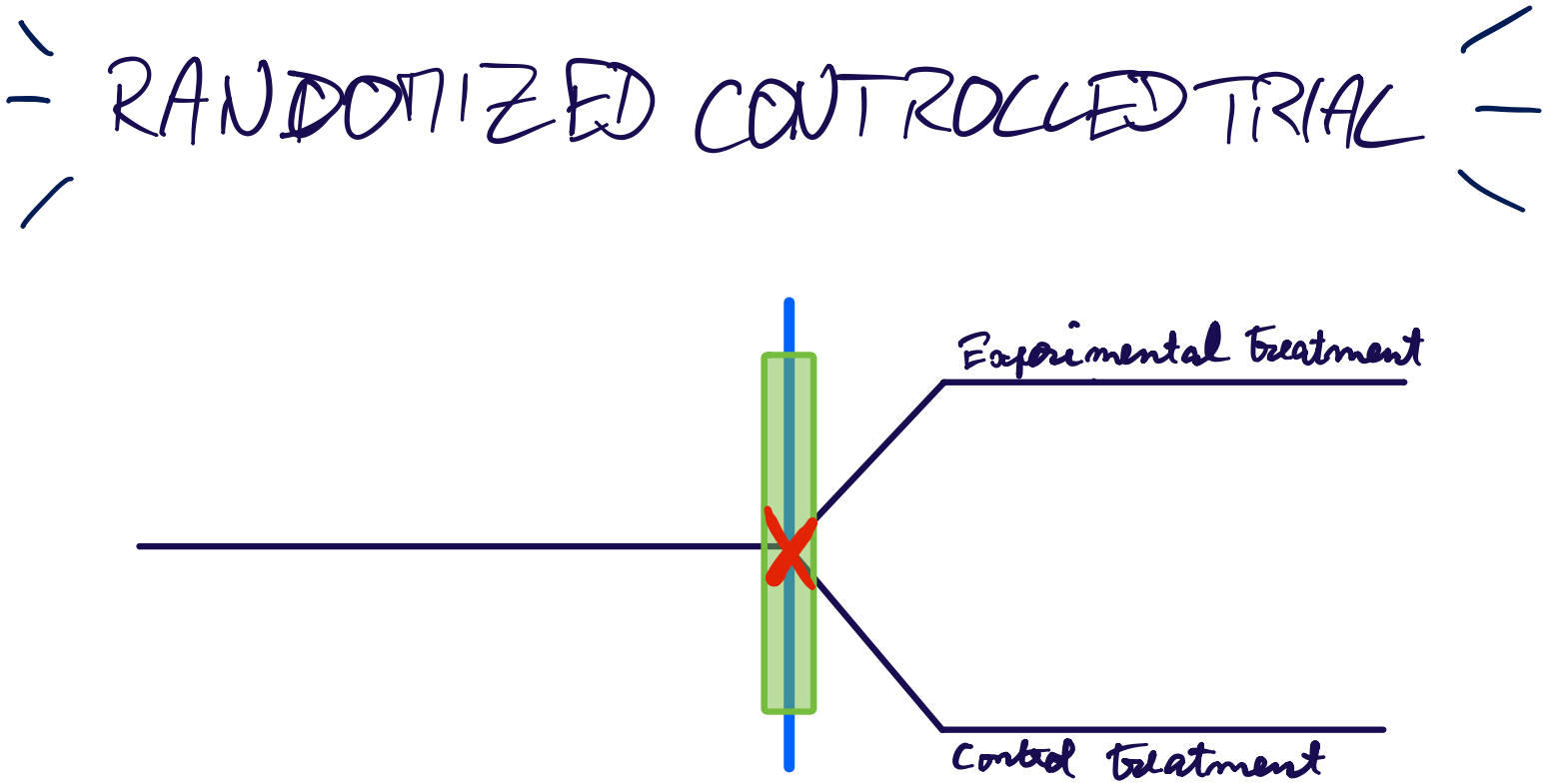
2

Time Zero Related Biases

What Happens in a Randomized Trial?

Some Misalignment Related Biases

What Happens in a Randomized Trial ? I

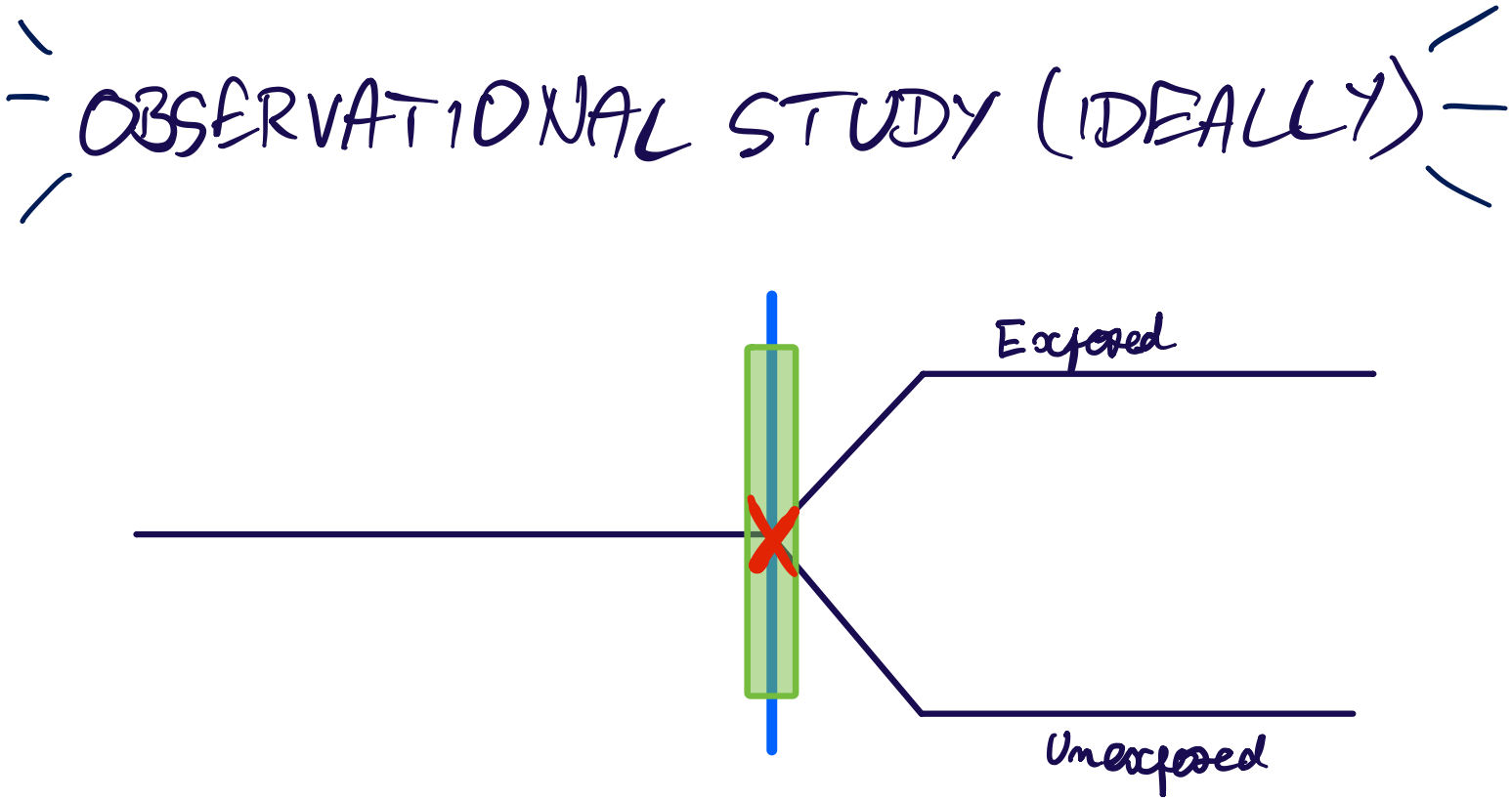


What Happens in a Randomized Trial ? II

Alignment of three key moments at baseline

- Eligibility criteria met (E)
- Treatment assignment (A)
- Start of follow-up (T_0)

What Should Happen in an Observational Study? I



What Should Happen in an Observational Study? II

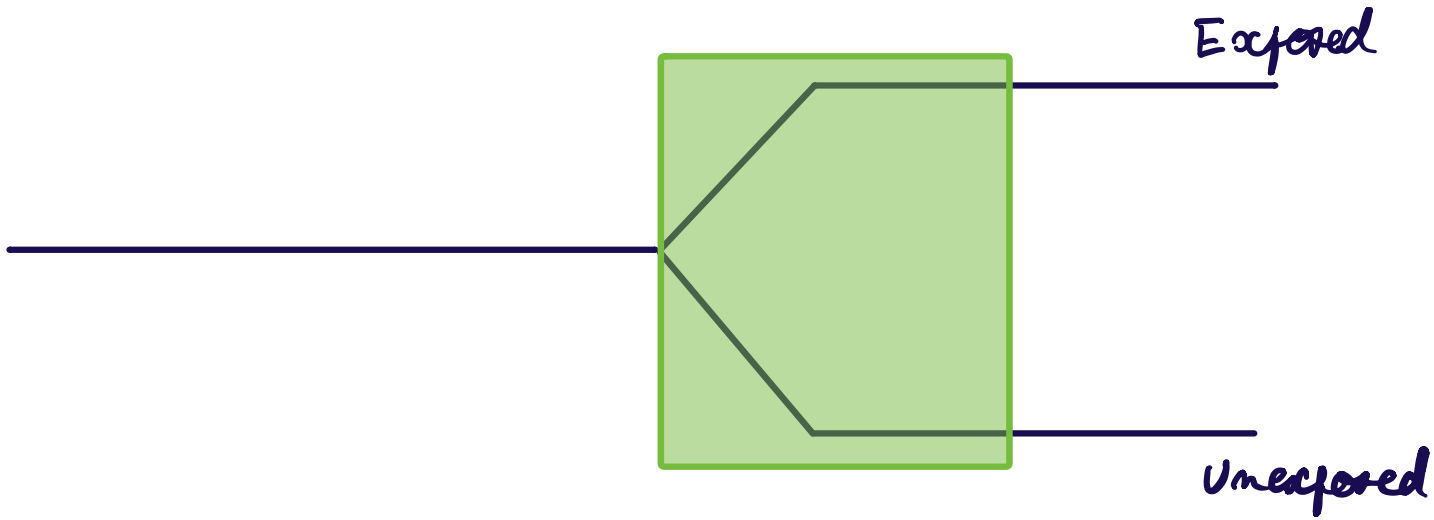
(If we wanted to emulate the trial)

Alignement of three key moments at baseline:

- Eligibility criteria met (E)
- Treatment initiation (A)
- Start of follow-up (T_0)

What Should Happen in an Observational Study? III

BUT OFTEN WE HAVE...



| To (start of follow up)
X Eligibility criteria met
▯ Treatment initiation

SANTÉ SORBONNE UNIVERSITÉ

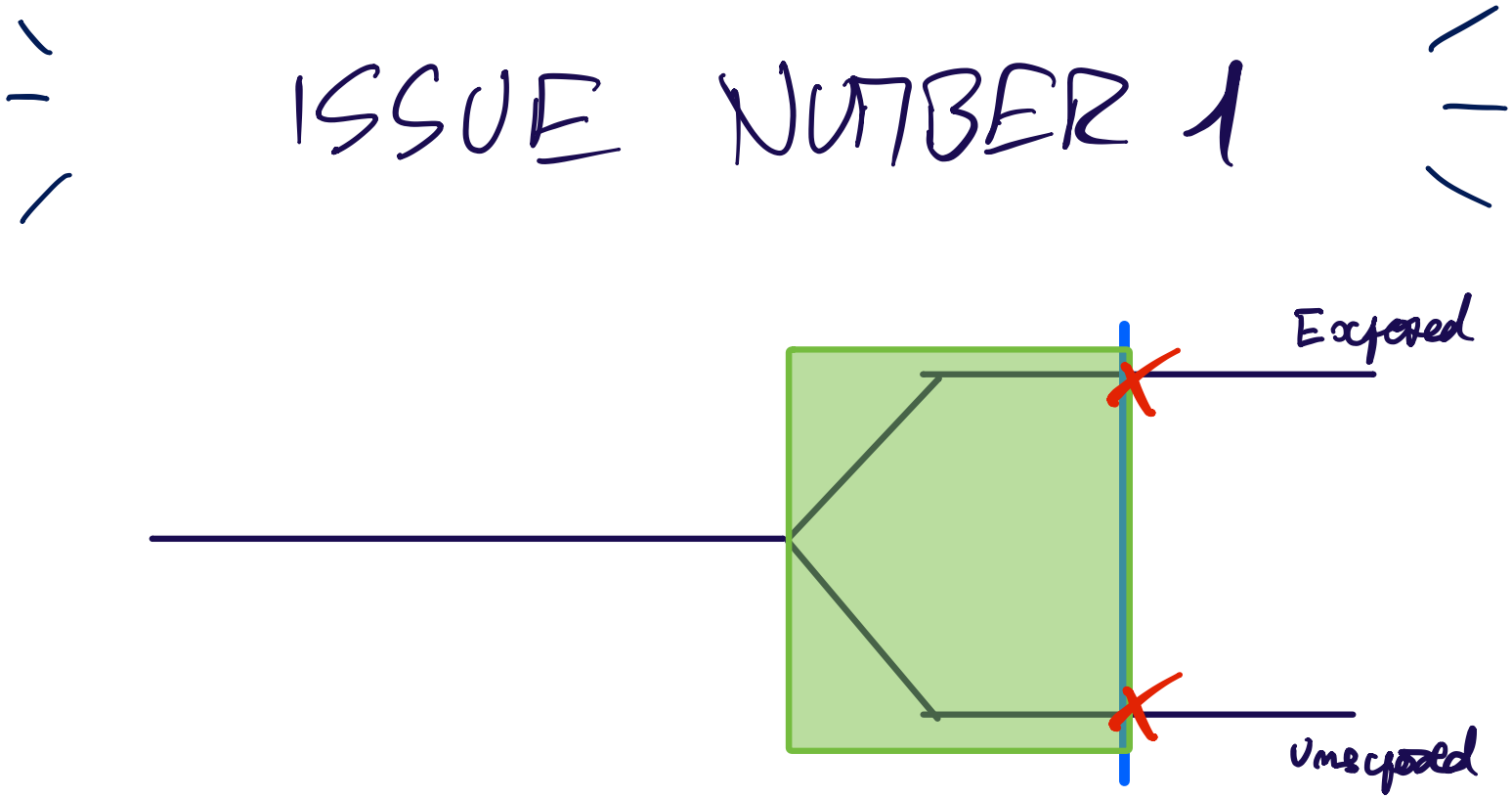
2

Time Zero Related Biases

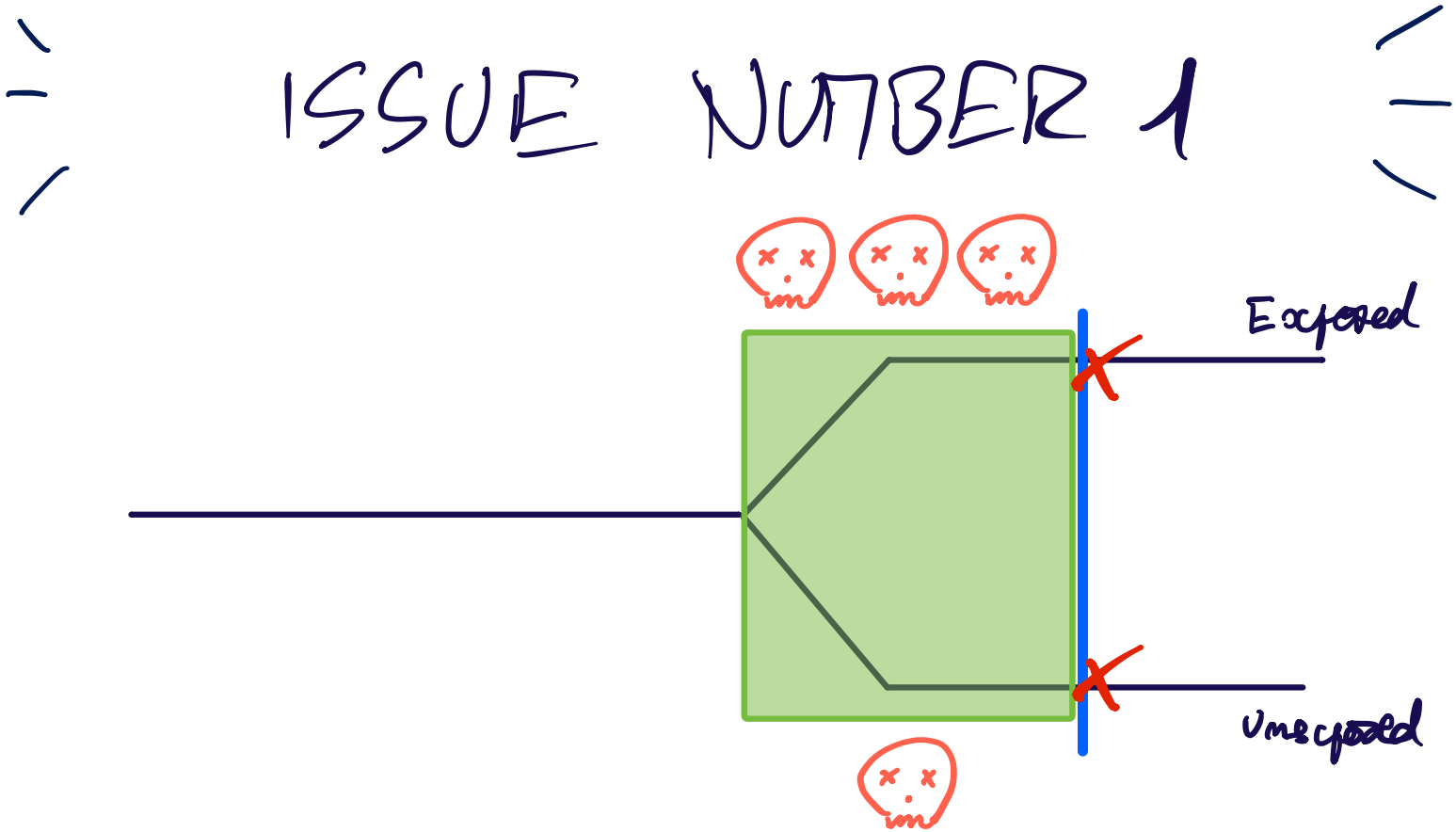
What Happens in a Randomized Trial?

Some Misalignment Related Biases

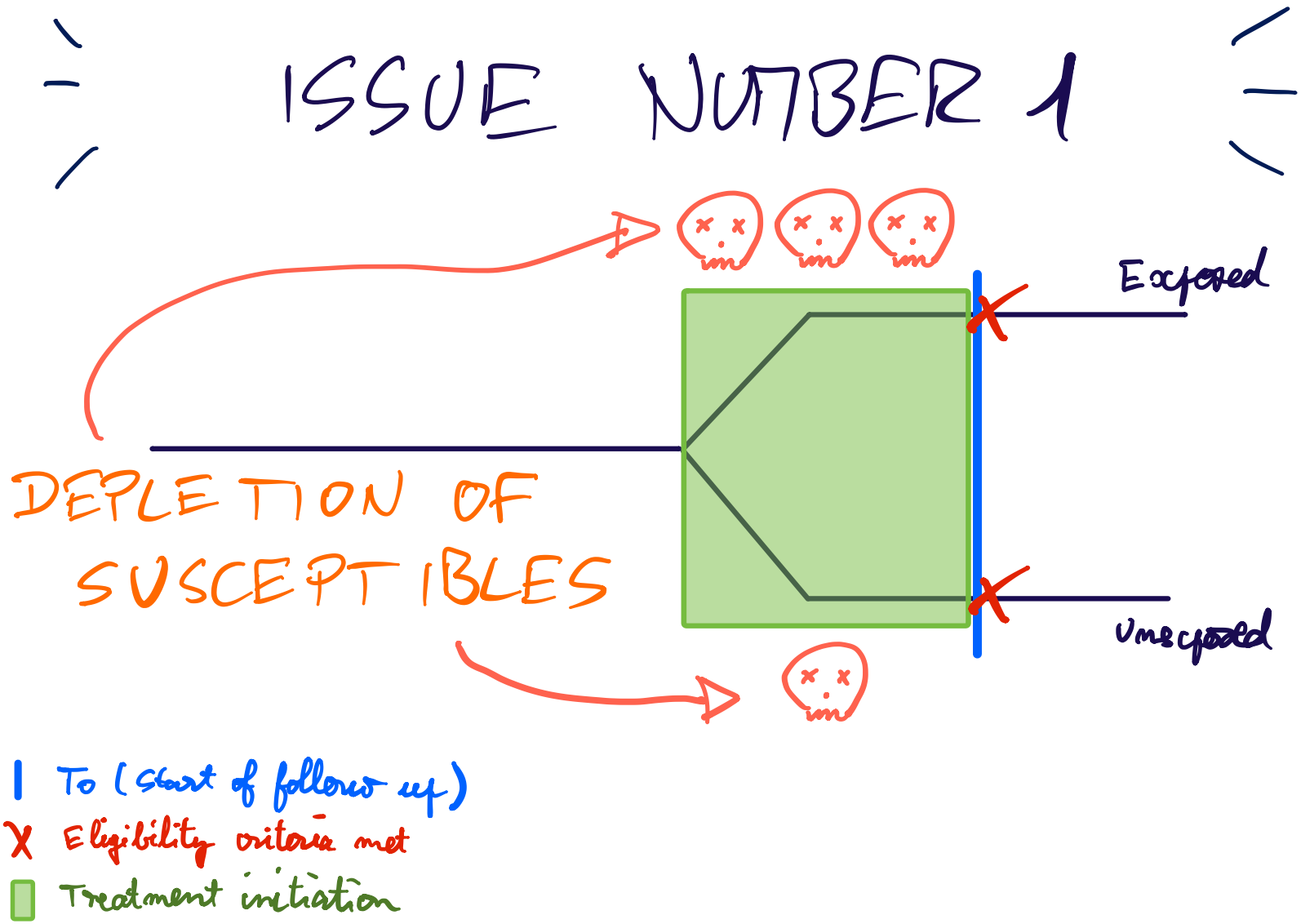
Issue number 1 |



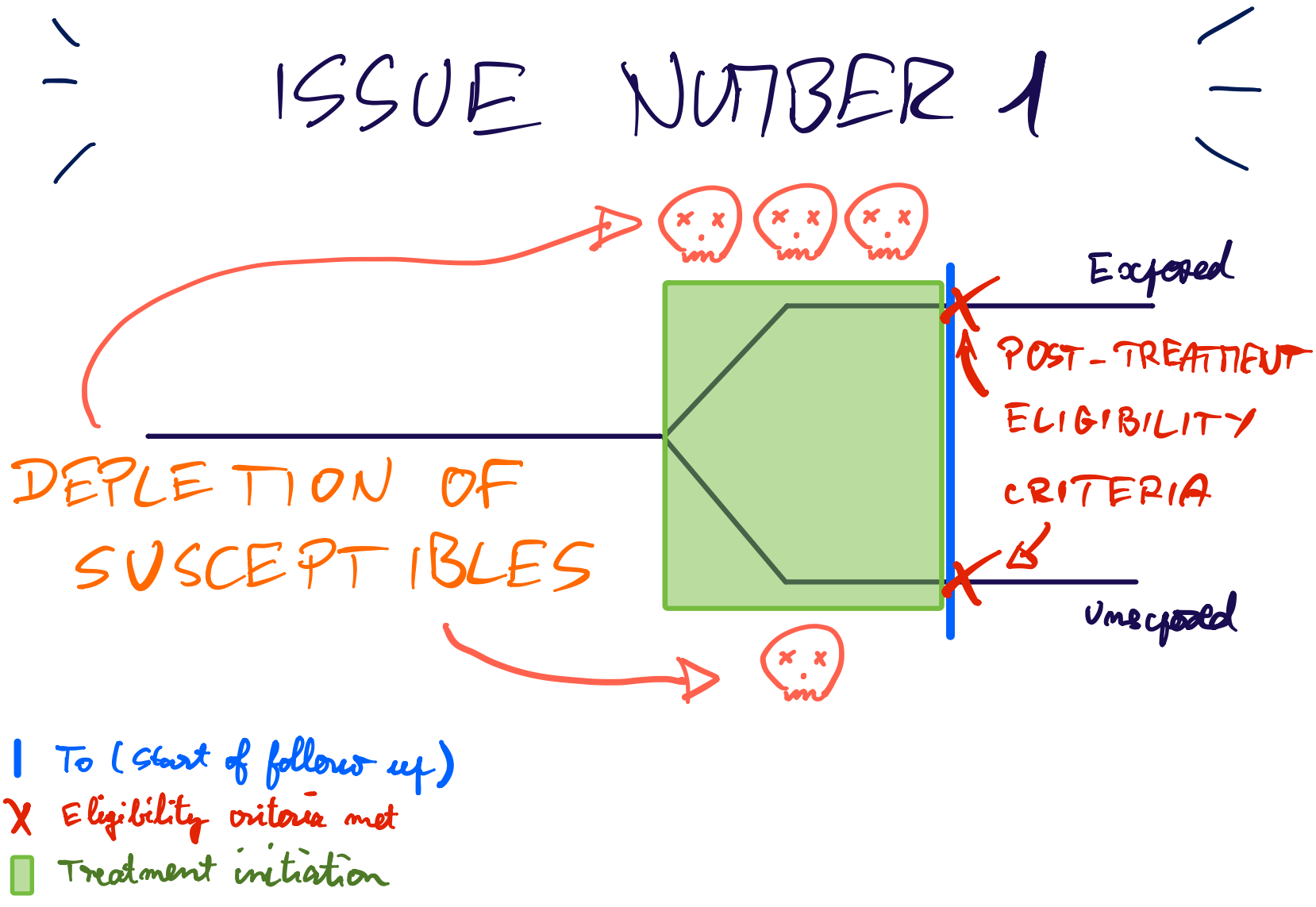
Issue number 1 II



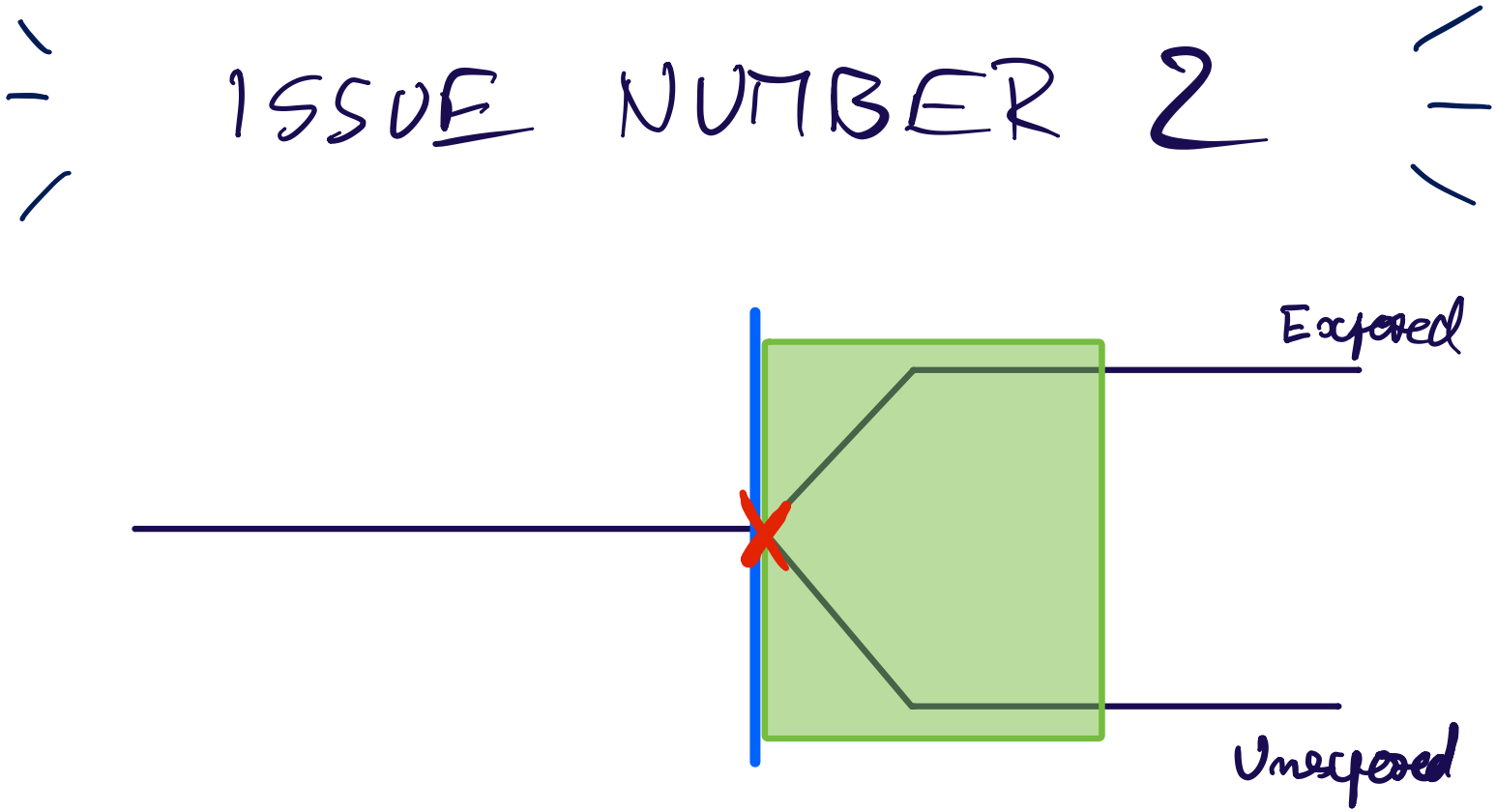
Issue number 1 III



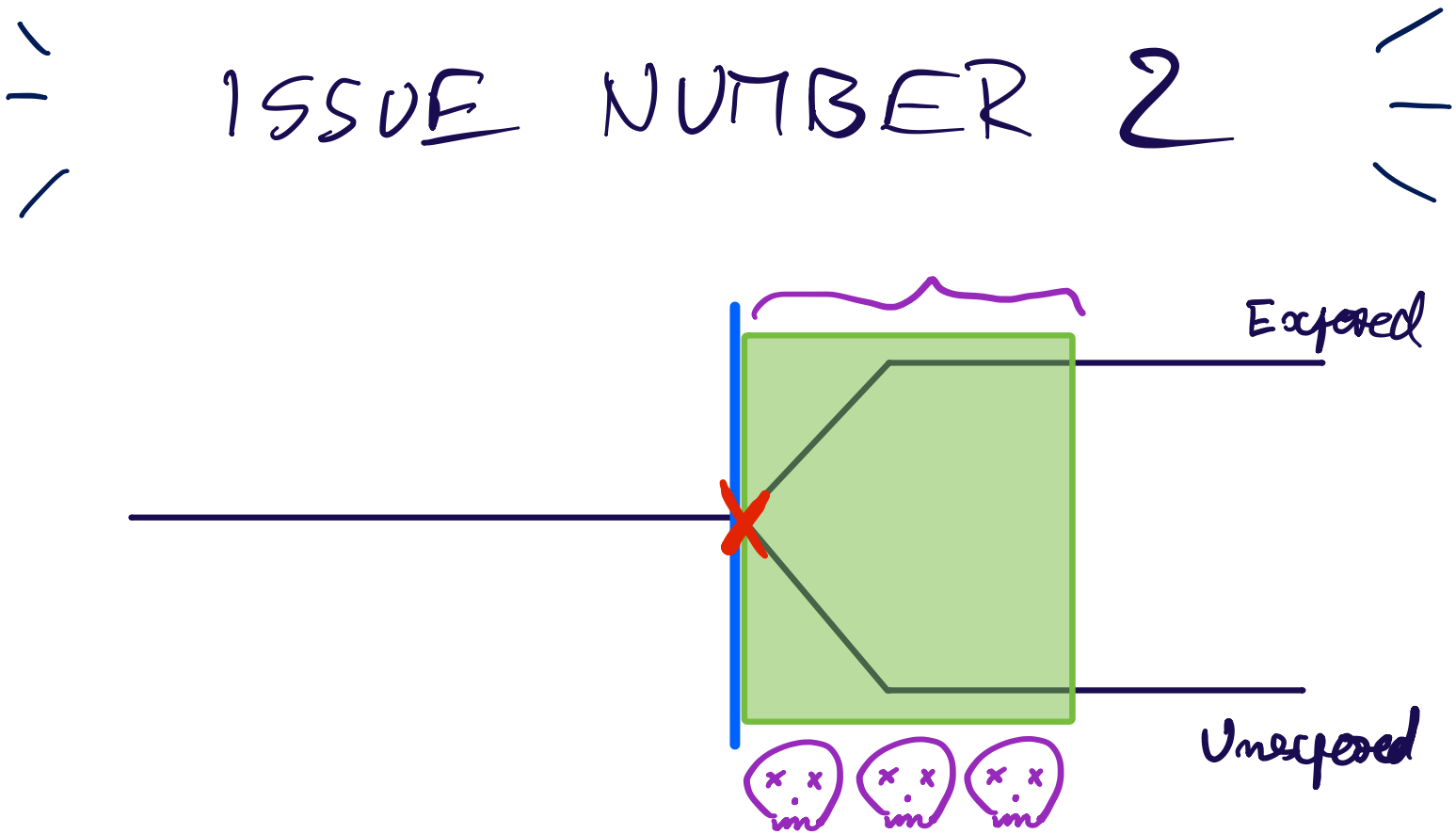
Issue number 1 IV



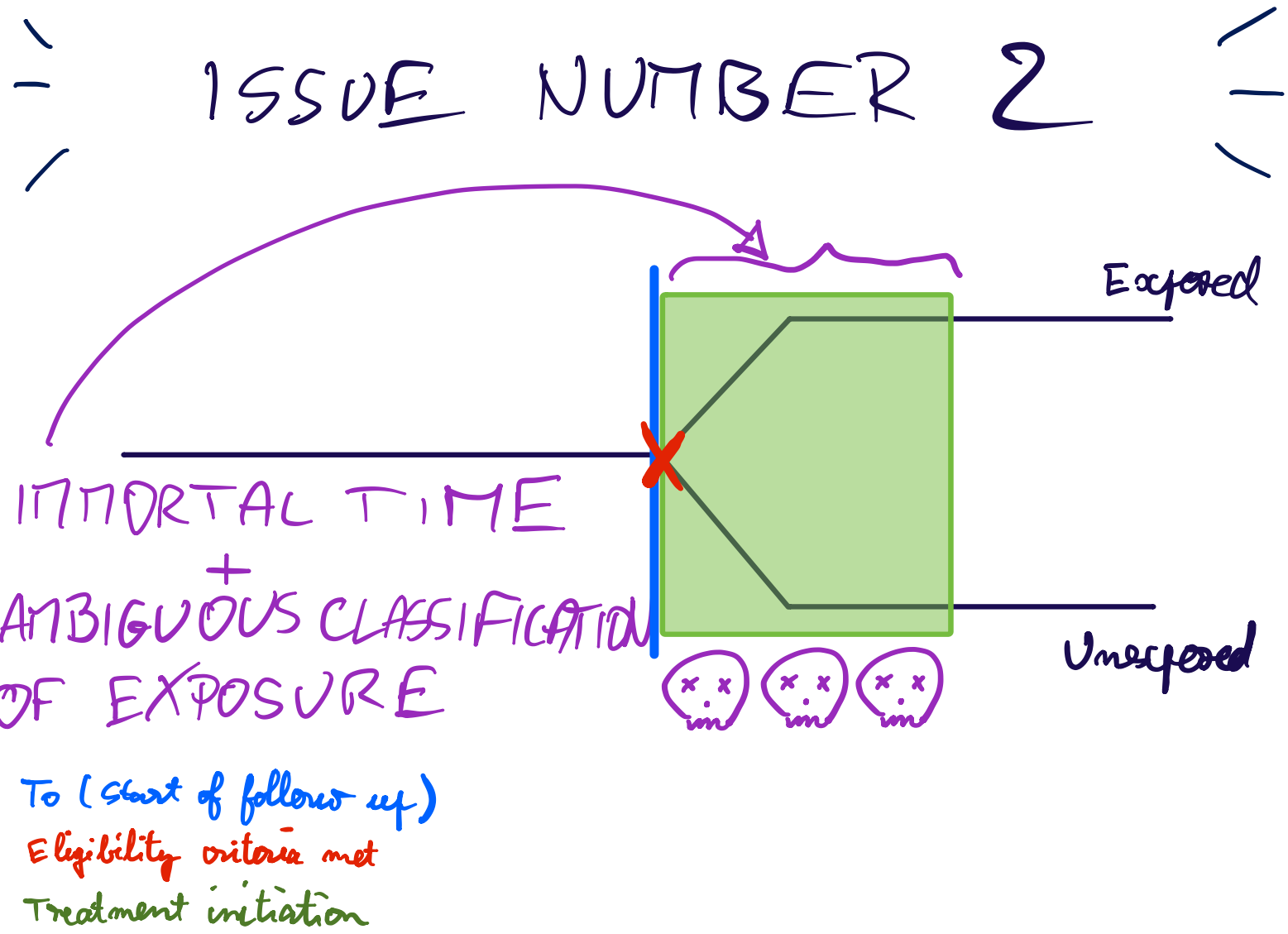
Error number 2 I



Error number 2 II



Error number 2 III



Immortal time bias

Generally occurs when information about treatment obtained after the start of follow-up is used to assign the treatment strategy.

Immortal time bias

Generally occurs when information about treatment obtained after the start of follow-up is used to assign the treatment strategy.

Recent example :

> *Here at IHU MI, you know that we treat people with HCQ and AZI. And of the 4,600 people who received their treatment, which was evaluable that is to say at least 3 days of treatment, there were 19 deaths.*

SANTÉ SORBONNE UNIVERSITÉ

3

Frequent Problems and Some Remedies

- Patients already under treatment at time zero
- Problem of Non-Adherence to the Assigned Strategy
- Strategies not defined at time zero
- Strategy too rare at time zero
- Time zero difficult to define uniquely

SANTÉ SORBONNE UNIVERSITÉ

3

Frequent Problems and Some Remedies

Patients already under treatment at time zero

Problem of Non-Adherence to the Assigned Strategy

Strategies not defined at time zero

Strategy too rare at time zero

Time zero difficult to define uniquely

New user (active comparator) study design

- Select incident users ('new user')
- Select a control group exposed to another active treatment (if possible)

  ⇒ Facilitates the selection of incident users in this group too!

  ⇒ Reduces confounding bias (better comparability between the two groups, including on unmeasured confounders).

  ⇒ Approaches the question asked in a randomized clinical trial (comparison of two therapeutic options in patients with the same clinical indication).

SANTÉ SORBONNE UNIVERSITÉ

3

Frequent Problems and Some Remedies

Patients already under treatment at time zero

Problem of Non-Adherence to the Assigned Strategy

Strategies not defined at time zero

Strategy too rare at time zero

Time zero difficult to define uniquely

Problem

Individuals do not adhere to the strategy defined at time zero: a frequent issue in the case of long-term treatments

Examples

- Question: What is the benefit of long-term aspirin use in reducing the risk of recurrence after colorectal cancer surgery?
- Problem: Individuals often discontinue their treatment after only a few weeks

- Question: What is the benefit of taking HCQ for at least three days in the case of COVID-19?
  Problem: Individuals often stop their treatment after just one day

- Question: Comparison of two long-term HRT strategies during menopause.
  Problem: Individuals often discontinue their treatment (in both groups).

Two Valid Approaches

- Do nothing
  - Estimation of an effect analogous to the ITT effect

Two Valid Approaches

- Do nothing
  - Estimation of an effect analogous to the ITT effect
- Artificially censor individuals at the moment they deviate from the strategy assigned at time zero
  - Estimation of an effect analogous to the per-protocol effect
  - This method requires addressing informative censoring:
    - Artificial censoring is not random; it likely depends on the individual's health status at the time of censoring
    - This introduces time-dependent selection bias
    - The 'inverse probability of censoring weights' (IPCW) method can be used to address this issue, accounting for time-dependent prognostic factors (complexity + requires collecting these data throughout follow-up)

SANTÉ SORBONNE UNIVERSITÉ

3

Frequent Problems and Some Remedies

Patients already under treatment at time zero

Problem of Non-Adherence to the Assigned Strategy

Strategies not defined at time zero

Strategy too rare at time zero

Time zero difficult to define uniquely

Example

Question : In individuals whose hemoglobin is greater than 11 g/dL at time zero, the benefit of: initiating erythropoietin if hemoglobin falls below 10 g/dL vs. initiating erythropoietin if hemoglobin falls below 11 g/dL vs. never initiating erythropoietin during follow-up

Example

Question : In individuals whose hemoglobin is greater than 11 g/dL at time zero, the benefit of: initiating erythropoietin if hemoglobin falls below 10 g/dL vs. initiating erythropoietin if hemoglobin falls below 11 g/dL vs. never initiating erythropoietin during follow-up

- Problem: To which strategy do individuals whose hemoglobin never falls below 10g/dL during follow-up belong?
- It is not always possible to distinguish therapeutic strategies unambiguously at the beginning and during follow-up
- (And doing so anyway may introduce immortal time bias)

Two Valid Approaches

- Assign a strategy by randomization $\Rightarrow$ unbiased, but not very efficient from a statistical point of view

Two Valid Approaches

- Assign a strategy by randomization $\Rightarrow$ unbiased, but not very efficient from a statistical point of view
- Cloning-censoring-weighting approach
  - Create identical copies (clones) of each individual and assign each clone to one of the strategies.
  - Censor a clone when their data cease to be consistent with the strategy assigned to them.
  - Requires an appropriate estimation of the variance. (to account for the fact that the same individual is duplicated).
  - Requires taking into account informative censoring (potential selection bias after time zero), usually using IPCW.
  - Estimation of an effect analogous to the per-protocol effect (impossible to estimate an analogue of the ITT effect).

SANTÉ SORBONNE UNIVERSITÉ

3

Frequent Problems and Some Remedies

Patients already under treatment at time zero

Problem of Non-Adherence to the Assigned Strategy

Strategies not defined at time zero

Strategy too rare at time zero

Time zero difficult to define uniquely

Example

- Question : benefit of initiating long-term aspirin on the risk of recurrence after digestive surgery for colorectal cancer
- Problem : only five individuals initiate aspirin precisely at time zero, even if 5,000 start therapy in the following 3 months
  - Estimation is bound to fail with only 5 individuals.
  - 5,000 individuals assigned to the strategy without aspirin... but they start it shortly after time zero

Valid Approach

Introduce a grace period

- Emulate a target trial in which individuals assigned to treatment are not required to start it exactly at randomization, but within a reasonable time called the grace period
  - Example : start aspirin in the first three months
- Advantages :
  - More efficient from a statistical point of view (here, 5,000 individuals in the exposed group)
  - Sometimes more realistic: even in a randomized trial, individuals are not always required to start treatment as soon as they meet the eligibility criteria and are randomized
- Disadvantages :
  - Therapeutic strategies not always defined at time zero (e.g., aspirin initiated at 2 months), which exposes to immortal time bias
  - $\Rightarrow$ Cloning-censoring-weighting approach

Cloning-censoring-weighting approach

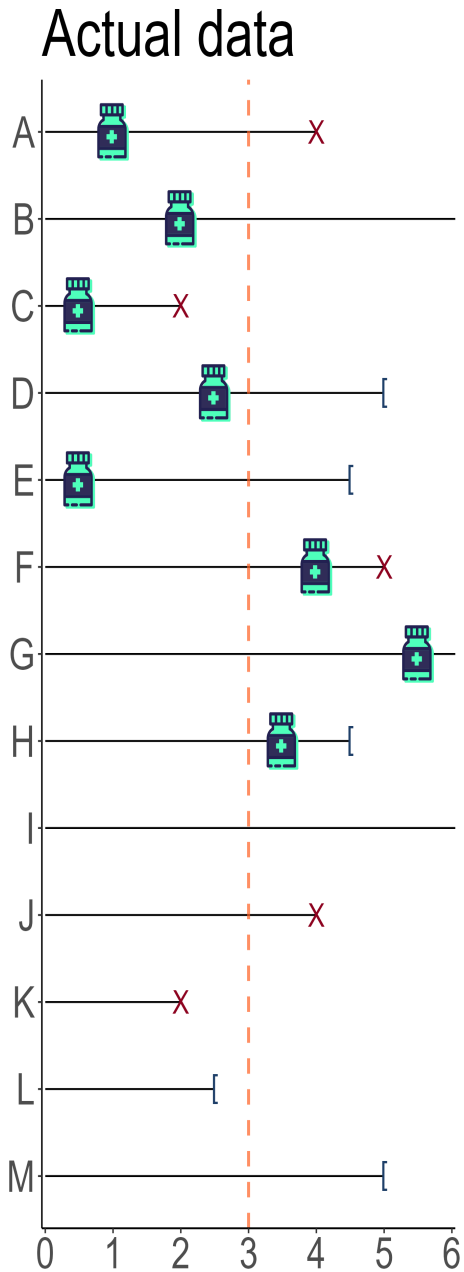
From:

International Journal of Epidemiology (2020)

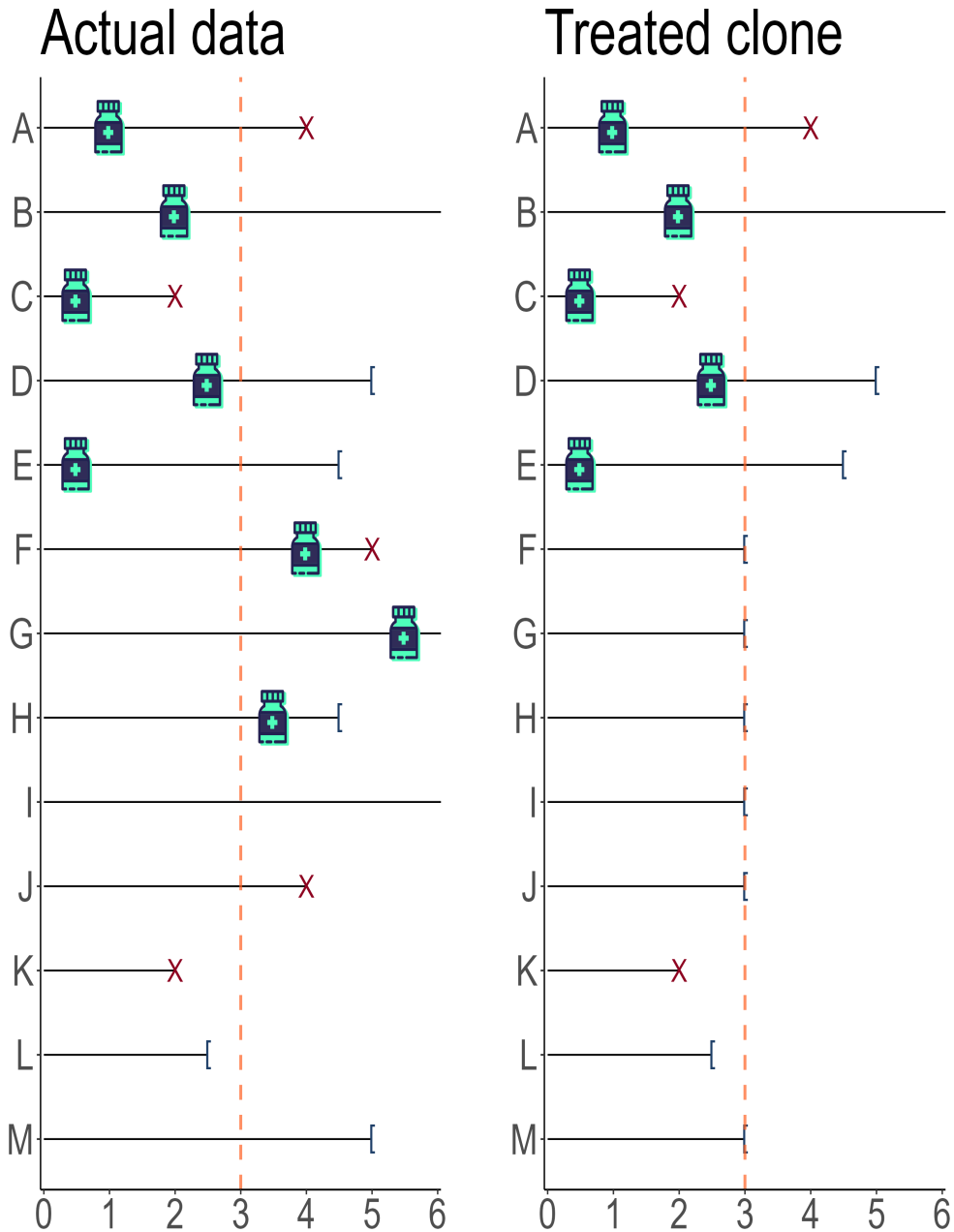
Reflection on modern methods: trial emulation in the presence of immortal-time bias. Assessing the benefit of major surgery for elderly lung cancer patients using observational data

Maringe, Camille; Benitez Majano, Sara; Exarchakou, Aimilia; Smith, Matthew; Rachet, Bernard; Belot, Aurélien; Leyrat, Clémence

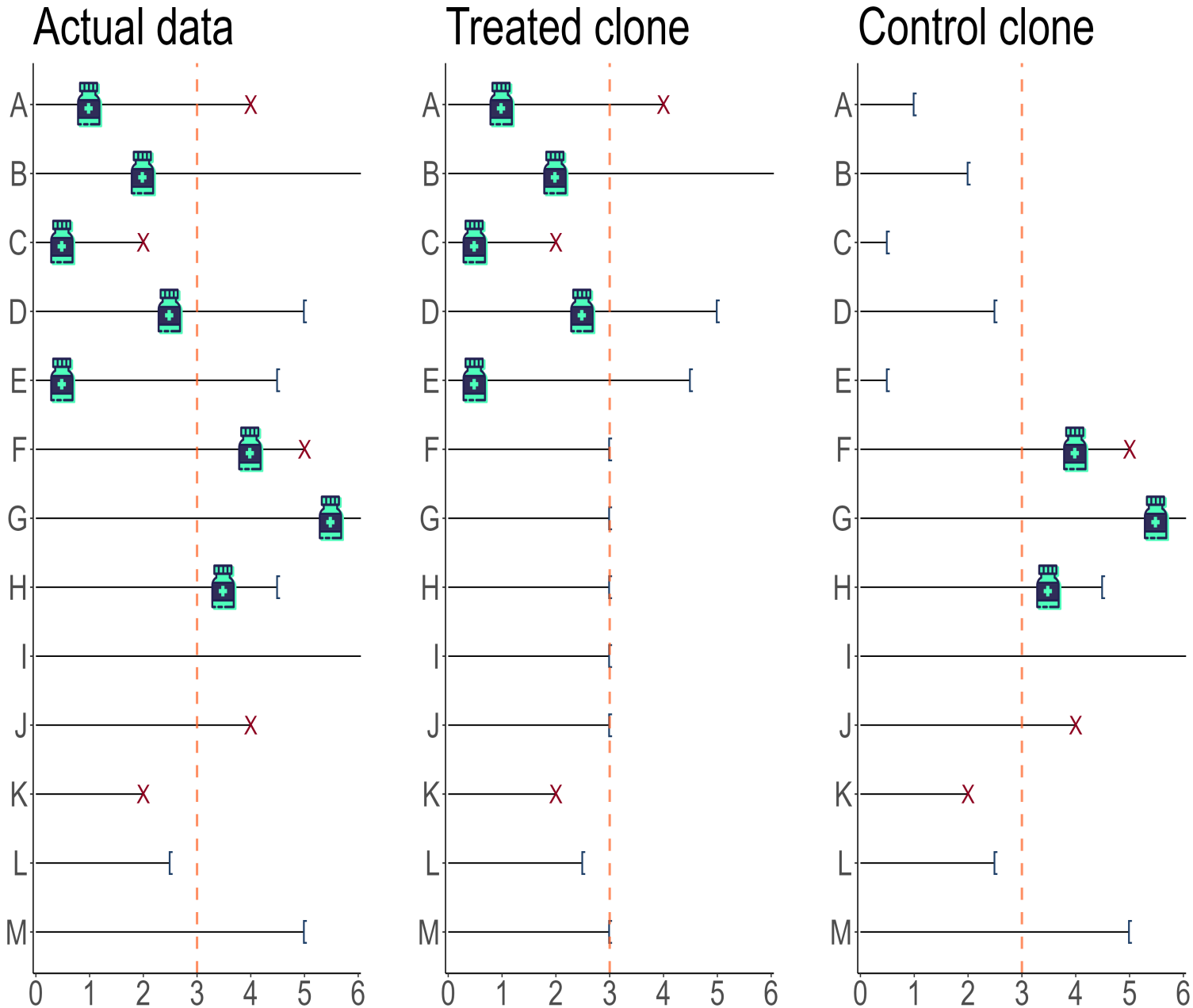
Cloning-censoring-weighting approach



Cloning-censoring-weighting approach



Cloning-censoring-weighting approach



SANTÉ SORBONNE UNIVERSITÉ

3

Frequent Problems and Some Remedies

Patients already under treatment at time zero

Problem of Non-Adherence to the Assigned Strategy

Strategies not defined at time zero

Strategy too rare at time zero

Time zero difficult to define uniquely

Example Without This Issue

- Question: long-term benefit of aspirin on the risk of recurrence after digestive surgery for colorectal cancer.
- Start of follow-up and eligibility coincide: digestive surgery for cancer.

When eligibility is determined by an event that occurs only once, the time zero is easy to determine

Example With This Issue

- Question: benefit of initiating HRT on the risk of breast cancer in menopausal women.
- Start of follow-up?
  - A 52-year-old menopausal woman without breast cancer, and never treated with HRT is eligible.
  - This woman remains eligible at 53, 54, etc., as long as she does not start her treatment (and does not get breast cancer).
  - If she starts HRT, it is the last time she is eligible.

$\Rightarrow$ Which time zero to choose?

Example With This Issue

- Question: benefit of initiating HRT on the risk of breast cancer in menopausal women.
- Start of follow-up?
  - A 52-year-old menopausal woman without breast cancer, and never treated with HRT is eligible.
  - This woman remains eligible at 53, 54, etc., as long as she does not start her treatment (and does not get breast cancer).
  - If she starts HRT, it is the last time she is eligible.

$\Rightarrow$ Which time zero to choose?

When eligibility is determined by a recurring event or a chronic condition, the time zero is not easy to choose

Two valid approaches

- Choose anyway!
  - The first time the eligibility criteria are met
  - Or a time zero chosen at random among the available times

    $\Rightarrow$ Ununbiased, but not necessarily very efficient from a statistical point of view

Two valid approaches

- Choose anyway!
  - The first time the eligibility criteria are met
  - Or a time zero chosen at random among the available times
    $\Rightarrow$ Ununbiased, but not necessarily very efficient from a statistical point of view
- Choose all available time zeros
  - Consider each individual at each eligible moment as a different individual.
  - This approach involves emulating a sequence of nested trials with an increasing time zero.
  - It also accounts for immortal time bias.
  - Requires an appropriate estimation of the variance. (to account for the fact that an individual can be duplicated several times).

A Reference to Start with...

Statistics in medicine (2020)

Matching with time-dependent treatments: a review and look forward

Thomas, Laine E; Yang, Siyun; Wojdyla, Daniel; Schaubel, Douglas E

A Reference to Go Further...

Statistics in medicine (2023)

Causal inference in survival analysis using longitudinal observational data: Sequential trials and marginal structural models

Keogh, Ruth H; Gran, Jon Michael; Seaman, Shaun R; Davies, Gwyneth; Vansteelandt, Stijn

SANTÉ SORBONNE UNIVERSITÉ

4

Step by Step Example

Data presentation

Impact of Initial Treatment

Comparison of "Treatment at All Visits" vs. "No Treatment at Any Visit"

Comparison of the 'Treatment at Least One Visit' Strategy versus the 'No Treatment at Any Visit' Strategy

SANTÉ SORBONNE UNIVERSITÉ

4

Step by Step Example

Data presentation

- Simulated cohort study
- Time-dependent treatment (denoted as A) at three visits (baseline, 1 year, and 2 years)
- Two confounding factors (X and L), one of which (L) is time-dependent and measured at each visit

Variable	Description
`id`	Individual identifier
`visit`	Visit number
`A`	Treatment (`A = 0` or `A = 1`, time-dependent)
`X`	Baseline confounder (binary)
`L`	Time-dependent confounder (continuous)
`T.start`	Start of follow-up (same values as `visit`)
`T.stop`	End of follow-up
`D`	Event indicator at `T.stop`

Data presentation

```r
df <- read.csv("df.csv")
df[1:6, ]
```

```
##   id visit T.start T.stop D X          L A
## 1  1     0       0   1.00 0 1 -0.27124046 0
## 2  1     1       1   1.91 1 1  2.08302877 0
## 3  2     0       0   0.12 1 0 -0.01680837 0
## 4  3     0       0   1.00 0 0 -1.15625842 0
## 5  3     1       1   2.00 0 0  1.65868909 1
## 6  3     2       2   2.12 0 0 -0.64862844 0
```

SANTÉ SORBONNE UNIVERSITÉ

4

Step by Step Example

Objective

The objective of this section is to estimate survival curves under two scenarios: if everyone initiated treatment A at time 0 ($a_0 = 1$) and if no one did ($a_0 = 0$). The corresponding estimands are denoted as:

- $S_{a_0=1}(t) = P(T_{a_0=1} > t)$
- $S_{a_0=0}(t) = P(T_{a_0=0} > t)$

Note that individuals initially exposed are allowed to stop the treatment later, and that individuals initially unexposed are allowed to start the treatment later. These strategies are "analogous to intention-to-treat" (ITT) analyses in randomized clinical trials.

Creating Variables $A0$ and $L0$

For this objective, we only need the baseline exposure and confounders (their values at the first visit). One option is to create a new dataset with only one row per individual, structured as follows:

```
library(dplyr)
df <- df %>%
  group_by(id) %>%
  mutate(A0 = first(A), L0 = first(L))
head(as.data.frame(df))
```

```
##   id visit T.start T.stop D X           L A A0          L0
## 1  1     0       0   1.00 0 1 -0.27124046 0  0 -0.27124046
## 2  1     1       1   1.91 1 1  2.08302877 0  0 -0.27124046
## 3  2     0       0   0.12 1 0 -0.01680837 0  0 -0.01680837
## 4  3     0       0   1.00 0 0 -1.15625842 0  0 -1.15625842
## 5  3     1       1   2.00 0 0  1.65868909 1  0 -1.15625842
## 6  3     2       2   2.12 0 0 -0.64862844 0  0 -1.15625842
```

Propensity Score Analysis I

We aim to estimate marginal Kaplan-Meier survival curves corresponding to the strategies `A0 = 1` for everyone and `A0 = 0` for everyone. This can be achieve using propensity score weighting.

1. Use a logistic model to estimate the weights:

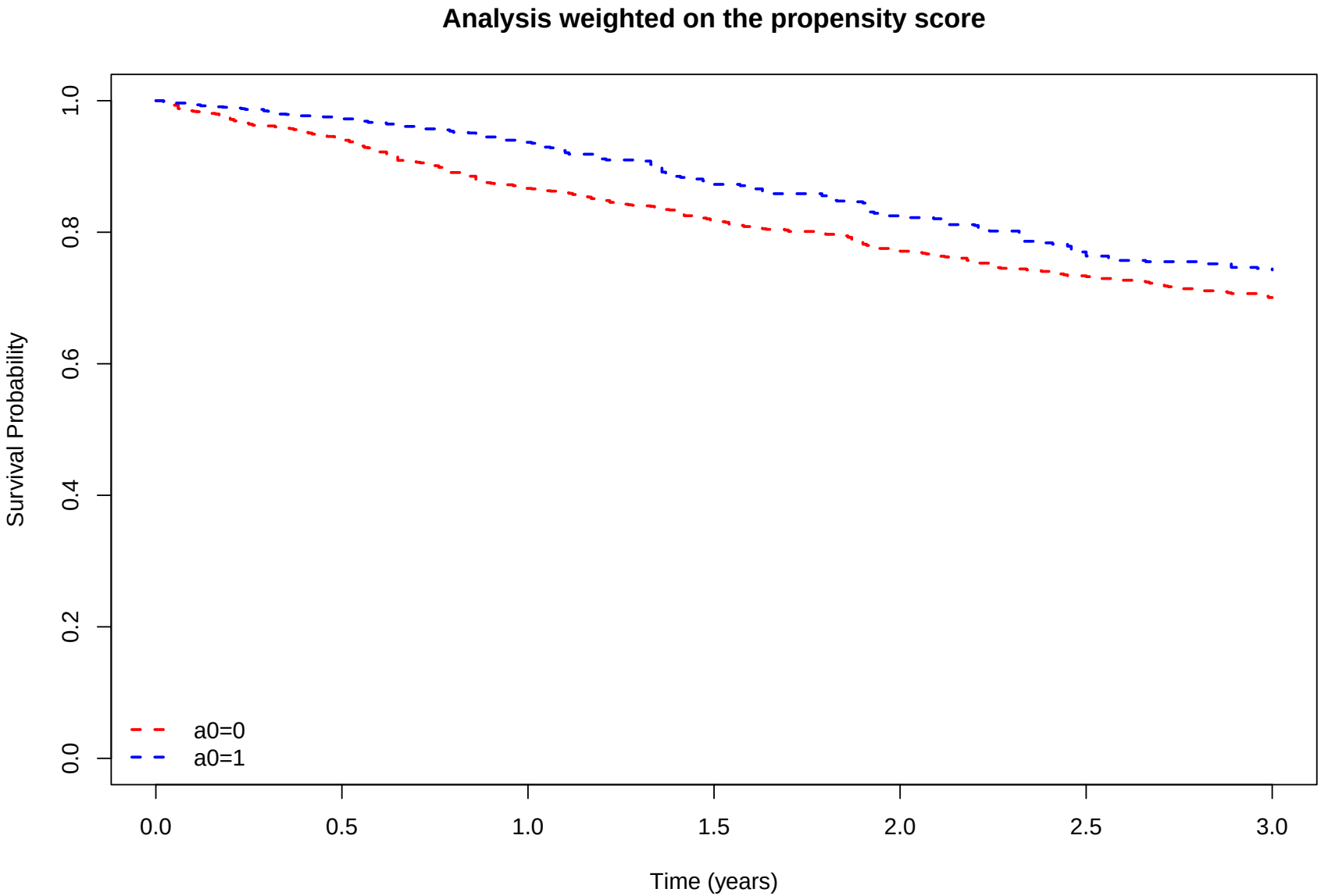
```
mod.treat <- glm(A0 ~ X + L0, data = df[df$T.start == 0, ], family = "binomial")
df$ps <- predict(mod.treat, newdata = df, type = "response")
df$iptw <- (df$A0 == 1)/df$ps + (df$A0 == 0)/(1 - df$ps)
```

Propensity Score Analysis II

2. Perform a weighted Kaplan-Meier analysis based on the initial treatment A0, and plot the estimated survival curves.

```
library(survival)
km.iptw <- survfit(Surv(T.start, T.stop,D) ~ A0, data = df, weights = iptw)
plot(km.iptw, xlab = "Time (years)", ylab = "Survival Probability",
     col = c("red", "blue"), lty = 2, lwd = 2, conf.int = FALSE,
     main = "Analysis weighted on the propensity score")
legend(x = "bottomleft", c("a0=0", "a0=1"),
       col = c("red", "blue"), lty = 2, lwd = 2, bty = "n")
```

Propensity Score Analysis III



SANTÉ SORBONNE UNIVERSITÉ

4

Step by Step Example

Data presentation

Impact of Initial Treatment

Comparison of "Treatment at All Visits" vs. "No Treatment at Any Visit"

Comparison of the 'Treatment at Least One Visit' Strategy versus the 'No Treatment at Any Visit' Strategy

Objective

The objective of this section is to compare survival curves under two treatment strategies:

1. Receive treatment at all visits ($a_0 = a_1 = a_2 = 1$), referred to as "always treated."
2. Receive no treatment at any visit ($a_0 = a_1 = a_2 = 0$), referred to as "never treated."

The corresponding estimands are:

- $S_{a_0=a_1=a_2=1}(t) = P(T_{a_0=a_1=a_2=1} > t)$
- $S_{a_0=a_1=a_2=0}(t) = P(T_{a_0=a_1=a_2=0} > t)$

These strategies are "analogous to per-protocol" (PP) analyses in randomized clinical trials.

Inverse Probability of Censoring Weighting I

For this, we will once again define the groups based on the initial exposure and artificially **censor individuals who deviate from their initial treatment strategy**. Since this artificial censoring is very likely informative, we will use the IPCW approach to account for the time-dependent selection bias introduced by this censoring.

1. Select the individuals initiating treatment at time zero (`df$A0 == 1`), on one hand, and those not initiating treatment at time zero (`df$A0 == 0`), on the other hand. Store these two groups of individuals in different objects named `df.1` and `df.0`, respectively.

```
df.1 <- df[df$A0 == 1, ]
df.0 <- df[df$A0 == 0, ]
```

Inverse Probability of Censoring Weighting II

2. In the `df.1` dataframe, create a new variable indicating whether a person has deviated from the strategy $a_0 = a_1 = a_2 = 1$. Keep all data lines up to and including the line where the individual first deviated from the strategy $a_0 = a_1 = a_2 = 1$, and remove the lines after the one where the individual first deviated. Name this new variable `switchA`. *NB: Later, we will remove the deviation line, but we must keep it for now in order to estimate the artificial censoring weights.*

Inverse Probability of Censoring Weighting III

```r
library(dplyr)
df.1 <- df.1 %>% group_by(id) %>% mutate(
  cumsumA = cumsum(A == 1), ## Cumulative sum of treatment status A = 1
  switchA = if_else(cumsumA == T.start+1, 0, 1), ## Indicator variable of
                                                 ## deviation from status
                                                 ## treatment A=1
  # NB: If treatment A is taken at all visits,
  ## then `cumsumA` must be equal to T.start+1
  switchA = cumsum(switchA) ## Cumulative sum of the indicator variable
                            ## of deviation
) %>% filter(switchA ⩽ 1) ## Removal of individuals where switchA > 1
```

Inverse Probability of Censoring Weighting IV

3. Repeat the entire previous step with the dataframe `df.0`, modifying the code to adapt it to the strategy $a_0 = a_1 = a_2 = 0$.

```
df.0 <- df.0 %>% group_by(id) %>% mutate(
  cumsumA = cumsum(A == 0), ## Cumulative sum of treatment status A = 0
  switchA = if_else(cumsumA == T.start+1, 0, 1), ## Indicator variable of
                                                 ## deviation from status
                                                 ## treatment A=0
  ## NB: If treatment A is never taken at all visits,
  ## then `cumsumA` must be equal to T.start+1
  switchA = cumsum(switchA) ## Cumulative sum of the indicator variable
                            ## of deviation
) %>% filter(switchA ⩽ 1) ## Removal of individuals where switchA > 1
```

Inverse Probability of Censoring Weighting V

4. From the `df.1` dataframe, fit a logistic regression model with `switchA` as the variable to be explained and the variables X and L of the same line as explanatory variables. Include `visit` (`T.start`) as a categorical variable. This model can be fit using all the lines in the dataframe even if some belong to the same individual (this type of model is called "pooled logistic regression"). From this, derive the estimate of the probability of NOT deviating. Note that this model is estimated from individuals who initiated treatment at the first visit only.

```
## Logistic model on the probability of deviating from initial treatment A=1
wt.mod.denom.1 <- glm(switchA ~ as.factor(T.start) + X + L, family = "binomial",
                      data = df.1)
## Probability of not deviating (complementary to the probability of deviating)
df.1$wt.denom <- 1 - predict(wt.mod.denom.1, type = "response", newdata = df.1)
df.1 <- df.1 %>% filter(switchA == 0) ## Remove the lines where `switchA = 1`
## (we are now done with this variable)
```

Inverse Probability of Censoring Weighting VI

6. Use the predictions from step 4) to obtain the artificial censoring weights (name the new variable `wt`). The weight in a given line is the probability that a person remains uncensored (i.e. they have not changed treatment strategy):

$$W(t) = \prod_{k=0}^{t} \frac{1}{P(switchA = 0|X, L_k)}, t = 0, 1, 2$$

```
df.1 <- df.1 %>% group_by(id) %>% mutate(wt = cumprod(1/wt.denom))
```

Inverse Probability of Censoring Weighting VII

7. Repeat steps 4) to 6) with the `df.0` dataframe, then stack the dataframes `df.1` and `df.0` using the `rbind( )` function. Name the new dataframe `dfpp` ('pp' for 'per protocol').

```
## Logistic model on the probability of deviating from initial treatment A=0
wt.mod.denom.0 <- glm(switchA ~ as.factor(T.start) + X + L, family = "binomial",
                      data = df.0)
## Probability of not deviating (complementary to the probability of deviating)
df.0$wt.denom <- 1 - predict(wt.mod.denom.0, type = "response", newdata = df.0)
df.0 <- df.0 %>% filter(switchA == 0)
df.0 <- df.0 %>% group_by(id) %>% mutate(wt = cumprod(1/wt.denom))
## Stacking of the two data.frames
dfpp <- rbind(df.1, df.0)
```

Inverse Probability of Censoring Weighting VIII

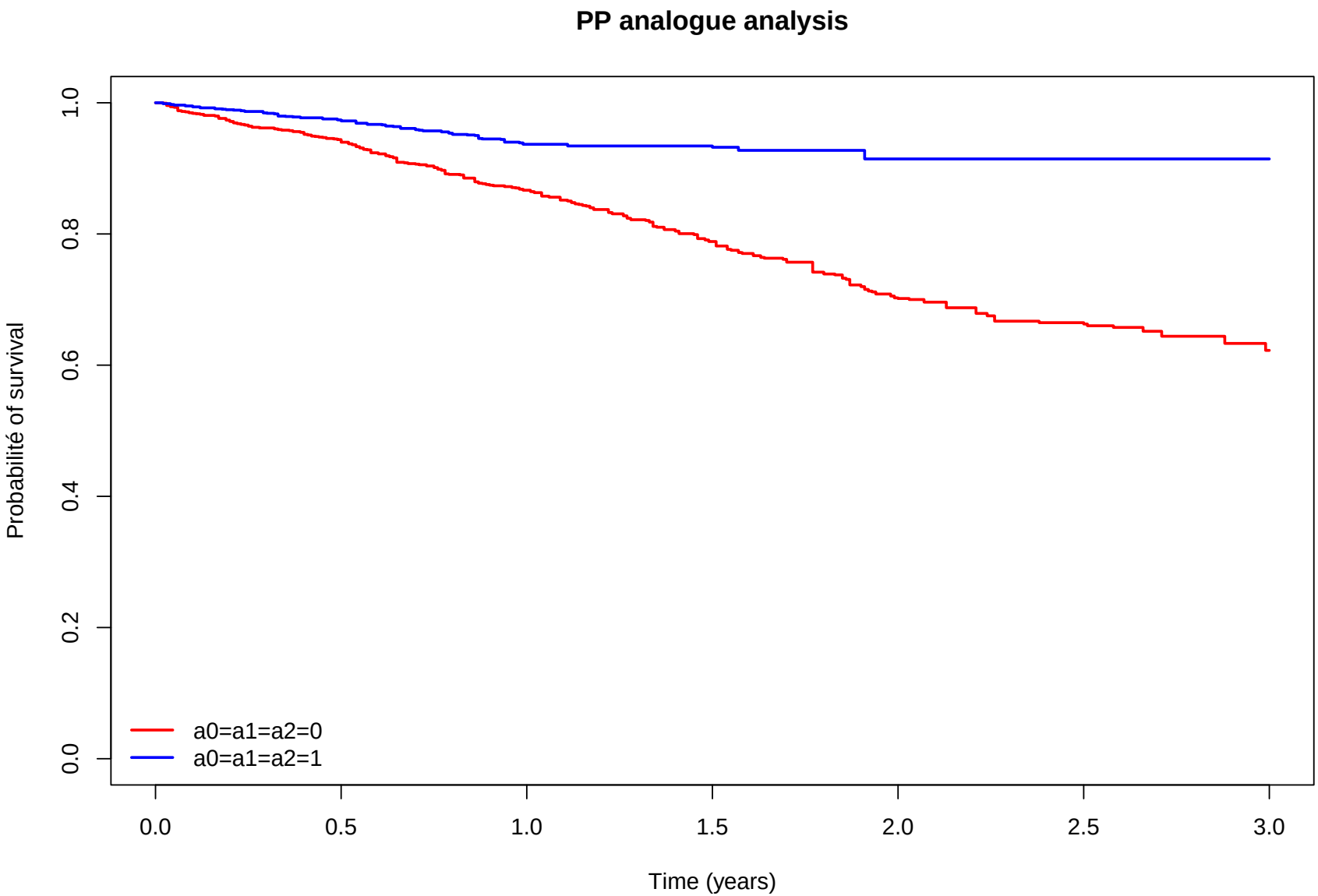
Up to this point, we have selected individuals based on their initial treatment, and artificially censored each individual when they deviated from the initial strategy. We have also estimated the probability of not deviating, and derived artificial censoring weights from this. These weights will allow us to overrepresent the individuals who have not yet deviated in order to "replace" the individuals who have already been censored, in order to compensate for any time-dependent selection bias.

Inverse Probability of Censoring Weighting IX

8. Generate a “combined” weight, multiplying the IPTW weight estimated in the previous part by the IPCW weight estimated in this part, and redo the survival analysis.

```
## Combined weight
dfpp$comb.wt <- dfpp$iptw*dfpp$wt
km.dfpp2 <- survfit(Surv(T.start, T.stop, D) ~ A0, data = dfpp, weights = comb.wt)
plot(km.dfpp2, xlab = "Time (years)", ylab = "Probabilité of survival",
     col = c("red", "blue"), lwd = 2, conf.int = FALSE, main = "PP analogue analysis"
legend(x = "bottomleft", c("a0=a1=a2=0", "a0=a1=a2=1"), col = c("red", "blue"),
       lty = 1, lwd = 2, bty = "n")
```

Inverse Probability of Censoring Weighting X



SANTÉ SORBONNE UNIVERSITÉ

4

Step by Step Example

Objective

The objective of this section is to estimate the survival curves if everyone had received treatment at least once at any of the visits ($a_0 = 1 \cup a_1 = 1 \cup a_2 = 1$) and if no one had received treatment at any visit ($a_0 = a_1 = a_2 = 0$). The corresponding estimands are denoted as :

- $S_{a_0=1 \cup a_1=1 \cup a_2=1}(t) = P(T_{a_0=1 \cup a_1=1 \cup a_2=1} > t)$
- $S_{a_0=a_1=a_2=0}(t) = P(T_{a_0=a_1=a_2=0} > t)$

In this situation, the first three visits constitute a grace period.

Cloning-censoring-weighting I

The first step of the cloning-censoring-weighting approach "mimics" this aspect of the objective.

1. Create two carbon copies of the data contained in `df`, each corresponding to one of the two treatment strategies of interest. Name these copies `clone.1` (for the strategy $a_0 = 1 \cup a_1 = 1 \cup a_2 = 1$) and `clone.0` (for the strategy $a_0 = a_1 = a_2 = 0$). In each of these two dataframes, create a new variable named `Atheo` representing the strategy theoretically followed by each clone (`Atheo = 1` in `clone.1`, and `Atheo = 0` in `clone.0`).

```
clone.1 <- df
clone.1$Atheo <- 1
clone.0 <- df
clone.0$Atheo <- 0
```

Cloning-censoring-weighting II

Technically, the following is similar in every respect to what we have already done in the previous part: we are going to censor the clones when they deviate from their assignated strategy, and then use IPCW.

2. In the `clone.1` dataframe, create a new variable indicating whether a person has deviated from the strategy $a_0 = 1 \cup a_1 = 1 \cup a_2 = 1$. Keep all data lines up to and including the line where the clone first deviated from the strategy $a_0 = 1 \cup a_1 = 1 \cup a_2 = 1$, and remove the lines after the one where the clone first deviated (note that these clones can't deviate from this strategy until the last visit). Name this new variable `switchA`. *NB: Later, we will remove the deviation line, but we must keep it for now in order to estimate the artificial censoring weights.*

Cloning-censoring-weighting III

```r
library(dplyr)
clone.1 <- clone.1 %>% group_by(id) %>% mutate(
  cumsumA = cumsum(A == 1), ## Cumulative sum of treatment status A = 1
  switchA = if_else(cumsumA == 0 & T.start == 2, 1, 0), ## Indicator variable
                                                        ## of deviation
  ## NB: If treatment A is never taken at all visits,
  ## then `cumsumA` will be equal to 0 at T.start == 2
  switchA = cumsum(switchA)
) %>%
  filter(switchA <= 1)
```

Cloning-censoring-weighting IV

3. Repeat the entire previous step with the dataframe `clone.0`, modifying the code to adapt it to the strategy $a_0 = a_1 = a_2 = 0$.

```
clone.0 <- clone.0 %>% group_by(id) %>% mutate(
  cumsumA = cumsum(A == 0), ## Cumulative sum of treatment status A = 0
  switchA = if_else(cumsumA == T.start+1, 0, 1), ## Indicator variable of
                                                 ## deviation from status
                                                 ## treatment A=0
  ## NB: If treatment A is never taken at all visits,
  ## then `cumsumA` must be equal to T.start+1
  switchA = cumsum(switchA) ## Cumulative sum of the indicator variable
  ## of deviation
) %>%
  filter(switchA ⩽ 1) ## Removal of individuals where switchA > 1
```

Cloning-censoring-weighting V

4. From the `clone.1` dataframe, fit a logistic regression model with `switchA` as the variable to be explained and the variables X and L from the same line as explanatory variables. Include visit (`T.start`) as a categorical variable. This model can be fit using all lines in the dataframe. Deduce from it the estimate of the probability of NOT deviating.

```
## Logistic model on the probability of deviating from initial treatment A=1
wt.mod.denom.1 <- glm(switchA ~ as.factor(T.start) + X + L, family = "binomial",
                      data = clone.1)
## Probability of not deviating (complementary to the probability of deviating)
clone.1$wt.denom <- 1 - predict(wt.mod.denom.1, type = "response", newdata = clone.1)
clone.1 <- clone.1 %>% filter(switchA == 0) ## Remove the lines where `switchA = 1`
                                            ## (we are now done with this variable)
```

Cloning-censoring-weighting VI

5. Use the predictions from step 4) to obtain the artificial censoring weights (name the new variable `wt`). The weight in a given line is the probability that a person remains uncensored (i.e. they have not changed treatment strategy):

$$W(t) = \prod_{k=0}^{t} \frac{1}{P(switchA = 0|X, L_k)}, t = 0, 1, 2$$

```
clone.1 <- clone.1 %>% group_by(id) %>% mutate(wt = cumprod(1/wt.denom))
```

Cloning-censoring-weighting VII

6. Repeat steps 4) to 6) with the dataframe `clone.0`, then stack the two dataframes `clone.1` and `clone.0` using `rbind()`. Name the new dataframe `clones`.

```
## Logistic model on the probability of deviating from initial treatment A=0
wt.mod.denom.0 <- glm(switchA ~ as.factor(T.start) + X + L, family = "binomial",
                      data = clone.0)
## Probability of not deviating (complementary to the probability of deviating)
clone.0$wt.denom <- 1 - predict(wt.mod.denom.0, type = "response", newdata = clone.0)
clone.0 <- clone.0 %>% filter(switchA == 0)
clone.0 <- clone.0 %>% group_by(id) %>% mutate(wt = cumprod(1/wt.denom))
## Stacking of the two data.frames
clones <- rbind(clone.1, clone.0)
```

Cloning-censoring-weighting VIII

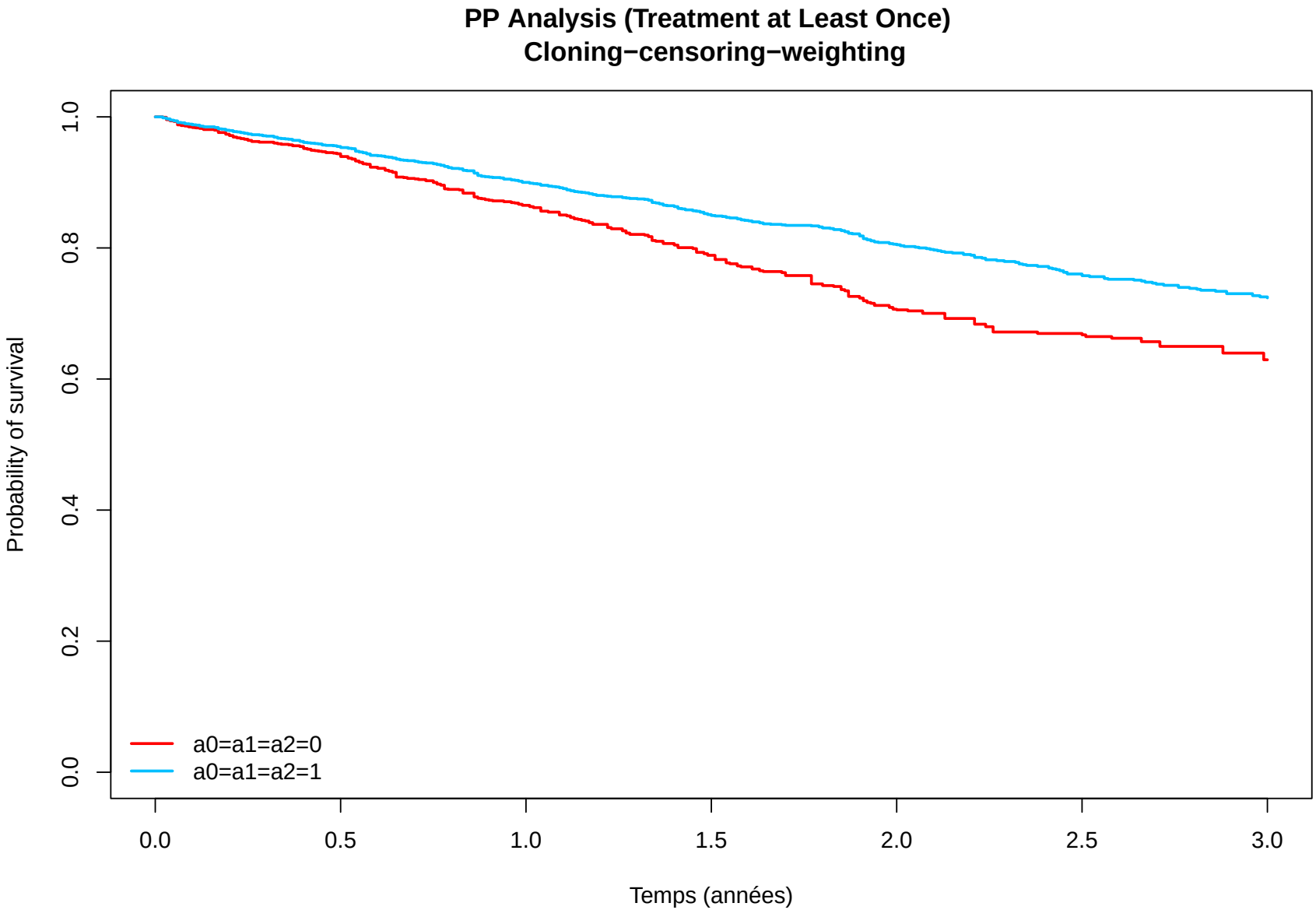
Up to this point, we have cloned all the individuals according to two treatment strategies, and artificially censored each clone when it deviated from the strategy that was assigned to it. We have also estimated the probability of not deviating, and derived artificial censoring weights from this. The removal of the deviation lines results, in the end, in the same selected individuals as in the `df.1` and `df.0` bases. But unlike the previous analysis, the models from which the weights were derived were estimated with all individuals (cloned from the original database). These weights will allow us to overrepresent the clones who have not yet deviated in order to "replace" the clones who have been censored for deviation from the strategy that was assigned to them, in order to compensate for a time-dependent selection bias.

Cloning-censoring-weighting IX

7. Perform a weighted Kaplan-Meier analysis using the `wt` variable, based on the treatment strategy followed by each clone (`Atheo`), using the `clones` dataframe. Plot the estimated survival curves. Obtain the estimated survival probabilities at time 3, as well as the difference in survival. What do you think? Does the analysis require an additional step?

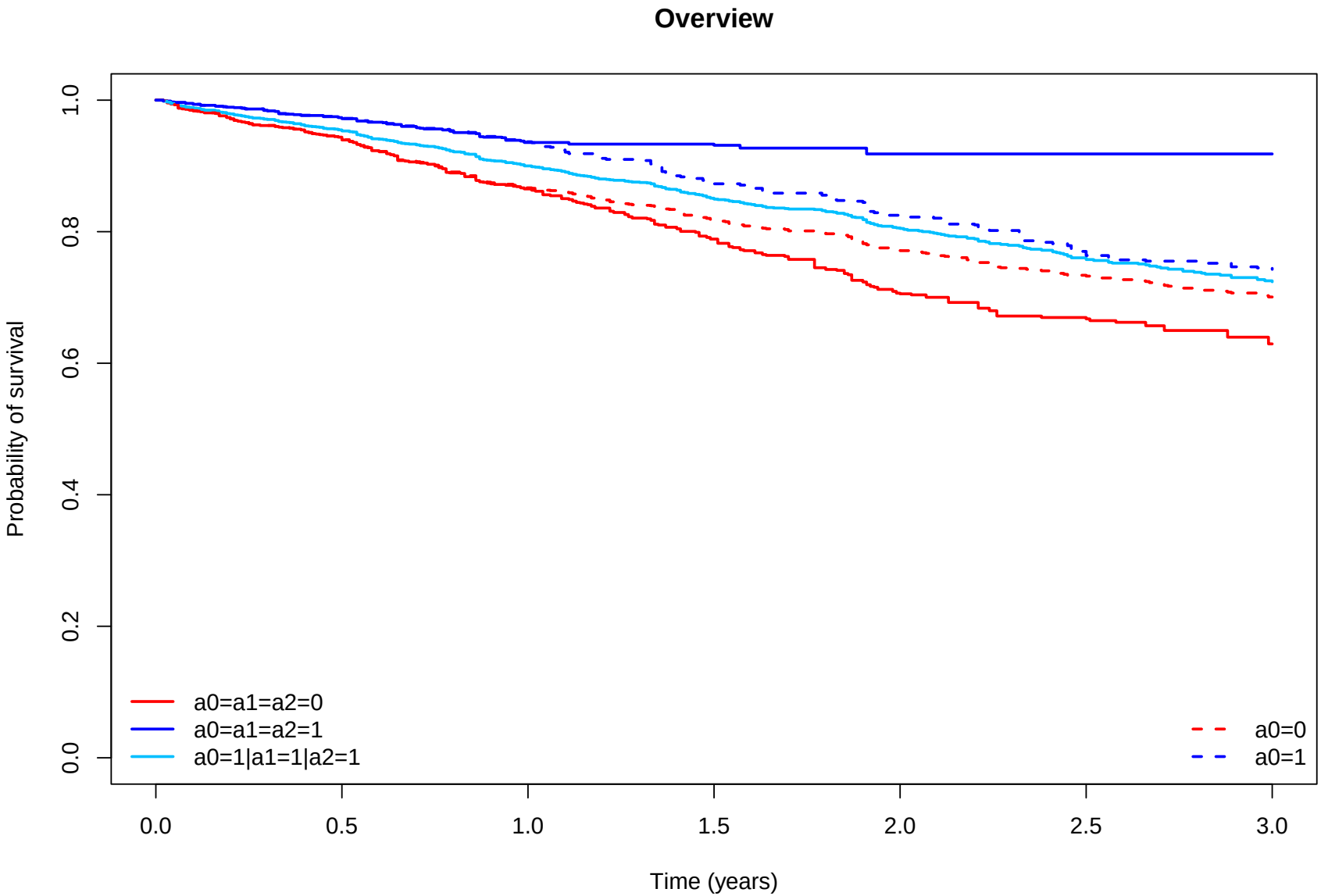
```
km.clones <- survfit(Surv(T.start, T.stop, D) ~ Atheo, data = clones, weights = wt)
plot(km.clones, xlab = "Temps (années)", ylab = "Probability of survival",
     col = c("red", "deepskyblue"), lwd = 2, conf.int = FALSE,
     main = "PP Analysis (Treatment at Least Once) \n Cloning-censoring-weighting")
legend(x = "bottomleft", c("a0=a1=a2=0", "a0=a1=a2=1"),
       col = c("red", "deepskyblue"), lty = 1, lwd = 2, bty = "n")
```

Cloning-censoring-weighting X



Comparison of the 'Treatment at Least One Visit' Strategy versus the 'No Treatment at Any Visit' Strategy

Overview of all answered objectives



Comparison of the 'Treatment at Least One Visit' Strategy versus the 'No Treatment at Any Visit' Strategy

What Did We Learn?

Answering the question “Does treatment improve overall survival?” does not have a single, simple answer. It requires clearly defining the objective of the research, the treatment strategies being compared, and then applying a statistical approach that specifically targets that objective.

SANTÉ
SORBONNE
UNIVERSITÉ

5

Conclusion

What Does Emulating Clinical Trials Allow?

- Makes the research question and the target trial explicit (population, treatment strategies, endpoint).
- Aligns the analytical methods with the research question (the method answers the question of interest).
- Even for complex research questions such as “when to treat” with the cloning-censoring-weighting method.
- Avoids “self-inflicted” biases (immortal time bias, biases related to prevalent users).
- Facilitates the methodological evaluation of observational studies.

What Does Emulating Clinical Trials Not Allow?

- Correct unmeasured confounding biases.
- Requires collecting sufficient information on confounding factors and then using them appropriately.
- The limitations of observational studies remain (but by emulating, we do not add additional biases).
- Impossible to reproduce a closely monitored placebo-controlled trial from observational data.

$\Rightarrow$ Pragmatic trials only

Conclusion

Observational analyses remain 'Option B'!

We use observational data:

- because conducting the target randomized controlled trial is not feasible,
- or while waiting to conduct it,
- or to confirm its results.

THANKS FOR WATCHING!

SANTÉ
SORBONNE
UNIVERSITÉ