On censoring and truncation in target trial emulation

Thomas Alexander Gerds

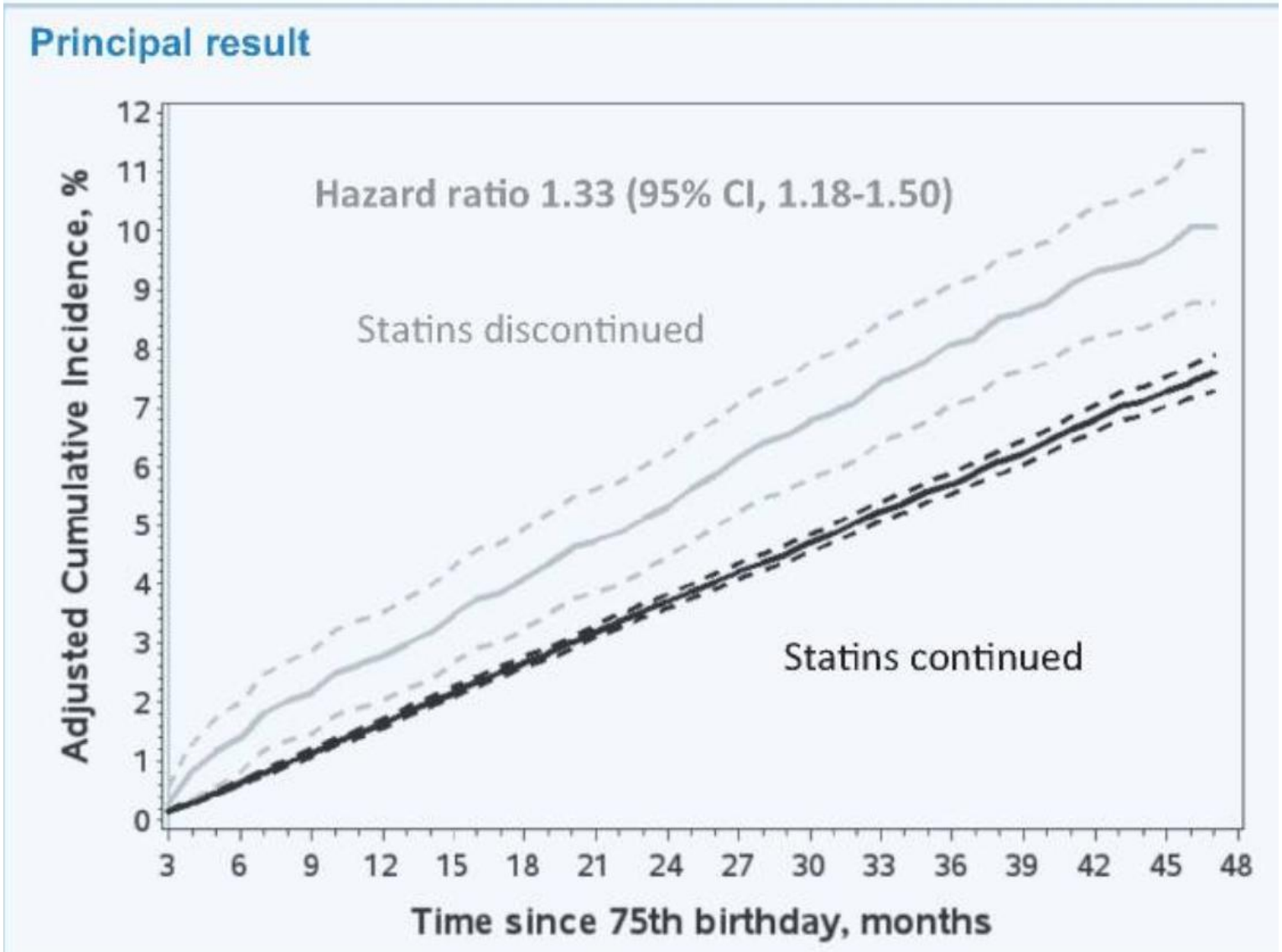
Example

Cardiovascular effect of discontinuing statins for primary prevention at the age of 75 years: a nationwide population-based cohort study in France [1]

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.

[1] 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?

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!

~~Censoring~~ Intervening on intermediate events

"More precisely, we will report four different causal analyses."[2]

The first analysis

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

$\Rightarrow$ survival

The second analysis

Regard a subject as dependently censored by the minimum of time to loss to follow-up and time to treatment crossover.

$\Rightarrow$ survival if the possibility of crossover to the other treatment arm had been eliminated from the treatment protocol.

[2] 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).

$\Rightarrow$ survival if no subject voluntarily stopps 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.

$\Rightarrow$ survival if all subjects were forced to stay on their assigned therapy, i.e., if (hypothetically) the toxicities that lead to medically indicated termination of therapy could be ameliorated with appropriate palliative therapy.

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)

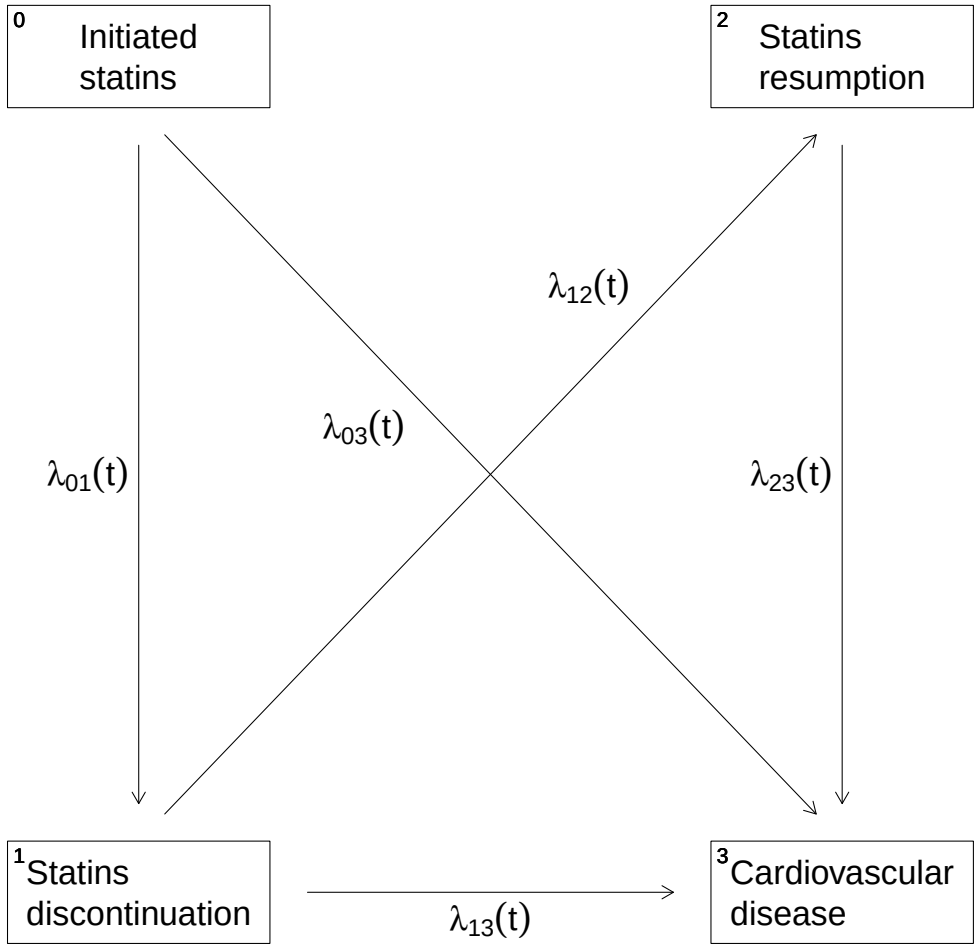
With the study design adopted, time-related bias, in particular immortal time bias, was avoided by . . .

. . . 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
- 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!

Censoring and competing risks in LTMLE

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

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

- C_k: Censoring is multivariate (end of follow-up, loss to follow-up, treatment cross-over, voluntarily treatment stop/break, medical indication for treatment stop/break).
- To predict censoring probabilities, we need to collect (time-updated) predictors for the different reasons of censoring.
- Censoring other than <end of and loss to followup> are part of the (stochastic) interventions.

Trial emulation: the definition of protocols is challenging

For the design of a target trial the estimation-related terms (~~dependent censoring~~, ~~artificial censoring~~) have to be translated to interventions.

The protocols must dictate the treatment in order to achieve well-defined estimands.

Trial emulation: the definition of protocols is challenging

For the design of a target trial the estimation-related terms (~~dependent censoring~~, ~~artificial censoring~~) have to be translated to interventions.

The protocols must dictate the treatment in order to achieve well-defined estimands.

Example:

Challenges for the protocol of an emulated target trial:

What is the probability of a medical indication for treatment stop/pause if, hypothetically, a diabetes patient who in the real world initatied treatment DPP4, had initiated GLP1-RA?

How likely is it a break, a stop, a cross-over?

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 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 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

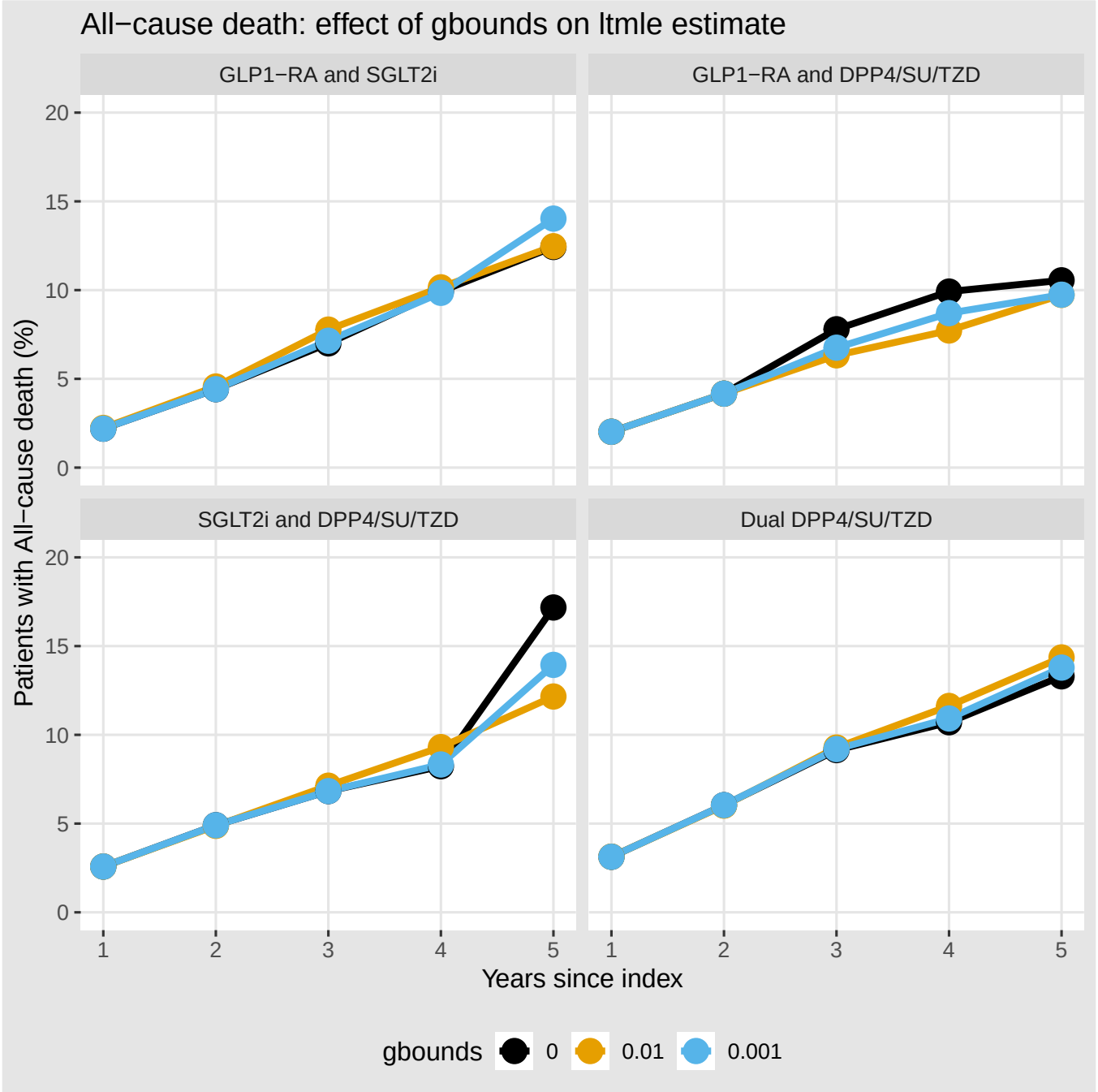
The default behavior of the R-package `ltmle` makes sure that (the product of) the estimated propensity scores and censoring probabilities cannot become smaller than (by default)

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

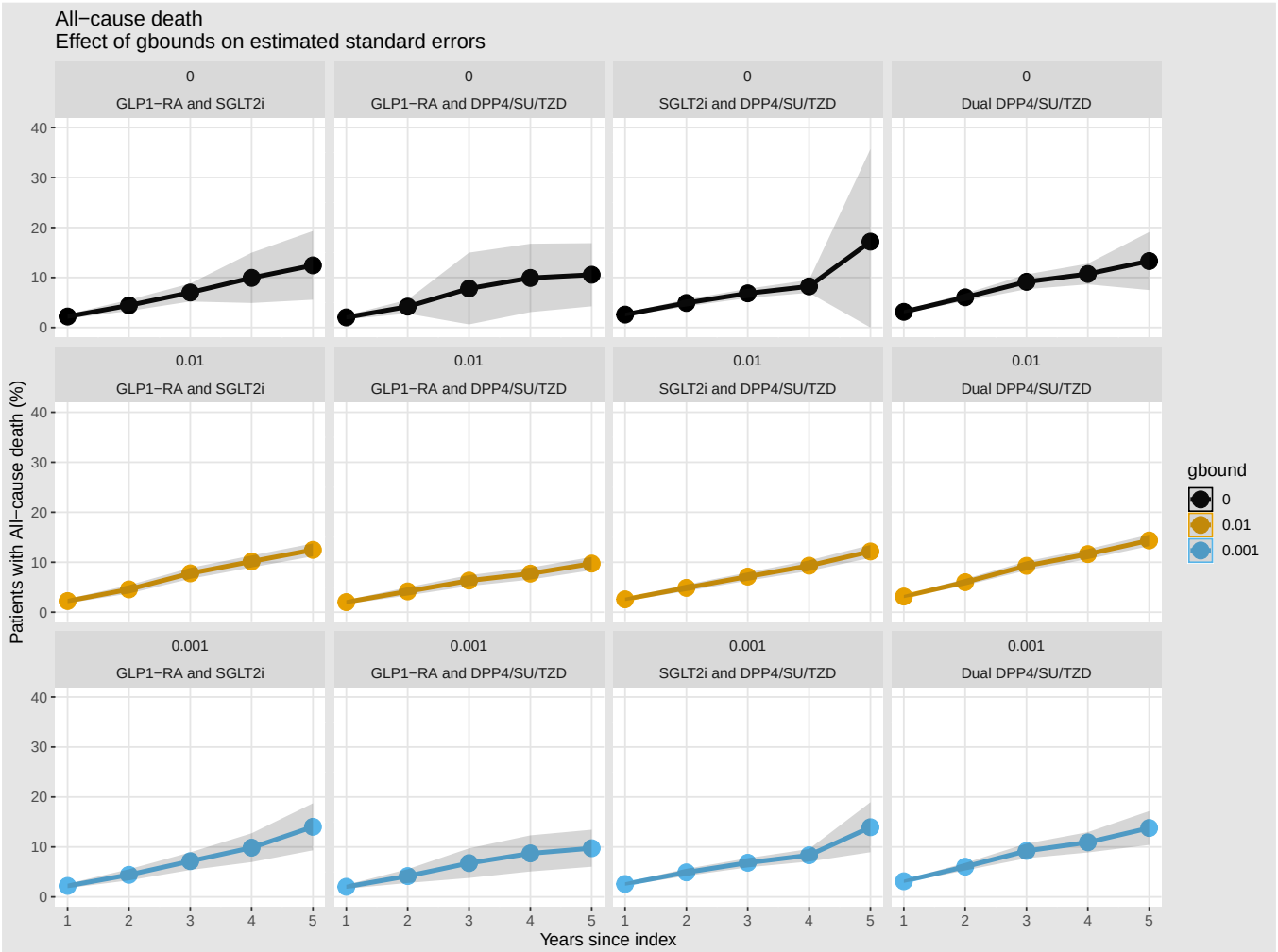
There is no good reason for us to use that particular value, or any other value, hence we subsequently set it to `0`.

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)



Discussion

- Explicitly emulating a target trial prevents flawed and results with unclear interpretation.[3]
- Are stochastic interventions at all realistic? Or just a compropise between positivity and unobserved confounding?
- The treatment protocol has to dictate multiple dimensions including what should happen in case of side effects!
- Should/can we predict side-effects of contrafactual treatments?
- The truncation hyper parameter is sometimes very powerful ...
- Non-adaptive deterministic truncation changes the population and hence the estimand.

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