Bias of the additive hazard model in the presence of causal effect heterogeneity

Richard A. J. Post[1] · Edwin R. van den Heuvel[1] · Hein Putter[2,3]

Abstract

Hazard ratios are prone to selection bias, compromising their use as causal esti-mands. On the other hand, if Aalen's additive hazard model applies, the hazard dif-ference has been shown to remain unaffected by the selection of frailty factors over time. Then, in the absence of confounding, observed hazard differences are equal in expectation to the causal hazard differences. However, in the presence of effect (on the hazard) heterogeneity, the observed hazard difference is also affected by selection of survivors. In this work, we formalize how the observed hazard differ-ence (from a randomized controlled trial) evolves by selecting favourable levels of effect modifiers in the exposed group and thus deviates from the causal effect of interest. Such selection may result in a non-linear integrated hazard difference curve even when the individual causal effects are time-invariant. Therefore, a homogene-ous time-varying causal additive effect on the hazard cannot be distinguished from a time-invariant but heterogeneous causal effect. We illustrate this causal issue by studying the effect of chemotherapy on the survival time of patients suffering from carcinoma of the oropharynx using data from a clinical trial. The hazard difference can thus not be used as an appropriate measure of the causal effect without making untestable assumptions.

Keywords Causal inference · Causal hazard difference · Individual effect modification · Selection bias · Aalen additive hazard model

Mathematics Subject Classification 62D20 · 62N02

✉ Richard A. J. Post
r.a.j.post@tue.nl

1 Department of Mathematics and Computer Science, Eindhoven University of Technology, Eindhoven, The Netherlands

2 Department of Biomedical Data Sciences, Leiden University Medical Center, Leiden, The Netherlands

3 Mathematical Institute, Leiden University, Leiden, The Netherlands

1 Introduction

Hazard ratios, often obtained by fitting a Cox proportional hazards model (Cox 1972), are the most common effect measures when dealing with time-to-event data. However, the hazard ratio is prone to selection bias due to conditioning on survival and therefore not suitable for causal inference (Hernán 2010; Aalen et al. 2015; Stensrud et al. 2017). It has been recommended to use other, better interpretable, estimands when interested in causal effects (Hernán 2010; Stensrud et al. 2018; Bartlett et al. 2020; Young et al. 2020). Alternatively, using the additive hazard model can avoid interpretation issues (Aalen et al. 2015; Martinussen et al. 2020). In the nonparametric model proposed by Aalen (1989), the hazard rate at time t for individual i with (possibly time-dependent) covariates $\mathbf{x}_i(t)$ (of dimension p) is determined by the values of $\mathbf{x}_i$ up until time t and equals

$$\lambda\big(t \mid \{\mathbf{x}_i(s)\}_{s \leq t}\big) = \beta_0(t) + \beta_1(t)x_{i1}(t) + \ldots + \beta_p(t)x_{ip}(t),$$

where the parameters $\beta_j(t)$ are arbitrary regression functions, allowing time-varying effects (Aalen et al. 2008). Restricted versions have been proposed by Lin and Ying (1994) and McKeague and Sasieni (1994), where all or some $\beta_j(t)$ are assumed to be constant over time. The cumulative regression function, $B_j(t) = \int_0^t \beta_j(s)ds$, may reveal changes in effect over time, see for example Aalen et al. (2008, pp. 160-162).

For cause-effect relations that can be accurately described with Aalen's additive hazard model, the hazard difference is a collapsible measure (Martinussen and Vansteelandt 2013). Then, in the absence of confounding, the hazard difference can be appropriately used to estimate the causal effect, even in the case of unmeasured risk factors (Aalen et al. 2015). For this, it is necessary that the exposure effect on the hazard does not depend on unmeasured individual features (modifiers) and thus is the same for all individuals. However, due to the fundamental problem of causal inference, the effect homogeneity assumption is untestable (Holland 1986). In our companion paper (Post et al. 2024), we showed that next to unmeasured risk factors, i.e. frailty (Aalen et al. 2008, Chapter 6), effect heterogeneity at the level of the individual hazard results in selection bias of observed hazard ratios.

In this work, we extend the additive hazard model studied in Aalen et al. (2015) by allowing heterogeneity of the effect (on the hazard) and quantify the bias of using the observed hazard difference when estimating the causal effect. In Sect. 2, we introduce notation to describe the cause-effect relations using a structural causal model for which we can define the causal hazard difference. In practice, when applying an additive hazard model the observed hazard difference is modelled. We show that (in absence of confounding) the expected value of the observed hazard equals a causal hazard marginalized over survivors. The expectation of the observed hazard difference thus equals a difference between hazards marginalized over survivors in the exposed and unexposed universe. By selection of individuals with favourable values of the effect modifier, this difference can deviate from the causal hazard difference as we formalize in Sect. 3. We present numerical examples to illustrate how this selection can result in a non-linear integrated hazard difference curve, that can be interpeted as reflecting a time-varying causal effect, while the actual individual causal effects are time-invariant. To emphasize

why it is important to be aware of such a difference between the expected observed hazard difference and the causal hazard difference, we reflect on the analysis of the effect of treatment on survival with carcinoma of the oropharynx from the clinical trial in Sect. 4. Finally, we present some concluding remarks in Sect. 5.

2 Notation and hazard differences

Probability distributions of factual and counterfactual outcomes are defined in the potential outcome framework (Neyman 1923; Rubin 1974). Let T_i and A_i represent the (factual) stochastic outcome and exposure assignment level of individual i. Let T_i^a equal the potential outcome of individual i under the intervention of the exposure to level a (counterfactual when $A_i \neq a$). For those more familiar with the do-calculus, T^a is equivalent to $T \mid do(A{=}a)$ as e.g. derived in Pearl (2009a, Equation 40) and Bongers et al. (2021, Definition 8.6). Throughout this paper, we will assume causal consistency, i.e. if $A_i{=}a$, then $T_i^a = T_i^{A_i} = T_i$. Causal consistency implies that potential outcomes are independent of the assigned exposure levels of other individuals. The hazard rate of the potential outcome can vary among individuals due to heterogeneity in risk factors U_0 as also considered by Aalen et al. (2015). The hazard difference of the potential outcomes with and without ($a = 0$) exposure might also vary among individuals due to an effect modifier U_1. Therefore, the hazard rate of individual i at time t of the potential outcome under exposure to level a is a function of U_{0i} and U_{1i} and thus random and equals

$$\lambda_i^a(t) = \lim_{h \to 0} h^{-1} \mathbb{P}\big(T_i^a \in [t, t+h) \mid T_i^a {\geq} t, U_{0i}, U_{1i} \big). \tag{1}$$

The hazard of the potential outcome T_i^a can be parameterized with a function that depends on U_{0i}, U_{1i} and a.

We describe cause-effect relations with a structural causal model (SCM) which is commonly used in the causal graphical literature, see e.g. Pearl (2009b, Chapter 1.4) and Peters et al. (2018, Chapter 6), to model observations. Instead, we include details on individual effect modifier U_1 as well as the latent common cause of the outcomes U_0, so that the SCM consists of a joint probability distribution of (N_A, U_0, U_1, N_T) and a collection of structural assignments (for more details, see Post et al. (2024, Section 2)) such that

$$
\begin{aligned}
A_i &:= f_A(N_{Ai}) \\
\lambda_i^a(t) &:= f_0(t, U_{0i}) + f_1(t, U_{1i}, a) \\
T_i^a &:= \min\{t > 0 \colon \exp(-\Lambda_i^a(t)) \leq N_{Ti}\},
\end{aligned}
\tag{2}
$$

where $\forall t$: $f_1(t, U_{1i}, 0) = 0$, f_A is the inverse cumulative distribution function of A, $N_{Ai}, N_{Ti} \sim \text{Uni}[0,1]$, $\Lambda_i^a(t) = \int_0^t \lambda_i^a(s)ds$, and $U_{0i}, U_{1i} \perp\!\!\!\perp N_{Ti}$ so that $\lambda_i^a(t)$ equals the hazard rate of the potential outcome given U_0 and U_1.

If in SCM (2), $N_A \!\not\perp\!\!\!\perp U_0, U_1, N_T$, then there exists confounding as the distributions of (U_0, U_1, N_T) are not exchangeable between exposed and non-exposed individuals. However, in this work we focus on the distribution of data observed from a properly executed RCT, where by the randomization $N_A \perp\!\!\!\perp U_0, U_1, N_T$ so that there is no confounding. It is important to realize that a SCM cannot be validated with data as it describes potential outcomes from different universes. For each individual the outcome can only be observed in one of the universes, and only the fit of the distribution of the outcomes in the factual world can be verified. In SCM (2), we did not restrict the distribution of U_0 and U_1 and only restricted f_λ and f_A to be properly defined hazard and inverse cumulative distribution functions respectively, so that the structural model is quite general. However, in this work we limit ourselves to cause-effect relations where the causal effect of the exposure is described by $\lambda_i^a(t) - \lambda_i^0(t) = f_1(t, U_{1i}, a)$.

2.1 Hazard differences

If SCM (2) applies and $\lambda_i^a(t) - \lambda_i^0(t) = f_1(t, a)$, then the causal effect is equal for each individual, i.e. effect homogeneity, and Aalen's additive hazard model applies. Otherwise, when $\lambda_i^a(t) - \lambda_i^0(t) = f_1(t, U_{1i}, a)$, the difference differs among individuals so that $\mathbb{E}[\lambda_i^a(t) - \lambda_i^0(t)]$ will typically be the estimand of interest. The latter contrast equals the difference between the expected hazard rate in the world where everyone is exposed to a and the world without exposure, and will therefore be referred to as the causal hazard difference (CHD) defined in Definition 1.

Definition 1 Causal hazard difference The CHD for cause-effect relations that can be parameterized with SCM 2 equals

$$
\mathbb{E}\left[\lambda_i^a(t)\right] - \mathbb{E}\left[\lambda_i^0(t)\right] = \mathbb{E}[f_1(t, U_1, a)]
$$
$$
= \int \lim_{h \to 0} h^{-1} \mathbb{P}\left(T^a \in [t, t+h] \mid T^a \geq t, U_0, U_1\right) dF_{U_0, U_1}
$$
$$
- \int \lim_{h \to 0} h^{-1} \mathbb{P}\left(T^0 \in [t, t+h] \mid T^0 \geq t, U_0\right) dF_{U_0}.
$$

Throughout this paper, we abbreviate the Lebesque-Stieltjes integral of a function g with respect to probability law F_X, $\int g(x) dF_X(x)$, as $\int g(X) dF_X$.

The CHD thus equals the difference of hazard rates of the potential outcomes marginalized over the population distribution of (U_0, U_1). However, the distribution of (U_0, U_1) among survivors will differ over time (in all worlds), i.e. $(U_0, U_1) \overset{d}{\neq} (U_0, U_1) \mid T^a \geq t$. In turn, $(U_0, U_1) \mid T^a \geq t$ and $(U_0, U_1) \mid T^0 \geq t$ can differ in distribution. As a consequence, the hazard rates for the observed (factual) outcomes are affected by these conditional distributions of U_0 and U_1, and the observed hazard difference may reflect both the causal effect of exposure and a difference in distributions of (U_0, U_1) between exposed and unexposed individuals. Since practioners are typically interested in the causal effect alone, it

is important to understand this mixture. Since the focus of this work is on estimands, for readability, we refer to the expected value of the difference of the observed hazards as the observed hazard difference (OHD) presented in Definition 2.

Definition 2 Observed hazard difference The OHD at time t equals

$$\lim_{h\to 0} h^{-1}\mathbb{P}(T \in [t, t+h] \mid T{\geq}t, A{=}a) - \lim_{h\to 0} h^{-1}\mathbb{P}(T \in [t, t+h] \mid T{\geq}t, A{=}0)$$

$$= \lim_{h\to 0} \int h^{-1}\mathbb{P}\big(T \in [t, t+h] \mid T{\geq}t, A{=}a, U_0, U_1\big) dF_{U_0, U_1 \mid T{\geq}t, A{=}a}$$

$$- \lim_{h\to 0} \int h^{-1}\mathbb{P}\big(T \in [t, t+h] \mid T{\geq}t, A{=}0, U_0\big) dF_{U_0 \mid T{\geq}t, A{=}0}.$$

To be precise, at time t the hazard rate can only be observed for non-censored individuals at that time ($C(t) = 0$). However, in this work we will assume independent censoring, so that $\mathbb{P}(T \mid T{\geq}t, A{=}a)$ is equal to $\mathbb{P}(T \mid T{\geq}t, A{=}a, C(t){=}0)$.

To compare the OHD to the CHD of interest, the OHD should be expressed in terms of potential outcomes. By causal consistency,

$$\mathbb{P}(T \in [t, t+h] \mid T{\geq}t, A{=}a) = \mathbb{P}(T^a \in [t, t+h] \mid T^a{\geq}t, A{=}a). \tag{3}$$

For a randomized controlled trial (RCT), where by design of the trial $A \perp\!\!\!\perp T^a$ (in SCM (2) equivalent to $N_A \perp\!\!\!\perp U_0, U_1, N_T$,

$$\mathbb{P}(T^a \in [t, t+h] \mid T^a{\geq}t, A{=}a) = \mathbb{P}(T^a \in [t, t+h] \mid T^a{\geq}t). \tag{4}$$

The OHD at time t is then equal to

$$\lim_{h\to 0} h^{-1}\mathbb{P}(T^a \in [t, t+h] \mid T^a{\geq}t) - \lim_{h\to 0} h^{-1}\mathbb{P}\big(T^0 \in [t, t+h] \mid T^0{\geq}t\big). \tag{5}$$

We refer to (5) as the survivor marginalized causal hazard difference (SMCHD) that is rewritten in Definition 3.

Definition 3 Survivor marginalized causal hazard difference The SMCHD at time t for cause-effect relations that can be parameterized with SCM 2 equals

$$\lim_{h\to 0} \int h^{-1}\mathbb{P}\big(T^a \in [t, t+h] | T^a{\geq}t, U_0, U_1\big) dF_{U_0, U_1 \mid T^a{\geq}t}$$

$$- \int h^{-1}\mathbb{P}\big(T^0 \in [t, t+h] | T^0{\geq}t, U_0\big) dF_{U_0 \mid T^0{\geq}t}.$$

As the integration in Definition 1 is with respect to the population distribution of (U_0, U_1) (instead of that of the survivors in the exposed and unexposed universe respectively), the SMCHD is thus affected by the difference in distribution between (U_0, U_1) and $(U_0, U_1) \mid T^a \geq t$ as well as the actual causal effect. Therefore, the SMCHD can deviate from the CHD. Nevertheless, Aalen et al. (2015) explained that $U_0 \perp\!\!\!\perp A | T \geq t$ so that for degenerate U_1, the SMCHD is only affected by the causal

effect and equals the CHD, so that the latter can be unbiasedly estimated from RCT data. In the next section, we formalize the SMCHD in case of effect heterogeneity (non-degenerate U_1) and show that $U_1 \not\perp\!\!\!\perp A | T \geq t$, so that then the SMCHD deviates from the CHD.

3 Results

In the remainder of the paper, we will focus on binary exposures such that $a \in \{0, 1\}$. In this section we quantify how the SMCHD describes both the causal effect and the difference in distribution of (U_0, U_1) between survivors in the exposed and unexposed universes.

For cause-effect relations where SCM (2) applies with $\lambda_i^a(t) = f_0(t, U_{0i}) + f_1(t, a)$, it is known from Aalen et al. (2015) that for an RCT $U_{0i} \perp\!\!\!\perp A_i | T_i \geq t$, so that U_0 remains exchangeable between exposed ($U_0 | T \geq t, A=1$) and nonexposed ($U_0 | T \geq t, A=0$) survivors. This independence, causal consistency, and the absence of confounding in an RCT ($T_a \perp\!\!\!\perp A$) imply

$$U_0 | T \geq t \stackrel{d}{=} U_0 | T \geq t, A=a$$
$$\stackrel{d}{=} U_0 | T^a \geq t, A=a$$
$$\stackrel{d}{=} U_0 | T^a \geq t.$$

Thus in absence of effect heterogeneity of the hazard difference, U_0 is exchangeable between survivors in the exposed ($U_0 | T^1 \geq t$) and unexposed universes ($U_0 | T^0 \geq t$), so that the OHD from an RCT (that equals the SMCHD) equals the CHD and describes the causal effect.

However, if heterogeneity exists, there will also be a selection of the modifier (U_1) in the exposed universe, where individuals with more favourable levels of U_1 are more likely to survive. As a result of this selection, the SMCHD over time no longer represents the (population) average effect. For the main result of this paper, we consider hazard functions that satisfy Condition 1.

Condition 1 Hazard without infinite discontinuity

$\forall t>0$: $\exists \tilde{h}>0$ such that $\forall h^* \in (0, \tilde{h})$: $\mathbb{E}\left[f_0(t + h^*, U_0) + f_1(t + h^*, U_1, a) | T^a \geq t\right] < \infty$

In Theorem 1, we show that the SMCHD, for a hazard function that satisfies Condition 1, can be expressed in terms of conditional expectations of $f_1(t, U_1, 1)$ and $f_0(t, U_0)$. In presence of effect heterogeneity, the SMCHD thus deviates from the CHD equal to $\mathbb{E}[f_1(t, U_1, 1)]$.

Theorem 1 *If the cause-effect relations of interest can be parameterized with SCM* (2), *where*

$$\lambda_i^a(t) := f_0(t, U_{0i}) + f_1(t, U_{1i}, a),$$

and Condition 1 *applies*, *then the SMCHD at time t equals*

$$\mathbb{E}[f_1(t, U_1, 1) \mid T^1 \geq t] + \mathbb{E}[f_0(t, U_0) \mid T^1 \geq t] - \mathbb{E}[f_0(t, U_0) \mid T^0 \geq t].$$

To illustrate how the SMCHD can deviate from the CHD we continue by presenting some examples and apply Theorem 1. All programming codes used for these examples can be found online at https://github.com/RAJP93/CHD. First, we consider cause-effect relations for which $U_0 \perp\!\!\!\perp U_1$.

3.1 Independent U_0 and U_1

As discussed at the start of this section, Aalen et al. (2015) implicitly showed that $U_0 \mid T^1 \geq t \overset{d}{=} U_0 \mid T^0 \geq t$ in absence of effect heterogeneity of the hazard difference. Based on similar arguments, Lemma 1 states that the additive frailty is also exchangeable in the presence of effect heterogeneity at the hazard scale that is independent of the frailty.

Lemma 1 *If the cause-effect relations of interest can be parameterized with SCM* (2), *where*

$$\lambda_i^a(t) := f_0(t, U_{0i}) + f_1(t, U_{1i}, a),$$

and $U_{0i} \perp\!\!\!\perp U_{1i}$ *then*,

$$\mathbb{E}\big[f_0(t, U_0) \mid T^1 \geq t\big] = \mathbb{E}\big[f_0(t, U_0) \mid T^0 \geq t\big].$$

Note that while $\mathbb{E}\big[f_0(t, U_0) \mid T^1 \geq t\big] = \mathbb{E}\big[f_0(t, U_0) \mid T^0 \geq t\big]$, $\mathbb{E}\big[f_0(t, U_0) \mid T^a \geq t\big] \neq \mathbb{E}\big[f_0(t, U_0)\big]$ as the conditional expectations will decrease over time representing the survival of less susceptible individuals. If $U_0 \perp\!\!\!\perp U_1$, as for the case of effect homogeneity, U_0 is thus exchangeable between survivors in the exposed ($U_0 \mid T^1 \geq t$) and unexposed ($U_0 \mid T^0 \geq t$) universes. By Theorem 1 and Lemma 1, the SMCHD at time t now equals $\mathbb{E}[f_1(t, U_1, 1) \mid T^1 \geq t]$.

Let us consider cause-effect relations for which SCM (2) applies with

$$f_1(t, U_{1i}, a) = U_1 f_1(t, a), \tag{6}$$

where $f_1(t, 0) = 0$, then $\mathbb{E}\big[f_1(t, U_1, 1) \mid T^1 \geq t\big] = f_1(t, a)\mathbb{E}[U_1 \mid T^1 \geq t]$. By Definition 1, the CHD equals $f_1(t, a)\mathbb{E}[U_1]$, so that the difference with the SMCHD varies over time and equals

$$f_1(t, a)\big(\mathbb{E}[U_1] - \mathbb{E}[U_1 \mid T^1 \geq t]\big). \tag{7}$$

For this multiplicative case, the conditional expectation $\mathbb{E}[U_1 \mid T^1 \geq t]$ can be expressed in terms of the Laplace transform of U_1, as stated in Lemma 2.

Lemma 2 *If the cause-effect relations of interest can be parameterized with SCM* (2), *where*

$$f_1(t, U_{1i}, 1) = U_{1i} f_1(t, 1),$$

and $U_{0i} \perp\!\!\!\perp U_{1i}$, *then*

$$\mathbb{E}[U_1 \mid T^1 \geq t] = -\frac{\mathcal{L}'_{U_1}(\int_0^t f_1(s, 1) ds)}{\mathcal{L}_{U_1}(\int_0^t f_1(s, 1) ds)},$$

where $\mathcal{L}_{U_1}(c) = \mathbb{E}[\exp(-cU_1)]$ *with derivative* $\mathcal{L}'_{U_1}(c)$.

We continue to illustrate how effect heterogeneity can affect the integrated hazard difference when the causal effect is time-invariant for each individual. To do so, we let the additive hazard effect modifier U_1 equal μ_1 (≤ 0, for individuals that benefit) with probability p_1, μ_2 (≥ 0, for individuals that are harmed) with probability p_2 or 0 (for individuals that are not affected). We denote this distribution as the Benefit-Harm-Neutral, BHN(p_1, μ_1, p_2, μ_2), distribution. Note that it is necessary that $\forall t: \mathbb{P}(f_0(t, U_0) < |\mu_1|) = 0$ to guarantee that the hazard rate is always positive for each individual. By Theorem 1 and Lemma 2 with $f_1(t, 1) = 1$, the SMCHD is equal to $\mathbb{E}[U_1 \mid T^1 \geq t] = \frac{\mu_1 p_1 \exp(-t\mu_1) + \mu_2 p_2 \exp(-t\mu_2)}{p_1 \exp(-t\mu_1) + p_2 \exp(-t\mu_2) + (1 - p_1 - p_2)}$, and deviates from the constant CHD equal to $\mathbb{E}[U_1] = p_1 \mu_1 + p_2 \mu_2$. For an RCT, the OHD equals the SMCHD, so that the integrated OHD equals

$$B(t) = \int_0^t \mathbb{E}[U_1 \mid T^1 \geq s] ds = -\log(p_1(\exp(-t\mu_1) - 1) + p_2(\exp(-t\mu_2) - 1) + 1).$$

Thus, although the CHD is time-invariant, due to the selection (of U_1) effect over time $B(t)$ will not be linear and deviates from the function $g(t) = t\mathbb{E}[U_1]$. Three types of curves could be observed as shown in Fig. 1 where for illustration, $p_1 = p_2 = 0.5$, so there exist only two levels for the modifier U_1.

First of all, let's consider the case where the exposure harms some individuals (for which $U_{1i} = 1$) while others do not respond to the exposure at all ($U_{1i} = 0$); see the orange line in Fig. 1. Initially, $B(t)$ evolves as $t\mathbb{E}[U_1] = 0.5t$. However, the individuals harmed by the exposure are less likely to survive over time, so the curve's derivative decreases. In the end, only individuals with $U_{1i} = 0$ are expected to survive so that $B(t)$ remains constant. Concluding that the exposure initially harms but loses effect over time is false for this case as the effect is time-invariant for each individual.

Secondly, when some individuals do benefit ($U_{1i} = -0.25$) while others are not affected ($U_{1i} = 0$), the derivative of $B(t)$ evolves from -0.125 to -0.25 at the moment only those that benefit are expected to survive, as illustrated with the purple line in Fig. 1. The effect for an individual is again constant and does not become more beneficial over time.

Finally, different individuals in the population might have opposite effects ($U_{1i} = 1$ or $U_{1i} = -0.1$), as illustrated with the pink line in Fig. 1. Initially, the integrated hazard differences increase as the expected effect is harmful. However, over time those

Fig. 1 $\int_0^t \mathbb{E}\left[U_1 \mid T^1 \geq s\right] ds$, when $U_1 \sim \mathrm{BHN}(p_1, \mu_1, p_2, \mu_2)$ (solid), and $g(t) = t\mathbb{E}[U_1]$ (dashed)
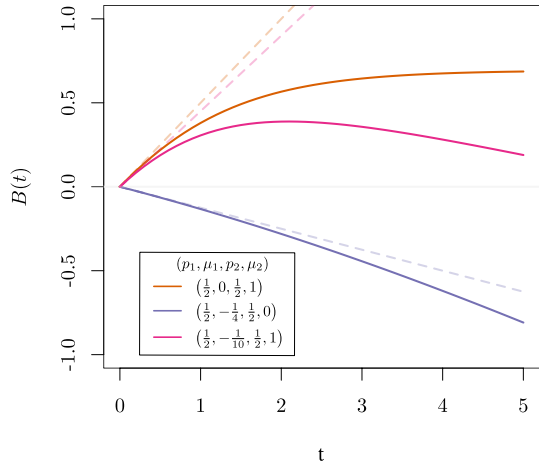
individuals with $U_{1i} = -0.1$ are more likely to survive so that $\mathbb{E}[U_1 \mid T^1 \geq 1]$ changes sign. Finally, only those that benefit are expected to survive, and the curve decreases with a derivative equal to -0.1. For this example, it would be false to conclude that the exposure first harms but becomes beneficial over time.

Similar patterns can be observed for a continuous U_1 distribution, in which case the $\mathbb{E}\left[U_1 \mid T^1 \geq t\right]$ will keep decreasing as for example presented in Appendix B.

In summary, if $U_0 \perp\!\!\!\perp U_1$, then the SMCHD will be less or equal to the CHD due to the selection of U_1. Therefore, decreasing or constant $B(t)$ curves that at some point increase again can not be explained by the selection of U_1 since individuals with less beneficial values of U_1 are expected to survive shorter. However, if $U_0 \not\!\perp\!\!\!\perp U_1$, such a pattern of the $B(t)$ curve can still occur when the CHD is time-invariant as we will show next.

3.2 Dependent U_0 and U_1

The bivariate joint distribution function of U_0 and U_1, $F_{(U_0, U_1)}$, can be written using the marginal distribution functions and a copula C (Sklar 1959). As such,

$$F_{(U_0, U_1)}(u_0, u_1) = C\left(F_{U_0}(u_0), F_{U_1}(u_1)\right)$$

and the Kendall's τ correlation coefficient of U_0 and U_1 can be written as a function of the copula (Nelsen 2006). For the next example, we consider cause-effect relations for which SCM (2) applies with

$$f_0(t, U_{0i}) = \ell + U_{0i}t^2, \tag{8}$$

and again

$$f_1(t, U_{1i}, a) = U_{1i}a, \tag{9}$$

while $U_{0i} \sim \Gamma(1,1)$ and $(U_{1i} + \ell) \sim \Gamma(1,1)$, so that the hazard is nonnegative for each individual. To illustrate how the dependence can affect the integrated SMCHD for the settings presented in Fig. 5 in Appendix B, we use a Gaussian copula

$$C(x,y) = \Phi_{2,\rho}(\Phi^{-1}(x), \Phi^{-1}(y)),$$

where Φ and $\Phi_{2,\rho}$ are the standard normal and standard bivariate normal with correlation ρ cumulative distribution functions, respectively. In Fig. 2, for $\rho \in \{-1, \sin(-0.25\pi), 0, \sin(0.25\pi), 1\}$ (such that $\tau \in \{-1, -0.5, 0, 0.5, 1\}$) and $\ell \in \{0, 0.5, 1\}$, we present the integrated SMCHD at time t that equals $\int_0^t \left(\mathbb{E}[U_1 \mid T^1 \geq s] + \mathbb{E}[U_0 s^2 \mid T^1 \geq s] - \mathbb{E}[U_0 s^2 \mid T^0 \geq s] \right) ds$ by Theorem 1. The conditional expectations are derived empirically from simulations (n=10,000), and the integral is approximated by taking discrete steps of size 0.1. For completeness, the survival curves of the potential outcomes can be found in Fig. 6 in Appendix C.

The difference between the integrated OHD and SMCHD increases when $\tau > 0$ (compared to $\tau = 0$). On the other hand, for $\tau < 0$, the difference is smaller most of the time than for $\tau = 0$ since favourable U_1 are expected to occur with unfavourable levels of U_0. Moreover, for $\tau = -1$, at larger t, we observe that the difference can even change sign. For $\tau \neq 0$, the SMCHD might thus be larger than the CHD, so the SMCHD is not a theoretical lower bound for the CHD. Note that if $U_0 \perp\!\!\!\perp U_1$, the integrated SMCHD depends on the functional form of f_0. In Fig. 7 in Appendix C, the results for $f_0(t, U_{0i}) = \ell + U_{0i} \frac{t^2}{20}$ are presented where the effect of the dependence is limited and the corresponding survival curves of the potential outcomes are presented in Fig. 8.

4 Case study: the Radiation Therapy Oncology Group trial

With the findings of the previous section we will reflect on a data analysis of an actual case study to illustrate why it is important for a practioner to be aware of the possible difference between the SMCHD and CHD. We consider a large clinical trial
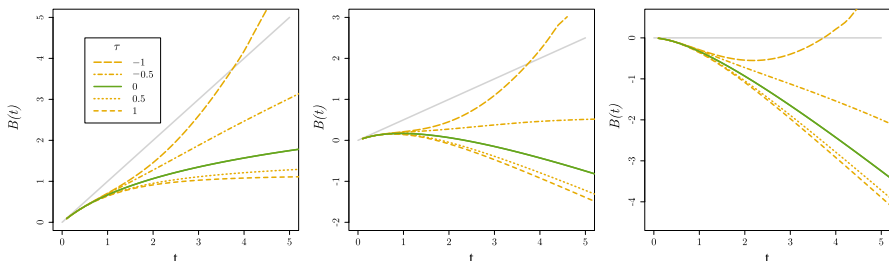
Fig. 2 Integrated hazard difference, $B(t)$, when $f_0(t, U_{0i}) = \ell + U_{0i}t^2$, $U_{0i} \sim \Gamma(1,1)$, $(U_{1i} + \ell) \sim \Gamma(1,1)$ for ℓ equal to 0 (left), $\frac{1}{2}$ (middle) and 1 (right) and different Kendall's τ for U_{0i} and U_{1i}. The lines for $\tau = 0$ were already presented in Fig. 5 in Appendix B. Furthermore, $g(t) = t\mathbb{E}[U_1]$ are presented as gray lines

carried out by the Radiation Therapy Oncology Group as described by Kalbfleisch and Prentice (2002, Section 1.1.2 and Appendix A) and also presented by Aalen (1989). From the patients with squamous cell carcinoma (a form of skin cancer) of 15 sites in the mouth and throat from 16 participating institutions, our focus is only on two sites (faucial arch and pharyngeal tongue) and patients from the six largest institutions. All participants were randomly assigned to radiation therapy alone or combined with a chemotherapeutic agent. So, we are interested in the causal effect of the chemotherapeutic agent in addition to radiation therapy on survival. If the causal mechanism can be parameterized with SCM (2) without effect heterogeneity, i.e.

$$\lambda_i^a(t) = f_0(t, U_{0i}) + f_1(t, a),$$

and the randomization was properly executed, implying $N_{Ai} \perp\!\!\!\perp U_{0i}$, then, by Theorem 1, the OHD equals the CHD. Moreover, the CHD can be unbiasedly estimated by fitting Aalen's additive hazard model. We did so by using the `aalen()` function from the package `timereg` in R. The estimated cumulative regression function (and a corresponding 95% confidence interval) of treatment combined with a chemotherapeutic agent is presented by the black lines in Fig. 3.

In the absence of effect heterogeneity (ignoring the statistical uncertainty), one could now conclude that initially adding the chemotherapy is expected to harm a patient as $B(t)$ takes on positive values and that the exposure loses its harmful effect over time as the derivative of $B(t)$ decreases over time. Following a similar reasoning, Aalen et al. (2008, pp. 160-161) discuss a conclusion on the effect of N-stage (an index of lymph node metastasis) on survival that may be drawn by practitioners based on the same dataset (while also including patients with a tumour located at the tonsillar fossa): *"The regression plot shows that this* [non-significant P-value for a zero-effect of N-stage from a Cox analysis] *is due to a strong initial positive effect being "watered down" by a lack of, or even a slightly negative effect after one year. Hence, not taking into consideration the change in effect over time may lead to missing significant effects."*. However, if in reality

Fig. 3 Estimated $B(t)$ and corresponding 95% confidence interval (black). Furthermore, the expected evolution of $B(t)$ when $\lambda_i^a(t) = f_0(t, U_{0i}) + U_{1i}a$ and $U_{1i} \sim \text{BHN}(0.5, -0.1, 0.5, 0.4)$ is presented (green)
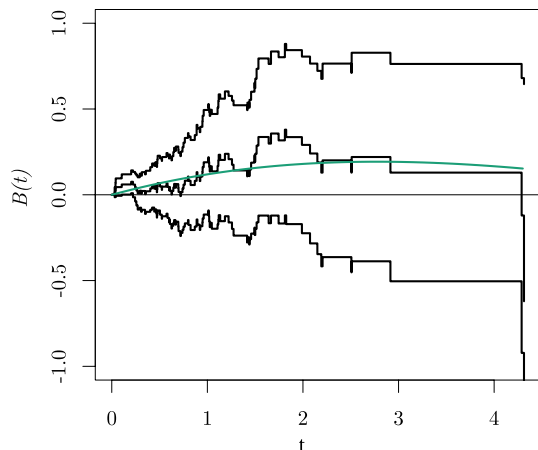
$$\lambda_i^a(t) = f_0(t, U_{0i}) + f_1(t, U_{1i}, a),$$

where $f_1(t, U_{1i}, 0) = 0$, the observed time-varying effect can also result from the modifier U_{1i} selection. For example, when $\lambda_i^a(t) = f_0(t, U_{0i}) + U_{1i}a$, and $U_{1i} \sim \mathrm{BHN}(0.5, -0.1, 0.5, 0.4)$, by Theorem 1, this pattern is expected (see the green line in Fig. 3) while the actual causal effect is time-invariant for each individual. The CHD equals 0.15 at each time point, but over time individuals that are harmed by the chemotherapy ($U_{1i} = 0.4$) are less likely to survive so that the SMCHD converges towards -0.1 (the effect for individuals that benefit from the chemotherapy). When we perform a stratified analysis by site in the oropharynx (where randomization remains), we observe that the effect of chemotherapy might have opposite effects for tumours located in the faucial arch and on the pharyngeal tongue, see Fig. 4. The tumour location could thus be the individual modifier underlying the BHN distribution. For this case study, we cannot be sure whether the effect of chemotherapy depends on the tumour location due to statistical uncertainty. However, it became clear that when statistical uncertainty is not the issue, it will be impossible to distinguish between a time-varying causal effect and a selection effect (of an unmeasured modifier) from data. Both phenomena can give rise to the same $B(t)$.

5 Discussion

The additive hazard model gives better interpretable estimates of causal effects than the proportional hazard model (Aalen et al. 2015). As discussed by Aalen et al. (2015), the model assumes that the additive part of the hazard involving the exposure (or treatment) is not affected by any other individual feature. Otherwise, if such effect heterogeneity at the hazard scale exists, we have shown that the SMCHD deviates from the CHD of interest. For an RCT, and independent censoring, a time-varying observed hazard difference can be the result of either an actual time-varying causal effect or of the selection of favourable effect-modifier levels over time.
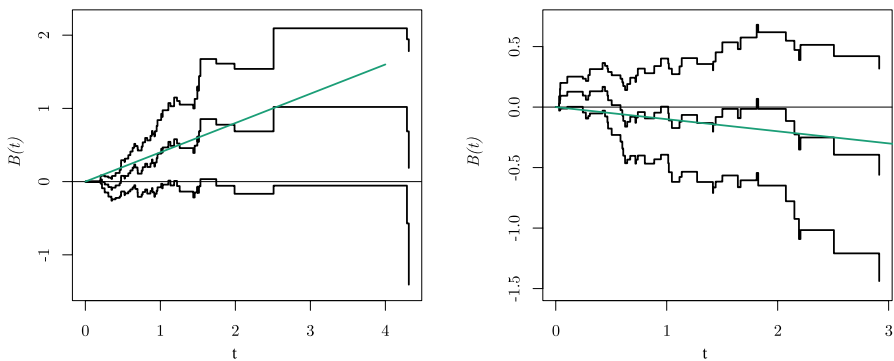
Fig. 4 Estimated $B(t)$ and corresponding 95% confidence interval (black) for patients with tumours located at the faucial arch (left) and pharyngeal tongue (right), respectively. Furthermore, the $B(t)$ for a homogeneous population is presented (green) equal to $0.4t$ (left) and $-0.1t$ (right) for comparison

Therefore, it is impossible to distinguish these scenarios based on data without making untestable assumptions. It is important to remark that for cause-effect relations that can be parameterized with SCM (2) where U_1 is degenerate (in which case the OHD equals the CHD), contrary to the individual hazard differences, the difference of the potential survival times, $T^1 - T^0$ can be heterogeneous. So, heterogeneous effects can still exist under Aalen's additive hazard model.

In the presented examples and the case study, we have illustrated that one should be very careful when concluding that the effect decreases over time based on the cumulative regression function, as this might result from the selection. The size of the bias depends on how much the distribution $F_{U_1|T^1 \geq t}$ changes over time. If the U_1 is low in variability, the bias will be small. When analyzing data from an RCT with an additive hazard model, it can thus be helpful to adjust for potential effect modifiers to reduce the remaining variability of unmeasured effect modifiers. We want to remark that for cause-effect relations that cannot be described by SCM (2), the CHD is not the appropriate estimand to quantify the causal effect, which is then a more serious concern than that the complicated causal interpretation of the observed hazard difference.

Even in the absence of confounding, the hazard difference and the hazard ratio (as discussed in Post et al. (2024)) have a difficult causal interpretation. Instead, contrasts of the survival probabilities, the median, or the restricted mean survival time, have clear causal interpretations and should thus be used to quantify the causal effect on time-to-event outcomes as suggested by others (Hernán 2010; Stensrud et al. 2018; Bartlett et al. 2020; Young et al. 2020). Nevertheless, (additive) hazard models can still be used for causal inference to derive these appropriate estimands (Ryalen et al. 2018).

Appendix A: Proofs

Appendix A.1: Proof of Theorem 1

Proof By causal consistency (Hernán and Robins 2020),

$$\lim_{h \to 0} h^{-1} \mathbb{P}(T \in [t, t+h] \mid T \geq t, A=a) = \lim_{h \to 0} h^{-1} \mathbb{P}(T^a \in [t, t+h] \mid T^a \geq t, A=a)$$

As $N_A \perp\!\!\!\perp U_0, U_1, N_T$, there is no confounding and $T^a \perp\!\!\!\perp A$, so that

$$\lim_{h \to 0} h^{-1} \mathbb{P}(T \in [t, t+h] \mid T \geq t, A=a) = \lim_{h \to 0} h^{-1} \mathbb{P}(T^a \in [t, t+h] \mid T^a \geq t,).$$

By the law of total probability,

$$\lim_{h \to 0} h^{-1} \mathbb{P}(T^a \in [t, t+h] \mid T^a \geq t)$$

$$= \lim_{h \to 0} \int h^{-1} \mathbb{P}\big(T^a \in [t, t+h] \mid T^a \geq t, U_0, U_1\big) dF_{U_0, U_1 | T^a \geq t}$$

First, we focus on the integrand,

$$h^{-1}\mathbb{P}\big(T^a \in [t, t+h) \mid T^a {\geq} t, U_0, U_1\big)$$

$$= h^{-1}\frac{\mathbb{P}\big(T^a {\geq} t \mid U_0, U_1\big) - \mathbb{P}\big(T^a {\geq} t + h \mid U_0, U_1\big)}{\mathbb{P}\big(T^a {\geq} t \mid U_0, U_1\big)}$$

$$= h^{-1}\left(1 - \frac{\mathbb{P}\big(T^a {\geq} t + h \mid U_0, U_1\big)}{\mathbb{P}\big(T^a {\geq} t \mid U_0, U_1\big)}\right)$$

$$= h^{-1}\left[1 - \frac{\exp\left(-\int_0^{t+h} f_0(s, U_0) + f_1(s, U_1, a)ds\right)}{\exp\left(-\int_0^{t} f_0(s, U_0) + f_1(s, U_1, a)ds\right)}\right]$$

$$= h^{-1}\left(1 - \exp\left(-\int_t^{t+h} f_0(s, U_0) + f_1(s, U_1, a)ds\right)\right)$$

For monotonic conditional hazard functions, if $h_2 < h_1$, then

$$h_1^{-1}\left(1 - \exp\left(-\int_t^{t+h_1} f_0(s, U_0) + f_1(s, U_1, a)ds\right)\right)$$

$$\leq h_2^{-1}\left(1 - \exp\left(-\int_t^{t+h_2} f_0(s, U_0) + f_1(s, U_1, a)ds\right)\right)$$

or

$$h_1^{-1}\left(1 - \exp\left(-\int_t^{t+h_1} f_0(s, U_0) + f_1(s, U_1, a)ds\right)\right)$$

$$\geq h_2^{-1}\left(1 - \exp\left(-\int_t^{t+h_2} f_0(s, U_0) + f_1(s, U_1, a)ds\right)\right)$$

as the average integrated conditional hazard over the interval increases (or decreases). Moreover,

$$\lim_{h \to 0} h^{-1}\mathbb{P}\big(T^a \in [t, t+h) \mid T^a {\geq} t, U_0, U_1\big) = f_0(s, U_0) + f_1(s, U_1, a) \geq 0.$$

Then, the limit and integral can be interchanged by directly applying the monotone convergence theorem.

For non-monotone conditional hazard functions, when Condition 1 applies, for every t, there exist a $\tilde{h}$ so that $\forall h^* \in (0, \tilde{h})$: $\mathbb{E}\big[f_0(t + h^*, U_0) + f_1(t + h^*, U_1, a) \mid T^a {\geq} t\big] < \infty$. Moreover, let $t^* = \arg\max_{s \in (t, t+\tilde{h})} f_0(s, U_0) + f_1(s, U_1, a)$, so that for $h < \tilde{h}$:

$$h^{-1}\mathbb{P}\big(T^a \in [t, t+h) \mid T^a {\geq} t, U_0, U_1\big) \leq h^{-1}\big(1 - \exp\big(-h\big(f_0(t^*, U_0) + f_1(t^*, U_1, a)\big)\big)\big).$$

Using the power series definition of the exponential function,

$$h^{-1} \mathbb{P}\big(T^a \in [t, t+h) \mid T^a \geq t, U_0, U_1\big)$$

$$\leq h^{-1} \left(1 - \frac{1}{\sum_{k=0}^{\infty} h^k (f_0(t^*, U_0) + f_1(t^*, U_1, a))^k \frac{1}{k!}} \right)$$

$$= h^{-1} \frac{\sum_{k=1}^{\infty} h^k (f_0(t^*, U_0) + f_1(t^*, U_1, a))^k \frac{1}{k!}}{\sum_{k=0}^{\infty} h^k (f_0(t^*, U_0) + f_1(t^*, U_1, a))^k \frac{1}{k!}}$$

$$= \big(f_0(t^*, U_0) + f_1(t^*, U_1, a)\big) \frac{\sum_{k=1}^{\infty} h^{k-1} (f_0(t^*, U_0) + f_1(t^*, U_1, a))^{k-1} \frac{1}{k!}}{\sum_{k=0}^{\infty} h^k (f_0(t^*, U_0) + f_1(t^*, U_1, a))^k \frac{1}{k!}}$$

$$= \big(f_0(t^*, U_0) + f_1(t^*, U_1, a)\big) \frac{\sum_{k=0}^{\infty} h^k (f_0(t^*, U_0) + f_1(t^*, U_1, a))^k \frac{1}{(k+1)!}}{\sum_{k=0}^{\infty} h^k (f_0(t^*, U_0) + f_1(t^*, U_1, a))^k \frac{1}{k!}}$$

$$< f_0(t^*, U_0) + f_1(t^*, U_1, a).$$

Moreover, $\mathbb{E}\big[f_0(t^*, U_0) + f_1(t^*, U_1, a) \mid T^a \geq t\big] < \infty$ when $\mathbb{E}[f_0(t+h, U_0) + f_1(t+h, U_1, a) \mid T^a \geq t] < \infty$ for all $h \in (0, \tilde{h})$. By application of the dominated convergence theorem, we can change the order of the limit and integral and conclude,

$$\lim_{h \to 0} h^{-1} \mathbb{P}(T^a \in [t, t+h) \mid T^a \geq t) = \mathbb{E}\big[f_0(t, U_0) + f_1(t, U_1, a) \mid T^a \geq t\big].$$

If $U_0 \perp\!\!\!\perp U_1$, by Lemma 1,

$$\mathbb{E}[f_0(t, U_0) + f_1(t, U_1, 1) \mid T^1 \geq t] - \mathbb{E}[f_0(t, U_0) \mid T^0 \geq t] = \mathbb{E}\big[f_1(t, U_1, 1) \mid T^1 \geq t\big],$$

and

$$\lim_{h \to 0} h^{-1} \mathbb{P}(T \in [t, t+h) \mid T \geq t, A=1) - \lim_{h \to 0} h^{-1} \mathbb{P}(T \in [t, t+h) \mid T \geq t, A=0)$$
$$= \mathbb{E}\big[f_1(t, U_1, 1) \mid T^1 \geq t\big].$$

$\square$

Appendix A.2: Proof of Lemma 1

Proof

$$\mathbb{E}\big[f_0(t, U_0) \mid T^a{\geq}t\big]$$

$$= \int f_0(t, u_0) dF_{U_0 \mid T^a \geq t}(u_0)$$

$$= \int f_0(t, U_0) \frac{\mathbb{P}(T^a{\geq}t \mid U_0{=}u_0)}{\mathbb{P}(T^a{\geq}t)} dF_{U_0}(u_0)$$

$$= \int f_0(t, U_0) \frac{\int \exp\Big(-(\int_0^t f_0(s, U_0)ds + \int_0^t f_1(s, U_1, a)ds)\Big) dF_{U_1 \mid U_0 = u_0}(u_1)}{\int \int \exp\Big(-(\int_0^t f_0(k_0, s)ds + \int_0^t f_1(k_1, a, s)ds)\Big) dF_{U_1 \mid U_0 = k_0}(k_1) dF_{U_0}(k_0)} dF_{U_0}(u_0)$$

$$= \int f_0(t, U_0) \frac{\exp\Big(-\int_0^t f_0(s, U_0)ds\Big)\Big(\int \exp\Big(-\int_0^t f_1(s, U_1, a)ds\Big) dF_{U_1 \mid U_0 = u_0}(u_1)\Big)}{\int \exp\Big(-\int_0^t f_0(k_0, s)ds\Big)\Big(\int \exp\Big(-\int_0^t f_1(k_1, a, s)ds\Big) dF_{U_1 \mid U_0 = k_0}(k_1)\Big) dF_{U_0}(k_0)} dF_{U_0}(u_0).$$

Moreover, if $U_0 \perp\!\!\!\perp U_1$, then

$$\mathbb{E}\big[f_0(t, U_0) \mid T^a{\geq}t\big]$$

$$= \int f_0(t, U_0) \frac{\exp\Big(-\int_0^t f_0(s, U_0)ds\Big)\Big(\int \exp\Big(-\int_0^t f_1(s, U_1, a)ds\Big) dF_{U_1}(u_1)\Big)}{\int \exp\Big(-\int_0^t f_0(k_0, s)ds\Big)\Big(\int \exp\Big(-\int_0^t f_1(k_1, a, s)ds\Big) dF_{U_1}(k_1)\Big) dF_{U_0}(k_0)} dF_{U_0}(u_0)$$

$$= \int f_0(t, U_0) \frac{\exp\Big(-\int_0^t f_0(s, U_0)ds\Big)\Big(\int \exp\Big(-\int_0^t f_1(s, U_1, a)ds\Big) dF_{U_1}(u_1)\Big)}{\Big(\int \exp\Big(-\int_0^t f_0(k_0, s)ds\Big) dF_{U_0}(k_0)\Big)\Big(\int \exp\Big(-\int_0^t f_1(k_1, a, s)ds\Big) dF_{U_1}(k_1)\Big)} dF_{U_0}(u_0)$$

$$= \int f_0(t, U_0) \frac{\exp\Big(-\int_0^t f_0(s, U_0)ds\Big)}{\int \exp\Big(-\int_0^t f_0(k_0, s)ds\Big) dF_{U_0}(k_0)} dF_{U_0}(u_0)$$

$$= \mathbb{E}\big[f_0(t, U_0) \mid T^0{\geq}t\big].$$

$\square$

Appendix A.3: Proof of Lemma 2

Proof If $U_0 \perp\!\!\!\perp U_1$, by Bayes rule, the probability density of U_1 given $T^1 \geq t$, $f(u_1 \mid T^1{\geq}t)$, equals

$$f(u_1|T^1 \geq t)$$

$$= \frac{\mathbb{P}(T^1 \geq t \mid U_1 = u_1)f(u_1)}{\int \mathbb{P}(T^1 \geq t \mid U_1 = u_1)dF_{U_1}(u_1)}$$

$$= \frac{\int \exp\left(-\left(\int_0^t f_0(s, U_0) + u_1 f_1(a, s)ds\right)\right)dF_{U_0}(u_0)f(u_1)}{\int \int \exp\left(-\left(\int_0^t f_0(k_0, s) + k_1 f_1(a, s)ds\right)\right)dF_{U_1}(k_1)dF_{U_0}(k_0)}$$

$$= \frac{\exp\left(-\int_0^t u_1 f_1(a, s)ds\right) \int \exp\left(-\int_0^t f_0(s, U_0)ds\right)dF_{U_0}(u_0)f(u_1)}{\int \exp\left(-k_1 f_1(a, s)ds\right)dF_{U_1}(k_1) \int \exp\left(-\int_0^t f_0(k_0, s)ds\right)dF_{U_0}(k_0)}$$

$$= \frac{\exp\left(-\int_0^t u_1 f_1(a, s)ds\right)f(u_1)}{\int \exp\left(-k_1 f_1(a, s)ds\right)dF_{U_1}(k_1)}.$$

So that the Laplace transform of $f(u_1|T^1 \geq t)$ can be written as

$$\mathcal{L}_{U_1|T^1 \geq t}(c) = \mathbb{E}[\exp\left(-U_1 c\right) \mid T^1 \geq t]$$

$$= \int \exp\left(-u_1 c\right)dF_{U_1|T^1 \geq t}(u_1)$$

$$= \int \exp\left(-u_1 c\right)\frac{\exp\left(-\int_0^t u_1 f_1(s, 1)ds\right)}{\int \exp\left(-k_1 f_1(s, 1)ds\right)dF_{U_1}(k_1)}dF_{U_1}(u_1)$$

$$= \frac{\int \exp\left(-u_1(c + \int_0^t f_1(s, 1)ds)\right)}{\int \exp\left(-k_1 f_1(s, 1)ds\right)dF_{U_1}(k_1)}dF_{U_1}(u_1)$$

$$= \frac{\mathbb{E}\left[\exp\left(-U_1(c + \int_0^t f_1(s, 1)ds)\right)\right]}{\mathbb{E}\left[\exp\left(-U_1 f_1(s, 1)ds\right)\right]}$$

$$= \frac{\mathcal{L}_{U_1}(c + \int_0^t f_1(s, 1)ds)}{\mathcal{L}_{U_1}(\int_0^t f_1(s, 1)ds)}.$$

Since for a random variable X, $\mathbb{E}[X] = -\mathcal{L}_X'(0)$,

$$\mathbb{E}[U_1 \mid T^1 {\geq} t] = -\mathcal{L}_{U_1 \mid T^1 \geq t}'(0) = -\frac{\mathcal{L}_{U_1}'(\int_0^t f_1(s,1)ds)}{\mathcal{L}_{U_1}(\int_0^t f_1(s,1)ds)}.$$

$\square$

Appendix B: Continuous U_1 distribution ($U_0 \perp\!\!\!\perp U_1$)

Let us consider cause-effect relations for which SCM (2) applies with

$$f_1(t, U_{1i}, a) = U_{1i}a, \tag{10}$$

and $U_0 \perp\!\!\!\perp U_1$. Moreover, let $(U_1 + \ell) \sim \Gamma(k, \theta)$, then $\mathcal{L}_{U_1+\ell}(c) = (1 + \theta c)^{-k}$. By lemma 2,

$$\mathbb{E}\big[U_1 \mid T^1 {\geq} t\big] = \frac{\theta k}{\theta t + 1} - \ell,$$

and

$$B(t) = k \log(\theta t + 1) - \ell t.$$

The $B(t)$ are presented over time in Fig. 5 for $\theta = k = 1$ and $\ell \in \{0, \frac{1}{4}, \frac{1}{2}, 1\}$.

Fig. 5 $\int_0^t \mathbb{E}\big[U_1 \mid T^1 {\geq} s\big] ds$ (thick lines), when $U_1 + \ell \sim \Gamma(1, 1)$ and $g(t) = t\mathbb{E}[U_1]$ (transparent lines)
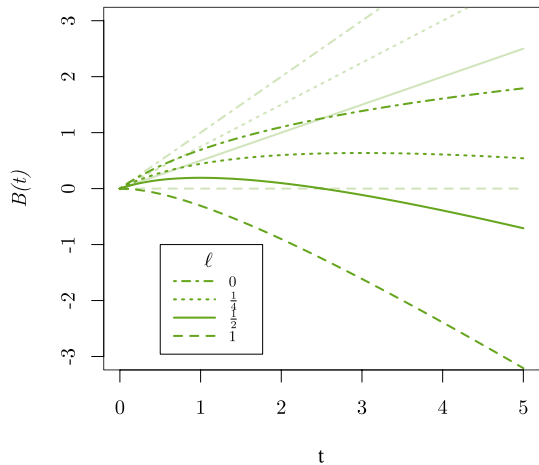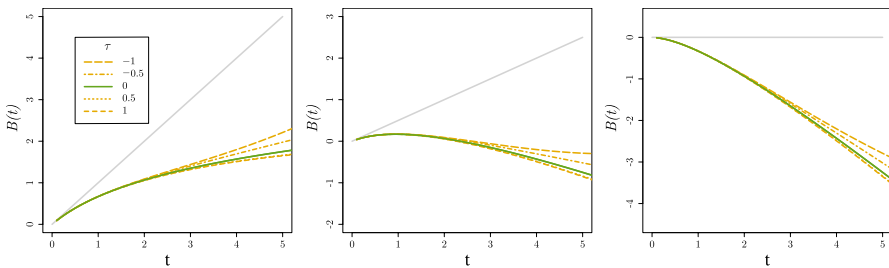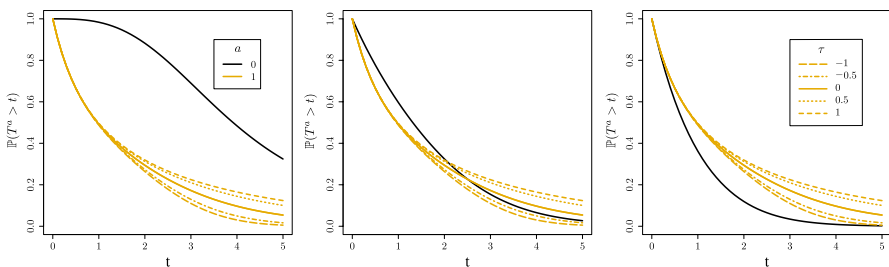
Appendix C: Figures dependent U_0 and U_1

The additional figures for $f_0(t, U_{0i}) = \ell + U_{0i}\frac{t^2}{20}$.
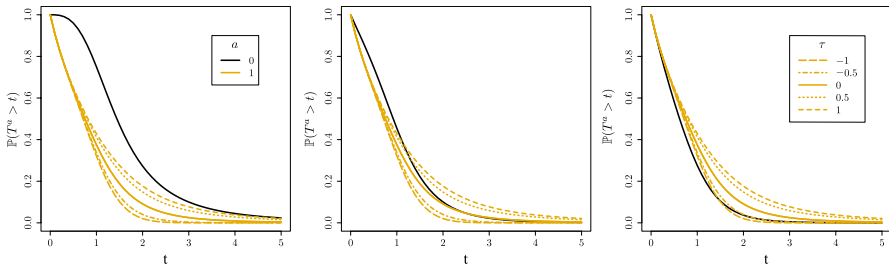
Fig. 6 Survival curves for Y^1 and Y^0 when $f_0(t, U_{0i}) = \ell + U_{0i}t^2$, $U_{0i} \sim \Gamma(1, 1)$, $(U_{1i} + \ell) \sim \Gamma(1, 1)$ for ℓ equal to 0 (left), $\frac{1}{2}$ (middle) and 1 (right) and different Kendall's τ for U_{0i} and U_{1i}

Fig. 7 $B(t)$, when $f_0(t, U_{0i}) = \ell + U_{0i}\frac{t^2}{20}$, $U_{0i} \sim \Gamma(1, 1)$, $(U_{1i} + \ell) \sim \Gamma(1, 1)$ for ℓ equal to 0 (left), $\frac{1}{2}$ (middle) and 1 (right) and different Kendall's τ for U_{0i} and U_{1i}. The lines for $\tau = 0.5$ and $\tau = 1$ do overlap, and the lines for $\tau = 0$ were already presented in Fig. 5. Furthermore, $g(t) = t\mathbb{E}[U_1]$ are presented as gray lines

Fig. 8 Survival curves for Y^1 and Y^0 when $f_0(t, U_{0i}) = \ell + U_{0i}\frac{t^2}{20}$, $U_{0i} \sim \Gamma(1, 1)$, $(U_{1i} + \ell) \sim \Gamma(1, 1)$ for ℓ equal to 0 (left), $\frac{1}{2}$ (middle) and 1 (right) and different Kendall's τ for U_{0i} and U_{1i}

References

Aalen OO (1989) A linear regression model for the analysis of life times. Stat Med 8(8):907–925

Aalen OO, Borgan Ø, Gjessing HK (2008) Survival and event history analysis, 1st edn. Springer, Berlin

Aalen OO, Cook RJ, Røysland K (2015) Does Cox analysis of a randomized survival study yield a causal treatment effect? Lifetime Data Anal 21(4):579–593

Bartlett JW, Morris TP, Stensrud MJ, Daniel RM, Vansteelandt SK, Burman CF (2020) The hazards of period specific and weighted hazard ratios. Stat Biopharm Res 12(4):518–519

Bongers S, Forré P, Peters J, Mooij JM (2021) Foundations of structural causal models with cycles and latent variables. Ann Stat 49(5):2885–2915

Cox DR (1972) Regression models and life-tables. J R Stat Soc Ser B 34(2):187–220

Hernán MA (2010) The hazards of hazard ratios. Epidemiology 21(1):13–15

Hernán MA, Robins JM (2020) Causal inference: what if. Chapman & Hall/CRC, Boca Raton, Florida, Boca Raton

Holland PW (1986) Statistics and causal inference. J Am Stat Assoc 81(396):945–960

Kalbfleisch JD, Prentice RL (2002) The statistical analysis of failure time data, 2nd edn. Wiley

Lin DY, Ying Z (1994) Semiparametric analysis of the additive risk model. Biometrika 81(1):61–71

Martinussen T, Vansteelandt S (2013) On collapsibility and confounding bias in Cox and Aalen regression models. Lifetime Data Anal 19(3):279–296

Martinussen T, Vansteelandt S, Andersen P (2020) Subtleties in the interpretation of hazard contrasts. Lifetime Data Anal 26(4):833–855

McKeague IW, Sasieni PD (1994) A partly parametric additive risk model. Biometrika 81(3):501–514

Nelsen RB (2006) An introduction to copulas, 2nd edn. Springer, Berlin

Neyman J (1923) On the application of probability theory to agricultural experiments. Essay on principles. Stat Sci 5(4):465–472

Pearl J (2009) Causal inference in statistics: an overview. Stat Surv 3:96–146

Pearl J (2009) Causality: models, reasoning, and inference, 2nd edn. Cambridge University Press, Cambridge

Peters J, Janzing D, Schölkopf B (2018) Elements of causal inference: foundations and learning algorithms. The MIT Press, Cambridge, Massachusetts

Post RAJ, van den Heuvel ER, Putter H (2024) The built-in selection bias of hazard ratios formalized using structural causal models. Lifetime Data Anal (2024). https://doi.org/10.1007/s10985-024-09617-y

Rubin DB (1974) Estimating causal effects of treatments in randomized and nonrandomized studies. J Educ Psychol 66(5):688–701

Ryalen PC, Stensrud MJ, Røysland K (2018) Transforming cumulative hazard estimates. Biometrika 105(4):905–916

Sklar A (1959) Fonctions de répartition à n dimensions et leurs marges. Publications de l'Institut de statistique de l'Université de Paris 8:229–231

Stensrud MJ, Valberg M, Røysland K, Aalen OO (2017) Exploring selection bias by causal frailty models: the magnitude matters. Epidemiology 28(3):379–386

Stensrud MJ, Aalen JM, Aalen OO, Valberg M (2018) Limitations of hazard ratios in clinical trials. Eur Heart J 40(17):1378–1383

Young JG, Stensrud MJ, Tchetgen Tchetgen EJ, Hernán MA (2020) A causal framework for classical statistical estimands in failure-time settings with competing events. Stat Med 39(8):1199–1236

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.