

# On treating right-censoring events like treatments

Lan Wen<sup>1</sup>, Aaron L. Sarvet<sup>2</sup>, and Jessica G. Young<sup>3,4</sup>

<sup>1</sup>*Department of Statistics and Actuarial Science, University of Waterloo*

<sup>2</sup>*Department of Biostatistics & Epidemiology, University of Massachusetts, Amherst, USA*

<sup>3</sup>*Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Health Care Institute, Boston, Massachusetts*

<sup>4</sup>*Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, Massachusetts*

## Abstract

In causal inference literature, potential outcomes are often indexed by the “elimination of all right-censoring events,” leading to the perception that such a restriction is necessary for defining well-posed causal estimands. In this paper, we clarify that this restriction is not required: a well-defined estimand can be formulated without indexing on the elimination of such events. Achieving this requires a more precise classification of right-censoring events than has historically been considered, as the nature of these events has direct implications for identification of the target estimand. We provide a framework that distinguishes different types of right-censoring events from a causal perspective, and demonstrate how this framework relates to censoring definitions and assumptions in classical survival analysis literature. By bridging these perspectives, we provide a clearer understanding of how to handle right-censoring events and provide guidance for identifying causal estimands when right-censored events are present.

## 1 Introduction

Right-censoring events are ubiquitous in randomized trials and observational studies, and they pose significant challenges for estimating the average causal effect of an exposure or treatment on an outcome of interest. The means in which right-censoring events are addressed or handled has important implications for both study design (ICH, 2019) and the validity of subsequent statistical inferences.

In the causal inference literature, a frequent approach to handling right-censoring is to conceptualize an intervention that removes all right-censoring events and by indexing potential outcomes under that hypothetical elimination (Neugebauer *and others*, 2014; Richardson and Rotnitzky, 2014; Young *and others*, 2020; Hernán and Robins, 2020; Wen *and others*, 2021, 2022, 2023). While useful for certain identification arguments, this framing can be misconstrued as asserting that removal of all right-censoring events is necessary to define a well-posed or well-defined causal estimand.

By contrast, classical survival analysis typically does not invoke hypothetical interventions that eliminate right-censoring events. Instead, the standard approach assumes conditionally independent right-censoring (Fleming and Harrington, 1991; Kalbfleisch and Prentice, 2002), an assumption that has likewise been adopted in several causal inference settings (Cai and van der Laan, 2020; Rytgaard *and others*, 2021; Ying *and others*, 2022; Cui *and others*, 2023; Li *and others*, 2025; Bu *and others*, 2025). The differences of approaches (treating right-censoring event as preventable in causal framework, versus assuming conditional independence in classical survival analysis) reflects a conceptual gap that can lead to confusion about what estimands are being defined and when they are identifiable.

In this paper we close that gap. We show that a well-defined causal estimand need not be indexed on the elimination of right-censoring events: one can formulate estimands that do not require positing a joint intervention to remove all of these events. Doing so, however, requires a more careful classification of right-censoring events. Different kinds of right-censoring events (e.g., administrative end-of-study censoring, loss to follow-up due to relocation, or right-censoring that is a direct consequence of treatment) can have distinct causal and identification implications. Consequently, handling these events should be driven by the (causal) estimand of interest: the estimand determines which right-censoring events require elimination and, in turn, which identification and estimation methods are appropriate.

Our contributions are threefold. First, we introduce a framework that distinguishes types of right-censoring events and characterizes their roles relative to the target estimand. Second, we show when and under what assumptions various estimands are identifiable, so that right-censoring events are handled transparently and appropriately for the scientific question at hand. Third, we relate this framework to classical survival-analysis assumptions. In doing so, we show that classical formulations of independent censoring do not explicitly engage with causal models, while modern causal inference approaches often imply that estimands should be defined through interventions on right-censoring events, even when right-censoring events has no causal effect on the outcome. Neither is one-size-fits-all; rather, each provides tools appropriate for different estimands and right-censoring mechanisms.

The remainder of the paper is organized as follows. Section 2 and 3 introduces notation and estimands considered herein. Section 4 gives formal definitions of different right-censoring event types within the causal framework. Section 5 states the identification results under various estimands and discusses practical solutions to violations of assumptions. Section 6 connects these results to classical survival assumptions, and Section 7 illustrates the ideas with examples via simulation studies.

## 2 Longitudinal data structure

Consider a longitudinal study in which measurements on a random sample of individuals from a population of interest are collected over a follow-up period. Let  $k=0, \dots, K$  denote equally spaced intervals within this follow-up, where  $k=0$  the baseline interval and  $K$  the last interval in which individual measurements are available. The length of these intervals will depend on the study design. For instance, in a study using electronic health records with dates of visits, services, labs, or prescriptions,  $k$  might denote a single day of the follow-up.

By contrast, in a trial or cohort study with scheduled follow-up visits no more frequent than weekly,  $k$  would represent a week.

Following [Robins and Richardson \(2010\)](#), we define *factual (or realized) variables* as those that could, in principle, be recorded in the actual world for subjects participating in the study. In turn, let  $\mathbf{L}_k$  denote a vector of time-varying covariates relative to time interval  $k$ ,  $A_k$  the *treatment of substantive interest* to the investigator in time interval  $k$ , and  $Y_k$  an indicator of survival by time interval  $k$  ( $Y_k=1$  if the individual remains alive at time  $k$  and  $Y_k=0$  otherwise), for  $k \geq 0$ . We will use the convention of an overline and underline to denote history and future of a random variable, respectively; e.g.,  $\overline{Y}_k=(Y_0, Y_1, \dots, Y_k)$  and  $\underline{Y}_{k+1}=(Y_{k+1}, Y_{k+1}, \dots, Y_K)$ . Correspondingly, let  $C_k$  denote an indicator of an event by time  $k$  ( $k=1, \dots, K$ ) that terminates the observation process of realized variables such that all information on the participant is unknown to the investigator (i.e. missing) beyond time  $k$ . In general, there can be multiple of these events such that  $\mathbf{C}_k=(C_{k1}, C_{k2}, \dots)$ , for  $k=1, \dots, K$ , each of which may affect the outcome differently or not at all. We refer to such events collectively as “*observation-terminating events*”.

Examples of observation-terminating events include deletion of a study app from a participant’s electronic device ([Adu and others, 2020](#); [Amagai and others, 2022](#)), or a formal withdrawal from a trial ([Gabriel and Mercado, 2011](#); [Sofrygin and others, 2019](#); [Wilson and others, 2021](#)), after which no further data are recorded for that individual. Events arising from designs analogous to Type I and Type II stopping rules in survival analysis ([Fleming and Harrington, 1991](#); [Kalbfleisch and Prentice, 2002](#); [Andersen and others, 2012](#); [Klein and others, 2014](#)) likewise constitute observation-terminating events. Throughout, we will rely on the assumption that either the temporal ordering  $(\mathbf{C}_k, Y_k, \mathbf{L}_k, A_k)$  holds in each interval  $k$  or that the length of each interval  $k$  is selected to be so small that there is no dependence between events within an interval (akin to the “no ties” assumption in survival analysis). We define  $\mathbf{C}_0 \equiv Y_0 \equiv 0$ ; that is, we restrict the population to those who are alive and have not yet experienced an event that precludes observation of future survival processes.

By the above, we can summarize the observed data for any individual who remains alive and under observation through the end of the study period (i.e.,  $\overline{\mathbf{C}}_K=0, \overline{Y}_K=1$ ) as a vector of factual variables given by  $(\overline{\mathbf{L}}_{K-1}, \overline{A}_{K-1}, \overline{\mathbf{C}}_K, \overline{Y}_K)$ . For an individual who died in an interval  $k < K$  without any observation-terminating event occurring prior to that time (i.e.,  $\overline{\mathbf{C}}_k=0, \overline{Y}_{k-1}=1, \underline{Y}_k=0$ ), we observe  $(\overline{\mathbf{L}}_{k-1}, \overline{A}_{k-1}, \overline{\mathbf{C}}_K, \overline{Y}_K)$ . Otherwise, for an individual whose observation was terminated before prior to  $k$  while still alive (i.e.,  $\overline{Y}_{k-1}=1, \overline{\mathbf{C}}_{k-1}=0, C_k=1$ ), we observe  $(\overline{\mathbf{L}}_{k-1}, \overline{A}_{k-1}, \overline{\mathbf{C}}_K, \overline{Y}_{k-1})$ .

### 3 Two distinct estimands

Suppose the investigators who conducted the study described above communicated their goal as interest in the “cumulative survival probability by some  $k$  under treatment regime  $g$ ”, where  $g=(g_0, \dots, g_{K-1})$  denotes a deterministic sequence of interventions on  $A_k$  (see [Young and others, 2011, 2019](#); [Wen and others, 2023](#) for a formal definition). More generally, we can formally define an average causal effect as a contrast in average potential outcomes between two regimes  $g=g'$  and  $g=g''$ . For ease of exposition, we consider the average

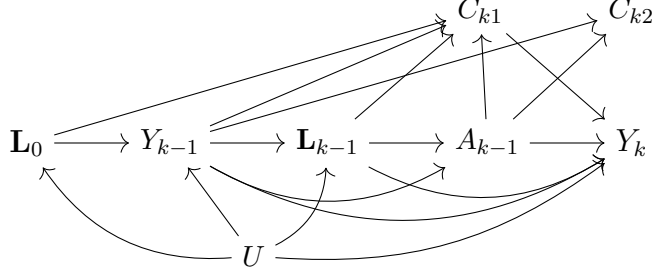


Figure 1: Directed Acyclic Graph (DAG) depicting the factual data. For simplicity, we define two observation-terminating event  $\mathbf{C}_k = (C_{k1}, C_{k2})$ . Note that we omit other arrows (e.g.,  $\mathbf{L}_0$  to  $Y_{k-1}$ ) to reduce clutter, as adding omitted edges from past to future variables does not affect our assumptions.

causal effect given by a contrast between the cumulative probabilities of survival under two treatment regimes  $g' = (g'_0, \dots, g'_{K-1})$  and  $g'' = (g''_0, \dots, g''_{K-1})$ , where  $g'_k = a'_k$  and  $g''_k = a''_k$  for each  $k = 0, \dots, K-1$ , with  $\bar{a}'_K \neq \bar{a}''_K$ . These results extend to any treatment strategies  $g$  that may depend on past covariates.

For any  $k = 1, \dots, K$ , first consider the potential outcome  $Y_k^{g, \bar{\mathbf{c}}=0}$ , the survival status an individual would have at time  $k$  had that individual followed treatment regime  $g$  and, further, had (somehow) all observation-terminating events been eliminated through  $k$ . This potential outcome can be more explicitly defined relative to a causal model. Figure 1 is a causal directed acyclic graph (DAG) depicting a possible underlying dependency structure among the factual variables in this study. Throughout we will assume that statistical independencies in the data are faithful to the graph under study (Verma and Pearl, 2022).

Based on the exposition above, the Single World Intervention Graph (SWIG; Richardson and Robins, 2013) in Figure 2 is a node-splitting transformation of the causal DAG in Figure 1, explicitly depicting  $(g, \bar{\mathbf{c}})$  potential variables. In turn, one possible formalization of the “cumulative survival probability by some  $k$  under treatment regime  $g$ ” is given by:

$$\Psi_1^g := E(Y_k^{g, \bar{\mathbf{c}}=0}) \quad (1)$$

Alternatively, consider the potential outcome  $Y_k^g$ , the survival status an individual would have at time  $k$  had that individual followed treatment regime  $g$  but with no additional interventions to eliminate observation-terminating events  $\mathbf{C}_k$ . This distinct potential outcome is more explicitly represented by the SWIG in Figure 3, which clarifies that  $Y_k^g$  corresponds to then individual’s survival status at  $k$  had they followed treatment regime  $g$  and had they experienced their *natural* observation termination process under  $g$  (Richardson and Robins, 2013; Young and others, 2020). Correspondingly, an alternative formalization of the “cumulative survival probability by  $k$  under treatment regime  $g$ ” is a given by:

$$\Psi_2^g := E(Y_k^g) \quad (2)$$

The estimands (1) and (2) are generally not equal. As shown in Section 5, they coincide in two special cases: (i) when no individual’s observation process is terminated before time  $K$ ;

or (ii) when the observation-terminating events in  $\mathbf{C}_k$  have no direct causal effect on  $\underline{Y}_k$ . The latter corresponds to the setting illustrated in Figure 1 if, like  $C_{k2}$ ,  $C_{k1}$  has no arrow directed into  $Y_k$ .

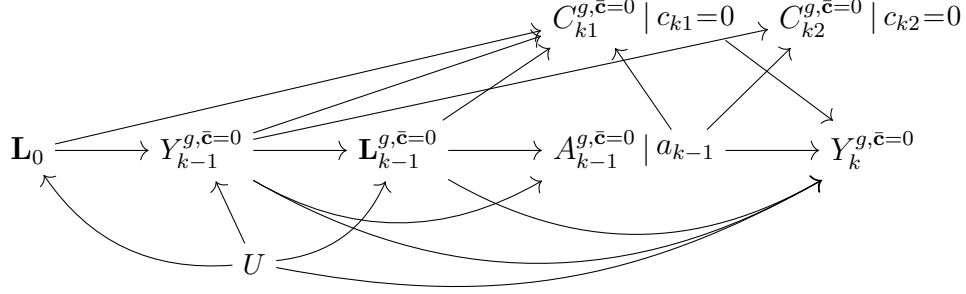


Figure 2: Single World Intervention Graph (SWIG) depicting  $(g, \bar{c})$  potential variables (counterfactuals). For simplicity, we define two observation-terminating event  $\mathbf{C}_k = (C_{k1}, C_{k2})$ . Note that we omit other arrows (e.g.,  $\mathbf{L}_0$  to  $Y_{k-1}$ ) to reduce clutter, as adding omitted edges from past to future variables does not affect our assumptions.

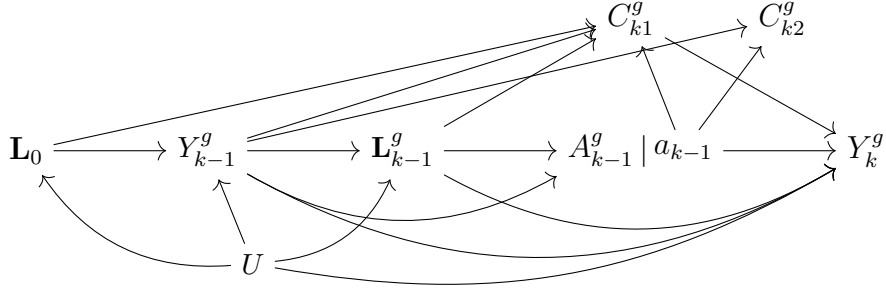


Figure 3: Single World Intervention Graph (SWIG) depicting  $g$  potential variables (counterfactuals). For simplicity, we define two observation-terminating event  $\mathbf{C}_k = (C_{k1}, C_{k2})$ . Note that we omit other arrows (e.g.,  $\mathbf{L}_0$  to  $Y_{k-1}$ ) to reduce clutter, as adding omitted edges from past to future variables does not affect our assumptions.

## 4 Defining and classifying right-censoring events

To classify the observation-terminating events in  $\mathbf{C}_k$  ( $k=1, \dots, K$ ) introduced in the previous section, we draw on the definition of a “right-censoring event” introduced in [Young and others \(2020\)](#):

“A *[right-]censoring event* is any event occurring in the study by [time interval]  $k$ ,  $k=1, \dots, K$ , that ensures the values of *all future potential outcomes of interest* under [a point treatment level]  $a$  are unknown even for an individual receiving the intervention  $a$ .”

Referring back to the estimands in Section 3, as noted by [Young and others \(2020\)](#), events in  $\mathbf{C}_k$  meets this definition of a right-censoring event: whenever any event in  $\mathbf{C}_k$  occurs by time

$k$ , it prevents us from observing all subsequent potential outcomes under the point treatment level  $a$  for that individual. As discussed later, how these events are addressed depends on the estimand of interest. Building on the work of [Young and others \(2020\)](#), we propose a general definition of right-censoring that applies uniformly to any outcome of interest and that facilitates the formulation of clear identification assumptions.

#### 4.1 A unified definition of right-censoring event

The definition proposed by [Young and others \(2020\)](#) is inherently tied to the outcomes of interest and requires that the causal estimand be clearly specified prior to data analysis. As a result, the choice of estimand automatically identifies which events are treated as part of the censoring mechanism and how they should be handled. Adopting this perspective, we present a more general definition of right-censoring event that is not limited to potential outcomes of interest under  $a$ :

**Definition 1** (Generalized definition of right-censoring event). *For any subject in a study, a right-censoring event refers to any event occurring in the study by time interval  $k$ , for  $k=1, \dots, K$ , that renders the values of their future outcomes of interest unobservable.*

To highlight the differences between  $\Psi_1^g$  and  $\Psi_2^g$ , we further classify different right-censoring events into *nuisance right-censoring events*, and all other right-censoring events as *non-nuisance right-censoring events*.

**Definition 2** (Nuisance and non-nuisance right-censoring events). *Any right-censoring event in which the investigator’s causal question or ideal trial does not invoke the elimination or prevention of the event is referred to as a “nuisance right-censoring event.” All other censoring mechanisms—those that are integral to the definition of the estimand—will be considered “non-nuisance right-censoring events”.*

An example of non-nuisance right-censoring event is non-adherence to  $g$ , when the outcome of interest is defined under a regime where non-adherence is entirely prevented. In contrast, nuisance right-censoring events include those such as loss to follow-up or study withdrawal, when those events are not part of the intervention requiring their elimination. In light of our definitions, let  $\Delta_k$  denote the *non-observation state* ( $\Delta_k=1$  if  $\underline{Y}_k$  is unobservable and 0 otherwise) in time interval  $k$ ,  $k=1, \dots, K$ . Specifically,

$$\Delta_k = \max(\mathbf{C}_k).$$

Figure 4 extends the DAG in Figure 1 to include  $\Delta_k$ . Non-nuisance right-censoring events should be defined in terms of the substantive causal question/story and the specific hypothetical interventions that would prevent them; only with such specification does it make sense to invoke a consistency assumption that links potential outcomes under intervention on those events to the factual data ([Young, 2024](#)). In contrast, nuisance right-censoring events are not targeted by the estimand, and identification proceeds without invoking such a consistency assumption for each nuisance event.

Since events in  $\mathbf{C}_k$  are right-censoring events, each component of  $\mathbf{C}_k$  is classified as belonging either to the set of nuisance right-censoring events or to the set of non-nuisance right-

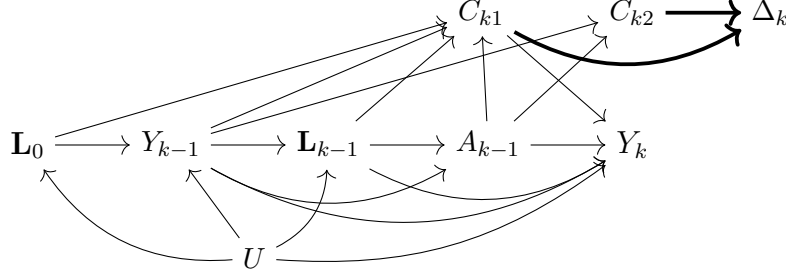


Figure 4: Directed Acyclic Graph (DAG) depicting the factual data, with right-censoring state added to Figure 1. Bold arrows are used to emphasize deterministic relationships assumed herein.

censoring events, relative the the estimand of interest (see Figure 5 and the next section for a detailed discussion).

**Proposition 1.** *All nuisance right-censoring events are a subset of the set of observation-terminating events  $\mathbf{C}_k$ .*

A proof can be found in Appendix B. Proposition 1 states that if there is a nuisance right-censoring event, then this event has to be a observation-terminating event (i.e., part of  $\mathbf{C}_k$ ). Moreover, if all observation-terminating events in the study are nuisance right-censoring events, it is sufficient to specify assumptions about the non-observation state  $\Delta_k$  itself – such as conditional independent  $\Delta_k$  (see Sections 5 and 6) – to achieve identification.

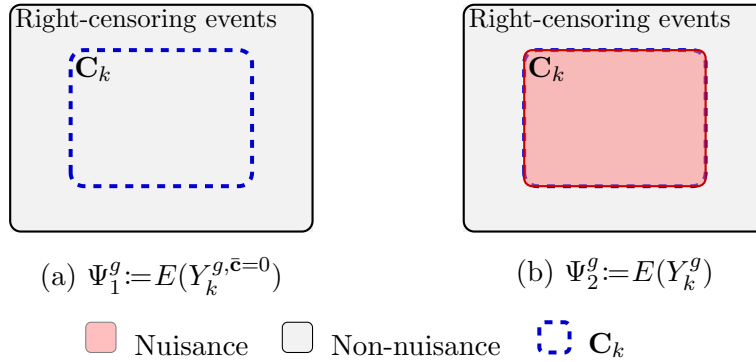


Figure 5: Classification of right-censoring events relative to the estimands considered herein.

#### 4.2 Relating examples from Section 3 to the right-censoring event definition

Returning to the estimands in Section 3, suppose for concreteness that the study involves two disjoint observation-terminating events, so that  $\mathbf{C}_k = (C_{k1}, C_{k2})$ . Based on this setup, consider the two estimands we defined in Section 3:

*Example 1.*  $\Psi_1^g := E(Y_k^{g, \bar{c}=0})$ ;

*Example 2.*  $\Psi_2^g := E(Y_k^g)$ .

In the next section we study identification of  $\Psi_1^g$  in the presence of only non-nuisance right-censoring events. The subsequent section considers identification of  $\Psi_2^g$  in the complementary setting with nuisance right-censoring events.

## 5 Identification of causal estimands

### 5.1 Example 1. Identification in the presence of only non-nuisance right-censoring events

In Example 1, the outcome of interest is a potential variable defined under a joint intervention  $(g, \bar{c}=0)$ . More precisely,  $\Psi_1^g$  is the cumulative survival probability at time  $k$  if every individual followed treatment regime  $g$  and all observation-terminating events ( $C_{k1}$  and  $C_{k2}$ ) were eliminated. In this example the non-nuisance right-censoring events at time  $k$  include  $I(A_{k-1} \neq g_{k-1})$ ,  $C_{k1}$  and  $C_{k2}$ ; there are no nuisance right-censoring events. For this estimand,  $I(A_{k-1} \neq g_{k-1})$  is a non-nuisance right-censoring event at time  $k$  because, for any individual whose treatment history at time  $k$  does not equal  $\bar{g}_{k-1}$ , their future potential outcomes  $\underline{Y}_k^g$  are unobservable and therefore  $\underline{Y}_k^{g, \bar{c}=0}$  are unobservable. Similar reasoning can be made about  $C_{k1}$  and  $C_{k2}$ .

At this stage it is helpful to describe a hypothetical trial whose unadjusted outcome contrast can be interpreted as the causal estimand of interest. Doing so not only clarifies the objective and makes the estimand explicit (Young, 2024; Moreno-Betancur and others, 2024), but also helps differentiate it from other possible estimands. Investigators aiming to identify  $\Psi_1^g$  would describe a trial in which all participants are perfectly randomized to one of two treatment regimes,  $g'$  and  $g''$ . In each arm every participant adheres fully to their assigned regime until the end of follow-up, and any observation-terminating events ( $C_{k1}$  and  $C_{k2}$ ) are prevented. In other words, all non-nuisance right-censoring events (as defined above) are eliminated. We now turn to the formal identification of  $\Psi_1^g$ .

#### 5.1.1 Identifying assumptions for $\Psi_1^g$

Consider the DAG in Figure 1 and the corresponding SWIG in Figure 2, which represent the intervention defined in Example 1. For  $k=1, \dots, K$ , we consider the following set of assumptions that are sufficient for identification of  $\Psi_1^g := E(Y_k^{g, \bar{c}=0})$ :

*Assumption 1.1* (Conditional  $\mathbf{C}_k$ -exchangeability):

$$\underline{Y}_k^{g, \bar{c}=0} \perp\!\!\!\perp \mathbf{C}_k \mid (\bar{\mathbf{L}}_{k-1}, \bar{A}_{k-1} = \bar{g}_{k-1}, \bar{Y}_{k-1} = 1, \bar{\mathbf{C}}_{k-1} = 0); \quad (3)$$

*Assumption 1.2* (Conditional treatment exchangeability):

$$\underline{Y}_k^{g, \bar{c}=0} \perp\!\!\!\perp A_k \mid (\bar{\mathbf{L}}_{k-1}, \bar{A}_{k-1} = \bar{g}_{k-1}, \bar{Y}_{k-1} = 1, \bar{\mathbf{C}}_k = 0);$$

*Assumption 1.3* (Consistency):  $(\bar{\mathbf{L}}_k^{g, \bar{c}=0}, \bar{Y}_k^{g, \bar{c}=0}) = (\bar{\mathbf{L}}_k, \bar{Y}_k)$  if  $\bar{A}_{k-1} = \bar{g}_{k-1}$  and  $\bar{\mathbf{C}}_k = 0$ .

*Assumption 1.4* (Positivity):  $P(\bar{A}_{k-2} = \bar{g}_{k-2}, \bar{\mathbf{L}}_{k-1} = \bar{\mathbf{l}}_{k-1}, \bar{Y}_{k-1} = 1, \bar{\Delta}_{k-1} = 0) > 0 \longrightarrow$

$$P(A_{k-1} = g_{k-1}, \Delta_k = 0 \mid \bar{Y}_{k-1} = 1, \bar{\Delta}_{k-1} = 0, \bar{A}_{k-2} = \bar{g}_{k-2}, \bar{\mathbf{L}}_{k-1} = \bar{\mathbf{l}}_{k-1}) > 0.$$

**Definition 3** (Model  $\mathcal{M}_1$ ). Let  $\mathcal{M}_1$  denote a model that satisfy Assumptions 1.1–1.4.

**Proposition 2.** Under models  $\mathcal{M}_1$ ,  $\Psi_1^g$  is identified via equation (4):

$$\int_{\bar{\mathbf{l}}_{k-1}} E(Y_k | \bar{Y}_{k-1}=1, \bar{A}_{k-1}=\bar{g}_{k-1}, \bar{\Delta}_k=0, \bar{\mathbf{L}}_{k-1}=\bar{\mathbf{l}}_{k-1}) \left[ \prod_{j=1}^{k-1} dP(\mathbf{l}_j | \bar{\mathbf{l}}_{j-1}, \bar{Y}_j=1, \bar{A}_{j-1}=\bar{g}_{j-1}, \bar{\Delta}_j=0) \right. \\ \left. P(Y_j=1 | \bar{Y}_{j-1}=1, \bar{A}_{j-1}=\bar{g}_{j-1}, \bar{\Delta}_j=0, \bar{\mathbf{L}}_{j-1}=\bar{\mathbf{l}}_{j-1}) \right] dP(\mathbf{l}_0). \quad (4)$$

In the example above, the variables  $C_{k1}$  and  $C_{k2}$  should be defined with respect to the substantive causal question that would necessarily involve joint interventions on  $C_{k1}$  and  $C_{k2}$ . This substantive question would generally differ from the investigator’s original aim described at the start of Section 3. Although our derivation does not require direct observation of the values of  $C_{k1}$  and  $C_{k2}$ , it relies on knowing the nature of the censoring mechanisms to properly adjust for their shared causes with  $Y_k$ ; the same consideration applies to the next estimand.

## 5.2 Example 2. Identification in the presence of nuisance right-censoring events

In Example 2, the outcome of interest is a potential variable defined under intervention  $g$ . More precisely,  $\Psi_2^g$  is the cumulative survival probability at time  $k$  if every individual followed treatment regime  $g$ , and corresponds to the substantive target specified by the investigator at the beginning of Section 3. Implicit in this interpretation is that observation-terminating events  $\mathbf{C}_k$  are not intervened upon. In this setting, the non-nuisance right-censoring event at time  $k$  is  $I(A_{k-1} \neq g_{k-1})$ ; nuisance right-censoring events are  $C_{k1}$  and  $C_{k2}$ .

Investigators aiming to identify  $\Psi_2^g$  would describe a hypothetical trial in which all participants are perfectly randomized to one of two treatment regimes,  $g'$  and  $g''$ , and in each arm every participant adheres fully to their assigned regime throughout follow-up. In this trial, all outcomes are eventually observed, regardless of whether participants discontinue treatment, relocate, or otherwise become unreachable through routine follow-up. For individuals for whom an observation-terminating event in  $\mathbf{C}_k$  occurs, outcomes that are unobserved during the study are nonetheless recoverable post hoc (e.g., via linkage to national death indices; comprehensive digital trace and physical record mining, if resources are unlimited) even though events in  $\mathbf{C}_k$  are not eliminated. Implicitly, this corresponds to an intervention on the non-observation state  $\Delta_k$  which does not directly affect  $Y_k$  (i.e.,  $Y_k = Y_k^{\bar{\delta}=0} = Y_k^{\bar{\delta}=0, \bar{\mathbf{C}}}$ ; see Figure C.1 in Appendix C). As a result, the outcome becomes observable through mechanisms that are independent of any observation-terminating events.

### 5.2.1 Identifying assumptions for $\Psi_2^g$

Now consider the same DAG in Figure 1 together with the SWIG in Figure 3, corresponding to the intervention in Example 2.

Suppose that, in Figure 1, there are no arrows from  $(C_{k1}, C_{k2})$  into  $Y_k$  at any time point  $k$ . Given this structure, we posit the following assumptions for each  $k=1, \dots, K$ , which are sufficient for identification of  $\Psi_2^g := E(Y_k^g)$ :

*Assumption 2.1a* (Conditional independent  $\Delta_k$ ):

$$\underline{Y}_k^g \perp\!\!\!\perp \Delta_k | (\bar{\mathbf{L}}_{k-1}, \bar{A}_{k-1} = \bar{g}_{k-1}, \bar{Y}_{k-1} = 1, \bar{\Delta}_{k-1} = 0); \quad (5)$$

In Section 6, we introduce a related but distinct conditional independent  $\Delta_k$  assumption based on factual variables, which is closely connected to the classical survival assumption of conditional independent censoring.

*Assumption 2.2a* (Conditional treatment exchangeability):

$$\underline{Y}_k^g \perp\!\!\!\perp A_k | (\bar{\mathbf{L}}_{k-1}, \bar{A}_{k-1} = \bar{g}_{k-1}, \bar{Y}_{k-1} = 1, \bar{\Delta}_k = 0);$$

*Assumption 2.3* (Consistency):  $(\bar{\mathbf{L}}_k^g, \bar{\mathbf{C}}_k^g, \bar{Y}_k^g) = (\bar{\mathbf{L}}_k, \bar{\mathbf{C}}_k, \bar{Y}_k)$  if  $\bar{A}_{k-1} = \bar{g}_{k-1}$ .

**Definition 4** (Model  $\mathcal{M}_2$ ). Let  $\mathcal{M}_2$  denote a model that satisfy right-censoring event irrelevance and Assumptions 1.4, 2.1a, 2.2a, 2.3.

**Proposition 3.** Under models  $\mathcal{M}_2$ ,  $\Psi_2^g$  is identified via equation (4). Under models  $\mathcal{M}_1 \cup \mathcal{M}_2$ ,  $\Psi_1^g$  is identified via equation (4).

Conditional independent  $\Delta_k$  implies that no event in  $\mathbf{C}_k$  can exert a direct effect on the outcome of interest (see a proof in Appendix D), including any effects mediated through  $\mathbf{L}_k$ . Consequently under model  $\mathcal{M}_2$ , both estimands are identified and yield the same value (i.e.,  $\Psi_1^g = \Psi_2^g$ ). By contrast, when the full-data law  $P$  lies in  $\mathcal{M}_1$  but not in  $\mathcal{M}_2$ , we generally expect  $\Psi_1^g \neq \Psi_2^g$  (see Section 7 for a further discussion). Importantly, even when  $P \notin \mathcal{M}_2$ ,  $\Psi_2^g$  remains a well-defined target of inference, although it is no longer identified by (4). In particular, when conditional independent  $\Delta_k$  in  $\mathcal{M}_2$  does *not* hold, then  $\Psi_2^g$  is not readily identified without making further assumptions. We discuss possible approaches below.

### 5.3 Generalization and summary of estimand identification

Thus far, we have considered two estimands  $\Psi_1^g$  and  $\Psi_2^g$  which are both based on potential outcomes. In Web Appendix A–D, we extend this to more general estimands, including those based on factual outcomes as well as estimands arising from randomized controlled trials. There we continue to use our classification of right-censoring events to derive clear assumptions and identifying formulas from the observed data. In Appendix F, we propose an extended definition of a right-censoring event to also capture events by time  $k$  after which the investigator deems an individual’s data unreliable or unusable, even if such events do not strictly prevent observation of future outcomes of interest.

We have outlined sufficient conditions under which the various estimands considered in this work are all identified by the same formula given in (4) (see Table 1 for a brief overview). In general, the choice of causal estimand should be guided by the underlying causal question. Even when some of the identification conditions are not satisfied, identification may still be achievable through the use of proxy variables (see [Ying and others, 2022](#); [Ying, 2024](#) for a proposal in a non-causal setting). In Web Appendix F, we present an approach that leverages proxy data to identify  $\Psi_2^g$  when conditional independent  $\Delta_k$  is violated for special point treatment settings. Alternatively, one can conduct sensitivity analyses to quantify the

extent to which conclusions might change under deviations from those assumptions to assess the robustness of the findings. We present one such analysis in Appendix E.

| <i>Models</i>   | <i>Estimand</i>                                             |                                                      |
|-----------------|-------------------------------------------------------------|------------------------------------------------------|
| $\mathcal{M}_1$ | $\Psi_1^g := E(Y_k^{g, \mathbf{c}=0})$<br>Identified by (4) | $\neq \Psi_2^g := E(Y_k^g)$<br>Not identified by (4) |
| $\mathcal{M}_2$ | $\Psi_1^g := E(Y_k^{g, \mathbf{c}=0})$<br>Identified by (4) | $= \Psi_2^g := E(Y_k^g)$<br>Identified by (4)        |

Table 1: Identification results under different models defined herein.

## 6 Connection to classical survival analysis

Classical survival analysis literature typically do not distinguish between types of right-censoring events; instead, assumptions are usually stated in terms of being right-censored. As described in [Andersen and others \(2012\)](#) (see also [Klein and Moeschberger, 2003](#); [Klein and others, 2014](#); [Tutz and others, 2016](#)):

“The best known example of such *incomplete observation* is *right-censoring* in classical survival analysis; here, not all of a set of independent lifetimes are observed, so that for some of them it is only known that they are larger than some specific value.”

This passage reflects a traditional survival-analysis view, which regards right-censoring as an observed data feature (the state of being right-censored) and formulates assumptions directly on that state. Our definition of right-censoring events provides insight into classical assumptions, namely, *conditional independent censoring*. In particular, a connection between this classical assumption and the set of assumptions we introduced above emerges when the events in  $\mathbf{C}_k$  are nuisance right-censoring events relative to a given estimand.

Specifically, consider the estimand  $\Psi_2^g$  whose identification we discussed in Section 5.2.1. Rather than imposing the condition in (5), we can formulate the following version of conditional independence for  $\Delta_k$  based solely on factual variables, which can also be used to identify  $\Psi_2^g$ . For each  $k=1, \dots, K$ , consider the following assumption:

*Assumption 2.1b.* (Conditional independent  $\Delta_k$ ):

$$(\underline{Y}_k, \underline{\mathbf{L}}_k, \underline{A}_k) \perp\!\!\!\perp \Delta_k \mid (\bar{\mathbf{L}}_{k-1}, \bar{A}_{k-1}, \bar{Y}_{k-1}, \bar{\Delta}_{k-1}=0). \quad (6)$$

By induction and the monotonicity of  $\Delta_k$ , (6) is equivalent to

$$(\underline{Y}_k, \underline{\mathbf{L}}_k, \underline{A}_k) \perp\!\!\!\perp \bar{\Delta}_k \mid (\bar{\mathbf{L}}_{k-1}, \bar{A}_{k-1}, \bar{Y}_{k-1}).$$

Note the difference between the (5) and (6): (6) involves only factual variables and thus connects directly to the missing-data and classical survival analysis literature, as we explain below. In contrast, (5) is stated in terms of the potential outcomes; this formulation allows  $\mathbf{C}_k$  to affect  $A_k$ , whereas (6) does not permit that dependence. This distinction would not

have been apparent if we had relied solely on the classical treatment of right-censoring, without engaging with the causal model.

Nonetheless, when all events in  $\mathbf{C}_k$  are nuisance right-censoring events with respect to the estimand, some practitioners may appeal to the conditional independence of  $\Delta_k$  as formulated in (6). This assumption is analogous to the *coarsening at random* assumption for monotone missing data in longitudinal studies. Specifically, if we temporarily treat  $A_k \in \mathbf{L}_k$ , then *conditional independent*  $\Delta_k$  2.1b can be written as  $(Y_k, \mathbf{L}_k) \perp\!\!\!\perp \Delta_k | (\bar{\mathbf{L}}_{k-1}, \bar{Y}_{k-1}=1, \bar{\Delta}_{k-1}=0)$ , which is equivalent to the usual *coarsening at random* or *conditional independent censoring* assumption for survival outcomes defined in Chapter 9.3 of Tsiatis (2006) (see also Laan and Robins, 2003; Rotnitzky, 2014).

**Remark 1.** Define  $T$  as the failure time and  $C$  as the observation-terminating time for a subject. In classical survival literature, conditionally independent censoring states that  $\lim_{\epsilon \rightarrow 0} P(C < t + \epsilon | C \geq t, T, \bar{\mathbf{L}}(T)) = \lim_{\epsilon \rightarrow 0} P(C < t + \epsilon | C \geq t, T \geq t, \bar{\mathbf{L}}(t))$ . Indeed if we let the time interval lengths shrink to zero, then  $(Y_k, \mathbf{L}_k) \perp\!\!\!\perp \Delta_k | (\bar{\mathbf{L}}_{k-1}, \bar{Y}_{k-1}=1, \bar{\Delta}_{k-1}=0)$  would indeed be equivalent with this classical definition.

To this end, consider how this alternative condition can be used to achieve identification of  $\Psi_2^g$ . Specifically, consider a modified version of the conditional treatment exchangeability assumption posed previously in 2.2a:

*Assumption 2.2b* (Conditional treatment exchangeability):

$$Y_k^g \perp\!\!\!\perp A_k | (\bar{\mathbf{L}}_{k-1}, \bar{A}_{k-1} = \bar{g}_{k-1}, \bar{Y}_{k-1}=1);$$

Assumption 2.2b corresponds to the standard conditional treatment exchangeability in the absence of any  $\mathbf{C}_k$  events. Framed this way, assumptions 2.1b and 2.2b separate the causal estimand from the handling of right-censored factual data, mirroring how missing data are often handled in causal inference literature (Ding and Geng, 2014; Zhang and others, 2016; Yang and others, 2019; Kennedy, 2020; Levis and others, 2025; Wen and McGee, 2025). Based on this, we can define a model  $\mathcal{M}'_2$  that encompasses Assumptions 1.4, 2.1b, 2.2b and 2.3. Proposition 3 then continues to hold when  $\mathcal{M}_2$  is replaced by  $\mathcal{M}'_2$ .

## 7 Identity slippages

Estimation of the g-formula in (4) has been extensively studied under various treatment regimes. For time-to-event outcomes, Young and others (2011) described the parametric g-formula, while Wen and others (2021) described iterative conditional expectation estimators. Related work has investigated inverse probability weighted and doubly robust estimators (Robins and Finkelstein, 2000; Hernán and others, 2000; Tran and others, 2019; Wen and others, 2022, 2023). Inverse probability weighted and doubly robust estimation has also been applied, for instance, to evaluate the effects of glucose-lowering strategies on the development or progression of albuminuria in patients with diabetes (Neugebauer and others, 2014; Sofrygin and others, 2019), and to assess grace-period treatment strategies in hypertension (Wanis and others, 2024). We refer readers to these studies for detailed estimation methods based on (4). A common feature in many of these works is that right-censoring events are

intervened upon through their elimination; consequently, any resulting inferences pertain to the effect of this joint intervention. This motivates the following.

Conditional independent  $\Delta_k$  is violated when either (1) nuisance right-censoring events affects the outcome of interest, and/or (2) there are unmeasured common causes between those events and the outcome of interest. Here we show via simulation studies that in general  $\Psi_1^g := E(Y_k^{g, \bar{c}=0}) \neq \Psi_2^g := E(Y_k^g)$  when conditional independent  $\Delta_k$  is violated (i.e.,  $P \notin \mathcal{M}_2$ ), even in absence of unmeasured common causes between right-censoring events and the outcome of interest ( $P \in \mathcal{M}_1$ ). Consequently, this discrepancy can lead to identity slippage (Sarvet and others, 2023), where researchers may inadvertently draw conclusions about the average treatment effect of an intervention on an outcome based on estimates for  $\Psi_1^g$  when  $\Psi_2^g$  was the initial target parameter of interest and  $\mathcal{M}_1$  was assumed in the analysis.

For simplicity, we consider a point treatment process where  $A_k = \emptyset$  for  $k > 0$  and a scalar observation-terminating event at each time  $k$ ,  $C_k$ . In this example, the only right-censoring event is  $C_k$ , and the treatment regime of interest is  $g = a_0$  with  $a_0 = 1$ . Consider the following data generating mechanism:  $(L_{10}, L_{20}, A_0, C_1, Y_1, L_{11}, L_{21}, C_2, Y_2, \dots, C_5, Y_5)$  where  $L_{1k}$  and  $L_{2k}$  are time-dependent confounders at time  $k$ . At baseline, we draw  $L_{10} \sim \text{Ber}(0.5)$ ,  $L_{20} \sim \text{Ber}(0.5)$  and assign treatment according to  $A_0 \sim \text{Ber}\{\text{expit}(1 - 2L_{10} + L_{20} + L_{10}L_{20})\}$ . Moreover,  $C_k$  and  $Y_k$  at each time  $k$  ( $k \geq 1$ ) are simulated from:

$$\begin{aligned} C_k &\sim \text{Ber}\{\text{expit}(\gamma - A_0 + L_{1,k-1} + L_{2,k-1})\} \text{ if } C_{k-1} = 0; \quad C_k = 1 \text{ if } C_{k-1} = 1 \\ Y_k &\sim \text{Ber}\{m(A_0, \bar{L}_k, \bar{C}_k)\} \text{ if } Y_{k-1} = 1; \quad Y_k = 0 \text{ if } Y_{k-1} = 0 \end{aligned}$$

where  $m(A_0, \bar{L}_k, \bar{C}_k) = \text{expit}(-1 + 2A_0 + 2L_{1,k-1} - 2L_{2,k-1} + 2L_{1,k-1}L_{2,k-1} + \alpha C_k)$ . When  $Y_k = 1$ , time-varying confounders at time  $k$  ( $k \geq 1$ ) are simulated from:  $L_{1k} \sim \text{Ber}\{\text{expit}(A_0 + L_{1,k-1} - L_{2,k-1})\}$ , and  $L_{2k} \sim \text{Ber}\{\text{expit}(-2 - 2A_0 - L_{1k})\}$  if  $L_{2,k-1} = 0$ , and is deterministically set to 1 if  $L_{2,k-1} = 1$ . For the estimand  $\Psi_1^g$ ,  $C_k$  is a non-nuisance right-censoring event, whereas for  $\Psi_2^g$ ,  $C_k$  is a nuisance right-censoring event.

In the data generating mechanism,  $\gamma$  controls the amount of censoring and  $\alpha$  quantifies the extent to which censoring affects the conditional probability of survival. A negative value of  $\alpha$  suggests that experiencing the right-censoring event reduces the probability of survival. We consider a range of values for  $\alpha$  and  $\gamma$  in simulated datasets of sample size  $N = 10^7$ .

The results are presented in Figures 6 and 7. Under conditionally independent  $\Delta_k$  (when  $\alpha = 0$ ), theory predicts that  $\Psi_1^g = \Psi_2^g$ , which is confirmed by our results in Figure 6. However, as the dependence of survival probability on censoring status increases, the two estimands begin to diverge. Furthermore, when  $\alpha \neq 0$ , this divergence is more pronounced under a greater amount of right-censoring. For instance, when  $\gamma = -2$  (approximately 56% observations are right-censored by time 5), the divergence is greater at all time points compared to when  $\gamma = -3$  (approximately 28% are right-censored by time 5). Conversely, in the absence of right-censoring event the estimands coincide ( $\Psi_1^g = \Psi_2^g$ ). Figure 7 highlights this trend for a fixed strong effect  $\alpha$  ( $\alpha = -2$ ): greater amount of right-censoring produces a larger gap between  $\Psi_1^g$  and  $\Psi_2^g$ .

Whereas  $\Psi_1^g$  remains identifiable from observed data via the g-formula,  $\Psi_2^g$  is not readily identifiable without making additional assumptions. To illustrate identifiability, we report results

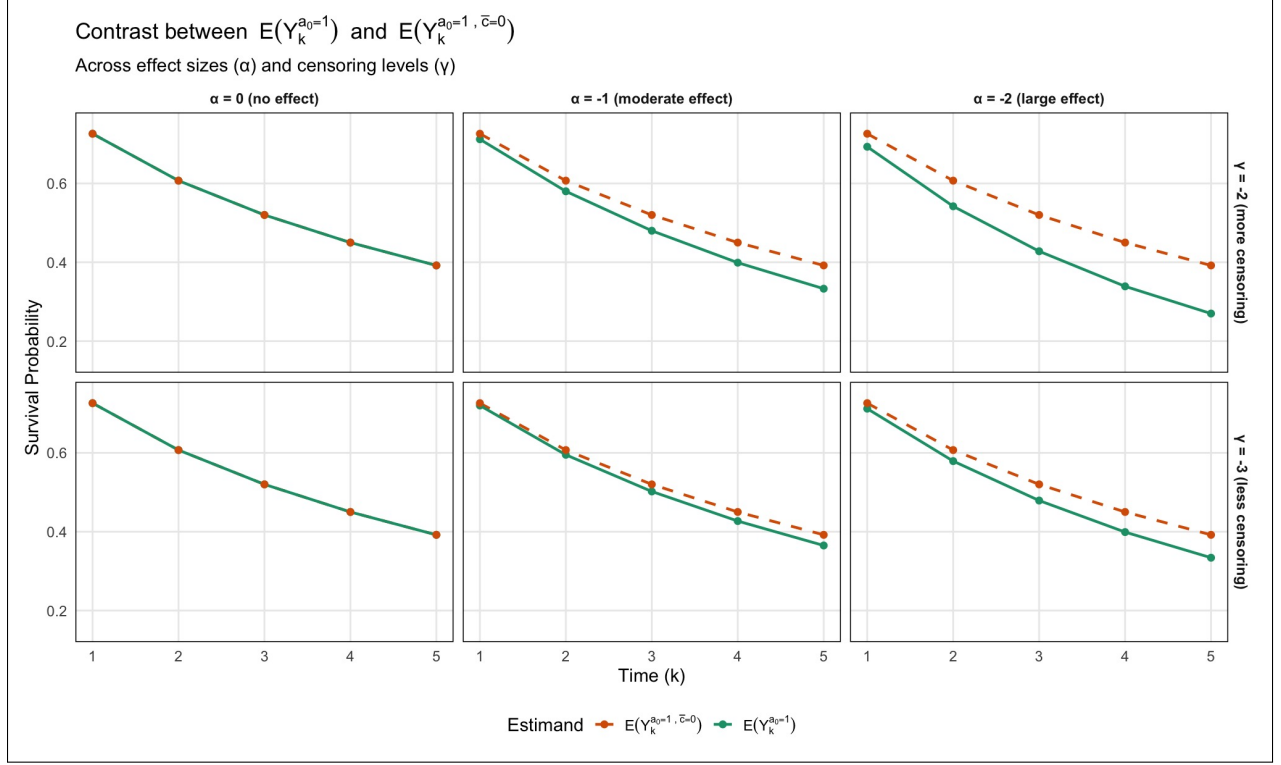


Figure 6: Survival probabilities of  $E(Y_k^{a_0=1})$  and  $E(Y_k^{a_0=1, \bar{c}=0})$  under different conditions of  $\alpha$  and  $\gamma$  across time points for sample size  $N=10^7$ .

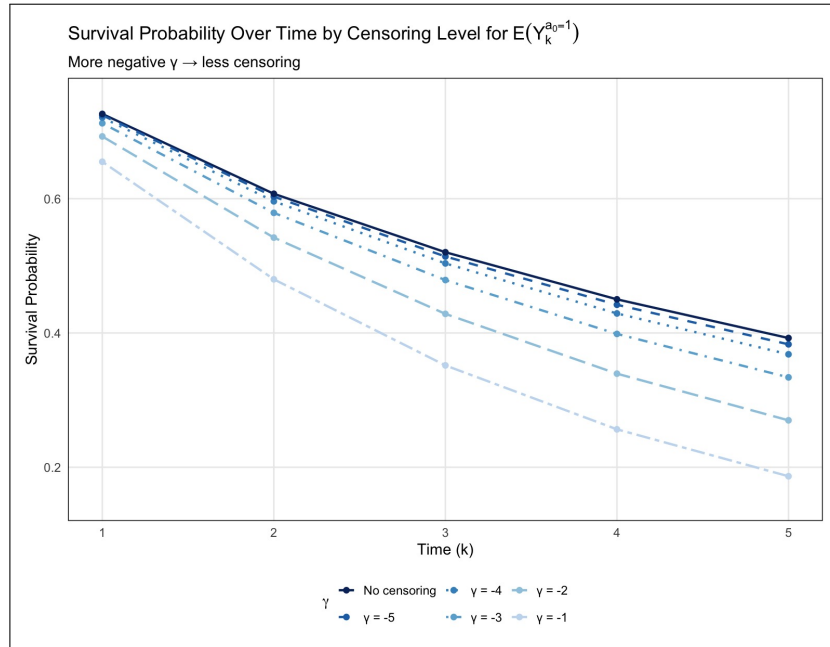


Figure 7: Survival probabilities of  $E(Y_k^{a_0=1})$  under different amounts of censoring across time points for sample size  $N=10^7$  when  $\alpha = -2$ . In “No censoring” setting,  $E(Y_k^{a_0=1}) = E(Y_k^{a_0=1, \bar{c}=0})$ .

from a parametric g-formula estimator based on (4) (see Web Appendix F for inverse probability weighted estimation results); more robust estimators (e.g., doubly robust estimators) may be used in practice. With a sample size of  $N=1000$  across 1000 simulated datasets, the parametric g-formula estimator produced nearly unbiased estimates of  $\Psi_1^g := E(Y_k^{a_0=1, \bar{c}=0}) = (0.7263, 0.6073, 0.5203, 0.4501, 0.3924)$  for  $k=1, \dots, 5$  under correct specification of all nuisance models. Specifically, the estimator yielded values of  $(0.7268, 0.6077, 0.5207, 0.4506, 0.3930)$  with corresponding standard errors  $(0.0125, 0.0152, 0.0175, 0.0192, 0.0204)$ . In Appendix E, we present a sensitivity analysis for  $\Psi_2^g := E(Y_k^{a_0=1})$ . We also use this simulation study to examine the performance of the method described in Web Appendix F for estimating  $\Psi_2^g$  under violations of conditional independent  $\Delta_k$ ; the results indicate that the proposed estimator performs well and is nearly unbiased in these settings.

Our simulation shows that  $\Psi_1^g$  coincides with  $\Psi_2^g$  when  $\mathbf{C}_k = \emptyset$  or when  $\mathbf{C}_k$  does not have a direct effect on  $Y_k$ . In our simulation setting,  $C_k$  is the only right-censoring event, and when this right-censoring event is associated with a lower probability of survival, we observe that  $\Psi_1^g \geq \Psi_2^g$ . In other words, the probability of survival under an intervention that sets  $A_0$  to  $a_0=1$  and eliminates the right-censoring event is at least as large as the probability of survival under an intervention that sets  $A_0$  to  $a_0=1$  without intervening on censoring. Web Appendix E presents results for causal contrasts, which align with the findings reported here.

## 8 Concluding remarks

The choice of estimand should be guided by underlying scientific objectives, as it determines how the right-censoring events are classified and consequently, how they should be handled. In this paper, we proposed a general definition of right-censoring events, and provided a classification that separates them into nuisance and non-nuisance categories. A nuisance right-censoring event is one that the investigator’s causal question and estimand does not aim to eliminate or prevent. In contrast, a non-nuisance right-censoring event is integral to the definition of the estimand.

Recent advances in causal inference propose that observation-terminating events  $\mathbf{C}_k$  need not be framed as a missing data problem, but rather as part of an intervention in which all right-censoring events are eliminated. We argue instead that whether to treat such observation-terminating events as interventions depends on the study objective. For instance, identification of the estimand given by  $E(Y_k^{g, \bar{c}=0})$  requires intervening on all observation-terminating events (considered non-nuisance right-censoring here) and adjusting for all common causes of these events and the outcome. In contrast, identification of  $E(Y_k^g)$  does not require intervening on observation-terminating events, but still requires adjustment for all common causes of nuisance right-censoring events and the outcome.

The subtlety in handling events in  $\mathbf{C}_k$  is that, whereas estimands such as intention-to-treat and per-protocol typically rest on different identifying assumptions and therefore different identification formulas, estimands defined in the presence of observation-terminating events can, under different sets of assumptions and models, have identical identifying formulas. However, as noted above and emphasized again here, identifying  $E(Y_k^g)$  relies on a different set of assumptions than those required for  $E(Y_k^{g, \bar{c}=0})$ , and consequently, these two quantities

may differ in practice. Ultimately, handling right-censoring events should follow from the estimand of interest rather than convenience or ease of identification. If a right-censoring event is not directly intervenable (see [Young and others, 2020](#)), consistency might not hold due to the ambiguity in counterfactuals arising from the multiple pathways or mechanisms involved. If this was indeed the case for observation-terminating events, then defining estimands based on the potential outcome  $Y_k^{g,\bar{c}=0}$  might be questionable.

Our findings highlight the importance of understanding right-censoring events in each study. Such knowledge is indispensable for accurately accounting for the shared common causes between the censoring mechanisms and the outcome of interest. Without this understanding, the validity of the estimand and the resulting inferences may be compromised. Thus, careful consideration of censoring mechanisms and their relationships with the outcome is essential for robust causal inference in the presence of right-censored data.

## References

- ADU, MARY D, MALABU, USMAN H, MALAU-ADULI, ADULI EO, DROVANDI, AARON AND MALAU-ADULI, BUNMI S. (2020). User retention and engagement with a mobile app intervention to support self-management in australians with type 1 or type 2 diabetes (my care hub): mixed methods study. *JMIR mHealth and uHealth* **8**(6), e17802.
- AMAGAI, SAKI, PILA, SARAH, KAAT, AARON J, NOWINSKI, CINDY J AND GERSHON, RICHARD C. (2022). Challenges in participant engagement and retention using mobile health apps: literature review. *Journal of medical Internet research* **24**(4), e35120.
- ANDERSEN, PER K, BORGAN, ORNULF, GILL, RICHARD D AND KEIDING, NIELS. (2012). *Statistical models based on counting processes*. Springer Science & Business Media.
- BU, QIUSHI, SU, WEN, ZHAO, XINGQIU AND LIU, ZHONGHUA. (2025). Semiparametric causal inference for right-censored outcomes with many weak invalid instruments. *arXiv preprint arXiv:2509.13176*.
- CAI, W. AND VAN DER LAAN, M. J. (2020). One-step targeted maximum likelihood estimation for time-to-event outcomes. *Biometrics* **76**(3), 722–733.
- CUI, YIFAN, KOSOROK, MICHAEL R, SVERDRUP, ERIK, WAGER, STEFAN AND ZHU, RUOQING. (2023). Estimating heterogeneous treatment effects with right-censored data via causal survival forests. *Journal of the Royal Statistical Society Series B: Statistical Methodology* **85**(2), 179–211.
- DING, PENG AND GENG, ZHI. (2014). Identifiability of subgroup causal effects in randomized experiments with nonignorable missing covariates. *Statistics in Medicine* **33**(7), 1121–1133.
- FLEMING, T. R. AND HARRINGTON, D. P. (1991). *Counting processes and survival analysis*. John Wiley & Sons.
- GABRIEL, ANDRÉ P AND MERCADO, CHARLES P. (2011). Data retention after a patient withdraws consent in clinical trials. *Open access Journal of clinical trials*, 15–19.

- HERNÁN, MIGUEL ÁNGEL, BRUMBACK, BABETTE AND ROBINS, JAMES M. (2000). Marginal structural models to estimate the causal effect of zidovudine on the survival of hiv-positive men.
- HERNÁN, MIGUEL A AND ROBINS, JAMES M. (2020). *What If?*. Chapman & Hall/CRC.
- ICH. (2019). ICH Harmonised Guideline: Addendum on Estimands and Sensitivity Analysis in Clinical Trials to the Guideline on Statistical Principles for Clinical Trials E9(R1).
- KALBFLEISCH, JOHN D AND PRENTICE, ROSS L. (2002). *The statistical analysis of failure time data*. John Wiley & Sons.
- KENNEDY, EDWARD H. (2020). Efficient nonparametric causal inference with missing exposure information. *The international journal of biostatistics* **16**(1).
- KLEIN, JOHN P. AND MOESCHBERGER, MELVIN L. (2003). *Survival Analysis: Techniques for Censored and Truncated Data*, 2 edition., Statistics for Biology and Health. New York, NY: Springer Science+Business Media.
- KLEIN, JOHN P, VAN HOUWELINGEN, HANS C, IBRAHIM, JOSEPH GEORGE AND SCHEIKE, THOMAS H. (2014). *Handbook of survival analysis*. CRC Press Boca Raton, FL:.
- LAAN, MARK J AND ROBINS, JAMES M. (2003). *Unified methods for censored longitudinal data and causality*. Springer.
- LEVIS, ALEXANDER W, MUKHERJEE, RAJARSHI, WANG, RUI AND HANEUSE, SEBASTIEN. (2025). Robust causal inference for point exposures with missing confounders. *Canadian Journal of Statistics* **53**(2), e11832.
- LI, KENDRICK QIJUN, LINDERMAN, GEORGE C, SHI, XU AND TCHETGEN, ERIC J TCHETGEN. (2025). Regression-based proximal causal inference for right-censored time-to-event data. *Epidemiology* **36**(5), 694–704.
- MORENO-BETANCUR, MARGARITA, WIJESURIYA, RUSHANI AND CARLIN, JOHN B. (2024). Ideal trials, target trials and actual randomized trials. *arXiv preprint arXiv:2405.10026*.
- NEUGEBAUER, R., SCHMITTDIEL, J.A. AND VAN DER LAAN, M.J. (2014a). Targeted learning in real-world comparative effectiveness research with time-varying interventions. *Statistics in Medicine* **33**, 2480–2520.
- NEUGEBAUER, ROMAIN, SCHMITTDIEL, JULIE A AND VAN DER LAAN, MARK J. (2014b). Targeted learning in real-world comparative effectiveness research with time-varying interventions. *Statistics in medicine* **33**(14), 2480–2520.
- RICHARDSON, THOMAS S AND ROBINS, JAMES M. (2013). Single world intervention graphs (swigs): A unification of the counterfactual and graphical approaches to causality. *Center for the Statistics and the Social Sciences, University of Washington Series. Working Paper* **128**(30), 2013.

- RICHARDSON, THOMAS S AND ROTNITZKY, ANDREA. (2014). Causal etiology of the research of james m. robins. *Statistical Science* **29**(4), 459–484.
- ROBINS, JAMES M AND FINKELSTEIN, DIANNE M. (2000). Correcting for noncompliance and dependent censoring in an aids clinical trial with inverse probability of censoring weighted (ipcw) log-rank tests. *Biometrics* **56**(3), 779–788.
- ROBINS, JAMES M AND RICHARDSON, THOMAS S. (2010). Alternative graphical causal models and the identification of direct effects. *Causality and psychopathology: Finding the determinants of disorders and their cures* **84**, 103–158.
- ROTNITZKY, ANDREA. (2014). Inverse probability weighting in survival analysis. *Wiley StatsRef: Statistics Reference Online* **87**.
- RYTGAARD, H., ERIKSSON, F. AND VAN DER LAAN, M. (2021). Estimation of time-specific intervention effects on continuously distributed time-to-event outcomes by targeted maximum likelihood estimation. *Biometrics*.
- SARVET, AARON L, STENSRUD, MATS J AND WEN, LAN. (2023). Interpretational errors in statistical causal inference. *arXiv preprint arXiv:2312.07610*.
- SOFRYGIN, OLEG, ZHU, ZHENG, SCHMITTDIEL, JULIE A, ADAMS, ALYCE S, GRANT, RICHARD W, VAN DER LAAN, MARK J AND NEUGEBAUER, ROMAIN. (2019). Targeted learning with daily ehr data. *Statistics in medicine* **38**(16), 3073–3090.
- TRAN, LINH, YIANNOUTSOS, CONSTANTIN, WOOLS-KALOUSTIAN, KARA, SIIKA, ABRAHAM, VAN DER LAAN, MARK AND PETERSEN, MAYA. (2019). Double robust efficient estimators of longitudinal treatment effects: comparative performance in simulations and a case study. *The international journal of biostatistics* **15**(2), 20170054.
- TSIATIS, A.A. (2006). *Semiparametric Theory and Missing Data*, 1 edition. New York, NY: Springer.
- TUTZ, GERHARD, SCHMID, MATTHIAS *and others*. (2016). *Modeling discrete time-to-event data*, Volume 3. Springer.
- VERMA, THOMAS S AND PEARL, JUDEA. (2022). Equivalence and synthesis of causal models. In: *Probabilistic and causal inference: The works of Judea Pearl*. pp. 221–236.
- WANIS, KEROLLOS NASHAT, SARVET, AARON L, WEN, LAN, BLOCK, JASON P, RIFAS-SHIMAN, SHERYL L, ROBINS, JAMES M AND YOUNG, JESSICA G. (2024). Grace periods in comparative effectiveness studies of sustained treatments. *Journal of the Royal Statistical Society Series A: Statistics in Society* **187**(3), 796–810.
- WEN, LAN, HERNÁN, MIGUEL A AND ROBINS, JAMES M. (2022). Multiply robust estimators of causal effects for survival outcomes. *Scandinavian Journal of Statistics* **49**(3), 1304–1328.
- WEN, LAN, MARCUS, JULIA L AND YOUNG, JESSICA G. (2023). Intervention treatment distributions that depend on the observed treatment process and model double robustness in causal survival analysis. *Statistical methods in medical research* **32**(3), 509–523.

- WEN, LAN AND MCGEE, GLEN. (2025). Estimating average causal effects with incomplete exposure and confounders. *arXiv preprint arXiv:2506.21786*.
- WEN, LAN, YOUNG, JESSICA G, ROBINS, JAMES M AND HERNÁN, MIGUEL A. (2021). Parametric g-formula implementations for causal survival analyses. *Biometrics* **77**(2), 740–753.
- WILSON, BROOKE E, NADLER, MICHELLE B, DESNOYERS, ALEXANDRA AND AMIR, EITAN. (2021). Quantifying withdrawal of consent, loss to follow-up, early drug discontinuation, and censoring in oncology trials. *Journal of the National Comprehensive Cancer Network* **19**(12), 1433–1440.
- YANG, SHU, WANG, LINBO AND DING, PENG. (2019). Causal inference with confounders missing not at random. *Biometrika* **106**(4), 875–888.
- YING, ANDREW. (2024). Proximal survival analysis to handle dependent right censoring. *Journal of the Royal Statistical Society Series B: Statistical Methodology* **86**(5), 1414–1434.
- YING, ANDREW, CUI, YIFAN AND TCHETGEN, ERIC J TCHETGEN. (2022). Proximal causal inference for marginal counterfactual survival curves. *arXiv preprint arXiv:2204.13144*.
- YOUNG, JESSICA G. (2024). Story-led causal inference. *Epidemiology* **35**(3), 289–294.
- YOUNG, J. G., CAIN, L. E., ROBINS, J. M., O'REILLY, E. J. AND HERNÁN, M. A. (2011). Comparative effectiveness of dynamic treatment regimes: an application of the parametric g-formula. *Statistics in biosciences* **3**, 119–143.
- YOUNG, JESSICA G, LOGAN, ROGER W, ROBINS, JAMES M AND HERNÁN, MIGUEL A. (2019). Inverse probability weighted estimation of risk under representative interventions in observational studies. *Journal of the American Statistical Association*.
- YOUNG, JESSICA G, STENSRUD, MATS J, TCHETGEN TCHETGEN, ERIC J AND HERNÁN, MIGUEL A. (2020). A causal framework for classical statistical estimands in failure-time settings with competing events. *Statistics in medicine* **39**(8), 1199–1236.
- ZHANG, ZHIWEI, LIU, WEI, ZHANG, BO, TANG, LI AND ZHANG, JUN. (2016). Causal inference with missing exposure information: Methods and applications to an obstetric study. *Statistical methods in medical research* **25**(5), 2053–2066.