Doubly robust estimator of risk in the presence of censoring dependent on time-varying covariates: application to a primary prevention trial for coronary events with pravastatin

Authors:

Takuya Kawahara[1], Tomohiro Shinozaki[2], Yutaka Matsuyama[3]

1 Clinical Research Promotion Center, The University of Tokyo Hospital, 7-3-1, Hongo, Bunkyo-ku, Tokyo 113-8655, Japan (tkawahara-tky@umin.ac.jp)

2 Department of Information and Computer Technology, Graduate School of Engineering, Tokyo University of Science, Tokyo, Japan (shinozaki@rs.tus.ac.jp)

3 Department of Biostatistics, School of Public Health, The University of Tokyo (matuyama@epistat.m.u-tokyo.ac.jp)

Corresponding author:

Takuya Kawahara, Ph.D., M.P.H.,

Clinical Research Promotion Center, The University of Tokyo Hospital, 7-3-1, Hongo, Bunkyo-ku, Tokyo 113-8655, Japan

E-mail: tkawahara-tky@umin.ac.jp

Abstract

Background: In the presence of dependent censoring even after stratification of baseline covariates, the Kaplan–Meier estimator provides an inconsistent estimate of risk. To account for dependent censoring, time-varying covariates can be used along with two statistical methods: the inverse probability of censoring weighted (IPCW) Kaplan–Meier estimator and the parametric g-formula estimator. The consistency of the IPCW Kaplan–Meier estimator depends on the correctness of the model specification of censoring hazard, whereas that of the parametric g-formula estimator depends on the correctness of the models for event hazard and time-varying covariates.

Methods: We combined the IPCW Kaplan–Meier estimator and the parametric g-formula estimator into a doubly robust estimator that can adjust for dependent censoring. The estimator is theoretically more robust to model misspecification than the IPCW Kaplan–Meier estimator and the parametric g-formula estimator. We conducted simulation studies with a time-varying covariate that affected both time-to-event and censoring under correct and incorrect models for censoring, event, and time-varying covariates. We applied our proposed estimator to a large clinical trial data with censoring before the end of follow-up.

Results: Simulation studies demonstrated that our proposed estimator is doubly robust,

namely it is consistent if either the model for the IPCW Kaplan–Meier estimator or the models for the parametric g-formula estimator, but not necessarily both, is correctly specified. Simulation studies and data application demonstrated that our estimator can be more efficient than the IPCW Kaplan–Meier estimator.

Conclusions: The proposed estimator is useful for estimation of risk if censoring is affected by time-varying risk factors.

Keywords

Double robustness; dependent censoring; prediction; time-varying covariate

Background

Establishment of the long-term effectiveness of primary prevention treatments often requires large randomized controlled trials (RCTs) over a long time period. In such RCTs, survival functions and risks between randomized groups are compared using the Kaplan–Meier estimator because censoring before the end of the follow-up cannot be avoided. This approach assumes independent censoring, such that censoring occurs randomly in each treatment group. The standardization approach can provide a consistent estimate of risks in each group even if censoring is not unconditionally

independent, but the conditionally independence of potential survival time after stratification of treatment groups and baseline covariates [1–4]. In this paper, we call this type of censoring as baseline-conditional independent censoring.

Even a baseline-conditional independent censoring assumption can be dubious. Our motivating study is the Management of Elevated Cholesterol in the Primary Prevention Group of Adult Japanese (MEGA) study, which is a large primary prevention RCT for coronary heart disease (CHD) using pravastatin, where censoring before the end of follow-up occurred in about 10% of patients [5]. Patients enrolled in the MEGA study had hypercholesterolemia (total cholesterol [TC] level: 220–270 mg/dl), ~~and~~ were 40–70 years old, and received daily clinical care during the follow-up period. When a patient with hypercholesterolemia received a medical checkup and found that their plasma lipids were worsening (e.g., increasing TC), they may have required other drugs that were not allowed in the study protocol. Patients who observed worsening of their symptoms might go to see a doctor other than their primary care doctor. These cases may have led to censoring dependent on mid-course clinical characteristics, and the censoring was correlated with future CHD events. ~~Because they received daily clinical care, censoring can be dependent on mid-course clinical characteristics that were predictive of CHD prognosis.~~ If censoring is dependent on

potential survival time even after stratification of treatment groups and baseline covariates, the Kaplan–Meier estimator provides inconsistent estimates of survival function [6]. In such a situation, one possibility to mitigate the dependency is to use time-varying covariates measured during the follow-up period.

The inverse probability of censoring weighted (IPCW) Kaplan–Meier estimator is a semiparametric method for estimation of risk that adjusts for censoring that may depend on the observed past~~, but not on the future prognosis~~ [7]. It requires fitting a model for the probability of censoring at each time conditional on past covariates. Calculation of the IPCW Kaplan–Meier estimate needs to update censoring probability at each time and to weight each subject in the risk set. The weight depends on the time-varying covariates, but not on the future prognosis. The drawback of the IPCW estimator is that it can be statistically inefficient [8].

An alternative to IPCW methods is a g-formula-based estimator, which can be estimated using two different principles. First representation of the g-formula is an iterated conditional expectation, and targeted maximum likelihood estimation can be applied, which was first introduced by Bang and Robins [9]. Their method uses the weight of the IPCW method and regression models for the outcome process. It can produce doubly robust estimates, meaning that the estimator is consistent if either the

regression model for the hazard of censoring or a regression model for the outcome process is correctly specified, but necessarily both [10–12]. However, only a few researchers have applied this method. One of the reasons may be that they are unintuitive because it requires recursive regression models for an iterated conditional expectation; first, we regressed the outcomes measured at $t = K$ on the covariates measured up to $t = K - 1$, second, we regressed the predicted outcome on the covariates measured up to $t = K - 2$, and we continue these procedures until $t = 1$. The second representation of the g-formula is the generalized version of standardization [1, 2], and the parametric g-formula estimator (g-computation algorithm formula) ~~is an alternative method~~can be applied. The parametric g-formula estimator ~~that also~~ requires models for the outcome and covariate process [13]. It can be regarded as a sequential, non-recursive imputation-based methodology [14, 15], so it is intuitive for applied researchers. It is flexible because it can easily compare dynamic treatment regimens [16]. However, it requires a specification of full-model likelihood, and robustness regarding model correctness can be a concern. A doubly robust estimator for the parametric g-formula estimator, involving the time-varying covariates, has not been proposed.

In this paper, we propose an extension of the parametric g-formula estimator

that is more robust at modeling misspecification. The key idea is to combine the IPCW estimator and the parametric g-formula estimator into doubly robust estimators [9, 17–19] while incorporating time-varying covariates to adjust for dependent censoring.

The paper is organized as follows. In the next section, we briefly describe the MEGA study and introduce notations and assumptions. We also describe our proposed estimator, and we give settings and the results of simulation studies. Finally, the proposed estimator is applied to the MEGA study data.

Data, notations, and assumptions

The MEGA study is a prospective, randomized, open-label, blinded-endpoint-designed controlled trial conducted in Japan to evaluate the primary preventive effect of pravastatin against CHD in daily clinical practice. A total of 7832 men and postmenopausal women aged 40–70 years with hypercholesterolemia and no history of CHD or stroke were randomized to dietary therapy only (diet group) or dietary therapy plus 10–20 mg daily pravastatin (diet plus pravastatin group) between February 1994 and March 1999.

Table 1. Type and number of events within 5 years in the MEGA study

	Diet group		Diet + pravastatin group	
	n	%	*n*	%
CHD event	85	2.1	57	1.5
Follow-up completed	3498	88.2	3353	86.7
Refusal of follow-up	259	6.5	364	9.4
Death by causes other than CHD	60	1.5	42	1.1
Loss to follow-up	64	1.6	50	1.3
Total	3966	100.0	3866	100.0

After randomization, laboratory tests were conducted at months 1, 3, and 6, and annually thereafter. The follow-up period was initially scheduled for 5 years. Table 1 shows the types and number of events within 5 years. Although there were three reasons for censoring during the study period (refusal of follow-up, death by causes other than CHD, and loss to follow-up), we collectively treated them as censoring before the end of the follow-up period.

Let $t = 1, \ldots, T$ denote month of follow-up where $T + 1 = 60$ months is the follow-up of interest. There were 7832 patients at baseline, and observations of patients

were assumed to be independently identically distributed. R denotes the treatment assigned ($R = 1$ for assignment to the diet plus pravastatin group, and $R = 0$ for assignment to the diet group). C_t and Y_t denote the indicator of censoring and occurrence of a CHD event by time t, respectively, with $C_0 = Y_0 = 0$ by definition. L_t denotes time-varying covariates measured at time t, and V denotes baseline covariates that are time-independent (e.g. sex, current smoker). We assumed that baseline covariates V and L_0 are always observed. We denoted the history of a variable using overbars. For example, $\bar{L}_t = (L_0, \ldots, L_t)$ is the covariate history through time t. We assumed the order (C_t, Y_t, L_t) within each interval (t - 1, t); therefore, Y_t and following variables are missing if C_t =1. We defined C_{T+1} =1 if $C_T = Y_T = 0$ (follow-up completed).

We wanted to estimate the marginal event-free survival in each treatment group if any censoring was absent in the study population. However, ~~measured~~ observed data contains censoring, as in the MEGA study. The usual the Kaplan–Meier estimator assumes independent censoring, that is, the hazard of Y_t among subjects at risk is the marginal hazard of Y_t given the treatment group. The standardization approach, or a g-formula that adjusts for baseline covariates, assumes baseline-conditional independent censoring, that is, the hazard of Y_t among subjects at risk is the conditional hazard of Y_t

given the treatment group and baseline covariates [1–4].

Even when these two assumptions are attainable, estimators discussed in the next section provide a consistent estimate of the marginal survival in each treatment group if any censoring was absent in the study population. These estimators assume positivity (equation 1) and are conditionally independent of censoring (equation 2);

$$\Pr\left(C_t = 0 \middle| \bar{C}_{t-1} = 0, \bar{Y}_{t-1} = 0, R, V, \bar{L}_{t-1}\right) > 0 \text{ for } t = 1, 2, \ldots, T \quad (1)$$

$$\Pr\left(C_t = 0 \middle| \bar{C}_{t-1} = 0, \bar{Y}_T = 0, R, V, \bar{L}_{t-1}\right) = \Pr\left(C_t = 0 \middle| \bar{C}_{t-1} = 0, \bar{Y}_{t-1} = 0, R, V, \bar{L}_{t-1}\right)$$

$$\text{for } t = 1, 2, \ldots, T \quad (2).$$

The conditional independence of censoring assumption, equation (2), states that for $t = 1, 2, \ldots T$, the variables (Y_t, …, Y_T) is independent of C_t, in other words, the distribution of (Y_t, …, Y_T) is the same between $C_t = 1$ and $C_t = 0$ among subjects who had a similar history of the covariates. The conditionally independence of ~~the~~ censoring assumption is also referred to as no unmeasured confounders for the censoring assumption [7], which states that conditional on the treatment groups, baseline covariates, and the time-varying covariates measured until time t - 1, the hazard of censoring at time t does not further depend on ~~possibly unobserved CHD events after time t~~no unmeasured confounders for censoring and unobserved CHD. In the next section, we describe the existing estimators and our proposed estimator for the hazard of

Y_t.

Existing estimators and proposed estimator

Due to randomization, baseline factors are balanced between treatment groups. In this section, we focus on the diet plus pravastatin group ($R = 1$) and suppress R for notational simplicity. A similar argument holds for the diet group ($R = 0$).

Estimators of hazard of Y_1

At time $t = 1$, the observed data is n copies of $(V, L_0, C_1, (1 - C_1)Y_1)$. We show three types of estimators for $\Pr(Y_1 = 1)$; the IPCW estimator, the parametric g-formula estimator, and the doubly robust estimator.

To obtain the IPCW estimate, we need to fit a model for C_1 such as the logistic model $\Pr(C_1 = 0|V, L_0; \alpha) = e(V, L_0; \alpha) = \{1 + \exp(-\alpha_0 - \alpha_1 V - \alpha_2 L_0)\}^{-1}$. After fitting the model, the IPCW estimator for $\Pr(Y_1 = 1)$ is expressed as $n^{-1} \sum_i (1 - C_{i1})Y_{i1}/e(V_i, L_{i0}; \hat{\alpha})$. The consistency of the IPCW estimator relies on the correct specification of $e(V, L_0; \alpha)$. If no censoring is observed at $t = 1$, we set $e(V_i, L_{i0}; \hat{\alpha}) = 1$; therefore, the IPCW estimator equals the empirical risk.

To obtain the parametric g-formula estimate, we need to fit a model for Y_1 such

as the logistic model $\Pr(Y_1 = 1|C_1 = 0, V, L_0; \beta) = p(V, L_0; \beta) = \{1 + \exp(-\beta_0 - \beta_1 V - \beta_2 L_0)\}^{-1}$. After fitting the model ~~in~~ for the subjects not censored at $t = 1$ (subjects with $C_1 = 0$)~~$C_1 = 0$~~, the parametric g-formula estimator for $\Pr(Y_1 = 1)$ is expressed as $n^{-1} \sum_i p(V_i, L_{i0}; \hat{\beta})$. The consistency of the parametric g-formula estimator relies on the correct specification of $p(V, L_0; \beta)$.

To obtain the doubly robust estimate, we need to fit a model for C_1 and Y_1 similarly as conducted for the IPCW estimator and the parametric g-formula estimator, respectively. After fitting the models $e(V, L_0; \alpha)$ and $p(V, L_0; \beta)$, the doubly robust estimator for $\Pr(Y_1 = 1)$ is expressed as

$$n^{-1} \sum_i \left[\frac{(1-C_{i1})Y_{i1}}{e(V_i, L_{i0}; \hat{\alpha})} - \frac{(1-C_{i1}) - e(V_i, L_{i0}; \hat{\alpha})}{e(V_i, L_{i0}; \hat{\alpha})} p(V_i, L_{i0}; \hat{\beta}) \right] \quad (3).$$

The contributions of censored patients or patients with an event are different; for censored patients, their contribution is $p(V_i, L_{i0}; \hat{\beta})$ like the g-formula estimator, and for patients with an event, their contribution is $Y_{i1}/e(V_i, L_{i0}; \hat{\alpha}) - \{1 - e(V_i, L_{i0}; \hat{\alpha})\} p(V_i, L_{i0}; \hat{\beta})/e(V_i, L_{i0}; \hat{\alpha})$. The doubley robust estimator is consistent if either the model $e(V, L_0; \alpha)$ or $p(V, L_0; \beta)$ is correctly specified [9, 17–19]. Intuitively, when the model for censoring is correctly specified, the term $(1 - C_{i1}) - e(V_i, L_{i1}; \hat{\alpha})$ should be zero, so (3) reduces to the IPCW estimator and is, therefore, consistent. Inside the summation can be expressed as $(1 - C_{i1})\{Y_{i1} - p(V_i, L_{i0}; \hat{\beta})\}/e(V_i, L_{i0}; \hat{\alpha}) +$

$p(V_i, L_{i0}; \hat{\beta})$, and when the model for an event is correctly specified, the term $Y_{i1} - p(V_i, L_{i0}; \hat{\beta})$ should be zero, so (3) reduces to the g-formula estimator and is, therefore, consistent. Our proposed estimator utilizes this doubly robust estimator for the hazard of Y_1. In the next subsection, we show how to extend it to estimate the hazard of Y_t ($t > 1$) incorporating time-varying covariates.

We noted that with one categorical baseline covariate and no parametric model ~~is~~ needed for outcomes and censoring, it can be shown that the IPCW estimator, the g-formula estimator, and the doubly robust estimator ~~is~~are equivalent. Specifically, given n subjects, all of whom may be stratified into j levels of a baseline covariate, such that a_j, m_j, and n_j are the number ~~of~~ observed (i.e. not censored), number of events, and overall number at level j of the covariate, respectively. The IPCW estimator can be written as $(1/n)\Sigma_j m_j / (a_j / n_j) = (1/n) \Sigma_j n_j m_j / a_j$, because $\Pr(C_1 = 0|$ level $j) = a_j / n_j$. The g-formula estimator can be written as $(1/n) \Sigma_j n_j (m_j / a_j)$, because $\Pr(Y_1 = 1|$ level $j) = m_j / a_j$. Finally, the doubly robust estimator can be written as $(1/n) \Sigma_j [m_j / (a_j / n_j) - \{(n_j - a_j) (0 - a_j / n_j) / (a_j / n_j) -+ a_j (1 - a_j / n_j) / (a_j / n_j)\} (m_j / a_j)] = (1/n) \Sigma_j n_j m_j / a_j$, which is exactly a common form ~~of~~as the IPCW estimator and the g-formula estimator.

Estimators of hazard of Y_t ($t > 1$)

In this subsection, we show the estimators of the hazard of Y_t ($t > 1$), which are extended versions of the IPCW estimator and the parametric g-formula estimators for $\Pr(Y_1 = 1)$. Finally, we propose a doubly robust estimator that extends equation (3).

To obtain the IPCW Kaplan–Meier estimate, we need to fit a model for C_t such as the pooled logistic model,

$$\text{logit}\,\Pr\left(C_t = 0 \middle| \bar{C}_{t-1} = 0, \bar{Y}_{t-1} = 0, V, \bar{L}_{t-1}\right) = \alpha_{0t} + \alpha_1 V + \alpha_2 L_{t-1} \quad (4).$$

In the model, it is possible to include $L_0, \ldots, L_{t-2}$, but in some cases, it may cause multicollinearity due to the correlation between $L_0, \ldots, L_{t-1}$. After fitting the model using the maximum likelihood estimation, the IPCW Kaplan–Meier estimator for the hazard of Y_t is expressed as $\widehat{\Pr}\left(Y_t = 1 \middle| \bar{Y}_{t-1} = 0\right) = \sum_i Y_{it}\pi_{it}(\hat{\alpha}) \Big/ \sum_i X_{it}\pi_{it}(\hat{\alpha})$, where $\pi_t(\hat{\alpha})$ is obtained as

$$\pi_t(\hat{\alpha}) = \prod_{j=1}^{t} \Pr\left(C_{tj} = 0 \middle| \bar{C}_{tj-1} = 0, \bar{Y}_{tj-1} = 0, V, \bar{L}_{tj-1}; \hat{\alpha}\right),$$

and X_t is the at-risk indicator, which is 1 if the patient is at-risk at time t and is 0 otherwise. Finally, the risk at t can be obtained as $1 - \prod_{j=0}^{t} \left\{1 - \widehat{Pr}\left(Y_j = 1 \middle| \bar{Y}_{j-1} = 0\right)\right\}$. The consistency of the IPCW Kaplan–Meier estimator relies on the correct specification of the model for C_t (equation 4) [7]. Note that the IPCW Kaplan–Meier estimator reduces to the usual Kaplan–Meier estimator when α_1 and α_2

of equation (4) are 0, that is, the independent censoring assumption is true [20].

To obtain the parametric g-formula estimate, we need to fit a model for Y_t. Unlike baseline covariates, time-varying covariates will not be measured for patients who were censored before time t. Thus, we need to specify the full-model likelihood (likelihood for conditional event probability and time-varying covariates) by fitting models for Y_t and L_t such as

$$\text{logit}\,\Pr\left(Y_t = 1 \middle| \bar{C}_t = 0, \bar{Y}_{t-1} = 0, V, \bar{L}_{t-1}\right) = \beta_{0t} + \beta_1 V + \beta_2 L_{t-1} \quad (5), \text{ and}$$

$$E\left(L_t \middle| \bar{C}_t = 0, \bar{Y}_{t-1} = 0, V, \bar{L}_{t-1}\right) = \gamma_{0t} + \gamma_1 V + \gamma_2 L_{t-1} \quad (6).$$

After fitting the models using the maximum likelihood estimation, we sequentially imputed the conditional probability of CHD event and time-varying covariates from t = 1 to T. The parametric g-formula estimator for ~~the hazard of~~ the risk at t can be obtained as~~Y_t is expressed as~~ $n^{-1}\sum_{j=1}^{t}\sum_i m_{j,i}(\hat{\beta},\hat{\gamma})$, where $m_t(\hat{\beta},\hat{\gamma})$ is obtained as $m_t(\hat{\beta},\hat{\gamma}) = \Pr\left(Y_t = 1 \middle| \bar{Y}_{t-1} = 0, V, \bar{L}_{t-1}; \hat{\beta},\hat{\gamma}\right)\prod_{j=1}^{t-1}\left\{1 - \Pr\left(Y_j = ~~0~~1 \middle| \bar{Y}_{j-1} = 0, V, \bar{L}_{j-1}; \hat{\beta},\hat{\gamma}\right)\right\}$. The consistency of the parametric g-formula estimator relies on the correct specification of the model for Y_t (equation 5) and the model for L_t (equation 6) [16, 21, 22].

We propose an estimator of the hazard of Y_t that extends the doubly robust estimator (equation 3). To obtain the estimate, we need to fit models for C_t, Y_t, and L_t as conducted for the IPCW Kaplan–Meier estimator (equation 4) and the parametric g-

formula estimator (equations 5 and 6). After fitting these models, the proposed doubly robust estimator for the hazard of Y_t is expressed as,

$$\widehat{\Pr}\Pr\left(Y_t = 1 \middle| \bar{Y}_{t-1} = 0\right) = (\sum Z_t)^{-1} \sum_i \left[\frac{(1-C_{it})Y_{it}}{\pi_{it}(\hat{\alpha})} - \frac{(1-C_{it})-\pi_{it}(\hat{\alpha})}{\pi_{it}(\hat{\alpha})} \Pr\left(Y_{i,t} = 1 \middle| \bar{Y}_{i,t-1} = 0, V_i, \bar{L}_{i,t-1}; \hat{\beta}, \hat{\gamma}\right)\right] \quad (7)$$

~~$n^{-1} \sum_i \left[\frac{(1-C_{it})Y_{it}}{\pi_{it}(\hat{\alpha})} - \frac{(1-C_{it})-\pi_{it}(\hat{\alpha})}{\pi_{it}(\hat{\alpha})} m_t(\hat{\beta}, \hat{\gamma})\right] (7).$~~

where Z_t is the at-risk or censored indicator, which is 1 if the patient is at-risk or censored at time t and is 0 otherwise. The contributions of patients censored ~~patients~~ by t or patients with an event at t are different; for censored patients, their contribution is $\Pr\left(Y_t = 1 \middle| \bar{Y}_{t-1} = 0, V, \bar{L}_{t-1}; \hat{\beta}, \hat{\gamma}\right)$. For patients with an event, their contribution of an event is weighted by the inverse of probability uncensored until t. Finally, the risk at t is obtained as $1 - \prod_{j=0}^{t} 1 \left\{1 - \widehat{\Pr}\left(Y_j = 1 \middle| \bar{Y}_{j-1} = 0\right)\right\}$. The weights and predicted event probabilities are similar ~~as~~to the ones used in the IPCW Kaplan–Meier estimator and the parametric g-formula estimator, but we need to calculate the function (7) and risk at t.

As demonstrated in the Additional file (Appendix A), this estimator is consistent if either the model for C_t (equation 4) or models for Y_t and L_t (equation 5 and 6) is correctly specified. In the IPCW Kaplan–Meier estimator, patients with $C_t = 1$ were out of the risk set; therefore, they do not contribute to the estimation of the hazard of Y_t. On

the other hand, patients with $C_t = 1$ contribute to the estimation of the hazard of Y_t in equation (7), because the function inside the summation reduces to $m_t(\hat{\beta}, \hat{\gamma})$, which might lead to statistical efficiency. The variance estimate of the proposed estimator can be obtained through a nonparametric bootstrap [23]. We have provided SAS code for the proposed estimator in an additional file 2 ~~(Appendix B)~~.

Comparison with existing doubly robust estimators

In this subsection, we briefly compare our proposed estimator (7) with existing doubly robust estimators [9, 24, 25]. Zhang et al. [24] and Bai et al. [25] proposed doubly robust estimators for survival functions, which can be summarized as follows:

Confounding between treatment groups: present due to the observational study setting

Censoring mechanism: baseline-conditional independent censoring (censoring may depend only on the baseline covariates)

On the other hand, we proposed an estimator for survival functions,

Confounding between treatment groups: absent due to randomization

Censoring mechanism: conditional independent censoring (censoring may depend on time-varying covariates)

In RCT settings considered here, where no baseline confounding occurs between the treatment groups, the proposed estimator that specifies an empty set as L_t (thus models are unnecessary for the joint density of L_t) results in the existing doubly robust estimators provided in [24, 25]. In other words, these existing estimators assume a baseline-conditional independent censoring mechanism, although they also attempt to adjust for baseline-confounding between the groups in observational-study settings.

Bang et al. [9] proposed a doubly robust estimator for the g-formula represented by an iterated conditional expectation. The estimator needs recursive fitting of the iterative conditional expectation. However, as Bang et al. [9] noted, the parametric models can be incompatible with each other, so it is difficult to specify all the models correctly.

Simulation study

~~*Simulation designs*~~

To evaluate the performance of the proposed estimator, we carried out simulation studies with dependent censoring due to a time-varying covariate. We simulated data from two treatment groups, coded as $R = 0$ (control treatment) and $R = 1$ (test

316 treatment). The simulations were based on 1000 replications. We considered the
317 situation where baseline covariates were measured at time $t = 0$, and time-varying
318 covariate and censoring were investigated at time $t = 1$, … ~~4~~ 2, on the other hand, event
319 time was measured from time $t = 0$ to $t =$ ~~3~~5 on a continuous time scale. We were
320 interested in the treatment group-specific risks and the risk ratio at $t = 3$ and $t = 5$.

321 For each patient i (= 1, …, 1000), a baseline covariate V was generated from
322 the Bernoulli distribution of success probability 0.5. Independently, the time-varying
323 covariate at t (= 0, 1, 2) was generated from the following mixed effect model,

324 $$L_{it} = 2 - 0.1(1 - R_i)t - 0.5\text{~~1.1~~}R_i t + b_{i0} + b_{i1}t + \epsilon_{it}.$$

325 Random variables (b_{i0}, b_{i1}) were generated from a bivariate normal distribution with
326 means of 0 and variance of 1.0 and 0.5, respectively, with a covariance of 0.5. The
327 random error ϵ_{it} was generated from the standard normal distribution. Distributions of
328 L_t were the same in both treatment groups at $t = 0$ but declined more steeply in the test
329 treatment group such that L_t mimicked TC in the MEGA study.

330 First, we generated a time to event T_i from the piecewise exponential model,
331 whose hazard function was,

$$\lambda\left(t\middle|V,\bar{L}_t,R\right)=\exp\left(-5+1.5V+1.2U_t+1.2(1-R)\right)\begin{cases}\exp\left(-6+1.5V+1.2U_0+1.2(1-R)\right), & 0<t\leq 1\\ \exp\left(-5+1.5V+1.2U_1+1.2(1-R)\right), & 1<t\leq 2\\ \exp\left(-4+1.5V+1.2U_2+1.2(1-R)\right), & 2<t\end{cases}$$

where $U_t = 1$ if $L_t < 0$, otherwise $U_t = 0$. Therefore, potential event time was shorter in the control treatment group through the effect of group and time-varying covariate.

Next, we generated censoring C_t at t (= 1, …, 42) from the Bernoulli distribution, whose probability was generated using the following logistic model,

$$\text{logit Pr}\left(C_t=1\middle|\bar{C}_{t-1}=0,T>t,V,\bar{L}_{t-1},R\right)=\alpha_0-4+t+1.5V+1.2U_{t-1}(t-1)+1.2\alpha_R(1-R).$$

Similarly to the event model, censoring occurred more frequently in the control treatment group through group effect and the effect of the time-varying covariate. $C_t=C_2\bar{C}_{4-}=0$ and $T_i > 35$ indicates that the follow-up was completed. The direct dependence between the event and the censoring time is shown in additional file 1 (Appendix B).

We considered three scenarios for α_0 and α_R: censoring probabilities in the control and test treatment groups are both 30% (scenario 1), both 20% (scenario 2), and 9% and 12%, respectively (scenario 3). The probabilities in scenario 3 were derived from Table 1.

In this setting, the observed event probabilities, censoring probabilities, and

~~complete follow-up probabilities were approximately 20%, 60%, and 20% for the control treatment group, and 15%, 40%, 45% for the test treatment group.~~

We ~~used~~ created 20,000,000 simulated patients without censoring to calculate the true value of survival probability using their empirical distribution. To understand the performance of estimators, we considered eight situations: all combinations of correct or incorrect censoring models, event models, and covariate models. We defined correct models for censoring, event, and covariate as a model that specified the same covariates with the data-generating model. We defined incorrect models for censoring and event as a model that specified by replacing U_t by $\exp(L_t)$ without incorporating V. An incorrect covariate model was specified without incorporating the interaction term of b_{i1} and t.

Simulations were evaluated in terms of the bias (mean difference between estimated and true parameter value) and relative efficiency (the ratio of the Monte Carlo standard deviation of the IPCW Kaplan–Meier estimator to that of the estimator) of the estimated survival probabilities at time $t = 3$ and $t = 5$.~~–~~

Results

Simulation results

We present our simulation results in Table 2. In Table 2, if the bias exceeded half of the standard error of the estimates, the printed bias was is shown in bold. In scenario 1, the bias for each group at $t = 5$ was seen for the IPCW Kaplan–Meier estimator when the censoring model is incorrect, for the parametric g-formula estimator when one of the event model or covariate model is incorrect. However, our proposed estimator is unbiased when at least one of the censoring model or event model is correctly specified. This result reflected the double robustness of our proposed estimator; when the censoring model or set of event and covariate models are correct, the estimate is unbiased. Unexpectedly, our proposed estimator is less biased than the parametric g-formula estimator, even when the covariate model was incorrect. We consider that this property is only in this simulation because if the covariate model is incorrect, the estimated event probability is also incorrect for true probability. At $t = 3$, the parametric g-formula estimator showed less bias for the test treatment group even when the event model is incorrect. Regarding the bias, similar results can be seen in the other two scenarios.

Regarding the relative efficiency using the IPCW Kaplan–Meier estimator as the reference, both the parametric g-formula estimator and our proposed estimator were more efficient at $t = 3$ ~~As expected, the IPCW Kaplan–Meier estimator was biased when the~~

~~censoring model was incorrectly specified, and the parametric g-formula estimator was also biased when the event and/or covariate models were incorrectly specified. On the other hand, the proposed estimator was unbiased when the censoring model was correctly specified, or the event and covariate models were correctly specified. However, it was biased in the situation where both the IPCW Kaplan–Meier and parametric g-formula estimators were biased.~~ than the reference in scenarios 1 and 2.The parametric g-formula estimator was more efficient than the reference even at $t = 5$~~, h~~. However, our proposed estimator had a similar standard error ~~as~~to the reference. In scenario 3, where the censoring probability ~~i~~was the lowest among the scenarios, our proposed estimator had a similar standard error as the reference at both $t = 3$ and 5. The coverage probability of the proposed estimator using the bootstrap method with the correctly specified models was close to the nominal level of 95 %. In summary, the efficiency recovery of our proposed estimator may be affected by the censoring probabilities (comparing between the scenarios) and the number of time points (comparing $t = 3$ and $t = 5$). When the censoring probability is high but the number of time points is less than five, our proposed estimator might be more efficient than the IPCW Kaplan–Meier estimator.~~When comparing Monte Carlo standard deviations among these estimators, the parametric g-formula estimator was the most efficient. Our proposed estimator showed efficiency recovery from the~~

~~IPCW Kaplan–Meier estimator. The efficiency recovery was greater when estimating the risk in the control treatment group.~~

[Table 2 is placed at the end of the document]

Data applications

Our proposed estimator was applied to the MEGA study data to estimate treatment group-specific risks at 5 years after randomization. As baseline covariates, we included age (years), gender, body mass index, history of hypertension and diabetes, hypercholesterolemia medication history, current smoking, current alcohol drinking, triglyceride, high-density lipoprotein cholesterol, and low-density lipoprotein cholesterol. As the time-varying covariate, we used recent TC.

After transforming our data into one record per person-time, we estimated the survival curve using our proposed estimator. First, we fitted the models for censoring C_t, event Y_t, and covariate L_t. We fitted pooled logistic models for C_t and Y_t, where the time-varying intercept was included as a restricted cubic spline with 4 knots at 1–4 years after randomization. We fitted a linear model for L_t. By fitting the pooled logistic model for Y_t, classical risk factors for CHD (age, male, hypertension, and diabetes) were found to be the prognostic factors (Additional file, Appendix C). By fitting the model

for C_t, those without hypertension, diabetes, or no history of medication for hyperlipidemia, tended to be censored before the end of the follow-up period (Additional file, Appendix D). Unexpectedly, time-varying TC hardly affected the event or censoring after adjusted for those important baseline covariates (Additional file 1, Appendix C and D); therefore, baseline-conditional independence assumption rather than conditional independence assumption might be plausible in the MEGA study.

We estimated the risk of CHD incidence at 5 years from randomization using the Kaplan–Meier estimator, IPCW Kaplan–Meier estimator, the parametric g-formula estimator, and our proposed estimator. The results are shown in Table 3. In the MEGA study dataset, the risk of CHD estimated using the usual Kaplan–Meier estimator and the risk estimated by other estimators were very similar. This may be due to the small impact of dependent censoring in the MEGA study and correctness of model specification for censoring and events. Because the ordinal Kaplan–Meier estimator showed similar results as the other three estimators that adjust for the possible dependent censoring, the impact of dependent censoring must be very mild. If the censoring model or event model was mis-specified, the results from the other three estimators might be more different. Therefore, the results from the three estimators may indicate that the postulated models were nearly correctly specified. The estimated

confidence interval of the parametric g-formula estimator was narrower than the other estimators ~~was similar or narrower than that of the IPCW Kaplan–Meier estimator and our proposed estimator~~. The estimated confidence interval for the risk of diet + pravastatin group of our proposed estimator ~~was similar or narrower than that of~~was narrower than the IPCW Kaplan–Meier estimator.

Table 3. Risk of coronary heart diseases in the MEGA study at 5 years after randomization

	Diet group		Diet + ~~p~~pravastatin group			
Method	Risk (%)	95% CI	Risk (%)	95% CI	Risk Ratio	95% CI
Kaplan–Meier*	2.34	(1.90, 2.89)	1.63	(1.26, 2.11)		
IPCW Kaplan–Meier	2.39	(1.91, 2.95)	1.60	(1.19, 2.10)	0.68	(0.40, 1.06)

Parametric g-formula	2.36	(1.97, 3.01)	1.66	(1.30, 2.05)	0.71	(0.47, 0.98)
Proposed estimator	2.38	(1.91, 2.95)	1.61	(1.22, 2.06)	0.69	(0.42, 1.03)

* The confidence intervals of the Kaplan–Meier estimator was obtained using the Greenwood formula.

Discussion

In this paper, we proposed a doubly robust estimator of risk that adjusts for dependent censoring due to time-varying covariates in RCT settings. The novelty of our proposed estimator is as an extension of the existing estimator [9, 19] for more complex data with $t > 1$ and time-varying covariates. The IPCW Kaplan–Meier estimator is routinely used in the analysis of RCTs for the purpose of adjusting for dependent censoring with time-varying covariates measured throughout the follow-up period. ~~The novelty of the proposed estimator is that we extended a well-known doubly robust estimator that is able to adjust dependent censoring due to only baseline covariates [24, 25] to one that is able to adjust dependent censoring due to time-varying covariates, by combining the IPCW Kaplan–Meier estimator and the parametric g-formula estimator.~~ We have also provided SAS codes ~~in Appendix B~~ and an example of simulation data ~~with~~has additional files~~, which can be easily implemented~~. The important property of our proposed

estimator is the double protection against model misspecification. Because risk factors for the endpoints are often identified before the beginning of the RCT, by measuring them longitudinally at as many time points as possible and by using them when constructing the models, we are in a better position to approximate the true regression function. ~~able to approximate the true regression function.~~

The second property of our proposed estimator is the efficiency recovery over the IPCW Kaplan–Meier estimator, as shown in the simulation study. The degree of efficiency recovery could depend on either the censoring probability, event probability, the dependency of variables, or all of these factors combined. As studied previously [8], we considered that the censoring probability is an important factor. Further studies will involve understanding the factors that affect the degree of efficiency recovery using further simulations. In the simulation study and analysis of the MEGA study, the parametric g-formula estimator outperformed regarding efficiency. This phenomenon was expected because the asymptotic variance of the classical doubly robust estimator is no smaller than that of the g-formula estimator [26].

Our estimator relies on the assumption that censoring and event time are independent conditional on observed covariates including time-varying ones. However, in a situation that censoring and event time are not independent even if we condition on

time-varying covariates, our proposed estimator and other existing estimators cannot correct for selection bias. We also need the assumption of correct model specification. We need to incorporate the covariates that affect both event and censoring probabilities, and moreover, we need to specify the model form that approximates the true regression function.

In this study, we considered the estimation of a survival function in a specific group. If we compare two or more survival functions that may be observed with different interventions, we also need an additional exchangeability assumption (or the no-unmeasured confounders assumption) between the intervention groups [27]. In the simulations and data analysis, the exchangeability assumption is satisfied at baseline owing to the randomized design. In a future study, it will be interesting to extend our estimator into the observational study setting [24, 25].

All the estimators in this study can be applied to right-censored data. We consider that our proposed estimator cannot be applied to the data with interval or left-censored data in its current form. With those censoring, we know that an event has occurred only before a specific time. In this situation, how to predict event probability and how to weight uncensored subjects are not obvious. Note that the MEGA study corrects exact event time, so we consider that interval censoring or left censoring is

absent in the real data.

There are several reasons for censoring in the MEGA study, as shown in Table 1. We treated refusal of follow-up, death by causes other than CHD, and loss to follow-up as reasons for censoring in the censoring model. Three estimators, including our proposed estimator, assessed the hypothetical survival function when there was no censoring. It may be meaningful to consider whether a survival function can be obtained if refusal to follow-up and loss to follow-up did not occur. When we separately accounted for the two reasons for dropouts, the survival curve was similar to the one using the Kaplan–Meier method [28]. It may be meaningful to consider whether a survival function would be obtained if refusal to follow-up and loss to follow-up did not occur. However, death by causes other than CHD needs additional consideration, because it is difficult to cease such competing risks for CHD without lowering the risk of CHD. Therefore, if there was no death by causes other than CHD, the survival function would be slightly lower than we estimated. Because in the MEGA study the proportion of censoring due to death by causes other than CHD was less than 1.5%, we believe the estimated survival functions are close to the true survival function, which would be obtained if these censorings had not occurred.

There are two limitations in this study. First, we were not able to verify the

assumptions with the measured data. The positivity assumption will be satisfied unless the conditional probabilities of~~or~~ censoring are zero for all patients at $t = 1,\ldots, T$. In the analysis of the MEGA study data, there were no patients who had an estimated probability of censoring near 1 (data not shown); therefore, we considered that the positivity assumption is acceptable. Conditional independence assumption implies that the treatment group, measured baseline, and time-varying covariates can completely explain censoring. However, given a rich collection of measured prognostic factors, the conditional independence assumption can be approximated. Several clinically important prognostic factors were measured in the MEGA study, and we used all of the baseline covariates and a time-varying covariate, TC. We considered time-varying TC was important for event and censoring probability, but the hazard ratio was close to 1; therefore, the impact of dependent censoring was very mild. In the future, we need to apply our estimator to data with censoring dependent on time-varying factors. The second limitation was the range of the simulation study. Because we were interested in the statistical properties of the estimators with fitted correct/incorrect models, the behavior of the estimators when other assumptions, such as positivity, were violated is unknown. We need further simulation studies to understand the performance of the estimators.

Conclusions

The proposed estimator is useful for the estimation of risk if censoring affected by time-varying risk factors occurred because of the doubly robust property and statistical efficiency over the IPCW Kaplan–Meier method.

List of abbreviations

CHD, coronary heart disease; CI, confidence interval; IPCW, inverse probability-of-censoring weighted; MEGA, Management of Elevated Cholesterol in the Primary Prevention Group of Adult Japanese; RCT randomized controlled trial; TC, total cholesterol.

Declarations

Ethics approval and consent to participate

Not applicable because this paper focuses on the development of statistical methods. Real example data were originally published in Nakamura et al. [5].

Consent for publication

Not applicable.

Availability of data and materials

The SAS code and an example of simulated dataset is available in Additional File 2~~1~~.

Competing interests

None declared.

Funding

This work was supported by JSPS KAKENHI Grant Number JP18K17314.

Authors' contributions

TK, TS, and YM designed the concept of this research. TK conducted the simulation study and analyzed the MEGA study data. TK and TS drafted the manuscript. YM supervised this study and critically reviewed the manuscript. All the authors have read and approved the manuscript.

Acknowledgments

575 The authors thank Daiichi Sankyo Co. Ltd for providing invaluable MEGA study data
576 and Dr. Koji Oba for reading an earlier draft of this article. Finally, we are grateful to the
577 MEGA study group.

578

579 **References**

580 1. Cole SR, Hernán MA. Adjusted survival curves with inverse probability weights.
581 Comput Methods Programs Biomed. 2004;75:45–9.
582 2. Xie J, Liu C. Adjusted Kaplan-Meier estimator and log-rank test with inverse
583 probability of treatment weighting for survival data. Stat Med. 2005;24:3089–110.
584 3. Shinozaki T, Matsuyama Y. Doubly robust estimation of standardized risk difference
585 and ratio in the exposed population. Epidemiology. 2015;26:873–7.
586 4. Komukai S, Hattori S. Doubly robust estimator for net survival rate in analyses of
587 cancer registry data. Biometrics. 2017;73:124–33.
588 5. Nakamura H, Arakawa K, Itakura H, Kitabatake A, Goto Y, Toyota T, et al. Primary
589 prevention of cardiovascular disease with pravastatin in Japan (MEGA Study): a
590 prospective randomised controlled trial. Lancet. 2006;368:1155–63.
591 6. Kleinbaum DG, Klein M. Survival Analysis—A Self Learning Text. 2nd edn. New
592 York: Springer; 2005.

593 7. Robins JM, Finkelstein DM. Correcting for noncompliance and dependent censoring
594 in an AIDS Clinical Trial with inverse probability of censoring weighted (IPCW) log-
595 rank tests. Biometrics. 2000;56:779–88. ~~doi:DOI 10.1111/j.0006-341X.2000.00779.x.~~
596 8. Stitelman OM, De Gruttola V, Van Der Laan MJ. A general implementation of
597 TMLE for longitudinal data applied to causal inference in survival analysis. Int J
598 Biostat. ~~2014~~2012;18.
599 9. Bang H, Robins JM. Doubly robust estimation in missing data and causal inference
600 models. Biometrics. 2005;61:962–72.
601 10. Schnitzer ME, van der Laan MJ, Moodie EEM, Platt RW. Effect of breastfeeding on
602 gastrointestinal infection in infants : a targeted maximum likelihood. Ann Appl Stat.
603 2014;8:703–25.
604 11. Petersen M, Schwab J, Gruber S, Blaser N, Schomaker M, van der Laan MJ.
605 Targeted maximum likelihood estimation for dynamic and static longitudinal marginal
606 structural working models. J Causal Inference. 201~~5~~4;2:147–85.
607 12. Schnitzer ME, Lok JJ, Bosch RJ. Double robust and efficient estimation of a
608 prognostic model for events in the presence of dependent censoring. Biostatistics.
609 2016;17:165–77.
610 13. Robins J. The control of confounding by intermediate variables. Stat Med.

611 1989;8:679–701.

612 14. Westreich D, Edwards JK, Cole SR, Platt RW, Mumford SL, Schisterman EF.

613 Imputation approaches for potential outcomes in causal inference. Int J Epidemiol.

614 2015;44:1731–7.

615 15. Wang A, Nianogo RA, Arah OA. G-computation of average treatment effects on the

616 treated and the untreated. BMC Med Res Methodol. 2017;17:1–5. ~~doi:10.1186/s12874-~~

617 ~~016-0282-4.~~

618 16. Taubman SL, Robins JM, Mittleman MA, Herna MA. Intervening on risk factors for

619 coronary heart disease : an application of the parametric g-formula. Int J Epidemiol.

620 2009;~~April~~38:1599–611.

621 17. Lunceford JK, Davidian M. Stratifcation and weighting via the propensity score in

622 estimation of causal treatment effects : a comparative study. Stat Med. 2004;~~2960–~~

623 ~~April~~23:2937–60.

624 18. Tsiatis AA. Semiparametric Theory and Missing Data. New York: Springer; 2006.

625 19. Kang JDY, Schafer JL. Demystifying double robustness: a comparison of

626 alternative strategies for estimating a population mean from incomplete data. Stat Sci.

627 2007;22:523–39.~~ doi:10.1214/07-STS227.~~

628 20. Satten GA, Datta S. The Kaplan-Meier estimator as an inverse-probability-of-

censoring weighted average. Am Stat. 2001;55:207–10.

21. Westreich D, Cole SR, Young JG, Palella F, Tien PC, Kingsley L, et al. The parametric g-formula to estimate the effect of highly active antiretroviral therapy on incident AIDS or death. Stat Med. 2012;31:2000–9.

22. Young JG, Cain LE, Robins JM, O'Reilly EJ, Hernán MA. Comparative effectiveness of dynamic treatment regimes: an application of the parametric g-formula. Stat Biosci. 2011;3:119–43.

23. Efron B, Tibshirani RJ. An Introduction to the Bootstrap. New York: Chapman and Hall; 1993.

~~24. Zhang M, Schaubel DE. Contrasting treatment-specific survival using double-robust estimators. Stat Med. 2012;31:4255–68.~~

~~25. Bai X, Tsiatis AA, O'Brien SM. Doubly-robust estimators of treatment-specific survival distributions in observational studies with stratified sampling. Biometrics. 2013;69:830–9.~~

~~264. Tan Z. Comment: understanding OR, PS and DR. Stat Sci. 2007;22:560-568.~~

~~275. Hernán MA, Robins JM. Causal Inference: What If. Boca Raton: Chapman & Hall/CRC; 2020.~~

24~~64~~. Zhang M, Schaubel DE. Contrasting treatment-specific survival using double-robust estimators. Stat Med. 2012;31:4255–68.

25~~75~~. Bai X, Tsiatis AA, O'Brien SM. Doubly-robust estimators of treatment-specific survival distributions in observational studies with stratified sampling. Biometrics. 2013;69:830–9.

26. Tan Z. Comment: understanding OR, PS and DR. Stat Sci. 2007;22:560-568.

27. Hernán MA, Robins JM. Causal Inference: What If. Boca Raton: Chapman & Hall/CRC; 2020.

28. Yoshida M, Matsuyama Y, Ohashi Y, for the MEGA Study Group. Estimation of treatment effect adjusting for dependent censoring using the IPCW method: an application to a large primary prevention study for coronary events (MEGA study). Clin Trials. 2007;4:318-28.

Table 2. Simulation results

Estimator	Model specification			Bias (× 100) at $t = 3$			Bias (× 100) at $t =$ Relative efficiency5y		
	Censoring	Event	Covariate	Control	Test	Log of Risk ratio	Control	Test	Log of Risk ratio
Scenario 1: 30% censoring in both control and test groups.									
IPCW	Correct	—	—	0.0-0.10.1	0.00.00.0	-0.2-0.40.8	- 0.11.001.00	- 0.01.001.00	- 0.41.001.00
Kaplan–Meier	Incorrect								
Kaplan–Meier				**0.51.9**5.3	**0.41.5**2.3	-0.2-0.912.2	**1.9**	**1.5**	-0.9
Parametric	—	Correct	Correct	0.0 (1.23)0.00.0	0.0 (1.22)0.00.0	-0.2 (1.23)- 0.30.2	0.0 (1.10)1.20	0.0 (1.081.04)	-0.3 (1.091.17)
g-formula	(relative efficiency)								
g-formula		Correct	Incorrect	0.1**1.2**2.1	0.1**0.4**2.7	0.1**3.9** 11.6	**1.2**	**0.4**	**3.9**
		Incorrect	Correct	**0.52.0**5.2	0.1**1.7**2.2	**3.4** 2.2211.8	**2.0**	**1.7**	-2.2

		Incorrect	Incorrect	**0.5 1.9** 5.1	0.1 **1.7** 2.4	**3.4** 2.4 9.8	**1.9**	**1.7**	-2.4
Proposed doubly robust	Correct	Correct	Correct	-0.1 (1.04) 0.3 0.1	0.0 (1.01) 0.1 0.1	-0.3 (1.02) 0.8 0.2	-0.3 (1.00) 1.09	-0.1 (0.99 1.05)	-0.8 (1.01 1.05)
		(relative efficiency)							
~~doubly robust~~		Correct	Incorrect	0.0 0.2 0.1	0.0 0.0 0.0	-0.1 0.8 0.5	0.2	0.0	0.8
		Incorrect	Correct	0.0 0.0 0.2	0.0 0.0 0.1	-0.2 -0.4 0.4	0.0	0.0	-0.4
		Incorrect	Incorrect	0.0 0.0 0.1	0.0 0.0 0.1	-0.2 -0.4 0.2	0.0	0.0	-0.4
	Incorrect	Correct	Correct	-0.1 -0.3 0.1	0.0 -0.1 0.0	-0.2 -0.7 0.5	-0.3	-0.1	-0.7
		Correct	Incorrect	0.0 0.2 0.4	0.0 0.0 0.0	-0.1 0.9 1.6	0.2	0.0	0.9
		Incorrect	Correct	**0.5 1.9** 1.3	**0.4 1.5** 0.5	-0.2 -1.0 3.6	**1.9**	**1.5**	-1.0
		Incorrect	Incorrect	**0.5 1.9** 1.3	**0.4 1.6** 0.4	-0.2 -0.9 4.4	**1.9**	**1.6**	-0.9
				Scenario 2: 20% censoring in both control and test groups.					
IPCW	Correct	—	—	0.0 0.1	0.0 0.0	-0.3 0.3	-0.1 1.00	0.0 1.00	-0.3 1.00

~~Kaplan–~~ ~~Meier~~Kaplan– Meier	Incorrect			**0.3~~1.3~~**	**0.2~~1.0~~**	-0.2~~-0.5~~	**1.3**	**1.0**	-0.5
Parametric g-formula	—	Correct	Correct	0.0 (1.25)~~0.0~~	0.0 (1.24)~~0.0~~	-0.2 (1.26)~~-0.3~~	0.0 (1.06)~~1.06~~	0.0 (1.04)~~1.04~~	-0.3 (1.04)~~1.04~~
~~g-formula~~	(relative efficiency)	Correct	Incorrect	0.0~~1.1~~	0.0~~0.3~~	0.2~~3.9~~	**1.1**	0.3	**3.9**
		Incorrect	Correct	0.1~~1.4~~	**-0.2~~1.2~~**	**4.2**~~-1.4~~	**1.4**	**1.2**	-1.4
		Incorrect	Incorrect	0.1~~1.3~~	**-0.2~~1.2~~**	**4.2**~~-1.6~~	**1.3**	**1.2**	-1.6
Proposed doubly robust	Correct	Correct	Correct	-0.1 (1.03)~~-0.2~~	0.0 (1.03)~~-0.1~~	-0.3 (1.05)~~-0.6~~	-0.2 (1.00)~~1.00~~	-0.1 (0.99)~~0.99~~	-0.6 (1.01)~~1.01~~
~~doubly robust~~	(relative efficiency)	Correct	Incorrect	0.0~~0.1~~	0.0~~0.0~~	-0.2~~0.5~~	0.1	0.0	0.5
		Incorrect	Correct	0.0~~0.0~~	0.0~~0.0~~	-0.3~~-0.3~~	0.0	0.0	-0.3
		Incorrect	Incorrect	0.0~~0.0~~	0.0~~0.0~~	-0.3~~-0.3~~	0.0	0.0	-0.3

	Incorrect	Correct	Correct	-0.1-0.2	0.0-0.1	-0.3-0.4	-0.2	-0.1	-0.4
		Correct	Incorrect	0.00.1	0.00.0	-0.20.5	0.1	0.0	0.5
		Incorrect	Correct	**0.31.3**	**0.21.0**	-0.2-0.5	**1.3**	**1.0**	-0.5
		Incorrect	Incorrect	**0.31.3**	**0.21.0**	-0.2-0.5	**1.3**	**1.0**	-0.5
Scenario 3: 9% censoring in control group and 12% censoring in test group									
IPCW Kaplan–MeierKaplan–Meier	Correct	—	—	0.0-0.1	0.00.0	-0.3-0.3	-0.11.00	0.01.00	-0.31.00
	Incorrect			0.1**0.5**	0.1**0.6**	-0.8-1.6	**0.5**	**0.6**	-1.6
Parametric g-formula g-formula	—	Correct	Correct	0.0 (1.26)0.0	0.0 (1.25)0.0	-0.2 (1.28)-0.3	0.0 (1.03)1.03	0.0 (1.02)1.02	-0.3 (1.02)1.02
	(relative efficiency)								
		Correct	Incorrect	0.0**1.0**	0.00.2	0.1**3.8**	**1.0**	0.2	**3.8**
		Incorrect	Correct	**-0.30.6**	**-0.50.8**	3.4-2.2	**0.6**	**0.8**	-2.2
		Incorrect	Incorrect	**-0.30.6**	**-0.50.7**	3.3-2.4	**0.6**	**0.7**	-2.4

Proposed doubly robust ~~doubly robust~~	Correct	Correct	Correct	0.0 (1.00)~~-0.1~~	0.0 (1.00)~~-0.1~~	-0.3 (1.00)~~-0.4~~	-0.1 (1.00)~~1.00~~	-0.1 (1.00)~~1.00~~	-0.4 (1.00)~~1.00~~
		(relative efficiency)							
		Correct	Incorrect	0.0~~0.0~~	0.0~~0.0~~	-0.3~~0.1~~	0.0	0.0	0.1
		Incorrect	Correct	0.0~~0.0~~	0.0~~0.0~~	-0.3~~-0.3~~	0.0	0.0	-0.3
		Incorrect	Incorrect	0.0~~0.0~~	0.0~~0.0~~	-0.3~~-0.3~~	0.0	0.0	-0.3
	Incorrect	Correct	Correct	0.0~~-0.1~~	0.0~~-0.1~~	-0.3~~0.0~~	-0.1	-0.1	0.0
		Correct	Incorrect	0.0~~0.0~~	0.0~~0.0~~	-0.3~~0.1~~	0.0	0.0	0.1
		Incorrect	Correct	0.1~~**0.5**~~	0.1~~**0.6**~~	-0.8~~-1.6~~	**0.5**	**0.6**	-1.6
		Incorrect	Incorrect	0.1~~**0.5**~~	0.1~~**0.6**~~	-0.8~~-1.6~~	**0.5**	**0.6**	-1.6

Numbers in parentheses are the relative efficiency compared with the IPCW Kaplan–Meier estimate with a correctly specified censoring model. If the bias exceeded half of the standard error of the estimates, the printed bias is shown in bold. True values calculated from a large simulated dataset were (0.89, 0.92, 0.69) (at $t = 3$) and (0.81~~0~~, 0.86, ~~and~~ 0.67~~7~~4) (at $t = 5$) for control group, test group, and risk ratio, respectively. The biases (× 100) from the method assuming the baseline-conditional independent censoring at $t = 5$ for the control and test groups were (0.5, 0.4) (scenario 1), (0.4, 0.3) (scenario 2), and (0.2, 0.2) (scenario 3).