Simulation study protocol: Bias after censoring concomitant users of test- and control medication when investigating long-term survival outcomes in observational studies of medication effects

Authors: Martin Russek, Britta Haenisch

Date: 29 October 2026

Introduction

Drug-based treatment of type 2 diabetes can have complex patterns, with switches between different agents and combination treatments across drug classes occurring frequently[1,2]. When trying to estimate the comparative effect between continuous use of two different medication classes, this can lead to methodological challenges. While there are models that are able to estimate counterfactual outcome probabilities of the strategies 'Always treat with drug A' versus 'Always treat with drug B'[3–5], handling of concomitant use of both drugs has received less attention so far. Frequently, individuals using both drugs at the same time are censored at the time the second treatment is initiated[6]. Study design approaches like the prevalent new-user design[7] assume that treatment addition only happens in one direction – users of drug A might add drug B, but users of drug B do not add drug A.

When the drugs of interest have immediate effects on outcomes, the censoring approach will not lead to considerable bias. However, when a drug has effects on longer-term health outcomes such as cancer diagnoses, censoring might lead to underestimation of the associated risk. An illustrative example is shown in Figure 1: Drug A causes the outcome after exactly 3 years, but all outcomes are missed because individuals start adding treatment before the outcome occurs and are thus censored.

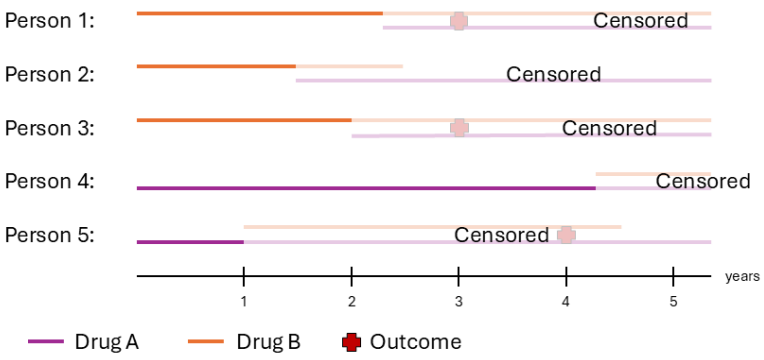


Figure 1: Illustration of risk of bias

The motivation for this simulation study stems from the research question whether SGLT2 inhibitors (SGLT2i), a class of type 2 diabetes medications, lead to a reduced risk of colorectal cancer compared with DPP4 inhibitors[8]. The effect of SGLT2i will take time to be identifiable in observational data sources only recording diagnosis codes. DPP4i

are considered to have a less strong effect on a patient's blood glucose level, but also fewer adverse events, which means that might be preferred to begin treatment with DPP4i rather than SGLT2i. If glycaemic control is not sufficient with DPP4i use, SGLT2i might be considered as an add-on treatment. If a patient starts with SGLT2i, DPP4i might be given as an add on. However, as SGLT2i, like the class of GLP-1 receptor agonists (GLP-1RA) have cardioprotective effects, it is possible that SGLT2i users are more likely to add GLP-1RA than DPP4i. Therefore, it is more likely that a person starting with DPP4i is censored for adding SGLT2i than vice versa. This would imply that more outcomes after DPP4 use would be missed than after SGLT2i use and any potential SGLT2i protective effect might be underestimated.

This simulation study aims to compare methods aiming to mitigate the bias from censoring concomitant users, namely variations of the landmarking method[9] and inverse-probability-of-censoring weighting[10] and a combination thereof. The simulation setup is informed by available literature on distributions and variables related to the detailed motivating research question. It follows the framework laid out in Keogh et al[11], using the additive hazards model and estimating true parameters through simulated randomised trials. It is planned following the clinical scenario evaluation approach by Benda et al[12] and Friede et al[13].

Assumptions

The simulations will generate individual participant data on time-varying exposure, covariates and outcome times, as well as potential censoring times. As laid out in the framework by Keogh et al[11], data is simulated sequentially, starting with baseline covariate information, followed by treatment assignment and outcome event indicators, repeated at each simulated time period k, corresponding to information at quarters k = 0, ..., m. This process enables simulating time-dependent confounding and treatment assignment.

Figure 2 illustrates the data generating mechanism as a DAG. It is similar to the setup in Keogh et al. An unobserved individual frailty U influenced the outcome risk Y and measured covariates L. L is potentially time-varying, with L_k being infuenced by L_{k-1} and A_{k-1}. Treatment assignment A_k, also potentially varying by quarter k, is influenced by previous treatment A_{k-1} and current covariates L_k. The outcome risk in quarter k, Y_k, is influenced by the history of treatment A_0, ..., A_k, and covariate history L_0, ..., L_k. Outcome hazards are additive across quarters.

U is a normally distributed random variable with mean 0 and standard deviation σ_U

L is a continuous variable, drawn from a normal distribution with mean

$\mu_k = L_{k-1} + [\alpha_1\ \alpha_2]\ A_{k-1} + U + \sigma_L$

A is a bidimensional vector, with the first element representing treatment with drug A and the second element representing treatment with drug B. The 2 elements are generated sequentially. At time 0, the first element of A is drawn from a Bernoulli distribution with success probability expit($\phi_0 + \beta_{1,0} L_0$), and the second element is defined as 0 if the first is one, or vice versa, such that initial treatment is binary (Either drug A or drug B). At subsequent times m, the first element is drawn from bernoulli distributions with success probability

expit($\phi_0 + \beta_{0,1} L_0 + \ldots + \beta_{m,1} L_m + \gamma_{0,1} A_0 + \ldots + \gamma_{m-1,1} A_{m-1}$)

and the second element as

expit(ϕ_0+ \beta0,2 L_0 + ... + \beta_m,2 L_m + \gamma0 A_0 + ... + \gamma_m-1 A_{m-1} + $\gamma_m A_m$)

where γ_k is 2-dimensional for all k lower than m, and γ_m and A_m are one-dimensional with A_m drawn just prior.

Hazards for the outcome are generated as

$\xi_m = \lambda_0 + \lambda_L L_m + \lambda_U U + \lambda_{A,0} A_0 + \ldots + \lambda_{A,m} A_m$

where $\lambda_{A,k}$ is bidimensional for all k.

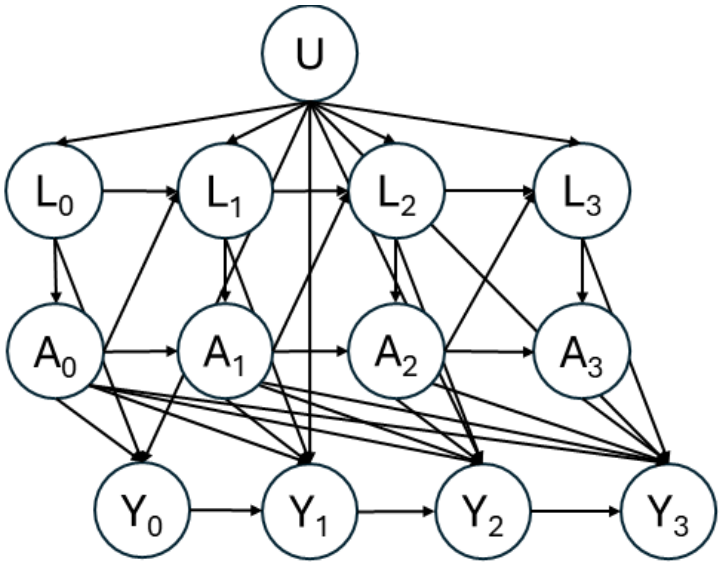


Figure 2: Directed acyclic graph of the simulation setup, simplified

For the simulated RCTs to estimate true values, A is set deterministically to either [1 0] or [0 1] identically at each time point for each person.

Options

The data generating mechanism is varied by the parameters used for drawing variables as described above. Several aspects are investigated here:

Ability to predict treatment assignment

σ_L takes values 0, 1, and 5, to represent easy prediction through L and A, some more randomness, and a situation where treatment assignment is relatively random.

Outcome probability

Incidence of the outcome is set at approximately 10, 50 and 100 per 100.000 person-years, reflecting roughly true values of colorectal cancer incidence in Germany [14]

Probability of concomitant treatment

γ_m varies between 2, 4 and 8 to recreate low, medium and high probabilities of receiving both drugs in the same quarter.

Treatment-outcome relationship

The influence of treatment on the outcome is simulated by varying λ_A. At time point k,

- $\lambda_{A,m} = 0$ for all m: no treatment effect
- $\lambda_{A,0} = ... = \lambda_{A,k-5} = 0$; $\lambda_{A,k-4} <= ... <= \lambda_{A,k}$: immediate, waning effect over time
- $\lambda_{A,0} = ... = \lambda_{A,k-5} = 0$; $\lambda_{A,k} = \lambda_{A,k-1} = 0$: Delayed treatment effect
- $\lambda_{A,0} <= ... <= \lambda_{A,k-5} >= ... >= \lambda_{A,k}$: Continuous effect, but peaking after 5 quarters

Random censoring

Random censoring at a rate of approximately 50 per 100.000 person years will be added to simulate realistic conditions of cohort drop-out.

Metrics

20 quarters are simulated. Each RCT is replicated 500 times to estimate true parameter values. Each simulation is repeated 10.000 times to enable accurate measurement of uncertainty.

For visual comparison, time-varying risk difference and relative risk curves are estimated and compared to the true curves. Additionally, the absolute bias in risk difference and relative risk at quarters 10 and 20 are compared. Uncertainty will be characterised by 95% bootstrap confidence intervals.

References

1. Pottegård A, Andersen JH, Søndergaard J, et al. Treatment trajectories for Danish individuals with type 2 diabetes in the era of emerging glucose-lowering therapies. *Diabetes, Obesity and Metabolism*. 2024;26(11):4996-5004. doi:10.1111/dom.15912

2. Lux TS, Lee D, Phillips JM, et al. Prescribing Trajectories in Type 2 Diabetes in the United States, 2019–2024. *Diabetes, Obesity and Metabolism*. 2026;28(9):8172-8188. doi:10.1111/dom.71013

3. Picciotto S, Hernán MA, Page JH, Young JG, Robins JM. Structural Nested Cumulative Failure Time Models to Estimate the Effects of Interventions. *J Am Stat Assoc*.

2012;107(499):10.1080/01621459.2012.682532. doi:10.1080/01621459.2012.682532

4. Hernán MÁ, Brumback B, Robins JM. Marginal Structural Models to Estimate the Causal Effect of Zidovudine on the Survival of HIV-Positive Men. *Epidemiology*. 2000;11(5):561.

5. Seaman S, Dukes O, Keogh R, Vansteelandt S. Adjusting for Time-Varying Confounders in Survival Analysis Using Structural Nested Cumulative Survival Time Models. *Biometrics*. 2020;76(2):472-483. doi:10.1111/biom.13158

6. Takeuchi Y, Shinozaki T, Kumamaru H, Hiramatsu T, Matsuyama Y. Analyzing intent-to-treat and per-protocol effects on safety outcomes using a medical information database: an application to the risk assessment of antibiotic-induced liver injury. *Expert Opinion on Drug Safety*. 2018;17(11):1071-1079. doi:10.1080/14740338.2018.1528224

7. Suissa S, Moodie EEM, Dell'Aniello S. Prevalent new-user cohort designs for comparative drug effect studies by time-conditional propensity scores. *Pharmacoepidemiology and Drug Safety*. 2017;26(4):459-468. doi:10.1002/pds.4107

8. Russek M, Wicherski J, Brandt A, Hermes U, Naumann-Winter F, Haenisch B. The effect of SGLT2 inhibitors among patients with Type 2 diabetes on the incidence of colorectal cancer. HMA-EMA Catalogues of real-world data sources and studies. Accessed October 8, 2026. https://catalogues.ema.europa.eu/node/4824/administrative-details

9. Anderson JR, Cain KC, Gelber RD. Analysis of Survival by Tumor Response and Other Comparisons of Time-to-Event by Outcome Variables. *J Clin Oncol*. 2008;26(24):3913-3915. doi:10.1200/JCO.2008.16.1000

10. Robins JM, Finkelstein DM. Correcting for Noncompliance and Dependent Censoring in an AIDS Clinical Trial with Inverse Probability of Censoring Weighted (IPCW) Log-Rank Tests. *Biometrics*. 2000;56(3):779-788. doi:10.1111/j.0006-341X.2000.00779.x

11. Keogh RH, Seaman SR, Gran JM, Vansteelandt S. Simulating longitudinal data from marginal structural models using the additive hazard model. *Biometrical Journal*. 2021;63(7):1526-1541. doi:10.1002/bimj.202000040

12. Benda N, Branson M, Maurer W, Friede T. Aspects of Modernizing Drug Development Using Clinical Scenario Planning and Evaluation. *Ther Innov Regul Sci*. 2010;44(3):299-315. doi:10.1177/009286151004400312

13. Friede T, Nicholas R, Stallard N, et al. Refinement of the Clinical Scenario Evaluation Framework for Assessment of Competing Development Strategies with an Application to Multiple Sclerosis. *Ther Innov Regul Sci*. 2010;44(6):713-718. doi:10.1177/009286151004400607

14. Oppelt KA, Luttmann S, Kraywinkel K, Haug U. Incidence of advanced colorectal cancer in Germany: comparing claims data and cancer registry data. *BMC Med Res Methodol*. 2019;19(1):142. doi:10.1186/s12874-019-0784-y