Emulating target trials in longitudinal register data with time-to-event outcomes

Thomas Alexander Gerds

University of Copenhagen

Register data

Every citizen in Denmark has a person identification number.

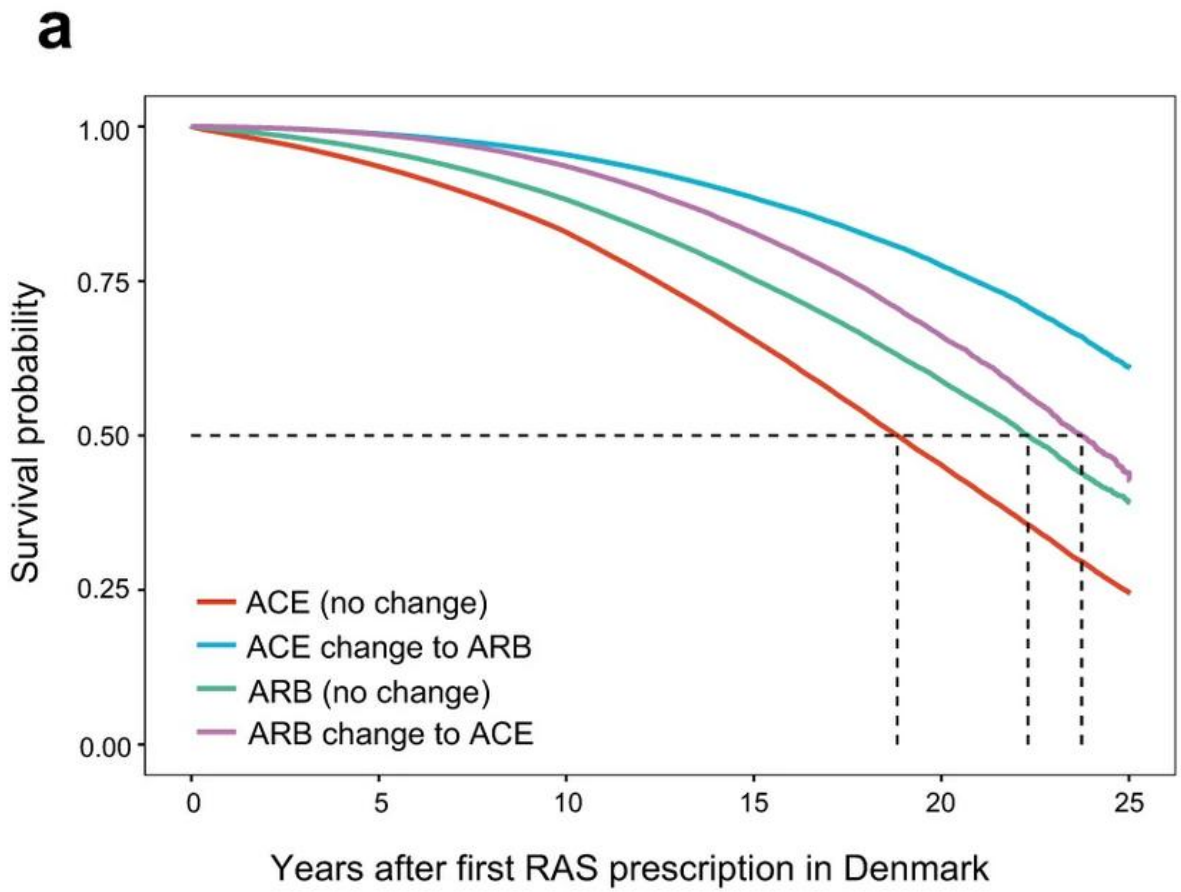
The national records include:

- birth date, immigration/emigration dates, date of death
- type (ATC) and redemption dates of medical prescriptions
- dates and reasons (ICD) for hospital admission
- blood tests, surgery ...
- education, income, household ...

Denmark Statistics: a playground for scientists ...

To be avoided: immortal time bias

From: Optimizing drug selection from a prescription trajectory of one patient



Number at risk

ACE (no change)	549,436	410,475	250,056	92,105	31,282	841
ACE change to ARB	229,216	200,730	136,798	54,393	19,223	493
ARB (no change)	287,488	169,789	82,178	43,741	15,545	190
ARB change to ACE	48,054	43,710	34,552	19,865	6,659	46

A wrong conclusion which could have been avoided

"Prescription trajectories can provide novel insights into how individuals' drug use change over time, identify suboptimal or futile prescriptions and suggest initial treatments different from first line therapies. Observations of this kind may also be important when updating treatment guidelines."[1]

[1] Aguayo-Orozco et al. NPJ digital medicine, 4(1):150, 2021.

A wrong conclusion which could have been avoided

"Prescription trajectories can provide novel insights into how individuals' drug use change over time, identify suboptimal or futile prescriptions and suggest initial treatments different from first line therapies. Observations of this kind may also be important when updating treatment guidelines."[1]

No question was asked!

The result was obtained by data mining!?

[1] Aguayo-Orozco et al. NPJ digital medicine, 4(1):150, 2021.

The Cox regression model with time-dependent covariates

The outcome-specific hazard rate at time t depends on the treatment and covariate history:

$$\lambda(t|\bar{A}(t), \bar{L}(t)) = \lambda_0(t) \exp\{\beta \bar{A}(t) + \gamma \bar{L}(t)\}$$

The hazard ratio e^{β} can be estimated via maximum partial likelihood, Poisson regression or conditional logistic regression (nested case control design).

The Cox regression model with time-dependent covariates

The outcome-specific hazard rate at time t depends on the treatment and covariate history:

$$\lambda(t|\bar{A}(t), \bar{L}(t)) = \lambda_0(t) \exp\{\beta \bar{A}(t) + \gamma \bar{L}(t)\}$$

The hazard ratio e^{β} can be estimated via maximum partial likelihood, Poisson regression or conditional logistic regression (nested case control design).

Can anyone ask a meaningful question such that e^{β} is the ~~answer~~ target parameter?

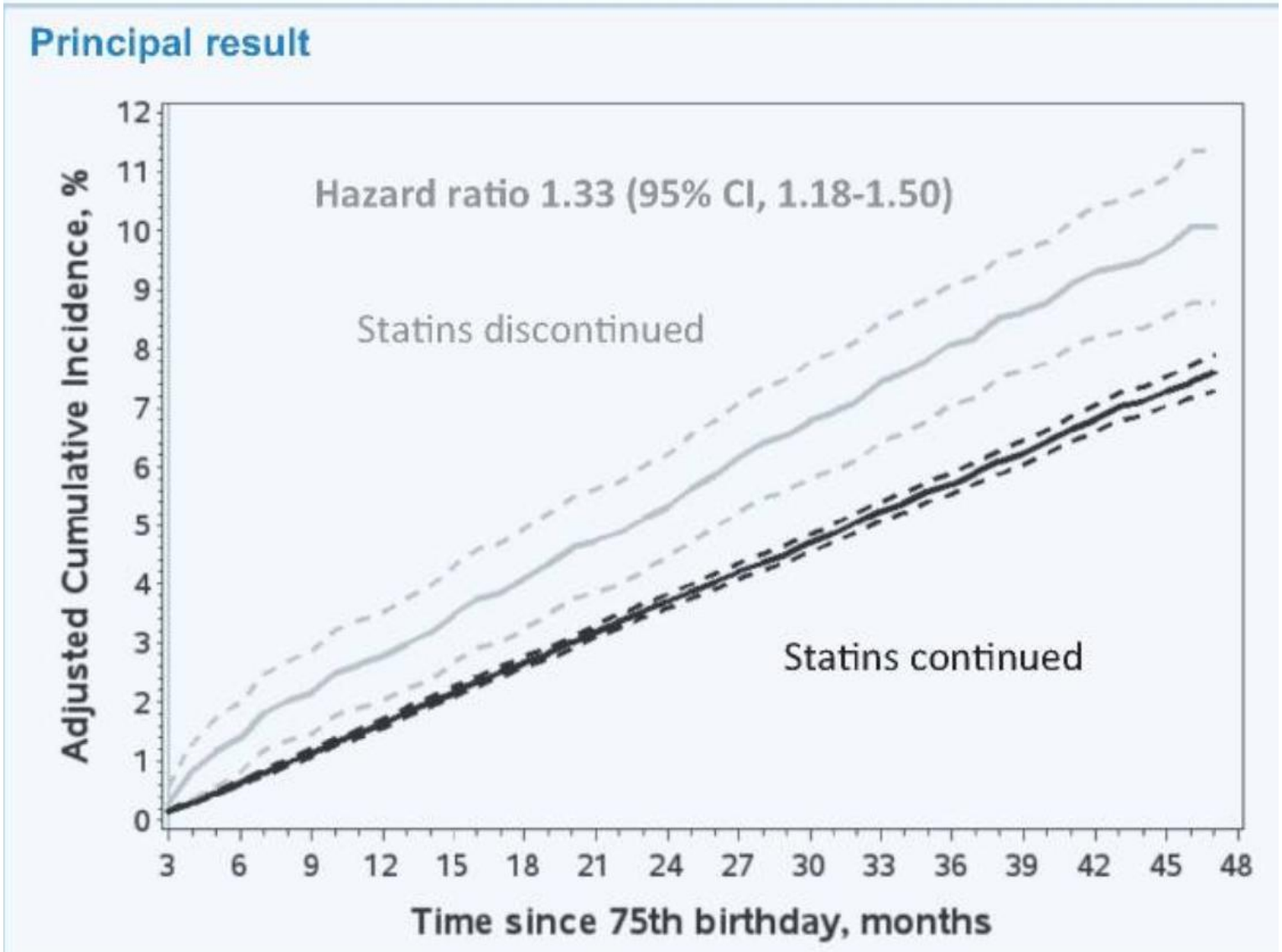
Another example: using Robin's g-methods

Cardiovascular effect of discontinuing statins for primary prevention at the age of 75 years: a nationwide population-based cohort study in France [2]

Aims	The effect of statin discontinuation on cardiovascular outcomes
Enrollment	All statin users who turned 75 in 2012-2014
Methods?	The hazard ratio comparing statin discontinuation with continuation was estimated using a marginal structural model adjusting for both baseline and time-varying covariates
Results??	Statin discontinuation was associated with a 33% (HR: 1.33) increased risk of admission for cardiovascular event in 75-year-old primary prevention patients.

[2] Giral et al. European heart journal, 40(43):3516–3525, 2019.

Adjusted cumulative incidence: Igitt?



The french study is half-baked

The authors define a proper framework for survival analysis:

- population
- time zero
- time to cardiovascular outcome

The authors use a marginal structural model and inverse probability weighting to deal with time-varying confounding.

But:

- What is the interpretation of the adjusted cumulative incidence curve and the adjusted hazard ratio?
- When is the time of statins discontinuation?
- What about the competing risk of death due to other causes?

Avoidable flaws in observational analyses[3]

The increasing availability of large healthcare databases is fueling an intense debate on whether real-world data should play a role in the assessment of the benefit-risk of medical treatments.

In many observational studies, for example, statin users were found to have a substantially lower risk of cancer than in meta-analyses of randomized trials. While such discrepancies are often attributed to a lack of randomization in the observational studies, they may be explained by flaws that can be avoided by

explicitly emulating a target trial.

[3] Dickerman, . . . , Hernan. Nature medicine, 25(10):1601–1606, 2019.

Emulating a target trial in register data[4]

Step 1 enrollment date (time zero for survival analysis, inclusion/exclusion)

Step 2 the target trial protocols dictate the treatment at any time during followup

Step 3 define the target parameters as contrasts of the expected outcomes – had all subjects ~~been randomized~~ adhered to the protocols

Step 4 estimation of target parameters

Step 5 communication of results and limitations

[4] Following the roadmap of targeted learning (van der Laan et al.)

The register data look like this:

id	sex	age at enrollment	date of enrollment
1	woman	74	2019-07-02
2	man	69	2016-01-13
3	woman	88	2021-11-27

Treatment:

id	date	drug
1	2019-07-02	GLP1-RA
1	2019-10-08	GLP1-RA
1	2021-01-07	GLP1-RA
2	2016-01-13	SGLT2i
2	2016-08-27	SGLT2i
3	2021-11-28	GLP1-RA
3	2021-12-03	SGLT2i

Hospital diagnoses:

id	hospital admission date	diagnosis
2	2018-02-15	hypertension
2	2018-09-03	hypertension
3	2022-01-02	arterial fibrillation

Concomitant medical treatment:

id	redemption date	drug
2	2018-12-24	ACE/ARP
3	2022-02-11	NOAC
3	2022-08-10	NOAC

Outcome

id	cardiovascular disease	date of death	end of followup
1	NA	NA	2024-09-17
2	NA	2019-03-03	2024-09-17
3	2023-02-07	2023-02-07	2024-09-17

Emulating trials in register data

In register studies there is often no natural control group of patients who are not exposed to one of the drugs.

- Matching can be used (e.g., exposure density sampling).
- A medical diagnosis (biomarker above threshold) can be used.
- Comparative effectiveness study: active comparator arms.

Emulating trials in register data

In register studies there is often no natural control group of patients who are not exposed to one of the drugs.

- Matching can be used (e.g., exposure density sampling).
- A medical diagnosis (biomarker above threshold) can be used.
- Comparative effectiveness study: active comparator arms.

Example: We include all diabetes patients in Denmark who initiated medical treatment with one of the following anti-diabetic drugs: GLP1-RA, SGLT2i between 2015 and 2022:

- Time zero is at (or 30 days after) the first purchase of the drug
- We follow the patients through the registers until comorbidity, death, emigration, or 2024, whatever comes first.

Treatment assignment in the target trial

The protocols of the target trial dictate the treatment(s) given at any time (doctor visit) during the target trial period.

Notation:

$$\pi^*(t \mid L(t-), A(t-))$$

Examples:

Protocol	Type of intervention	$\pi^*(t \mid L(t-), A(t-))$
Never treat	Static	1
Always treat	Static	0
Treat for 2 years	Static	$1\{t \leq 2\}$
Treat if $L(t-) > \xi$	Dynamic	$1\{L(t-) > \xi\}$
Treat with probability 0.8	Stochastic	0.8
If $L(t-) > \xi$ treat with probability 0.8	Stochastic, Dynamic	$1\{L(t-) > \xi\}0.8$

Treatment assignment in the target trial

Example of treatment regimens:

- **Protocol 1** Patients should use GLP1-RA continuously for 3 years and not intensify with SGLT2i
- **Protocol 2** Patients should use SGLT2i continuously for 3 years and not intensify with GLP1-RA

Protocol 1 assigns 100% probability for GLP1-RA and 0% probability for SGLT2i:

$$\pi^{1*}(t) = (1, 0).$$

Protocol 2 is defined similarly:

$$\pi^{2*}(t) = (0, 1)$$

The target parameter (aka the estimand)

The analysis of the emulated target trial estimates the absolute risks of the outcome(s) if hypothetically all patients had followed the treatment protocols (per-protocol effects).

Example: 3-year risk of cardiovascular disease under π^{j*} and differences thereof:

$$\mathrm{P}_{\pi^{1*}}(Y(3) = 1) - \mathrm{P}_{\pi^{2*}}(Y(3) = 1)$$

The analyst uses the information of the time-varying covariates (comorbidity, co-medicine) to achieve a good compromise between:

- the available data
- the desire of the investigators
- the causal assumptions: *positivity*, *sequential coarsening at random (NUC)*, *consistency*.

Estimation on ~~discrete~~ discretized time scale

new R-package 'rtmle' (work in progress)

```
library(rtmle)
library(data.table)
x <- rtmle_init(intervals=6,
                name_id="id",
                name_outcome="cvd_death",
                name_competing="Dead",
                name_censoring="Censored",
                censored_label="censored")
baseline_data(x) <- data.table(id=1:3,
                               sex=c("woman","man","woman"),
                               age=c(74,69,88),
                               start_followup_date=c("2019-07-02","
    2016-01-13","2021-11-27"))
treatment_data(x) <- list("GLP1RA"=data.table(id=c(1,1,1,1,2),
                                              date=c("2019-07-02",
                                                     "2019-10-08",
                                                     "2021-01-07",
                                                     "2021-11-28")),
                          "SGLT2i"=data.table(id=c(2,3,3),
                                              date=c("2016-01-13",
                                                     "2016-08-27",
                                                     "2021-12-03")))
timevar_data(x) <- list("hypertension"=data.table(id=c(2,2),
                                                  date=c("2018-02-15",
                                                         "2018-09-03"),
                                                  value=c(1,1)),
                        "af"=data.table(id=3,date="2022-01-02",value=1))
```

new R-package 'rtmle' (work in progress)

```
x$long_data$outcome_data <- data.table(id = 1,
                                       date = "2023-02-07")
x$long_data$censored_data <- data.table(id = c(1,2,3),
                                        date = rep("2024-09-17",3))
x$long_data$competingrisk_data <- data.table(id = c(2,3),
                                             date = c("2019-03-03",
                                                      "2023-02-07"))
# protocol for emulated trial
protocol(x) <- list(name = "always GLP1RA",
                    treatment_variables = c("GLP1RA","SGLT2i"),
                    intervention = c(1,0))
target(x) <- list(name = "Outcome_risk",
                  strategy = "additive",
                  estimator = "tmle",
                  estimands = 3,
                  protocols = "always GLP1RA")
# 3 years followup in six 6 months long intervals
prepare_data(x) <- list(treatment_variables = "GLP1RA",
                        reset = TRUE,
                        intervals = seq(0,30.45*36,30.45*6))
# data in wide format
x$prepared_data[["always GLP1RA"]]$data
```

The register data forced onto a discretized time scale

```
       id    sex   age hypertension_0  af_0 GLP1RA_0 SGLT2i_0 Censored_1 cvd_death_1 Dead_
   <char> <char> <num>          <num> <num>    <num>    <num>     <fctr>       <num>  <nu
1:      1  woman    74              0     0        1        0 uncensored           0
2:      2    man    69              0     0        0        1 uncensored           0
3:      3  woman    88              0     0        0        0 uncensored           0
   hypertension_1  af_1 GLP1RA_1 SGLT2i_1 Censored_2 cvd_death_2 Dead_2 hypertension_2  af
            <num> <num>    <num>    <num>     <fctr>       <num>  <num>          <num> <nu
1:              0     0        1        0 uncensored           0      0              0
2:              0     0        0        0 uncensored           0      0              0
3:              0     1        0        1 uncensored           0      0              0
   GLP1RA_2 SGLT2i_2 Censored_3 cvd_death_3 Dead_3 hypertension_3  af_3 GLP1RA_3 SGLT2i_3
      <num>    <num>     <fctr>       <num>  <num>          <num> <num>    <num>    <num>
1:        0        0 uncensored           0      0              0     0        0        0
2:        0        0 uncensored           0      0              0     0        0        0
3:        0        0 uncensored           0      1             NA    NA       NA       NA
   Censored_4 cvd_death_4 Dead_4 hypertension_4  af_4 GLP1RA_4 SGLT2i_4 Censored_5 cvd_dea
       <fctr>       <num>  <num>          <num> <num>    <num>    <num>     <fctr>       <
1: uncensored           0      0              0     0        1        0 uncensored
2: uncensored           0      0              0     0        0        0 uncensored
3:       <NA>           0     NA             NA    NA       NA       NA       <NA>
   Dead_5 hypertension_5  af_5 GLP1RA_5 SGLT2i_5 Censored_6 cvd_death_6
    <num>          <num> <num>    <num>    <num>     <fctr>       <num>
1:      0              0     0        1        0 uncensored           0
2:      0              1     0        0        0 uncensored           0
3:     NA             NA    NA       NA       NA       <NA>           0
```

Introduction to the notation

Uncensored data, one interval $[0, t_1]$ [5]

$$X = (L_0, A_0, Y_1)$$

Observed likelihood

$$\mathrm{P}X = \underbrace{\mathrm{P}_{Y_1|A_0,L_0}}_{F_1} \underbrace{\mathrm{P}_{A_0|L_0}}_{\pi_0} \underbrace{\mathrm{P}_{L_0}}_{H_0}$$

Likelihood in the target trial

$$P^* = F_1 \, \pi_0^* \, H_0$$

[5] $Y_1 = 1\{T \leq t_1\})$

Robins g-methods[6]

Data	$\{(L_{0i}, A_{0i}, Y_{1i})\}_{i=1}^n = \{X_i\}_{i=1}^n$
Outcome regression model	$\hat{F}_1$
Empirical distribution	$\hat{H}_0 = \frac{1}{n}\sum_{i=1}^n L_{0,i}$
Empirical distribution	$\mathbb{P}_n = \frac{1}{n}\sum_{i=1}^n X_i$
Propensity of treatment	$\hat{\pi}_0$

Estimator 1: g-formula

$$\hat{\Psi}_g^* = \hat{F}_1 \pi_0^* \hat{H}_0$$

Estimator 2: Inverse probability weighting

$$\hat{\Psi}_{\text{IPTW}}^* = \frac{\mathbb{P}_n \pi_0^*}{\hat{\pi}_0}$$

[6] Robins (1986) Mathematical modelling, 7(9-12):1393–1512

Longitudinal setting [7]

Discretized time scale:

$$[0 \cdots\cdots\cdots t_1 \cdots\cdots\cdots t_2]$$

Data for two time intervals:

$$X = (L_0, A_0, Y_1, L_1, A_1, Y_2).$$

The joint probability distribution:

$$\mathrm{P}X = \mathrm{P}_{Y_2|A_1,L_1,Y_1,A_0,L_0}\mathrm{P}_{A_1|L_1,Y_1,A_0,L_0}\mathrm{P}_{L_1|Y_1,A_0,L_0}\mathrm{P}_{Y_1|A_0,L_0}\mathrm{P}_{A_0|L_0}\mathrm{P}_{L_0}$$

[7] no censoring, no competing risks, $Y_1 = 1\{T \leq t_1\}$, $Y_2 = 1\{T \leq t_2\}$

Longitudinal setting [7]

Discretized time scale:

$$[0 \cdots\cdots\cdots t_1 \cdots\cdots\cdots t_2]$$

Data for two time intervals:

$$X = (L_0, A_0, Y_1, L_1, A_1, Y_2).$$

The joint probability distribution:

$$\mathrm{P}X = \underbrace{\mathrm{P}_{Y_2|A_1,L_1,Y_1,A_0,L_0}}_{F_2} \underbrace{\mathrm{P}_{A_1|L_1,Y_1,A_0,L_0}}_{\pi_1} \underbrace{\mathrm{P}_{L_1|Y_1,A_0,L_0}}_{H_1} \underbrace{\mathrm{P}_{Y_1|A_0,L_0}}_{F_1} \underbrace{\mathrm{P}_{A_0|L_0}}_{\pi_0} \underbrace{\mathrm{P}_{L_0}}_{H_0}$$

[7] no censoring, no competing risks, $Y_1 = 1\{T \leq t_1\}$, $Y_2 = 1\{T \leq t_2\}$

Longitudinal causal inference on discretized time scale

Uncensored data, two intervals [8]

$$X = (L_0, A_0, Y_1, L_1, A_1, Y_2)$$

Observed likelihood

$$\mathrm{P}X = \underbrace{\mathrm{P}_{Y_2|A_1,L_1,Y_1,A_0,L_0}}_{F_2}\underbrace{\mathrm{P}_{A_1|L_1,Y_1,A_0,L_0}}_{\pi_1}\underbrace{\mathrm{P}_{L_1|Y_1,A_0,L_0}}_{H_1}\underbrace{\mathrm{P}_{Y_1|A_0,L_0}}_{F_1}\underbrace{\mathrm{P}_{A_0|L_0}}_{\pi_0}\underbrace{\mathrm{P}_{L_0}}_{H_0}$$

Likelihood in the target trial

$$\mathrm{P}^* = F_2\,\pi_1^*\,H_1\,F_1\,\pi_0^*\,H_0$$

[8] Changes of A(t) and L(t) in last interval $(t_{1,t2}]$ are ignored.

Robins g-methods (continued)

Estimator 1: g-formula

$$\hat{\mathrm{P}}^*_g = \hat{F}_2 \pi^*_1 \hat{H}_1 \hat{F}_1 \pi^*_0 \hat{H}_0$$

Estimator 2: Inverse probability weighting

$$\hat{\mathrm{P}}^*_{\mathsf{IPTW}} = \frac{\mathbb{P}_n \, \pi^*_0 \pi^*_1}{\hat{\pi}_0 \hat{\pi}_1}$$

Robins g-methods (continued)

Estimator 1: g-formula

$$\hat{\mathrm{P}}^*_g = \hat{F}_2 \pi^*_1 \hat{H}_1 \hat{F}_1 \pi^*_0 \hat{H}_0$$

Estimator 2: Inverse probability weighting

$$\hat{\mathrm{P}}^*_{\mathsf{IPTW}} = \frac{\mathbb{P}_n \, \pi^*_0 \pi^*_1}{\hat{\pi}_0 \hat{\pi}_1}$$

Estimator 3: Iterative conditional expectations AKA Sequential regression

$$\begin{aligned}
\mathrm{E}[Y_2] &= \mathrm{E}[\mathrm{E}[Y_2|L_0]] \\
&= \mathrm{E}[\mathrm{E}[\mathrm{E}[Y_2|L_0, A_0]|L_0]] \\
&= \mathrm{E}[\mathrm{E}[\mathrm{E}[\mathrm{E}[Y_2|L_0, A_0, L_1]|L_0, A_0]|L_0]] \\
&= \mathrm{E}[\mathrm{E}[\mathrm{E}[\mathrm{E}[\mathrm{E}[Y_2|L_0, A_0, L_1, A_1]|L_0, A_0, L_1]|L_0, A_0]|L_0]]
\end{aligned}$$

Iterative conditional expectations: discretized time

Estimator 3: Robins (1999), Bang & Robins (2005)

$$\mathrm{E}[Y_2] = \mathrm{E}[\mathrm{E}[\mathrm{E}[\mathrm{E}[\underbrace{\mathrm{E}[Y_2|L_0, A_0, L_1, A_1]}_{Q_2(L_0,A_0,L_1,A_1)}|L_0, A_0, L_1]|L_0, A_0]\,|L_0]]$$
$$\underbrace{\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad}_{Q_1(L_0,A_0)}$$

Step 1 $\hat{Q}_2$: Regress Y_2 on L_0, A_0, L_1, A_1

Step 2 Integrate $\hat{Q}_2$ with respect to π_1^*

Step 3 $\hat{Q}_1$: Regress result of Step 2 on L_0, A_0, L_1, A_1

Step 4 Integrate $\hat{Q}_1$ with respect to π_0^*

Step 5 Average with respect to $\hat{H}_0$

Longitudinal targeted minimum loss based estimator

A motivation for the roadmap of targeted learning is the problem that the various nuisance parameter regression models could be misspecified.

The targeted minimum loss based estimator can be consistent even if some of the nuisance parameter models are misspecified.

For longitudinal data analysis, the sequential regression estimator is moved closer to the unknown truth by sequential updating with

"clever covariates"

which depend on the inverse propensity of treatment weights.[9]

[9] van der Laan & Gruber (2012) The international journal of biostatistics

Longitudinal targeted minimum loss based estimator

Under the usual causal assumptions and if the convergence rate of the estimators of the nuisance parameters is sufficiently fast we can estimate the target parameter:

$$\psi : \mathcal{M} \mapsto \mathbb{R},$$

where ψ is a suitably smooth functional defined on a set of probability measures, at the $\sqrt{n}$-rate

$$\sqrt{n}(\psi(\hat{\mathrm{P}}^*_{\mathsf{LTMLE}}) - \psi(\mathrm{P})) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \mathsf{IF}_{\mathrm{P}}(X_i) + o_{\mathrm{P}}(1)$$

Longitudinal targeted minimum loss based estimator

The efficient influence function for the target parameter (two intervals):

$$\mathrm{IF}_{\mathrm{P}}(X) = (Y_2 - Q_2)\frac{\pi_1^*\pi_0^*}{\pi_1\pi_0} + (Q_2 - Q_1)\frac{\pi_0^*}{\pi_0} + Q_1 - \psi(\mathrm{P})$$

Targeting algorithm:

- Step 1 Initial estimators: $\hat{\pi}_1, \hat{\pi}_0$
- Step 2 $\hat{Q}_2$: Regress Y_2 on L_0, A_0, L_1, A_1
- Step 3 $\hat{Q}_2^*$: TMLE update: Loss function and parametric fluctuation model to solve the current part of the efficient influence function
- Step 4 Integrate $\hat{Q}_2^*$ with respect to π_1^*
- Step 5 Regress result of Step 4 on A_0, L_0
- Step 6 ...

The many shades of censoring

Censoring clash

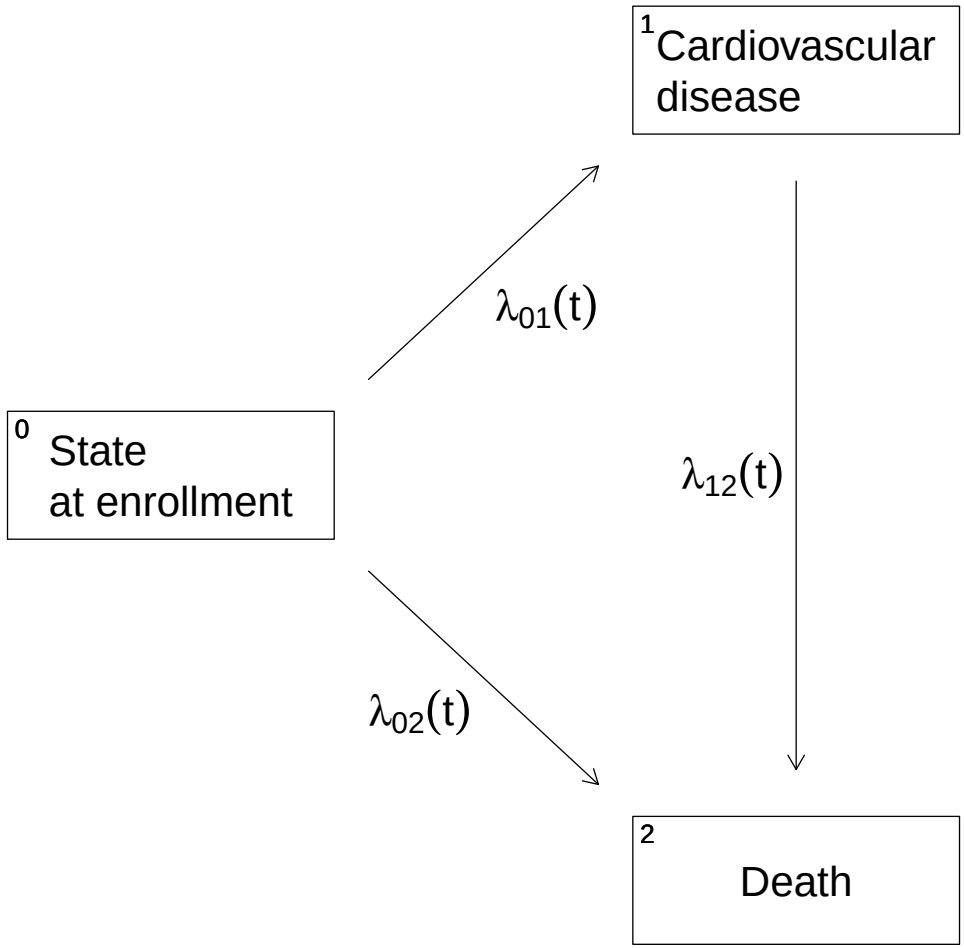
Censored data are the existence of survival analysis!

We define the target parameters in the uncensored world and then use the censored data to estimate them.

The target parameters do not depend on the censoring mechanism.

When investigators decide to censor patients at intermediate events (such as side effects of a drug or change of treatment) they implicitly change the protocol of the target trial and the interpretation of the target parameter!

Multi-state model



Multi-state model: at any time the patient is in one of the states.
Censoring is not a state!

Statistical methods (statins discontinuation paper)

A weighted Cox proportional hazard model was used to estimate the hazard ratio of statin discontinuation vs continuation, controlling for baseline and time-varying confounding. Weighting creates a pseudo-population, in which at each time the probability of being treated is unrelated to [redacted] time-varying confounders. Also, compared conventional multivariate Cox models, weighted Cox proportional hazard models are more flexible in that they do not rely on the assumptions of proportional hazards and no [redacted] In order to [redacted] potentially informative censoring at treatment resumption, the patient's initial weight was multiplied by the inverse probability of remaining uncensored up until month t, based on the patient's covariate history. The resulting weights are called "inverse probability of treatment and censoring weights" (IPTCW). Under the assumptions of no unmeasured confounding, positivity, correct model specification, and consistency, this approach estimates the parameter of a marginal structural model, which can be interpreted as average causal effect of treatment

Certain potential biases (statins discontinuation paper)

With the study design adopted, time-related bias, in particular immortal time bias, was avoided by . . .

. . . Potential bias due to treating competing events as censoring events as in our analyses is also small: in the main analysis, only 2.7% of patients were censored for death and, . . .

. . . The percentage of competing events was therefore much lower than 10%, the critical value reported in the literature (Austin et al., 2016).

Certain potential biases (statins discontinuation paper)

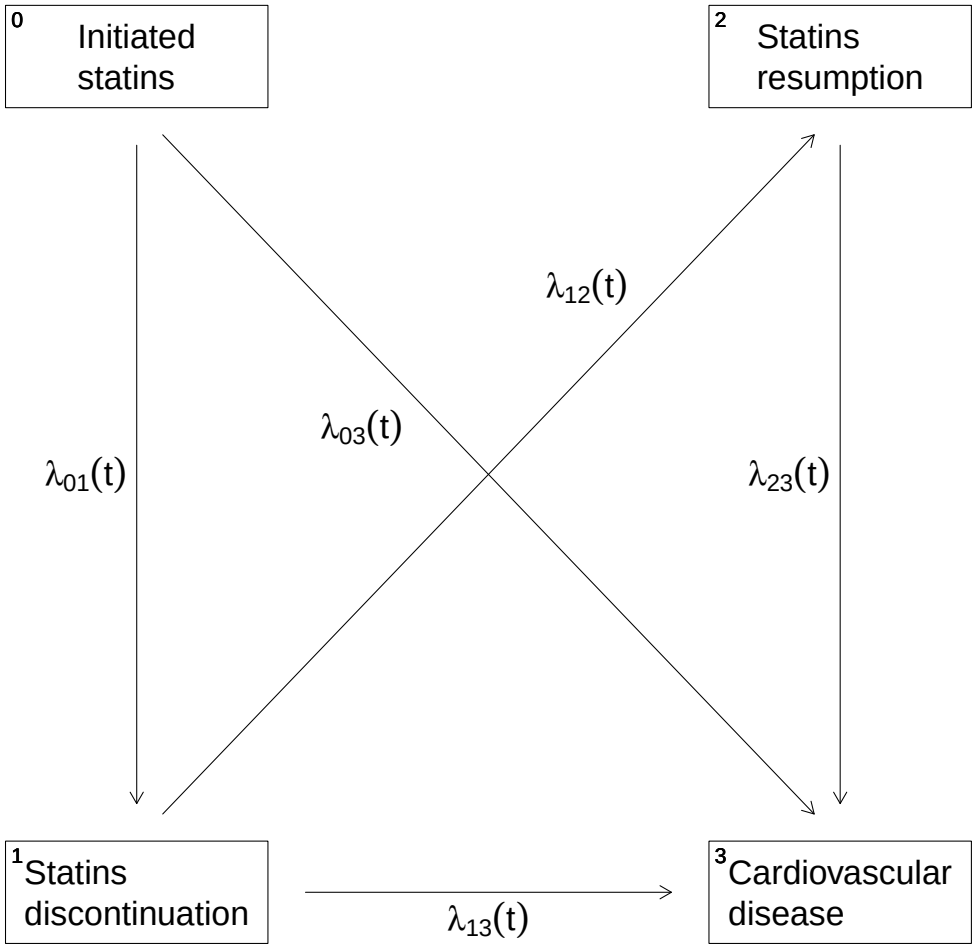
With the study design adopted, time-related bias, in particular immortal time bias, was avoided by . . .

. . . Potential bias due to treating competing events as censoring events as in our analyses is also small: in the main analysis, only 2.7% of patients were censored for death and, . . .

. . . The percentage of competing events was therefore much lower than 10%, the critical value reported in the literature (Austin et al., 2016).

Censor the nonsense!

Multi-state model



At any time each patient is in one of the states!?
Death?
Statins resumption?

Investigator defined censoring affects the interpretation

Review of statins discontinuation paper (Giral et al.):

- the target parameter is a hazard ratio [10]
- no explicit definition of target trial protocols
- censoring at statins resumption
- censoring at death

Possible protocols for the target trial:

Protocol 1 During the next 5 years, discontinue statins treatment with a probability of 50% at any day (or doctor visit). Do not resume statins treatment. Stay alive!

Protocol 2 Continue statins treatment for 5 years. Stay alive!

---

[10] Hernan (2010) The hazards of hazard ratios. Epidemiology.

Censoring and competing risks in LTMLE

The order of the data in wide format with censoring and competing risks:

$$X = (L_0, A_0, C_1, Y_1, D_1, L_1, A_1, C_2, Y_2, D_2, \ldots, Y_K)$$

- probabilities are modelled conditional on the history (variables to the left). We denote by G_k the conditional distribution of being uncensored.
- competing risk probabilities are not modelled but D_k is used to remove patients from being at risk. (However, the presence of competing risks is important for the clinical interpretation!)
- when outcome event Y_k and censoring C_k occur in same interval, then evidently outcome is not censored!

LTMLE algorithm

1. Prepare algorithm by estimating all propensity score and censoring models conditional on current history.
2. Fit model for outcome at last time point conditional on past
3. Get targeted predictions for previous time point via TMLE update step with fluctuation model and clever covariates:

$$\prod_{k=0}^{K} \frac{\pi_k^* 1\{C_k = 0\}}{\hat{\pi}_k \hat{G}_k}$$

4. Fit model with targeted predictions as outcome conditional on past ...

LTMLE algorithm

1. Prepare algorithm by estimating all propensity score and censoring models conditional on past.
2. Fit model for outcome at last time point conditional on past
3. Get targeted predictions for previous time point via TMLE update step with fluctuation model and clever covariates:

$$\texttt{gbounds:} \prod_{k=0}^{K} \frac{\pi_k^* 1\{C_k = 0\}}{\max(0.01, \hat{\pi}_k \hat{G}_k)}$$

1. Fit model with targeted predictions as outcome conditional on past ...

Truncation of weights

Statins discontinuation paper

Table S6. Sensitivity analysis: Impact of weight truncation on the estimated effect

Truncation percentiles 1-p, p[a]	HR of statin discontinuation vs continuation (95% CI)	Estimated weights	
		Mean (SD)	Minimum/ maximum
0, 100[b]	1.43 (1.23-1.66)	1.01 (0.30)	0.01/77.63
0.1, 99.9	1.37 (1.21-1.56)	1.00 (0.19)	0.07/3.65
0.5, 99.5	1.33 (1.18-1.50)	1.00 (0.15)	0.23/2.14
1, 99	1.31 (1.17-1.46)	1.00 (0.12)	0.43/1.66
5, 95	1.28 (1.15-1.42)	1.00 (0.04)	0.91/1.11
10, 90	1.27 (1.15-1.41)	0.99 (0.03)	0.94/1.05
25, 75	1.27 (1.14-1.41)	0.99 (0.02)	0.97/1.01
50, 50[c]	1.26 (1.14-1.40)	1.00 (0.00)	1.00/1.00

- [a] Weights were truncated by resetting the value of the weights greater (less) than the p (1-p) percentile to the value of the p (1-p) percentile.
- [b] No weight truncation.
- [c] Constant weights, i.e. no adjustment.

From the statins discontinuation paper . . .

This table illustrates the tradeoff between bias and variance: with increasing weight truncation, confidence intervals were shorter, but bias increased

Assuming estimation without weight truncation is unbiased.

In this case, it can be reasonably argued to report the result with the weights truncated at the 0.5% and 99.5% percentiles, on the basis of centering of the weights at a value of 1 and the reduction in the 1/minimum and maximum weights.

The gbounds issue

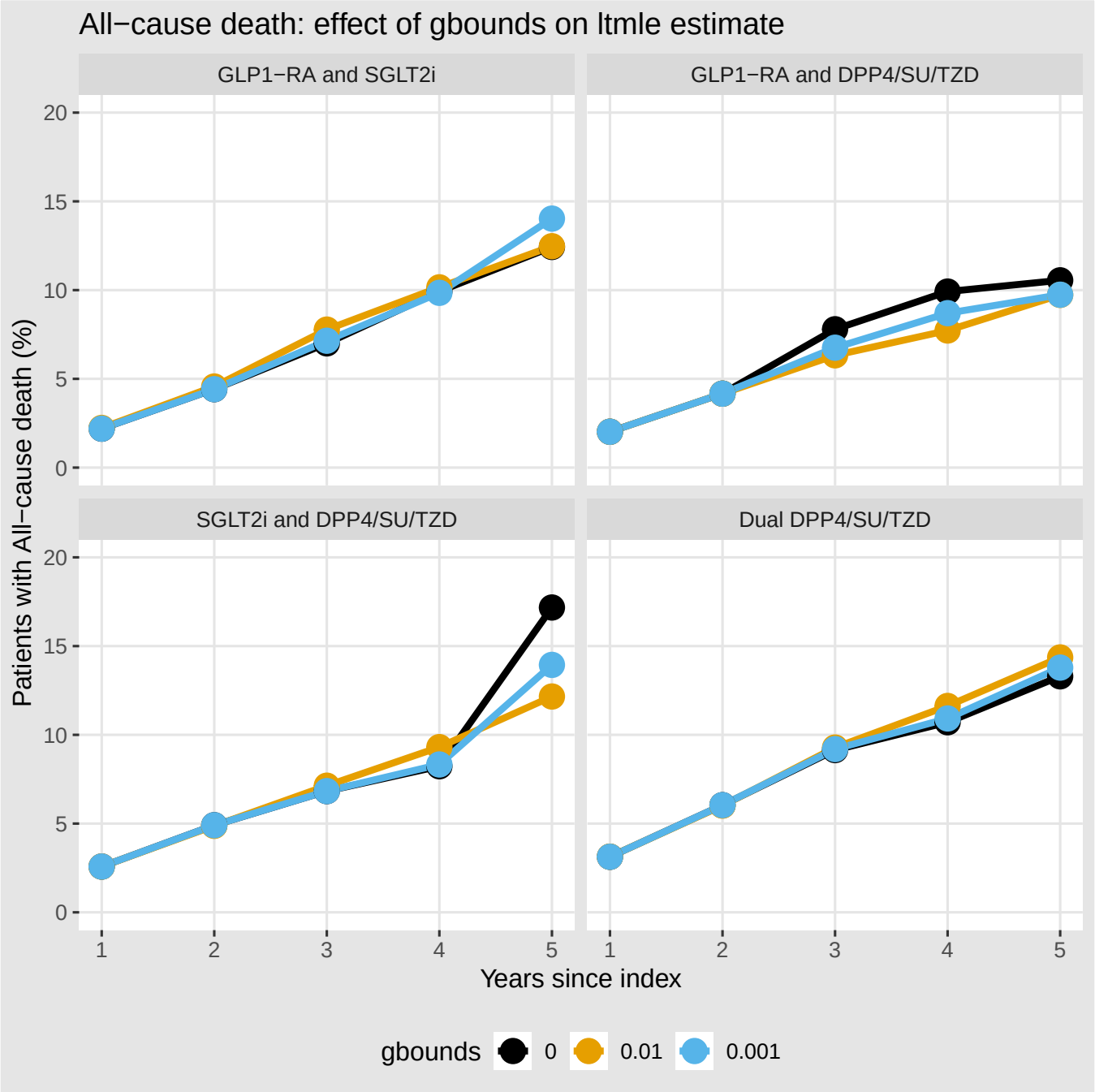
The default behavior of the R-package `ltmle` makes sure that (the product of) the estimated propensity scores and censoring probabilities cannot become smaller than (by default)

$$\texttt{gbounds} = 0.01.$$

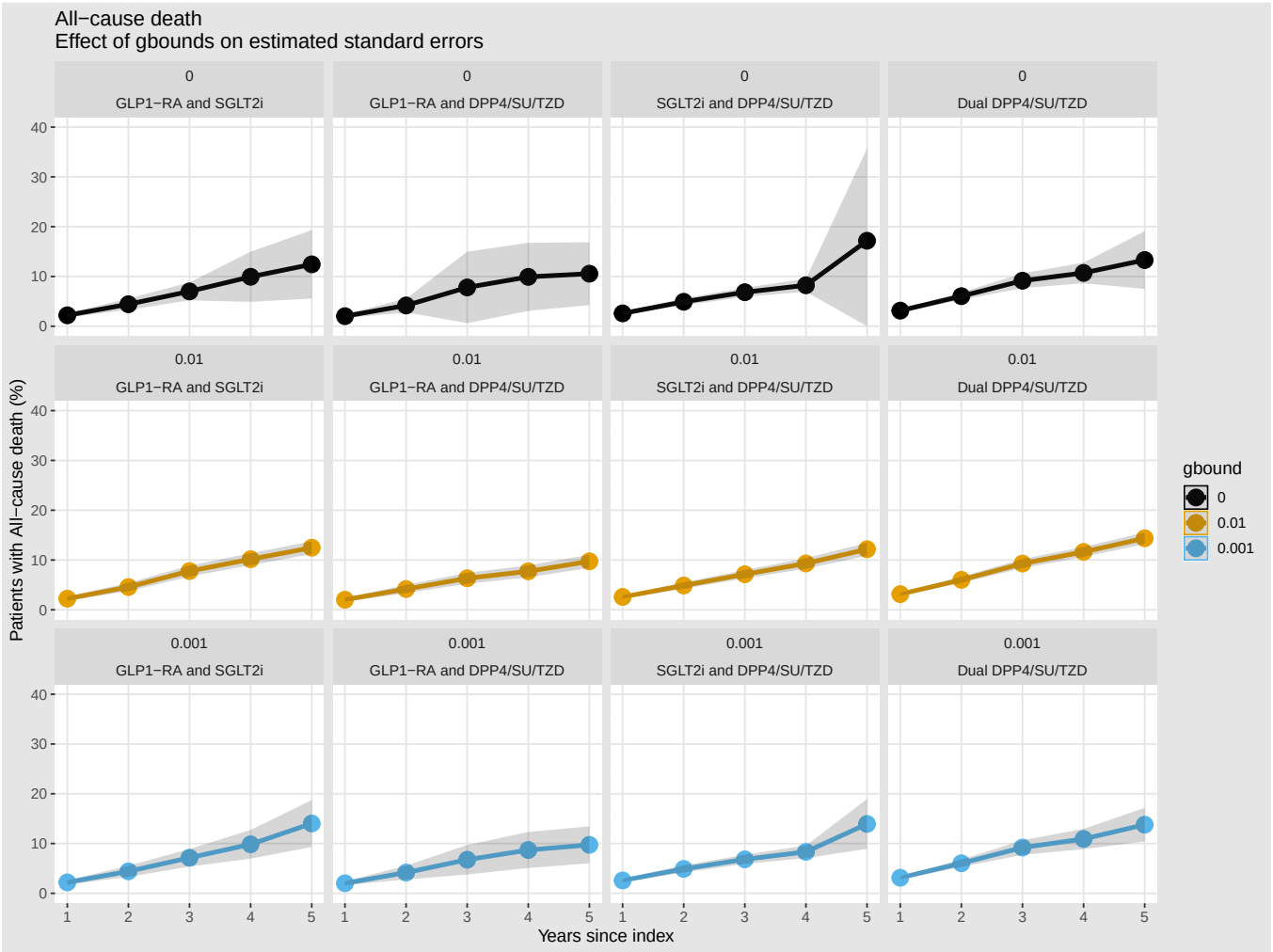
There is no good reason for us to use that particular value, or any other value, hence we subsequently set it to `0`.

The `gbounds` argument affects both the point estimate via the clever weights of the tmle-update step BUT even more the influence function based plug-in estimate of the standard errors.

Effect of gbounds on LTMLE point estimate



Effect of gbounds on standard errors (via influence function)



This ~~is~~ could be the last page

- The target trial helps to separate what we estimate from how we estimate it
- Hazard ratios are not causal and have no directly clinical interpretation, in particular not when there are competing risks
- The choice of which intermediate events are treated as censoring should be reflected in the protocol because it changes the interpretation of the result (see Robins & Finkelstein, 2000)
- Truncation of weights: a very powerful hyperparameter or simply a huge problem?
- Modeling of the process history of multicategorical treatment processes
- Superlearning treatment/comorbidity process histories
- Continuous time version of longitudinal targeted minimum loss based estimation (Rytgaard et al, 2022, Annals of Statistics)

~~Censoring~~ Intervening on intermediate events

"More precisely, we will report four different causal analyses."[11]

The first analysis

Compare mortality in the two treatment arms with death censored only by loss to or end of follow-up.

The second analysis

Regard a subject as dependently censored by the minimum of time to loss to follow-up and time to treatment crossover.
This analysis attempts to estimate what the survival curves would have been if the possibility of crossover to the other treatment arm after the development of PCP had been eliminated from the treatment protocol.

[11] Robins & Finkelstein (2000). Biometrics 56, pages 779–788

~~Censoring~~ Intervening on intermediate events

The third analysis

Regard subjects as dependently censored at the minimum of time to loss to follow-up, time to treatment crossover, and time to voluntarily stopping therapy (for nonmedically related reasons).
For public health purposes, this analysis may be preferred since it attempts to estimate the survival benefit had no subject voluntarily stopped their assigned therapy without medical indication.

The fourth analysis

Regard subjects as censored at the minimum of time to loss to follow-up, crossover, or stopping therapy for any reason.
In this analysis, we are attempting to compare the survival if all subjects were forced to stay on their assigned therapy.
Even if we could successfully estimate this parameter, it would only be of public health relevance if, as discussed above, the toxicities that led to medically indicated termination of therapy could be ameliorated with appropriate palliative therapy.

Cheap subsampling bootstrap confidence interval[12]

1. Draw without replacement $m < n$ times from the observations to obtain a subsample $\mathcal{D}_m^*$. Estimate $\Psi(P)$ on this subsample to obtain $\hat{\Psi}_m^*$ of $\Psi(P)$. Repeat this $B \geq 1$ times to obtain the resamples $\hat{\Psi}_{(m,1)}^*, \ldots, \hat{\Psi}_{(m,B)}^*$. Calculate also $\hat{\Psi}_n$ based on the full sample $\mathcal{D}_n$.
2. Calculate $S = \sqrt{\frac{1}{B}\sum_{b=1}^{B}\left(\hat{\Psi}_{(m,b)}^* - \hat{\Psi}_n\right)^2}$ and obtain the confidence interval

$$\mathcal{I}_{(m,n)} = \left(\hat{\Psi}_n - t_{B,1-\alpha/2}\sqrt{\frac{m}{n-m}}S, \hat{\Psi}_n + t_{B,1-\alpha/2}\sqrt{\frac{m}{n-m}}S\right), \tag{1}$$

where $t_{B,1-\alpha/2}$ is the $1-\alpha/2$ quantile of a t - distribution with B degrees of freedom.

[12] Work by Johan Ohlendorff (2024) based on the general approach of Henry Lam (2022, A Cheap Bootstrap Method for Fast Inference)