Subtleties in the interpretation of hazard contrasts

Torben Martinussen[1] · **Stijn Vansteelandt[2,3]** · **Per Kragh Andersen[1]**

Abstract

The hazard ratio is one of the most commonly reported measures of treatment effect in randomised trials, yet the source of much misinterpretation. This point was made clear by Hernán (Epidemiology (Cambridge, Mass) 21(1):13–15, 2010) in a commentary, which emphasised that the hazard ratio contrasts populations of treated and untreated individuals who survived a given period of time, populations that will typically fail to be comparable—even in a randomised trial—as a result of different pressures or intensities acting on different populations. The commentary has been very influential, but also a source of surprise and confusion. In this note, we aim to provide more insight into the subtle interpretation of hazard ratios and differences, by investigating in particular what can be learned about a treatment effect from the hazard ratio becoming 1 (or the hazard difference 0) after a certain period of time. We further define a hazard ratio that has a causal interpretation and study its relationship to the Cox hazard ratio, and we also define a causal hazard difference. These quantities are of theoretical interest only, however, since they rely on assumptions that cannot be empirically evaluated. Throughout, we will focus on the analysis of randomised experiments.

Keywords Causality · Cox regression · Hazard difference · Hazard ratio · Randomised study · Survival analysis

1 Introduction

The popularity of the Cox regression model has contributed to the enormous success of the hazard ratio as a concise summary of the effect of a randomised treatment on a survival endpoint. Notwithstanding this, use of the hazard ratio has been criticised over recent years. Hernán (2010) argued that selection effects caused by unobserved

---

✉ Torben Martinussen
   tma@sund.ku.dk

Extended author information available on the last page of the article

heterogeneity ("frailty") render a causal interpretation of the hazard ratio difficult when treatment affects outcome. While the treated and untreated people are comparable by design at baseline, the treated people who survive a given time t may then tend to be more "frail" (as a result of lower mortality if treatment is beneficial) than the untreated people who survive the given time t, so that the crucial comparability of both groups is lost at that time. Aalen et al. (2015) re-iterated Hernán's concern. They viewed the problem more as one of non-collapsibility (Martinussen and Vansteelandt 2013), which is a concern about the interpretation of the hazard ratio, though not about its justification as a causal contrast. In particular, they argued that the magnitude of the hazard ratio typically changes as one evaluates it in smaller subgroups of the population (e.g. frail people), even in the absence of interaction effects on the log hazard scale.

In this paper, we aim to develop more insight into these matters. We pay special attention to what can be learned about the treatment effect from the hazard contrast vanishing after a certain point in time, i.e. the hazard ratio becoming 1 or the hazard difference becoming 0. Throughout, we will assume that data are available from a randomised experiment on the effect of a dichotomous treatment A (coded 1 for treatment and 0 for control) on a survival endpoint T, so that issues of confounding can be ignored. Time-changing hazard ratios are often interpreted as a change in treatment effect. Two recent examples of this are Lederle et al. (2019) and Primrose et al. (2019). We argue that there is no causal support for such reporting, however.

A causal contrast compares the same population under the hypothetical scenarios 'everybody treated' versus 'everybody untreated' and we furthermore define a causal hazard ratio, which contrasts the intensities at time t between the treatment and control groups of subjects who would survive up to time t regardless of treatment. Its magnitude is indeed causally interpretable in the sense that it makes a contrast for the same principal stratum of subjects with $(T^0 \geq t, T^1 \geq t)$, where T^a, $a = 0, 1$, is the potential lifetime under treatment a. This is related to earlier work on defining causal effects based on principal strata, see Frangakis and Rubin (2002) and Bartolucci and Grilli (2011). We show that this causal hazard ratio, HR(t), is smaller than the Cox hazard ratio at all times $t > 0$ if the joint distribution of (T^0, T^1) is governed by a so-called Archimedean copula as generated by a frailty setting. This corroborates intuition, also given in Hernán (2010), that the Cox hazard ratio underestimates the causal effect (on the hazard scale) of a beneficial treatment as a result of the treatment group then containing relatively more frail subjects compared to the control group at each time $t > 0$. We surprisingly find this intuition to break down in a more general setting where there is still positive correlation between T^0 and T^1, but the joint distribution of (T^0, T^1) is not governed by an Archimedean copula. We also define a causal hazard difference, however, both this measure and the suggested causal hazard ratio have the drawback that their causal interpretations rely on untestable assumptions.

If the Cox model (or an extended version with a time-changing coefficient) is correctly specified then the parameters of the Cox model are estimated consistently by use of the Cox partial score function and the Breslow estimator. Also the standard estimators of the variability of these estimators are consistent. Therefore, inference based on these classical tools, such as the log-rank test, is valid. Our point with this paper, however, is that the obtained hazard ratio cannot generally be given a causal interpretation when interpreted as a hazard ratio, except in the trivial case where there

is no causal effect, i.e. a hazard ratio of 1 or if an unverifiable assumption holds to which we return later.

2 The Cox model and causal reasoning

2.1 Hazard ratios

Analyses of time-to-event endpoints in randomised experiments are commonly based on the Cox model

$$\lambda(t; a) = \lambda_0(t)e^{\beta a},$$

where $\lambda(t; a)$ denotes the hazard function of T given $A = a$, evaluated at time t, and $\lambda_0(t)$ is the unspecified baseline hazard function. This model implies that for all t

$$P(T > t|A = 1) = P(T > t|A = 0)^{\exp(\beta)}, \tag{1}$$

so that $\exp(\beta)$ can be interpreted as

$$\exp(\beta) = \frac{\log P(T > t|A = 1)}{\log P(T > t|A = 0)},$$

a ratio between cumulative hazards. This represents a causal contrast (i.e., it compares a functional of the survival time distribution for the same population under different interventions). Since randomisation ensures that $T^a \perp\!\!\!\perp A$ for $a = 0, 1$, we have by consistency, $T = T^a$ if $A = a$, that

$$\exp(\beta) = \frac{\log P(T^1 > t)}{\log P(T^0 > t)}.$$

This shows that, under the proportional hazard assumption, $\exp(\beta)$ forms a valid one-number summary providing a relative and causal contrast of what the log survival probability would be at an arbitrary time t if everyone were treated, versus what it would be at that time if no one were treated.

The log-transformation makes the above interpretation of $\exp(\beta)$, while causal, difficult. It is therefore more common to interpret $\exp(\beta)$ as a hazard ratio

$$\exp(\beta) = \frac{\lim_{h\to 0} P(t \leq T < t + h|T \geq t, A = 1)}{\lim_{h\to 0} P(t \leq T < t + h|T \geq t, A = 0)}$$

$$= \frac{\lim_{h\to 0} P(t \leq T^1 < t + h|T^1 \geq t)}{\lim_{h\to 0} P(t \leq T^0 < t + h|T^0 \geq t)}.$$

Interpretation appears simpler now, but this is somewhat deceptive for two reasons. First, the righthand expression shows that $\exp(\beta)$ contrasts the hazard functions with and without intervention for two separate groups of individuals, those who survive time $t > 0$ with treatment ($T^1 \geq t$) and those who survive time $t > 0$ without treatment ($T^0 \geq t$). Those groups will typically fail to be comparable if treatment affects outcome (Hernán 2010). In particular, when treatment has a beneficial effect then, despite randomisation, the subgroup $T^1 \geq t$ in the numerator will generally contain more frail people than the subgroup $T^0 \geq t$ in the denominator, where the frailest people may have died already. When viewed as a hazard ratio, $\exp(\beta)$ therefore does not represent a causal contrast and it would be unwise to state that 'treatment works in the same way throughout time'. Second, the interpretation of $\exp(\beta)$ as a hazard ratio is further complicated by it being non-collapsible, so that its magnitude typically becomes more pronounced as one evaluates smaller subgroups of the study population (Martinussen and Vansteelandt 2013; Aalen et al. 2015).

2.2 MRC RE01 study

As an illustration, we reconsider the kidney cancer data described in White and Royston (2009). These data are from the MRC RE01 study which was a randomised controlled trial comparing interferon-α (IFN) treatment with the best supportive care and hormone treatment with medroxyprogesterone acetate (control) in patients with metastatic renal carcinoma. We use the same 347 patients as in White and Royston (2009). In this illustrative analysis we consider only the first 30 months of follow up. The median follow-up time was 242 days, and 85% of the patients died within the considered time frame. The two Kaplan-Meier estimates in Fig. 1 contain all the available information about the treatment effect. IFN treatment seems superior to the standard treatment (control), although a supremum test comparing the two survival curves results in a non-significant p value of 0.09. The score process plot of Lin et al. (1993) was calculated using the R-package `timereg` (Martinussen and Scheike 2006), giving no evidence against the proportional hazards assumption ($P = 0.43$ according to the supremum test). We therefore fitted the Cox model, giving the estimate $\hat{\beta} = -0.29$ (SE 0.12, $P = 0.01$). The corresponding hazard ratio of 0.75 expresses that IFN treatment reduces the log survival probability at each time by 25% (on a relative scale), as compared to control. This interpretation is causal, but not insightful. Results are therefore best communicated by visualising identity (1) in terms of estimated survival curves with versus without treatment (Hernán 2010). This has the advantage that it provides better insight into the possible public health impact of the intervention, but the drawback that it does not permit a compact way of reporting and that survival curves do not provide an understanding of a possible dynamic treatment effect. To enable a more in-depth understanding, it is tempting to interpret $\exp(\beta)$ as a hazard ratio, but then interpretation becomes subtle. We will demonstrate this in more detail in the next section, where we will investigate to what extent hazard ratios may provide insight into the dynamic nature of the treatment effect.
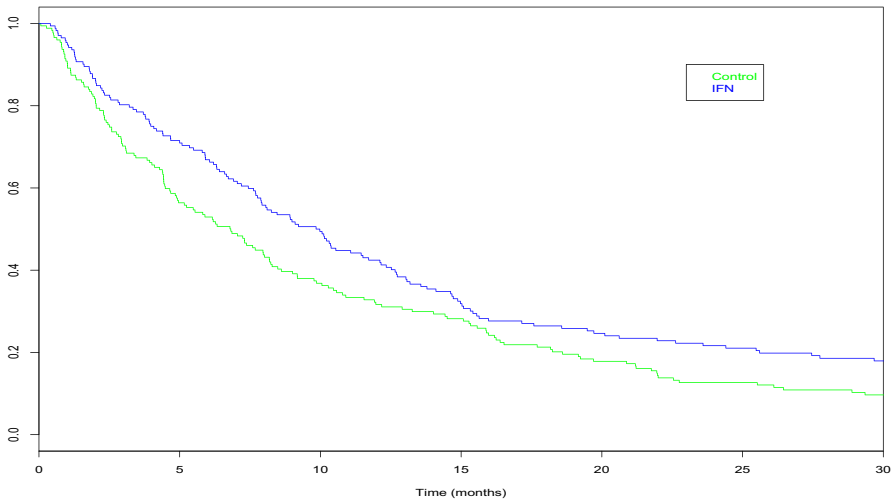
Fig. 1 MRC RE01 study. Kaplan Meier plot, control group (green curve) and IFN group (blue curve) (Color figure online)

3 Time-varying hazard contrasts are not causally interpretable

3.1 Hazard ratios

Consider now a study where the hazard ratio changes with time in the following sense:

$$\frac{\lambda(t; A = 1)}{\lambda(t; A = 0)} = \begin{cases} \exp(\beta_1) & \text{if } t \leq \nu \\ \exp(\beta_2) & \text{if } t > \nu \end{cases} \tag{2}$$

with $\beta_1 \neq \beta_2$, where $\nu > 0$ denotes the change point. Suppose in particular that $\beta_1 < 0$ and $\beta_2 = 0$, commonly interpreted as implying that the treatment is beneficial for $t < \nu$ but ineffective for $t > \nu$.

To develop a greater understanding whether indeed hazard ratios permit such dynamic understanding of the treatment effect, we will first study data-generating processes (DGP) that could give rise to (2); in "Appendix A.2", we formulate a more general DGP. Let Z represent the participants' unmeasured baseline frailty (higher means more frail), which affects T, but is independent of A by randomisation. Suppose that the hazard function $\lambda(t; a, z)$ of T given $A = a$ and $Z = z$ satisfies

$$\lambda(t; a, z) = z\lambda^*(t; a), \tag{3}$$

for some function $\lambda^*(t; a)$, and let Z be Gamma distributed with mean 1 and variance θ.

We will investigate what choices of $\lambda^*(t; a)$ give rise to model (2). With $\phi_Z(u) = E(e^{-Zu})$ the Laplace transform associated with the distribution of Z, the following relationship between the hazard function of interest $\lambda(t; a)$, and $\lambda^*(t; a)$, can be shown to hold:

$$\Lambda^*(t; a) = \phi_Z^{-1}(e^{-\Lambda(t;a)}) = \frac{1 - e^{-\theta\Lambda(t;a)}}{\theta e^{-\theta\Lambda(t;a)}},$$

where $\Lambda(t; a) = \int_0^t \lambda(s; a)\, ds$ and similarly with $\Lambda^*(t; a)$. Simple calculations then show that model (3) implies model (2) when

$$\lambda(t; a, Z) = \begin{cases} Z\lambda_0(t)e^{\beta_1 a}\exp\{\theta\Lambda_0(t)e^{\beta_1 a}\} & \text{if } t \le v \\ Z\lambda_0(t)e^{\beta_2 a}\exp\{\theta\Lambda_0(v)e^{\beta_1 a} + \theta\Lambda_0(v, t)e^{\beta_2 a}\} & \text{if } t > v \end{cases} \quad (4)$$

where $\Lambda_0(v, t) = \int_v^t \lambda_0(s)\, ds$. The specific $\lambda^*(t; a)$ that results in model (2) can be read off from (4). For subjects with given $Z = z$, it follows that the conditional hazard ratio is

$$\begin{aligned} \text{HR}_Z(t) &= \frac{\lambda(t; A = 1, Z)}{\lambda(t; A = 0, Z)} \\ &= \begin{cases} \exp(\beta_1)\exp\left[\theta\Lambda_0(t)\{\exp(\beta_1) - 1\}\right] & \text{if } t \le v \\ \exp(\beta_2)\exp\left[\theta\Lambda_0(v)\{\exp(\beta_1) - 1\} + \theta\Lambda_0(v, t)\{\exp(\beta_2) - 1\}\right] & \text{if } t > v \end{cases} \end{aligned}$$

For $\beta_2 = 0$, this simplifies to

$$\frac{\lambda(t; A = 1, Z)}{\lambda(t; A = 0, Z)} = \begin{cases} \exp(\beta_1)\exp\left[\theta\Lambda_0(t)\{\exp(\beta_1) - 1\}\right] & \text{if } t \le v \\ \exp\left[\theta\Lambda_0(v)\{\exp(\beta_1) - 1\}\right] & \text{if } t > v \end{cases}$$

which is less than 1 for all $t > 0$ since $\beta_1 < 0$. This model suggests that the treatment would be beneficial at all $t > 0$ for individuals with the same value of z, regardless of z. This contradicts the earlier, naïve interpretation that, across all individuals combined, treatment is ineffective from time v onwards.

The root cause of these contradictory conclusions is the fact that the hazard ratio at a given time does not express a causal effect (Hernán 2010). This has nothing to do with model misspecification, as the models considered either hold by construction or (in the MRC RE01 example) hold reasonably well by a goodness-of-fit examination. To appreciate that the treated and untreated risk sets indeed lose their exchangeability over time, note that

$$\frac{E(Z|T > t, A = 1)}{E(Z|T > t, A = 0)} = \begin{cases} \exp\{\Lambda_0(t)(1 - e^{\beta_1})\} & \text{if } t \le v \\ \frac{\exp\{\Lambda_0(v)\} + \Lambda_0(v, t)}{\exp\{\Lambda_0(v)e^{\beta_1}\} + \Lambda_0(v, t)} & \text{if } t > v. \end{cases}$$

when $\beta_2 = 0$ and $\theta = 1$. This shows that, while Z is independent of A by randomisation, this independence is not generally maintained within subgroups of survivors, where we are left with more frail subjects in the active treatment group: $E(Z|T > t, A = 1) > E(Z|T > t, A = 0)$. That selection takes place, does not rely on Z being Gamma-distributed as shown in "Appendix A.1", where we consider a situation with Z being binary.

3.2 Gastrointestinal tumour study

Stablein and Koutrouvelis (1985) presented survival data from a randomised clinical trial on locally unresectable gastric cancer. Half of the total of 90 patients were assigned to chemotherapy, and the other half to combined chemotherapy and radiotherapy.

It was suggested that there was superior survival for patients who received chemotherapy, but only in the first year or so. The same application was considered by Collett (2015), pp. 386–389. For illustrative purposes we consider here the first 720 days of follow up corresponding to the two first time periods considered by Collett (2015). The Kaplan Meier curves corresponding to the two groups are shown in Figure 1 in the supplementary material. It is seen that the survival curves come close at the end of the considered time interval. Applying a Cox regression model with time-by-treatment-interaction, allowing for separate HR's before and after 1 year of follow-up gives estimated HR's (Combined vs chemotherapy) of 2.40 (95% CI; 1.25, 4.63) in the first year and 0.78 (95% CI; 0.34,1.76) thereafter. The supremum score process test of Lin et al. (1993) gave no convincing evidence against these two models, with $P = 0.07$ in the first interval and $P = 0.26$ in the second interval. It is tempting to conclude that the chemotherapy is beneficial in the first year only, and that there is possibly a reverse effect afterwards; see Collett (2015) for a similar analysis and conclusion. However, such conclusion is not necessarily supported by the data. We demonstrate this using the two estimated hazard ratios and assuming the DGP (3) with the frailty variable being Gamma distributed with mean and variance equal to θ, corresponding to a Kendall's τ of 0.3. Figure 2 in the Supplementary Material displays the estimated hazard ratio $\lambda(t, A = 1, Z)/\lambda(t, A = 0, Z)$. It is seen to depend on time, being larger than 2.4 and increasing towards 3.7 in the first year and, after the change-point (1 year), starting at around 1.2 and then decreasing but being larger than 1 at all times. The reason this effect reversal does not occur under this model is that when chemotherapy is the more effective treatment, then there will also be relatively more and more frail subjects in that group over time.

3.3 Hazard differences

For an additive hazards model

$$\lambda(t; a) = \lambda_0(t) + \beta a,$$

following the arguments for the Cox models in Sect. 2, the hazard difference β may be written as

$$\beta = -\frac{1}{t}\big(\log P(T > t|A = 1) - \log P(T > t|A = 0)\big)$$

which equals

$$\beta = -\frac{1}{t}\big(\log P(T^1 > t) - \log P(T^0 > t)\big)$$

when the treatment is randomly assigned. This is a causal contrast, but suffers from the same interpretational difficulties as does the hazard ratio.

It is nonetheless tempting to think that hazard differences have a more appealing interpretation than hazard ratios, apart from them being collapsible (Martinussen and Vansteelandt 2013). Indeed, suppose that the additive hazards model

$$\lambda(t; A, Z) = \psi(t)A + \omega(t, Z), \tag{5}$$

holds for general functions ψ, ω and (possibly unmeasured) baseline covariates Z. Then the baseline exchangeability of treated and untreated individuals w.r.t. the covariates Z, as guaranteed by randomisation, extends to all risk sets, in the sense that $A \perp\!\!\!\perp Z|T > t$ (Vansteelandt et al. 2014; Aalen et al. 2015). The previous concerns about selection may therefore appear of less relevance for (time-varying) hazard differences. However, this is generally not the case. In Sect. 4.3, we show that an Aalen additive hazards analysis can easily result in the same misleading conclusion as what we obtained from the Cox regression analysis with treatment-by-time interaction in Sect. 3.1. In Sect. 4.3, we demonstrate that (time-varying) hazard differences can express causal contrasts, however, but only provided that additional unverifiable assumptions are made.

4 Towards causal contrasts for survival endpoints

4.1 Causal hazard ratios

In the preceding sections we have discussed difficulties in connection with interpreting hazard ratios or hazard differences causally and suggested alternative ways of illustrating time-dependent treatment contrasts. This section is devoted to a discussion of ways of defining contrasts based on the hazard function, which have a causal interpretation.

To remedy the selection problem, it seems intuitively of interest to evaluate conditional hazard ratios

$$\frac{\lim_{h\to 0} P(t \le T < t + h|T \ge t, A = 1, Z = z)}{\lim_{h\to 0} P(t \le T < t + h|T \ge t, A = 0, Z = z)}$$
$$= \frac{\lim_{h\to 0} P(t \le T^1 < t + h|T^1 \ge t, Z = z)}{\lim_{h\to 0} P(t \le T^0 < t + h|T^0 \ge t, Z = z)},$$

for a large collection of baseline variables Z, such that those who survive time $t > 0$ with treatment ($T^1 \ge t$) are comparable to those who survive time $t > 0$ without treatment ($T^0 \ge t$), but have the same covariate values z. Such comparability would be attained if

$$T^1 \perp\!\!\!\perp T^0|Z. \tag{6}$$

Under this assumption, it follows via Bayes' rule that the righthand side of the above identity equals

$$\frac{\lim_{h \to 0} P(t \leq T^1 < t + h | T^0 \geq t, T^1 \geq t, Z = z)}{\lim_{h \to 0} P(t \leq T^0 < t + h | T^0 \geq t, T^1 \geq t, Z = z)}. \tag{7}$$

This estimand expresses the instantaneous risk at time t on treatment versus control for the principal stratum of individuals with covariates z who would have survived up to time t, no matter what treatment. It represents a causal contrast, which is closely related to the so-called survivor average causal effect (Rubin 2000). We will refer to it as the conditional causal hazard ratio.

Unfortunately, assumption (6) is untestable and biologically implausible, as it is essentially impossible to believe that one can get hold of all predictors of the event time such that knowledge of the event time without treatment does not further predict the event time with treatment. Furthermore, even if one could get hold of all such predictors, then because Z would probably carry so much information about the event time, one would logically expect the numerator and denominator of (7) to be so close to 0 or 1 that it would render the conditional causal hazard ratio essentially meaningless. In the next section we will therefore focus on the marginal causal hazard ratio $\mathrm{HR}(t)$ obtained from (7) with Z empty:

$$\mathrm{HR}(t) = \frac{\lim_{h \to 0} P(t \leq T^1 < t + h | T^0 \geq t, T^1 \geq t)/h}{\lim_{h \to 0} P(t \leq T^0 < t + h | T^0 \geq t, T^1 \geq t)/h}. \tag{8}$$

4.2 Contrasting causal versus observed hazard ratios

Even if it seems impossible to estimate $\mathrm{HR}(t)$ from the observed data without making strong assumptions, it is of interest to gain some understanding of the causal hazard ratio $\mathrm{HR}(t)$ and its relationship to the standard Cox HR (which we also refer to as the observed HR). It is particularly of interest to know whether or not $\mathrm{HR}(t)$ is always smaller than HR if, as we expect, T^0 and T^1 are positively correlated. To develop insight, let us assume that the Cox model is correctly specified, in the sense that

$$P(T^0 > t) = \exp\{-\Lambda_0(t)\} \quad \text{and} \quad P(T^1 > t) = \exp\{-\phi\Lambda_0(t)\},$$

where $\phi = e^\beta$ is the Cox HR. Also, we still assume that treatment is randomised. A model for the joint distribution of T^0 and T^1, which obeys these marginal distributions can now be based on a copula C (Nelsen 2006). Specifically, we let $u_t = e^{-\Lambda_0(t)}$ and $v_t = e^{-\phi\Lambda_0(t)}$ and specify the joint distribution of (T^0, T^1) via

$$P(T^0 \geq t_0, T^1 \geq t_1) = P(U \leq u, V \leq v) = C(u, v)$$

where $U = u_{T^0}$, $u = u_{t_0}$, $V = v_{T^1}$ and $v = v_{t_1}$. Sklar's Theorem (Nelsen 2006) ensures that for any joint distribution of T^0 and T^1 there exists a copula C so that the latter display is fulfilled. It easy to show that

$$\mathrm{HR}(t) = \phi\Omega(t), \quad \Omega(t) = \frac{v_t \dot{C}_v(u_t, v_t)}{u_t \dot{C}_u(u_t, v_t)}, \tag{9}$$

where $\dot{C}_u(u, v) = \frac{\partial C(u,v)}{\partial u}$ and similarly with $\dot{C}_v(u, v)$. Display (9) gives the desired relation between $\mathrm{HR}(t)$ and the Cox HR, ϕ. To obtain further insight, let us for a moment restrict the copula C to belong to the class of Archimedean copulas. Bivariate distributions generated by frailty models are a subclass of Archimedean distributions (Oakes 1989).

For Archimedean copulas, there exists a function $\psi(v)$ so that

$$\psi(v) = \psi^*(t_0, t_1) \tag{10}$$

with $v = S(t_0, t_1) = P(T^0 > t_0, T^1 > t_1) \in]0, 1]$ (Oakes 1989) and

$$\psi^*(t_0, t_1) = \frac{\lambda_{T^0}(t_0|T^1 = t_1)}{\lambda_{T^0}(t_0|T^1 > t_1)} = \frac{\lambda_{T^1}(t_1|T^0 = t_0)}{\lambda_{T^1}(t_1|T^0 > t_0)}$$

is the so-called cross-ratio function (Oakes 1989). In the latter display, $\lambda_{T^0}(t_0|T^1 = t_1)$ denotes the conditional hazard function of T^0 given $T^1 = t_1$, and likewise with the other hazard functions in the latter display. This gives rise to the explicit expression

$$\Omega(t) = \exp\left(-\int_{e^{-\Lambda_0(t)}}^{e^{-\phi\Lambda_0(t)}} \frac{\psi(y)}{y} dy\right) \exp\{-\Lambda_0(t)(\phi - 1)\}.$$

Using the mean value theorem gives

$$\Omega(t) < 1 \iff \psi(e^{-H_t}) > 1,$$

for some H_t on the line segment between $\phi\Lambda_0(t)$ and $\Lambda_0(t)$. The right hand side of the latter display is fulfilled if $\psi(v) > 1$ for all v. If the copula is induced by a frailty model where the frailty distribution is not degenerate (at unity) then T^0 and T^1 are positively correlated and also $\psi(v) > 1$. This confirms the intuition layed out in Hernán (2010) who argues that the causal hazard ratio is smaller than the Cox hazard ratio at all times $t > 0$ based on a frailty setting.

Example Clayton-copula (Gamma-frailty). In this case, $\psi(v) = \theta + 1$ where $0 < \theta < \infty$. A direct calculation shows that

$$\mathrm{HR}(t) = \phi \exp\{-\theta\Lambda_0(t)(1 - \phi)\} < \phi,$$

for $t > 0$ so that $\Lambda_0(t) > 0$. Clearly,

$$0 < \mathrm{HR}(t) < \phi,$$

with the value of $\mathrm{HR}(t)$ depending on θ and $\Lambda_0(t)$.   □

Example Inverse Gaussian copula. In this case, $\psi(v) = 1 + \frac{1}{\eta - \log v}$, with $\eta > 0$. Independence of T_0 and T_1 is obtained when η tends to infinity, while η small results in high correlation. Again, a direct calculation shows that

$$\mathrm{HR}(t) = \phi \left\{ \frac{\eta + \phi \Lambda_0(t)}{\eta + \Lambda_0(t)} \right\} < \phi$$

when $\Lambda_0(t) > 0$. It is seen that $\lim_{\eta \to 0} \mathrm{HR}(t) = \phi^2$, again when $\Lambda_0(t) > 0$. So,

$$\phi^2 < \mathrm{HR}(t) < \phi,$$

for $t > 0$ with $\Lambda_0(t) > 0$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

These two examples show that $\mathrm{HR}(t) < \phi$ as the copulas used are indeed both Archimedean. However, they also show that it is not possible to bound $\mathrm{HR}(t)$ from below as we can see from the Clayton-copula case that it can go all the way down to zero, and obviously there is no way of checking the joint distribution of (T^0, T^1).

We now return to the general case, where the copula C may not be Archimedean. We see from (9) that

$$\mathrm{HR}(t) < \phi \iff u_t \dot{C}_u(u_t, v_t) > v_t \dot{C}_v(u_t, v_t)$$

since the first order partial derivatives of C are never negative (Nelsen 2006). Since $v_t > u_t$ as $\phi < 1$, it follows if the copula satisfies

$$u \dot{C}_u(u, v) - v \dot{C}_v(u, v) > 0, \text{ for } 0 \leq u < v \leq 1, \qquad\qquad (11)$$

that $\mathrm{HR}(t) < \phi$. Copulas obeying (11) then induce the positive correlation needed in order to have $\mathrm{HR}(t) < \phi$. We are not aware of previous work that has studied copulas where (11) holds. The condition (11) is local in the sense that it needs to hold for all $0 \leq u < v \leq 1$ whereas well known correlation measures such as Kendall's τ and Spearman's ρ are global, e.g. Spearman's ρ can be expressed in terms of the copula as

$$\rho = 12 \int_0^1 \int_0^1 \{C(u, v) - uv\} du dv.$$

Alternatively, there are various types of local correlation measures (Nelsen 2006), one being the so-called positively quadrant correlation (PQD): two random variables X and Y are said to have PQD if

$$P(X \leq x, Y \leq y) \geq P(X \leq x) P(Y \leq y), \text{ for all } x \text{ and } y,$$

which is equivalent to

$$C(u, v) \geq \Pi(u, v) = uv, \ (u, v) \in [0, 1]^2,$$

where C is the corresponding copula and Π is the independence copula. We now construct a copula that gives rise to PQD, but for which (11) does not hold. For any $\xi \in (1/4, 1/2)$ (Nelsen 2006), define

$$C^*(u, v) = \begin{cases} \xi M(\frac{u}{\xi}, \frac{v}{\xi}) & \text{if } (u, v) \in [0, \xi]^2 = J_1 \\ \xi + (1 - 2\xi)W(\frac{u-\xi}{1-2\xi}, \frac{v-\xi}{1-2\xi}) & \text{if } (u, v) \in [\xi, 1-\xi]^2 = J_2 \\ 1 - \xi + \xi M(\frac{u+\xi-1}{\xi}, \frac{v+\xi-1}{\xi}) & \text{if } (u, v) \in [1-\xi, 1]^2 = J_3 \\ M(u, v) & \text{otherwise} \end{cases}$$

with $W(u, v) = \max(u + v - 1, 0)$, $M(u, v) = \min(u, v)$ being the lower and upper Frechet-Hoeffding bounds, respectively. This copula C^*, which is the ordinal sum of $\{M, W, M\}$ with respect to the partition $\{[0, \xi], [\xi, 1-\xi], [1-\xi, 1]\}$, is PQD (Nelsen, Exercise 5.30). The suggested partition of $[0, 1]^2$ is illustrated in Fig. 2, where ξ is set to 0.35. The expression for the copula can be simplified to

$$C^*(u, v) = \begin{cases} M(u, v) & \text{if } (u, v) \notin J_2 \\ \xi + (u + v - 1)I(u + v > 1) & \text{if } (u, v) \in J_2 \end{cases}$$

Hence, for $(u, v) \in J_2$, we have

$$u\dot{C}^*_u(u, v) - v\dot{C}^*_v(u, v) = (u - v)I(u + v > 1),$$

which may be smaller than zero for $u < v$, and therefore (11) does not hold. Hence, PQD does not suffice to ensure (11), which may not be surprising as PQD, although a local dependence measure, is cumulative in that it depends on survival functions and not hazard functions as does the cross-ratio function.

By a direct calculation, Spearman's ρ for this copula is given by

$$\rho(\xi) = \xi(16\xi^2 - 24\xi + 12) - 1$$

that is an increasing function of ξ with $\rho(1/4) = 0.75$ and $\rho(1/2) = 1$. Surprisingly, this copula therefore gives rise to a situation where $HR(t) > \phi$ for some t's even when $\phi < 1$ and Spearman's $\rho \geq 0.75$.

In particular, we have that

$$HR(t) = \phi\Omega(t), \quad \text{with} \quad \Omega(t) = \frac{\exp\{-\phi\Lambda_0(t)\}}{\exp\{-\Lambda_0(t)\}},$$

when

$$\exp\{-\phi\Lambda_0(t)\} + \exp\{-\Lambda_0(t)\} > 1 \quad \text{and} \tag{12}$$

$$\xi < \exp\{-\Lambda_0(t)\} < \exp\{-\phi\Lambda_0(t)\} < 1 - \xi. \tag{13}$$

Condition (12) corresponds to the part of J_2 above the diagonal (the $v = 1 - u$ line, see Fig. 2) and condition (13) corresponds to the part of J_2 above the (anti-)diagonal ($v = u$). If we follow the path $(\exp\{-\Lambda_0(t)\}, \exp\{-\phi\Lambda_0(t)\})$ (starting at the point $(1,1)$ when $t = 0$) in $[0, 1]^2$ in such a way that we pass through the upper triangle of J_2 (with J_2 consisting of 4 triangles like the back of an envelope, see Fig. 2) then $HR(t)$ will be equal to 0 as long as we have not yet reached J_2 but then it jumps to $k(u, v)\phi$,
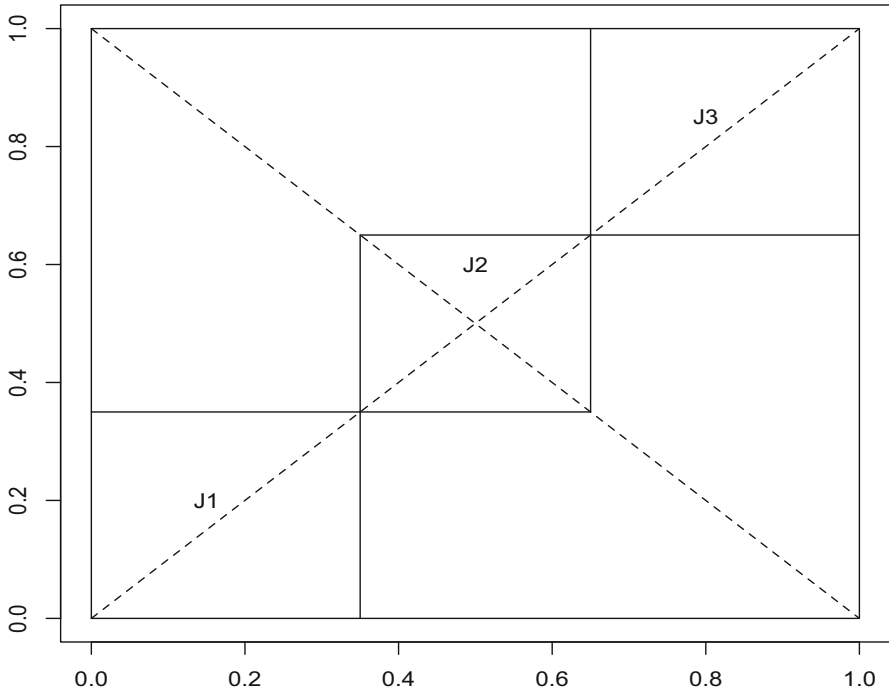
Fig. 2 Illustration of the regions used in the definition of the copula C^*, Sect. 4.1, showing the rectangular regions J_1, J_2 and J_3. The value of ξ is set to 0.35. The two broken lines are the lines $v = u$ and $v = 1 - u$, respectively

where $k(u, v) > 1$. When we cross the line $v = 1 - u$ in J_2, HR(t) is undefined as both partial derivatives of the copula are equal to zero, and when we escape J_2 then HR(t) = 0. In "Appendix A.4", we gain further insight by generating samples from this latter copula. The empirical result in Fig. 3 confirms the above reasoning that more than just positive correlation between T^0 and T^1 is needed to have that HR(t) < ϕ.

If one allows for negative correlation between T^0 and T^1 then it is easy to construct a situation where $\Omega(t) > 1$. We just need a copula C that satisfies

$$u\dot{C}_u(u, v) - v\dot{C}_v(u, v) < 0, \text{ for } 0 \leq u < v \leq 1,$$

Take for example the Farlie–Gumbel–Morgenstern Copulas, where

$$C_\theta(u, v) = uv - \theta uv(1 - u)(1 - v),$$

with $\theta \in [-1, 1]$. Specifically, let $\theta = -1$, corresponding to Spearman's $\rho = -1/3$, then $\Omega(t) > 1$ and we can even have that HR(t) = $\phi\Omega(t) > 1$ although $\phi < 1$. Negative correlation between T^0 and T^1 does not seem very likely, however.
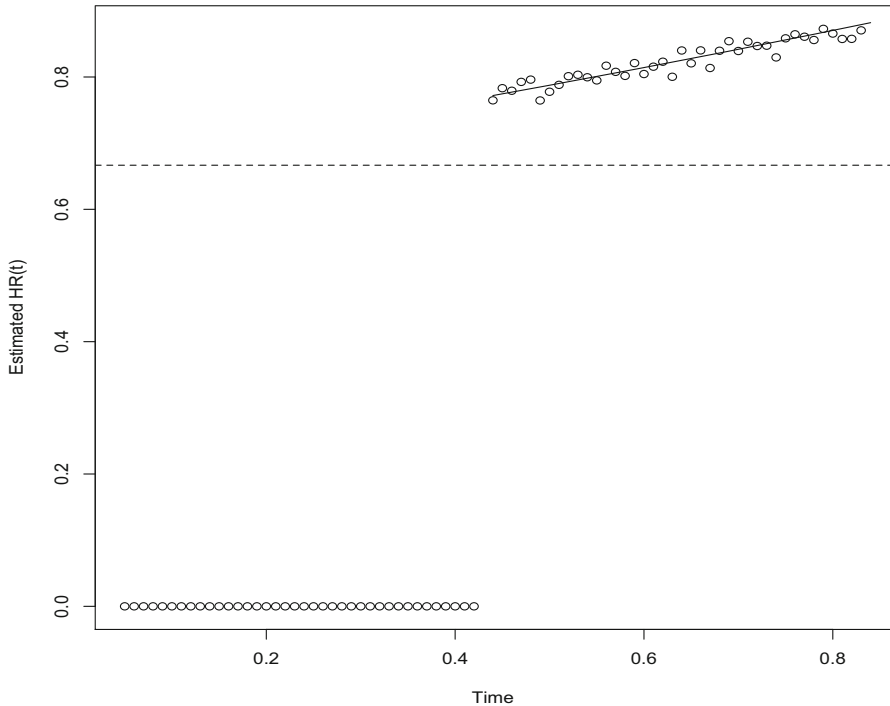
Fig. 3 Estimated (points) and theoretical (full line) values of HR(t). Broken line corresponds to $\phi = 2/3$

4.3 Hazard differences

Define

$$\mathrm{HD}(t) = \lim_{h \to 0} P(t \leq T^1 < t + h | T^0 \geq t, T^1 \geq t)/h$$
$$- \lim_{h \to 0} P(t \leq T^0 < t + h | T^0 \geq t, T^1 \geq t)/h,$$

which is a causal hazard function contrast. In the additive hazards model (5), $\psi(t)$ reduces to the causal contrast, $\psi(t) = \mathrm{HD}(t)$, provided that (6) holds (see "Appendix A.3"). The latter assumption, however, as discussed in Sect. 4.1 is both untestable and biologically implausible.

Since A is binary we can always write the hazard function $\lambda(t; A)$ as

$$\lambda(t; A) = \lambda_0(t) + \psi(t)A \tag{14}$$

Further, one may show the following relationship between $\mathrm{HD}(t)$ and $\psi(t)$,

$$\mathrm{HD}(t) = \Omega(t)\Psi(t)\{\lambda_0(t) + \psi(t)\} - \Psi(t)\lambda_0(t) \tag{15}$$

where

$$\Omega(t) = \frac{v_t \dot{C}_v(u_t, v_t)}{u_t \dot{C}_u(u_t, v_t)}, \quad \Psi(t) = \frac{u_t \dot{C}_u(u_t, v_t)}{C(u_t, v_t)},$$

and where we specify the joint distribution of (T^0, T^1) via

$$P(T^0 \geq t_0, T^1 \geq t_1) = P(U \leq u, V \leq v) = C(u, v)$$

with $U = u_{T^0}$, $u = u_{t_0}$, $V = v_{T^1}$ and $v = v_{t_1}$. We now show that an Aalen additive hazards analysis can easily lead to the same misleading conclusion as was the case for the Cox regression analysis with a treatment-by-time interaction that we considered in Sect. 3.1. If the DGP is given by $\lambda(t; A, Z)$ in (4) then the $\psi(t)$ is (14) has the form

$$\psi(t) = \lambda_0(t)(e^{\beta_1} - 1)I(t \leq v) + \lambda_0(t)(e^{\beta_2} - 1)I(t > v).$$

If $\beta_2 = 0$ we see that $\psi(t) = 0$ for $t > v$ and an Aalen additive hazards analysis would therefore also indicate that the treatment effect vanishes after time point v. We have illustrated this in the Supplementary Materials, where the marginal Aalen additive hazards model fits the data perfectly but also suggests a beneficial effect of the treatment in an initial period (first 4 years), which then disappears, see Figure 5 in the Supplementary Material. In this specific setting, we have, for $t > v$, that

$$\text{HD}(t) = \{\Omega(t) - 1\}\Psi(t)\lambda_0(t) < 0$$

since both $\Omega(t)$ and $\Psi(t)$ are within the interval $(0, 1)$. Note also, for this specific example, that $\lambda(t; A, Z)$ is not of the form (5).

One may also construct scenarios where the DGP is of the form (5) and where $\psi(t) \neq \text{HD}(t)$ if (6) does not hold. Hence, whether or not $\psi(t)$ in (14) has a causal interpretation in terms of a contrast between hazard functions depends on unverifiable assumptions.

4.4 Alternative causal contrasts

Given the subtle interpretation of hazard ratios, we believe that the effects of treatments on survival endpoints are best communicated via unconditional effect measures. For a correctly specified Cox proportional hazards model the relative risk function

$$\text{RR}(t) = \frac{P(T \leq t | A = 1)}{P(T \leq t | A = 0)} = \frac{P(T^1 \leq t)}{P(T^0 \leq t)},$$

can be estimated consistently by

$$\widehat{\text{RR}}(t) = \frac{1 - \exp\{-\hat{\Lambda}_0(t)e^{\hat{\beta}}\}}{1 - \exp\{-\hat{\Lambda}_0(t)\}},$$

where $\hat{\beta}$ is the Cox partial likelihood estimator and $\hat{\Lambda}_0(t)$ is the corresponding Breslow estimator. Or we may use the restricted mean survival time (RMST) at time t (Uno et al. 2014; Zhao et al. 2016), defined as $\mathrm{RMST}(t) = E\{\min(T, t)\}$. This is the area under the survival curve of T up to time t, which can easily be estimated using the corresponding Kaplan-Meier curve up to time t (see Zhao et al. (2016) for details on inference). Contrasts of the $\mathrm{RMST}(t)$ corresponding to different (randomised) treatment groups therefore carry a causal interpretation. One may also report the (restricted) mean time lost before time t, which is defined as $\mathrm{RMTL}(t) = t - \mathrm{RMST}(t)$.

However, the restricted mean residual survival time $E\{\min(T, t) - s | T > s\}$ for $s < t$ leads to the same subtleties as seen for the hazard function, which is due to the conditioning on $T > s$.

4.5 Analysis of the MRC RE01 study

For the MRC RE01 trial, we fitted Aalen's additive hazard model

$$\lambda(t; A, V) = \beta_0(t) + \psi(t)A + \beta_1(t)^T V, \tag{16}$$

where V includes days from metastasis to randomization (log-transformed), WHO performance status (0, 1 and 2; with group 0 and 1 collapsed into one group), and Haemoglobin (g/dl). As the treatment variable A is independent of the other covariates, the above Aalen model is collapsible meaning that the interpretation of $\psi(t)$ is the same in the conditional and marginal model. The Cox model does not have this property. Model (16) appeared to fit the data well, using the tools described in Chapter 5 in Martinussen and Scheike (2006); specifically no interaction between the treatment indicator and the baseline risk factors was found. We estimated $\hat{\Psi}(t) = \int_0^t \psi(s)ds$ both from the conditional model and the marginal model, and the two estimators were almost identical, which, as pointed out, should be the case if model (16) is correctly specified. Using model (16), we next tested the null hypothesis $\psi(t) = \psi$ of a constant effect ($P = 0.58$), which was subsequently estimated to be $\hat{\psi} = -0.02$ (SE, 0.009). If $T^0 \perp\!\!\!\perp T^1 | V$ (or if the addition of all variables Z conditional on which T^0 and T^1 become independent, does not change the additive structure of the model), then this is also the causal hazard difference. This would mean that over the course of the follow-up, an average of approximately 2 additional deaths will occur for each month of follow-up in each 100 persons under the control treatment alive at the start of the month and *who would also be alive under the IFN treatment*, compared with each 100 IFN treated persons alive at the start of the month and *who would also be alive under the control treatment*.

As suggested before, the assumption that $T^0 \perp\!\!\!\perp T^1 | V$ is implausible (just like the assumption that all variables Z for which $T^0 \perp\!\!\!\perp T^1 | Z$ do not interact with treatment on the additive hazard scale). In view of this, we additionally report the effects of treatment on the survival chances and restricted mean survival time. The relative risk function estimate, along with 95% pointwise and uniform confidence bands, is displayed in Fig. 4. We see that the estimated relative risk function is below 1 at all times, in favour of IFN treatment. For instance, it is seen that the relative risk at one year is estimated
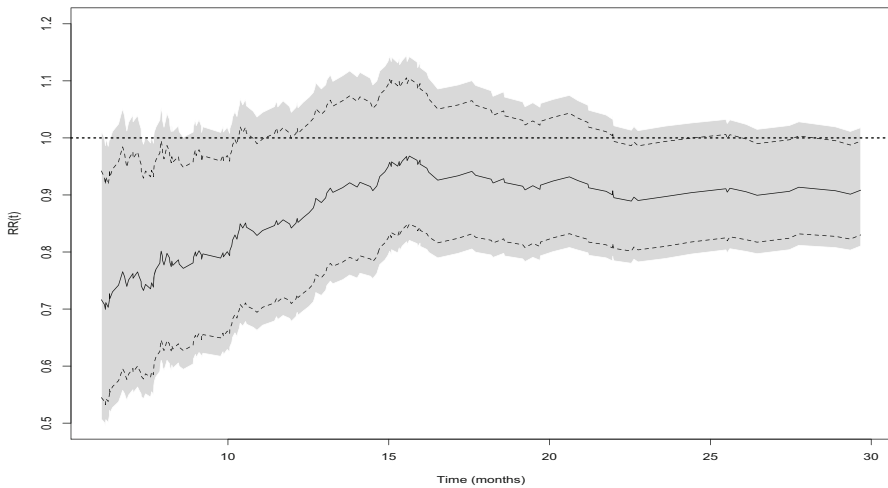
Fig. 4 MRC RE01 study. IFN treatment versus control treatment. Estimate of relative risk $RR(t)$ along with 95% pointwise confidence bands (dashed curves) and 95% uniform confidence bands (shaded area)

to be approximately 0.85. Judging from the 95% confidence bands (dashed curves), this is close to being significant. A uniform test over the considered time span is also close to being significant judging from the 95% uniform confidence bands.

Further, the 'months of life lost up to 30 months' RMTL(30) is estimated to be 17.3 for IFN treatment and 19.8 for control treatment. The ratio of these two (IFN vs control) is 0.87 (95% CI, 0.70 to 1.04). Thus, on IFN treatment there is a 13% reduced loss of lifetime compared to the control treatment during the first 30 months of follow up.

5 Concluding remarks

We have argued that the treatment effect in a proportional hazards model carries a causal interpretation, but that its interpretation is subtle. The proportional hazards assumption does not express, for instance, that treatment works equally effectively at all times, as the hazard ratio at a given time mixes differences between treatment arms due to treatment effect as well as selection. The danger of over interpreting hazard ratios become most pronounced when the hazard ratio is not constant over time (e.g. when the hazard ratio is below 1 for some time and then becomes 1). We have argued that this cannot be interpreted as implying that treatment effectiveness disappears after some time. In our opinion, this is the source of much confusion, and a real concern. Non-constant hazard ratios are indeed fairly common in real life because the proportional hazards assumption is a rather unstable assumption in the following sense. Even when valid in some population, this assumption is likely to fail in subgroups of that population (e.g. if one studies men and women separately), and vice versa. This makes the assumption, at best, an approximation in practice.

We have described hazard contrasts that are causally interpretable because they compare intensities at a given time t with and without treatment for the same patient population: those who would survive that time, no matter what treatment. Such hazard contrasts are not estimable without untestable assumptions concerning the joint distribution of (T^0, T^1) and a sensitivity analysis is therefore the only option. One may alternatively attempt to proceed under the monotonicity assumption that no one is harmed by treatment, i.e.

$$T^1 \geq T^0 \quad \text{with probability 1.}$$

Under this strong assumption, we can write

$$P(t \leq T < t + h|T \geq t, A = 0) = P(t \leq T^0 < t + h|T^0 \geq t)$$
$$= P(t \leq T^0 < t + h|T^0 \geq t, T^1 \geq t)$$

and

$$P(t \leq T < t + h|T \geq t, A = 1) = P(t \leq T^1 < t + h|T^1 \geq t)$$
$$= P(t \leq T^1 < t + h|T^0 \geq t, T^1 \geq t)\pi(t)$$
$$+ P(t \leq T^1 < t + h|T^0 < t, T^1 \geq t)\{1 - \pi(t)\},$$

with

$$\pi(t) \equiv P(T^0 \geq t|T^1 \geq t) = \frac{P(T^0 \geq t)}{P(T^1 \geq t)} = \frac{P(T \geq t|A = 0)}{P(T \geq t|A = 1)}.$$

However, if the distribution of (T^0, T^1) is assumed to be absolutely continuous then one may show that $\lim_{h \to 0} h^{-1} P(t \leq T^1 < t + h|T^0 \geq t, T^1 \geq t) = 0$ and thus $\text{HR}(t) = 0$ under the monotonicity assumption, see "Appendix A.5". A further disadvantage of the discussed hazard contrasts, which are causally interpretable, is also that they describe the effect for an unknown subgroup of the population.

The reason that the causal hazard ratio (8) is not identifiable without invoking strong assumptions is because it attempts to answer an overly ambitious question. Imagine a trial that randomises participants over an implanted medical device (e.g. a stent or a pacemaker) versus no treatment (or placebo). Suppose that the medical device gradually deteriorates and stops being operational after some time v. Then we would say that treatment no longer works from time v onwards. This would correspond with $\text{HR}(t) = 1$ for $t \geq v$. However, how could data from a randomised trial be informative about the effect of treatment after time v when no information is collected on the times at which the medical device is operational or not? To learn about the treatment effect at each time t, we should ideally need data A_t on whether ($A_t = 1$) or not ($A_t = 0$) the device is operational at that time. When the operation time is ignorable, then one may learn about the treatment effect at each time t through contrasts of the form

$$\frac{\lim_{h \to 0} P(s \leq T < s + h|T \geq s, \overline{A}_{t-} = 1, A_t = 1)}{\lim_{h \to 0} P(s \leq T < s + h|T \geq s, \overline{A}_{t-} = 1, A_t = 0)},$$

for all $s \geq t$, where $\overline{A}_{t-}$ is the information generated by all A_u, $u < t$, so $\overline{A}_{t-} = 1$ refers to "being on treatment" up to just before time t. It is unsurprising that without detailed data on the operation times of each device, strong assumptions are needed to develop insight into the dynamic nature of the treatment effect. A better strategy in practice, when interest lies in the dynamic aspects of a treatment, is therefore to design the study such that the collected data provide immediate insight into the dynamic aspects of treatment (e.g. by modifying treatment assignments over time).

While we have argued that time-varying hazard ratios are not causally interpretable, note that they continue to be valuable for descriptive purposes when addressing non-causal questions. Imagine for example a setting where the outcome is time-to death after onset of a certain disease, and say that there are two subclasses ($A = 1$ and $A = 0$) of the disease. Suppose we know from a previous study where model (2) was used, and was correctly specified, that $\exp(\beta_1) = 2$, $\beta_2 = 0$, for $\nu = 2$ years. Assume also that all other risk factors were equally balanced in the two subclasses at time 0 (onset of disease) or that proper adjustment for these factors was performed in the hazard model. Then, based on this analysis, we can say that patients in disease subclass 1 who survive the first two years from onset of the disease will have the same subsequent risk of dying as those in disease subclass 0—a statement that may be useful for consulting patients at time $\nu = 2$ years, or later.

The difficulties we have described concerning the causal interpretation of hazard ratios pertain for hazard differences. This problem is thus not only an issue for proportional hazard models but also for additive hazard models. The root cause of the problem is the interpretation of hazard function itself, more than its particular structure.

Appendix A

A.1 Binary frailty variable

Let Z be binary, for instance $P(Z = 0.2) = 0.2$ and $P(Z = 1.2) = 0.8$, low and high risk groups, then $E(Z) = 1$, and similar results as in the Gamma-distribution case (Sect. 3.2) are obtained. In this binary case, the Laplace transform is

$$\phi_Z(u) = 0.2e^{-0.2u} + 0.8e^{-1.2u}.$$

Therefore

$$\mathrm{HR}_Z(t) \frac{g_Z(e^{-\Lambda(t,a=1)})}{g_Z(e^{-\Lambda(t,a=0)})} = \frac{\lambda(t; a = 1)}{\lambda(t; a = 0)},$$

where

$$g_Z(u) = \{D \log(\phi_Z)\}\{\phi_Z^{-1}(u)\}, \tag{17}$$

which is an increasing function. Hence, if we take $\beta_1 < 0$ and $\beta_2 = 0$, then again

$$\mathrm{HR}_Z(t) < 1$$

for all t.

A.2 Frailty model arising by marginalization

The DGP (4) given by $\lambda(t; A, Z)$ with $\nu = \infty$ induces the marginal Cox model for the observed data where we do not get access to Z. It is tempting to interpret $\mathrm{HR}_Z(t)$ as a causal hazard ratio, but this only holds under further untestable assumptions as shown in Sect. 4.1. Below, we formulate a more general DGP involving an additional frailty variable so that both the marginal Cox model and the model $\lambda(t; A, Z)$ are correctly specified.

Rename Z to Z_1. We then show that similarly we can also pick a DGP $\lambda(t; A, Z_1, Z_2)$ so that it marginalizes to $\lambda(t; A, Z_1)$ that further marginalizes to $\lambda(t; A)$, the latter being the Cox model. For ease of calculations, let

$$\lambda(t; A, Z_1, Z_2) = Z_2 \lambda^*(t; A, Z_1)$$

with Z_2 being Gamma distributed with mean and variance equal to 1, and independent of Z_1 and A. Similar calculations as those in Sect. 3.1 gives the following expression

$$\lambda(t; A, Z_1, Z_2) = Z_1 Z_2 \lambda_0(t) e^{\beta A} \exp\left[\Lambda_0(t) e^{\beta A} + Z_1 \{\exp\{\Lambda_0(t) e^{\beta A} - 1\}\right] \quad (18)$$

Hence, if the DGP is governed by (18) then model (4), with $\nu = \infty$, and the marginal Cox model, only conditioning on A, are also correctly specified.

A.3 Hazard differences

Assume (5) and (6). Then

$$\lim_{h \to 0} P(t \le T^1 < t + h | T^0 \ge t, T^1 \ge t) = \psi(t) + \frac{E\{\omega(t, Z) e^{-2 \int_0^t \omega(s, Z)\, ds}\}}{E\{e^{-2 \int_0^t \omega(s, Z)\, ds}\}}$$

and

$$\lim_{h \to 0} P(t \le T^0 < t + h | T^0 \ge t, T^1 \ge t) = \frac{E\{\omega(t, Z) e^{-2 \int_0^t \omega(s, Z)\, ds}\}}{E\{e^{-2 \int_0^t \omega(s, Z)\, ds}\}}$$

and therefore

$$\psi(t) = \lim_{h \to 0} P(t \le T^1 < t + h | T^0 \ge t, T^1 \ge t)$$
$$- \lim_{h \to 0} P(t \le T^0 < t + h | T^0 \ge t, T^1 \ge t).$$

A.4 HR(t) under copulas not obeying (11)

We outline now how to generate samples from the copula C^* to gain further insight, Nelsen (2006), pp. 40–41,

1. Generate a standard uniformly distributed random variable u;
2. Set

$$v = \begin{cases} u & \text{if } u \leq \xi \text{ or } u \geq 1 - \xi \\ 1 - u & \text{if } \xi < u < 1 - \xi \end{cases}$$

3. Set $T^0 = \Lambda_0^{-1}\{-\log(u)\}$ and $T^1 = \Lambda_0^{-1}\{-\log(v)/\phi\}$ (assuming Λ_0 is strictly increasing).

As an illustration, we took $\Lambda_0(t) = t$, $\phi = 2/3$ and $\xi = 0.251$ and generated $n = 2,000,000$ samples. For these simulated data we obtained an estimated Cox HR of 0.666 and Spearman's ρ was estimated to equal 0.753, which fits nicely with the theoretical values. As mentioned, if we follow the path $(\exp\{-\Lambda_0(t)\}, \exp\{-\phi\Lambda_0(t)\})$ (starting at (1,1) when $t = 0$) in $[0,1]^2$ in such a way that we pass through the upper triangle of J_2 then HR(t) will be equal to 0 before we reach J_2 but then it jumps to $k(u,v)\phi$, where $k(u,v) > 1$. To illustrate this, we estimated HR(t) non-parametrically in time-points t until we cross the line $v = 1 - u$ in J_2. This was done by taking the observed frequency of T^1-failures in the time interval $[t, t + \delta]$, $\delta = 0.01$, among those where both $T^0 \geq t$ and $T^1 \geq t$; and similarly for T^0; and then taking the ratio of the two. This is seen in Fig. 3, where the theoretical values $\phi \exp\{-t(\phi - 1)\}$ are superimposed. The dashed line corresponds to $\phi = 2/3$. We see that the theoretical values fit nicely to the estimated ones. For this copula it can be shown that HR(t) is always below 1, but as we have just seen it can also be larger than ϕ! Hence, to have that HR(t) $< \phi$ we need more than just positive correlation between T^0 and T^1 as this example illustrates.

A.5 Monotonicity and HR(t)

Assume that $T^1 \geq T^0$ a.s. and that the distribution of (T^0, T^1) is absolutely continuous with density $f(t_0, t_1)$ with respect to the Lebesgue measure. Monotonicity is attained if $f(t_0, t_1) = 0$ for $t_0 > t_1$. We now look at

$$\begin{aligned} z(h) &= P(T^0 \geq t, t \leq T^1 \leq t + h) \\ &= P(T^0 \geq t, t \leq T^1 \leq t + h, T^0 \leq T^1) \\ &= \int_t^{t+h} \int_t^{t_1} f(t_0, t_1) dt_0 dt_1. \end{aligned}$$

Hence

$$\lim_{h \to 0} \frac{z(h)}{h} = \lim_{h \to 0} \frac{G(t + h) - G(t)}{h} = G'(t) = \int_t^t f(t_0, t) dt_0 = 0,$$

where

$$G(u) = \int_0^u \left\{ \int_t^{t_1} f(t_0, t_1) dt_0 \right\} dt_1.$$

But this means $\lambda_{T^1}(t|T^0 \geq t, T^1 \geq t) = 0$, and therefore $\mathrm{HR}(t) = 0$. To avoid having $\lambda_{T^1}(t|T^0 \geq t, T^1 \geq t)$ equal to zero we need for instance T^0 given T^1 to have a discrete distribution, but this seems unappealing.

References

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):579–93

Bartolucci F, Grilli L (2011) Modeling partial compliance through copulas in a principal stratification framework. J Am Stat Assoc 106(494):469–479

Collett D (2015) Modelling survival data in medical research. CRC Press, Boca Raton

Frangakis CE, Rubin DB (2002) Principal stratification in causal inference. Biometrics 58(1):21–9

Hernán M (2010) The hazards of hazard ratios. Epidemiology (Cambridge, Mass) 21(1):13–15

Lederle FA, Kyriakides TC, Stroupe KT, Freischlag JA, Padberg FT, Matsumura JS, Huo Z, Johnson GR (2019) Open versus endovascular repair of abdominal aortic aneurysm. N Engl J Med 380(22):2126–2135 PMID: 31141634

Lin DY, Wei LJ, Ying Z (1993) Checking the Cox model with cumulative sums of martingale-based residuals. Biometrika 80:557–572

Martinussen T, Scheike TH (2006) Dynamic regression models for survival data. Springer, New York

Martinussen T, Vansteelandt S (2013) On collapsibility and confounding bias in cox and aalen regression models. Lifetime Data Anal 19(3):279–96

Nelsen RB (2006) An introduction to copulas. Springer, New York

Oakes D (1989) Bivariate survival models induced by frailties. J Am Stat Assoc 84(406):487–493

Primrose JN, Fox RP, Palmer DH, Malik HZ, Prasad R, Mirza D, Anthony A, Corrie P, Falk S, Finch-Jones M, Wasan H, Ross P, Wall L, Wadsley J, Evans JTR, Stocken D, Praseedom R, Ma YT, Davidson B, Neoptolemos JP, Iveson T, Raftery J, Zhu S, Cunningham D, Garden OJ, Stubbs C, Valle JW, Bridgewater J, Primrose J, Fox R, Morement H, Chan O, Rees C, Ma Y, Hickish T, Falk S, Finch-Jones M, Pope I, Corrie P, Crosby T, Sothi S, Sharkland K, Adamson D, Wall L, Evans J, Dent J, Hombaiah U, Iwuji C, Anthoney A, Bridgewater J, Cunningham D, Gillmore R, Ross P, Slater S, Wasan H, Waters J, Valle J, Palmer D, Malik H, Neoptolemos J, Faluyi O, Sumpter K, Dernedde U, Maduhusudan S, Cogill G, Archer C, Iveson T, Wadsley J, Darby S, Peterson M, Mukhtar A, Thorpe J, Bateman A, Tsang D, Cummins S, Nolan L, Beaumont E, Prasad R, Mirza D, Stocken D, Praseedom R, Davidson B, Raftery J, Zhu S, Garden J, Stubbs C, Coxon F (2019) Capecitabine compared with observation in resected biliary tract cancer (bilcap): a randomised, controlled, multicentre, phase 3 study. Lancet Oncol 20(5):663–673

Rubin DB (2000) Causal inference without counterfactuals: comment. J Am Stat Assoc 95(450):435–438

Stablein DM, Koutrouvelis IA (1985) A two-sample test sensitive to crossing hazards in uncensored and singly censored data. Biometrics 41(3):643–52

Uno H, Claggett B, Tian L, Inoue E, Gallo P, Miyata T, Schrag D, Takeuchi M, Uyama Y, Zhao L, Skali H, Solomon S, Jacobus S, Hughes M, Packer M, Wei L-J (2014) Moving beyond the hazard ratio in quantifying the between-group difference in survival analysis. J Clin Oncol 32(22):2380–5

Vansteelandt S, Martinussen T, Tchetgen ET (2014) On adjustment for auxiliary covariates in additive hazard models for the analysis of randomized experiments. Biometrika 101(1):237–244

White IR, Royston P (2009) Imputing missing covariate values for the cox model. Stat Med 28(15):1982–98

Zhao L, Claggett B, Tian L, Uno H, Pfeffer MA, Solomon SD, Trippa L, Wei LJ (2016) On the restricted mean survival time curve in survival analysis. Biometrics 72(1):215–21

Affiliations

Torben Martinussen[1] (ID) **· Stijn Vansteelandt[2,3] · Per Kragh Andersen[1]**

Stijn Vansteelandt
stijn.vansteelandt@ugent.be

Per Kragh Andersen
pka@biostat.ku.dk

[1]   Section of Biostatistics, University of Copenhagen, Øster Farimagsgade 5B, 1014 Copenhagen K, Denmark

[2]   Department of Applied Mathematics, Computer Sciences and Statistics, Ghent University, Krijgslaan 281 (S9), 9000 Gent, Belgium

[3]   Centre for Statistical Methodology, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK