Assumption-lean Cox regression

Stijn Vansteelandt*
Department of Applied Mathematics, Computer Science and Statistics,
Ghent University,
and Department of Medical Statistics,
London School of Hygiene and Tropical Medicine
and
Oliver Dukes and Kelly Van Lancker
Department of Applied Mathematics, Computer Science and Statistics,
Ghent University
and
Torben Martinussen
Department of Biostatistics
University of Copenhagen

September 6, 2022

*The authors gratefully acknowledge support from grants G016116N and BOF.01P08419. Code for the proposed estimators and reproducing the simulation studies and data analysis is available on https://github.com/svsteela/MLCox.

Abstract

Inference for the conditional association between an exposure and a time-to-event endpoint, given covariates, is routinely based on partial likelihood estimators for hazard ratios indexing Cox proportional hazards models. This approach is flexible and makes testing straightforward, but is nonetheless not entirely satisfactory. First, there is no good understanding of what it infers when the model is misspecified. Second, it is common to employ variable selection procedures when deciding which model to use. However, the bias and uncertainty that imperfect variable selection adds to the analysis is rarely acknowledged, rendering standard inferences biased and overly optimistic. To remedy this, we propose a nonparametric estimand which reduces to the main exposure effect parameter in a (partially linear) Cox model when that model is correct, but continues to capture the (conditional) association of interest in a well understood way, even when this model is misspecified in an arbitrary manner. We achieve an assumption-lean inference for this estimand based on its influence function under the nonparametric model. This has the further advantage that it makes the proposed approach amenable to the use of data-adaptive procedures (e.g., variable selection, machine learning), which we find to work well in simulation studies and a data analysis.

Key words: conditional treatment effect; debiased machine learning; estimand; hazard ratio; model misspecification; post-selection inference.

1 Introduction

The hazard ratio has grown to become one of the most reported measures of association in epidemiology and medicine. It is by far the most popular measure of association in survival analysis, expressing, for instance, a contrast in survival distributions between treated and untreated individuals. Its popularity can be partly explained by the success of the Cox proportional hazards model (Cox, 1972), which uses the hazard ratio as a canonical parameter. Its success more fundamentally underlies a more substantive difficulty when it comes to measuring associations with respect to event times, namely that the strength of the association may be dependent upon the length of the time window over which risks are being evaluated. A (constant) hazard ratio accommodates this by evaluating risks over an infinitesimally small time window. This gives it some appeal relative to competing measures (e.g., restricted mean survival differences), which demand pre-specification of a time window.

The estimation of hazard ratios is routinely based on partial likelihood estimators (Cox, 1972) under a Cox proportional hazards model. This approach has great appeal. It can incorporate arbitrary types of exposures and covariates, does not require modelling the time effect, and returns (globally) efficient estimators when the model holds. The latter assumption is questionable, however. Hazard ratios have been argued to be non-proportional in many practical settings (Stensrud and Hernán, 2020) and it is unlikely that all covariate effects can be modelled correctly (Robins, 1999; van der Laan and Rose, 2011), especially when adjustment for many covariates is needed in order to control for confounding or informative censoring. This is a major concern as there is no good understanding of what the partial likelihood estimator of the hazard ratio converges to when the model is misspecified (Struthers and Kalbfleisch, 1986). It does not simply converge to a weighted average of the time-varying hazard ratios (unlike what is suggested in Stensrud and Hernán (2020)), but instead equals a complex functional of the observed data distribution. For instance, with a dichotomous exposure A (coded 0 or 1) and in the absence of covariates, this functional is the value θ that solves the non-linear integral equation:

$$0 = \int \frac{w(t,1)w(t,0)}{\theta w(t,1) + w(t,0)} \{\lambda(t|A=1) - \theta\lambda(t|A=0)\}\, dt,$$

where for $a = 0, 1$, $\lambda(t|A = a)$ denotes the hazard at time t in exposure group a, $w(t, a) \equiv P(T > t, C > t|A = a)P(A = a)$ and T and C denote the event and censoring time, respectively. In the presence of covariate data, the analogous functional is even more complicated, and delivers a solution that may not even represent a summary measure of the conditional association between the exposure and event time, given covariates. In particular, under model misspecification, the partial likelihood estimator of the hazard ratio may converge to a value different from 1 under the null hypothesis that the exposure is conditionally independent of the time-to-event endpoint, given covariates, thereby leading to invalid tests. This failure to summarise the conditional association of interest under model misspecification is typical of most statistical procedures. For instance, under model misspecification, ordinary least squares estimators of the exposure effect in linear regression models with a main effect of exposure and covariates typically do not converge to a summary measure of the conditional association between exposure and outcome, given covariates (e.g., a weighted average of conditional mean outcome differences between treated and untreated individuals) (Vansteelandt and Dukes, 2022).

The natural response to the above concerns would be to make the Cox proportional hazards model sufficiently complex by incorporating interactions and other non-linearities where needed. This reflects what is commonly done in practice, but only to a limited degree. First, the curse of dimensionality forces data analysts to keep models sufficiently simple. This is usually done by means of data-adaptive procedures, such as variable selection algorithms, which help steer the model's complexity, relative to the information in the data. Second, the primary purpose of a data analysis is to summarise and provide insight.

Data analysts must therefore strike a balance between keeping models sufficiently simple so that the key message can be communicated well, versus modelling the real complexity. Such compromises are often acceptable as we would not usually be interested in knowing exactly how exposure effects vary between covariate strata so long as the differences are small, and likewise we would usually not mind reporting a linear dose effect on the log hazard so long as the association is roughly linear on that scale.

While trade-offs between simplicity versus complexity are thus useful, and indeed often a necessity, they also raise concerns. The reason is that simplifications of the patterns in the data, whether by deliberate choice or in view of the curse of dimensionality, are often taken for granted and viewed as representing some a priori given, ground truth when it comes to drawing inference (van der Laan and Rose, 2011; Vansteelandt and Dukes, 2022). For instance, standard inference will commonly ignore that certain model simplifications (e.g., linear covariate effects, the absence of interactions, ...) were made for convenience rather than because they represent a known fact. The resulting potential for model misspecification is worrying in view of the earlier discussion. It adds concerns about bias, but also excess variability which commonly results when models are misspecified (Buja et al., 2019; Vansteelandt and Dukes, 2022). Likewise, standard inference will commonly ignore that the decision not to adjust for certain covariates was data-driven. In particular, the bias and uncertainty that imperfect variable selection adds to the analysis is rarely acknowledged, rendering standard inferences biased and overly optimistic.

An alternative, albeit less popular response to the above concerns would be to use less restrictive models that parameterize only the hazard's dependence on the primary exposure of interest, A. For instance, given an additional vector $\mathbf{L}$ of baseline covariates, the continuously stratified Cox model (Sasieni, 1992; Dabrowska, 1997) postulates that the conditional hazard obeys

$$\lambda(t|A, \mathbf{L}) = \lambda_0(t|\mathbf{L}) \exp(\beta A),$$

at each time $t \geq 0$, where $\lambda_0(t|\mathbf{L})$ is an unknown function. By focussing on this model, one can prevent misspecification of the hazard's dependence on time or residual covariates $\mathbf{L}$ to induce bias in the estimated exposure 'effect'. Semi-parametric (efficient) estimation has been considered under this model (Dabrowska, 1997) and under a number of submodels that assume additivity of $\lambda_0(t|\mathbf{L})$ in t and $\mathbf{L}$ (or linear functions thereof) (Sasieni, 1992; Zhong et al., 2022; Huang, 1999; Lu et al., 2006; Scheike and Zhang, 2002; Martinussen and Scheike, 2006). Though conceptually appealing, existing proposals have had limited success in practice. This is partly because they are computationally demanding, and often difficult for use in high-dimensional settings because of reliance on kernel weighting or splines. Furthermore, as with previous proposals, it is unclear where the suggested estimators converge to when the continuously stratified Cox model is misspecified (e.g., because of the hazard's dependence on A being modified by $\mathbf{L}$), in which case valid inference (even for the population limit of the estimator) is moreover difficult to attain.

To accommodate all of the above concerns, we will initiate the analysis of the conditional association between an exposure and a time-to-event endpoint, given covariates, without relating to a specific statistical model. Instead, we will start the analysis with specification of a so-called estimand, which captures that association in a model-free way. Letting the choice of an estimand be central to the analysis, is also typical of the causal inference literature (Hernán and Robins, 2021; van der Laan and Rose, 2011; Daniel et al., 2016). The estimand that we will propose surpasses the causal inference context and is well-defined regardless of whether the exposure is discrete or continuous. It has the further advantage that it reduces to the logarithm of the hazard ratio when the hazard ratio is constant, and remains well defined more generally; e.g., for a dichotomous exposure, it represents a well understood weighted average of time- and covariate-specific differences in the log cumulative hazard. This is important in view of the earlier discussion on partial likelihood estimators. We will next achieve an assumption-lean inference for the proposed estimand based on its influence function under the nonparametric model (Pfanzagl, 1990). This brings several further advantages. First, it justifies the use of flexible data-adaptive procedures (e.g., variable selection or machine learning procedures), while removing the so-called plug-in bias that would otherwise arise as a result of the bias-variance trade-off of such procedures being optimised for prediction rather than estimation (van der Laan and Rose, 2011; Naimi and Kennedy, 2017; Hines et al., 2022). Second, it delivers inferences which express the overall uncertainty about the considered estimand, including the uncertainty that arises from variable or model selection, while extracting information only from the data, and not from modelling assumptions.

The proposed developments build upon related work for generalised linear models in a discussion paper by Vansteelandt and Dukes (2022), which in turn is closely related to and leans upon recent developments on targeted maximum likelihood estimation (van der Laan and Rose, 2011) and debiased machine learning (Chernozhukov et al., 2018). For the analysis of survival data, to the best of our knowledge, all current developments focus on estimation of the marginal counterfactual probability of surviving a given time point (or functionals thereof, e.g., counterfactual marginal restricted mean survival (Díaz et al., 2019)) if all individuals were, say, treated, or untreated (e.g., Stitelman et al. (2012); Díaz (2019); Cai and van der Laan (2020); Westling et al. (2021)). Such causal estimands are appealing, and in certain settings arguably more appealing than the estimand we will propose. However, as will argue in the next section, the proposed approach has several important advantages: it lays foundations for statistical tests for a conditional association between an exposure and a time-to-event endpoint, it enables the analysis of arbitrary (e.g., continuous) exposures, limits extrapolation by eliminating covariate strata within which positivity violations (for exposure) occur, and remains relevant when no causal inferences are targeted (e.g., when characterising subgroups at greater risk of certain events).

2 A model-free hazard ratio estimand

2.1 Proposal

We begin with some notation. Let $R(t) = I(T \geq t)$ and $R_c(t) = I(C \geq t)$ denote the at risk indicators for the event time T and censoring time C. Further, let A be the exposure of interest, and $\mathbf{L}$ a vector of measured covariates. Finally, let $\lambda(t|A, \mathbf{L})$, $\Lambda(t|A, \mathbf{L})$ and $S(t|A, \mathbf{L})$ denote the conditional hazard, cumulative hazard and survival function of the event T at time t, given A and $\mathbf{L}$. Throughout, we will assume that data are available on $\tilde{T} = \min(T, C)$, $\Delta = I(T \leq C)$, A and $\mathbf{L}$ for each of n mutually independent subjects, and that censoring is non-informative in the sense that $C \perp\!\!\!\perp T|A, \mathbf{L}$. Further, let $S_c(t|A, \mathbf{L})$ denote $P(C > t|A, \mathbf{L})$ and let the counting process corresponding to the survival time T be denoted by $N(t) = I(\tilde{T} \leq t, \Delta = 1)$.

For pedagogic purposes, consider first a dichotomous treatment A, coded 0 (for unexposed, say) or 1 (for exposed, say), and suppose that the log cumulative hazard ratio

$$\beta(t, \mathbf{L}) = \log \frac{\Lambda(t|A = 1, \mathbf{L})}{\Lambda(t|A = 0, \mathbf{L})}$$

is constant in time t (as would be the case when hazards are proportional), but varies possibly between covariate strata $\mathbf{L}$; in that special case, $\beta(t, \mathbf{L})$ is also the log hazard ratio in stratum $\mathbf{L}$. There is then a distribution of log cumulative hazard ratios, governed by the distribution of $\mathbf{L}$. A worthwhile goal is to summarise that distribution, e.g. in terms of its central location and variability. In doing so, it is important to realise that certain covariate strata may carry little or no information about the conditional association between T and A, given $\mathbf{L}$, because most individuals in that stratum are either treated or untreated. We will therefore focus on the following weighted average of $\beta(t, \mathbf{L})$:

$$\beta = \frac{E\{w(\mathbf{L})\beta(t, \mathbf{L})\}}{E\{w(\mathbf{L})\}},$$

where

$$w(\mathbf{L}) = P(A = 1|\mathbf{L})P(A = 0|\mathbf{L})$$

and t is arbitrary (given that $\beta(t, \mathbf{L})$ is assumed constant in t for now). When the continuously stratified Cox model (Sasieni, 1992; Dabrowska, 1997)

$$\lambda(t|A, \mathbf{L}) = \lambda_0(t, \mathbf{L}) \exp(\gamma A),$$

holds for all $t \geq 0$, then

$$\beta(t, \mathbf{L}) = \log \frac{\Lambda(t|A = 1, \mathbf{L})}{\Lambda(t|A = 0, \mathbf{L})} = \log \frac{\exp(\gamma) \int_{\tau_0}^{t} \lambda_0(s, \mathbf{L}) ds}{\int_{\tau_0}^{t} \lambda_0(s, \mathbf{L}) ds} = \gamma,$$

so that β reduces to the log hazard ratio γ in that model. This model expresses that the hazard ratio of exposure does not depend on time and covariates, but leaves the hazard ratios corresponding to covariates $\mathbf{L}$ unrestricted. It thereby drastically relaxes the restrictions imposed by a standard Cox model. Further relaxing this model to

$$\lambda(t|A, \mathbf{L}) = \lambda_0(t, \mathbf{L}) \exp\{\gamma(\mathbf{L})A\}, \tag{1}$$

for certain functions $\gamma(\mathbf{L})$, we have that $\beta(t, \mathbf{L}) = \gamma(\mathbf{L})$. The proposed estimand β then reduces to the average log hazard ratio in a 'retargeted' population in which the probability of inclusion is proportional to $w(\mathbf{L})$; i.e.

$$\frac{E\left\{w(\mathbf{L})\gamma(\mathbf{L})\right\}}{E\left\{w(\mathbf{L})\right\}}.$$

Such retargeting by so-called propensity-overlap weights $w(\mathbf{L}) = P(A = 1|\mathbf{L})P(A = 0|\mathbf{L})$ has been considered in other contexts (Crump et al., 2006; Vansteelandt and Daniel, 2014; Kallus, 2021; Li et al., 2018; Vansteelandt and Dukes, 2022) with the aim to prevent extrapolation. The resulting retargeted population arguably characterises the type of individuals for whom there is sufficient uncertainty about treatment choice in current practice, so that learning the treatment effect is most relevant for them; in particular, it may roughly characterise the population of individuals who would be considered for inclusion in a randomised experiment. This is relevant not only from the viewpoint of statistical efficiency, in the sense that these individuals' data generally carry more information about the conditional association between T and A, given $\mathbf{L}$, but also because it may be too ambitious (and often scientifically less relevant) to infer this association for strata of individuals in which little or no treatment variability is seen. This retargeting is irrelevant when $\beta(t,\mathbf{L})$ does not vary with $\mathbf{L}$, but otherwise affects the interpretation of the proposed estimand; in future work, we will therefore also study the variability in $\beta(t,\mathbf{L})$.

Consider next a continuous exposure A. We will then redefine $\beta(t,\mathbf{L})$ to be the least squares projection (in the Hilbert space of functions of $A,\mathbf{L}$ and t, equipped with the covariance inner product) of $\log\Lambda(t|A,\mathbf{L})$ onto A in individuals from stratum $\mathbf{L}$:

$$\beta(t,\mathbf{L}) = \frac{\mathrm{Cov}\left\{A, \log\Lambda(t|A,\mathbf{L})|\mathbf{L}\right\}}{\mathrm{Var}\left(A|\mathbf{L}\right)},$$

and will moreover redefine $w(\mathbf{L})$ as $\mathrm{Var}\left(A|\mathbf{L}\right)$. This generalises the proposal previously made for a dichotomous exposure. To acknowledge that $\beta(t,\mathbf{L})$ may moreover vary with time t, we will further generalise the earlier proposal by taking a uniformly weighted average over time:

$$\beta = \frac{E\left\{w(\mathbf{L})\int_{\tau_0}^{\tau_0+\tau}\beta(t,\mathbf{L})dt\right\}}{E\{w(\mathbf{L})\tau\}} = \frac{E\left[\int_{\tau_0}^{\tau_0+\tau}\text{Cov}\left\{A,\log\Lambda(t|A,\mathbf{L})|\mathbf{L}\right\}dt\right]}{\tau E\left\{\text{Var}\left(A|\mathbf{L}\right)\right\}}, \tag{2}$$

where $\tau_0 > 0$ is a time point close to zero (to be discussed later) and $\tau_0 + \tau$ is for instance the end-of-study time.

The fact that the estimand (2) can be interpreted as the standard log (cumulative) hazard ratio when the partially linear Cox model holds, but continues to represent a known weighted average of time- and covariate-specific log cumulative hazard ratios more generally, is precisely what makes it appealing for practical use. Indeed, the high dimensionality of $\beta(t,\mathbf{L})$ makes it essentially impossible to report for each t and $\mathbf{L}$ separately, prompting the need to summarise. Summarising usually implies some loss of accuracy, in that only an approximation of reality is attained. But that is often acceptable, and even desired, as we may not be interested in knowing precisely how $\beta(t,\mathbf{L})$ changes with t or $\mathbf{L}$. At least, by working with a clearly defined estimand, we have the guarantee of it characterising the conditional association of interest. What is much more disputable is not knowing what summary measure we end up producing, and whether it even summarises the association of interest, which is what happens with standard (partial likelihood based) analyses. While this paper was under review, Huang et al. (2021) published a technical report which likewise considers model-free estimands that reduce to the exposure coefficient in an accelerated failure time model when that model holds; however, since their focus is on screening high-dimensional predictors, their proposal does not consider covariate adjustment.

2.2 Comparison with alternative estimands

As can be seen from the definition of the weights $w(\mathbf{L})$, the proposed estimand depends on the exposure distribution. This may be viewed as undesirable (Stitelman et al., 2011; Whitney et al., 2019) as it means that the estimand's magnitude may be partly determined by external, ancillary factors. The so-called marginal causal hazard ratio $\lambda_1(t)/\lambda_0(t)$ (Hernán and Robins, 2021) which contrasts the (counterfactual) hazard $\lambda_a(t), a = 0, 1$ at each time t if all individuals were treated versus if all were untreated (in the absence of censoring) undisputably lends itself to a simpler interpretation. However, this greater simplicity comes with a price. In particular, the appeal of the marginal causal hazard ratio as an estimand is somewhat deceptive. Indeed, for the same reasons as made clear for partial likelihood estimators in the introduction, it is not readily clear where standard estimators of it converge to when the assumption of a constant marginal causal hazard ratio is violated. In view of this, Whitney et al. (2019) focus on the average log marginal causal hazard ratio

$$\int_0^\infty \log\left\{\frac{\lambda_1(t)}{\lambda_0(t)}\right\} f_T(t)dt,$$

averaging over the marginal time-to-event distribution $f_T(t)$, in the special case of dichotomous exposures. Our choice to make a different proposal is motivated by the following concerns. First, the above proposal does not readily extend to continuous exposures. Second, the observed data may carry little information about marginal causal hazard ratios and/or the marginal time-to-event distribution when there are near positivity violations (i.e., when individuals with certain covariate values are very unlikely treated or untreated). The resulting lack of information may result in a few individuals becoming very influential in the analysis (as a result of inverse probability of treatment weighting). While this may in turn deliver corresponding estimators with large standard errors, thus flagging a lack of information, heuristic and routinely applied remedial measures such as weight truncation may also end up hiding it. There is thus a need for more extensive empirical evaluations of estimators of the estimand in Whitney et al. (2019) (which are currently limited to completely independent censoring and a single dichotomous confounder). Third, the magnitude of the marginal causal hazard ratio may be strongly dependent upon the covariate distribution of the study population. This needs not be problematic so long as we work with random samples from a carefully defined study population in which detailed insight is provided, but arguably this is rarely the case. It renders interpretation at least more subtle as it makes it unclear whether a reported marginal causal hazard ratio (and/or the considered marginal time-to-event distribution) is readily transportable to a different population (Pearl and Bareinboim, 2014), and in our opinion this is too often overlooked by consumers of study results. Issues of transportability are likewise a concern for the proposed estimand, but we believe it may well be less sensitive to changes in the covariate distribution, and thereby better transportable across populations, by only giving large weights to subpopulations that carry much information about the treatment effect (Kallus, 2021; Vansteelandt and Dukes, 2021); these subpopulations of individuals for whom there is much uncertainty whether to treat or not, may well be more similar across populations.

Our choice to focus on the cumulative hazard ratio may also be viewed as undesirable, but is partly motivated by the fact that its causal interpretation does not suffer from the same subtleties that affect hazard ratios (Hernán, 2010; Martinussen et al., 2020). The cumulative hazard ratio equals 1 over the entire course of follow-up when the exposure is (conditionally) independent of the event time. The proposed estimand can therefore be used as a basis for testing and decision making (see later). Such tests are more cumbersome with typical debiased machine learning approaches for survival analysis which focus on survival probabilities, as they demand simultaneously testing the null hypotheses of equal survival probabilities at all times. Further, the cumulative hazard ratio $\exp\{\beta(t,\mathbf{L})\}$ expresses how a given survival probability $S(t|A=0,\mathbf{L})$ for an untreated subgroup at time t translates to a survival probability of $S(t|A=0,\mathbf{L})^{\exp\{\beta(t,\mathbf{L})\}}$ for the corresponding treated subgroup at that time. This then makes the cumulative hazard ratio a relevant and compact summary, useful for scientific reporting.

3 Data-adaptive inference

3.1 A plug-in estimator

Having constructed a model-free estimand

$$\beta \equiv \frac{E\left[[\int_{\tau_0}^{\tau_0+\tau} \mathrm{Cov}\left\{A, \log\Lambda(t|A,\mathbf{L})|\mathbf{L}\right\} dt\right]}{\tau E\left\{\mathrm{Var}\left(A|\mathbf{L}\right)\right\}} = \frac{E\left[\int_{\tau_0}^{\tau_0+\tau} \left\{A-p(\mathbf{L})\right\}\left\{\log\Lambda(t|A,\mathbf{L})-q(t,\mathbf{L})\right\} dt\right]}{E\left[\tau\left\{A-p(\mathbf{L})\right\}^2\right]}. \tag{3}$$

a so-called plug-in estimator is readily obtained. In particular, upon substituting $p(\mathbf{L}) = E(A|\mathbf{L})$, $\Lambda(t|A,\mathbf{L})$ and $q(t,\mathbf{L}) = E\left\{\log\Lambda(t|A,\mathbf{L})|\mathbf{L}\right\}$ by consistent estimators $\hat{p}(\mathbf{L})$, $\hat{\Lambda}(t|A,\mathbf{L})$ and $\hat{q}(t,\mathbf{L})$ (which will be discussed in the next section), respectively, and moreover replacing population expectations by empirical analogs, we obtain

$$\frac{\sum_{i=1}^n \int_{\tau_0}^{\tau_0+\tau} \left\{A_i-\hat{p}(\mathbf{L}_i)\right\}\left\{\log\hat{\Lambda}(t|A_i,\mathbf{L}_i)-\hat{q}(t,\mathbf{L}_i)\right\} dt}{\sum_{i=1}^n \tau\left\{A_i-\hat{p}(\mathbf{L}_i)\right\}^2}. \tag{4}$$

Use of the above plug-in estimators, obtained by plugging consistent estimators of the unknown (conditional) expectations in (3), raises no specific challenges if these (conditional) expectations are estimated under pre-specified (semi-)parametric models. In that case, $\hat{p}(\mathbf{L}), \hat{q}(t,\mathbf{L})$ and $\hat{\Lambda}(t|A,\mathbf{L})$ can all (typically) be expected to have standard asymptotic behaviour, which propagates into a standard asymptotic behaviour of the plug-in estimators of β. In particular, under the standard Cox model

$$\lambda(t|A, \mathbf{L}) = \lambda_0(t) \exp(\tilde{\beta} A + \gamma' \mathbf{L}), \tag{5}$$

$q(t, \mathbf{L})$ reduces to $\log\{\Lambda_0(t)\} + \tilde{\beta} p(\mathbf{L}) + \gamma' \mathbf{L}$. Plug-in estimator (4) then reduces to the partial likelihood estimator of $\tilde{\beta}$ in model (5). Relying on pre-specified (semi-)parametric models would however defeat the purpose of the whole approach. Indeed, the likely misspecification of these models raises concerns about bias. In view of this, we will focus on data-adaptive procedures, in line with recent debiased machine learning developments (van der Laan and Rose, 2011; Chernozhukov et al., 2018). This may include variable selection procedures under Cox proportional hazards models, as well as more advanced machine learning methods. The use of such procedures brings new complications, however. First, their bias-variance trade-off is optimised towards minimal prediction error, but not towards the considered estimand (van der Laan and Rose, 2011). This may result in bias, e.g., as a result of not selecting certain variables, thereby inducing confounding bias, or selection bias due to informative censoring. Second, data-adaptive estimators $\hat{p}(\mathbf{L}), \hat{q}(t, \mathbf{L})$ and $\hat{\Lambda}(t|A, \mathbf{L})$ have a non-standard and typically unknown distribution. This makes it difficult or even impossible to account for the uncertainty in these estimators when drawing inference about β based on (4). This is for instance so when $\tilde{\beta}$ in model (5) is estimated via post-lasso (i.e., upon fitting the Cox model (5) for the subset of variables $\mathbf{L}$ that were selected via the lasso), in which case its distribution may be a mixture distribution that cannot be well approximated by a normal distribution, no matter what sample size. In view of this, we will make a proposal for accommodating these concerns.

3.2 Data-adaptive nuisance parameter estimators

Before making the suggested proposals, we explain what estimators $\hat{p}(\mathbf{L})$, $\hat{q}(t, \mathbf{L})$ and $\hat{\Lambda}(t|A, \mathbf{L})$ we will use in the simulation experiments in Section 4. To demonstrate the flexibility of the proposal, we will do this both for the case of variable selection in the Cox model (5), as well as for the case where no survival model is assumed. The former is computationally attractive and shows the additional usefulness of the proposal for acknowledging variable selection uncertainty, but the latter is generally recommended by not relying on the Cox model restrictions.

First, $P(A = 1|\mathbf{L})$ (or $E(A|\mathbf{L})$) will be estimated as $\hat{p}(\mathbf{L})$ via Super Learner (van der Laan et al., 2007), an ensemble learner which we based on a main effects logistic (linear) model with and without AIC-based stepwise covariate selection, logistic (linear) additive models and random forests regression. In the simulation studies on variable selection in the Cox model (5), we first fitted the model using post-lasso (via the function `glmnet` in R, with penalty chosen via 20-fold cross-validation to be the largest value of the penalty such that the prediction error was within 1 standard error of the minimum). We then estimated $S(t|A, \mathbf{L})$ and $S_c(t|A, \mathbf{L})$ (which will be needed in Section 3.3) using the function `predictSurvProb` in the R-package `pec`, and $\log \Lambda(t|A, \mathbf{L}) - q(t, \mathbf{L})$ as $\tilde{\beta} \{A - \hat{p}(\mathbf{L})\}$ under the considered model. In the more general simulation studies, we used survival random forests (Ishwaran et al., 2008) (using the function `survival_forest` in the package `grf`) for the estimation of $S(t|A, \mathbf{L})$ and $S_c(t|A, \mathbf{L})$, or the survival Super Learner (using the function `survSuperLearner`; see details later). Further, we estimated $q(t, \mathbf{L})$ for a dichotomous exposure based on the following identity

$$q(t, \mathbf{L}) = \log \Lambda(t|1, L) p(\mathbf{L}) + \log \Lambda(t|0, L) \{1 - p(\mathbf{L})\}.$$

In the data analysis, in which also survival random forests were fitted, but the exposure was continuous, we estimated $q(t, \mathbf{L})$ for each observed event time t via Super Learner prediction of the obtained estimates of $\log \Lambda(t|A, \mathbf{L}) = \log \{-\log S(t|A, \mathbf{L})\}$ onto $\mathbf{L}$.

3.3 Debiased machine learning

In view of the aforementioned problems with plug-in estimators, we will develop inference for β under a nonparametric model. This prevents inference being explicitly or implicitly reliant on modelling restrictions, which in turn could lead to overly optimistic inferences. Like other debiased machine learning proposals, this will be achieved by making use of the influence function of the estimand (Pfanzagl, 1990; Bickel et al., 1993). This is a mean zero functional of the observed data and the data-generating distribution, which characterises the estimand's sensitivity to arbitrary (smooth) changes in the data-generating law. Theorem 1, which is proved in Appendix A, gives the influence function of estimand (3).

Theorem 1. *The estimand (3) has influence function under the nonparametric model for the observed data* $(\tilde{T}, \Delta, A, \mathbf{L})$ *given by*

$$\begin{aligned} \phi(L, A, \tilde{T}, \Delta) \quad = \quad & \frac{\{A - p(\mathbf{L})\}}{\tau E\left[\{A - p(\mathbf{L})\}^2\right]} \int_{\tau_0}^{\tau_0+\tau} \Bigg[\log \Lambda(t|A, \mathbf{L}) - q(t, \mathbf{L}) - \beta \{A - p(\mathbf{L})\} \\ & + \frac{1}{\Lambda(t|A, \mathbf{L})} \int_0^t \frac{dM(s|A, \mathbf{L})}{S(s|A, \mathbf{L}) S_c(s|A, \mathbf{L})} \Bigg] dt, \end{aligned}$$

where β *is given by (3), and where* $p(\mathbf{L}) = E(A|\mathbf{L})$*,* $q(t, \mathbf{L}) = E\{\log \Lambda(t|A, \mathbf{L})|\mathbf{L}\}$*, and* $dM(t|A, \mathbf{L}) = R(t) R_c(t) \{dN(t) - d\Lambda(t|A, \mathbf{L})\}$*.*

If the infinite-dimensional nuisance parameters $p(t, \mathbf{L}), q(t, \mathbf{L}), \Lambda(t|A, \mathbf{L}), S(t|A, \mathbf{L})$ and $S_c(t|A, \mathbf{L})$ indexing the influence function were known, then it would follow from its mean zero property that a consistent estimator of β could be obtained as the value of β that makes the sample average of the influence functions zero. In practice, we will however substitute these nuisance parameters by the estimators discussed in the previous section. Using the resulting estimators, we obtain a closed-form estimator as the value β that sets the sample average of the influence functions equal to zero:

$$\begin{aligned}\hat{\beta} &= \left[\tau \sum_{i=1}^{n} \{A_i - \hat{p}(\mathbf{L}_i)\}^2\right]^{-1} \sum_{i=1}^{n} \{A_i - \hat{p}(\mathbf{L}_i)\} \int_{\tau_0}^{\tau_0+\tau} \left[\log \hat{\Lambda}(t|A_i, \mathbf{L}_i) - \hat{q}(t, \mathbf{L}_i)\right. \\ &\quad \left. + \frac{1}{\hat{\Lambda}(t|A_i, \mathbf{L}_i)} \int_0^t \frac{d\hat{M}_i(s|A_i, \mathbf{L}_i)}{\hat{S}(s|A_i, \mathbf{L}_i)\hat{S}_c(s|A_i, \mathbf{L}_i)}\right] dt, \qquad (6)\end{aligned}$$

where the hazard function estimator $d\hat{\Lambda}(t|A_i, \mathbf{L}_i)$ indexing $d\hat{M}_i(s|A_i, \mathbf{L}_i) = R_i(s)R_{i,c}(s) \times \left\{dN_i(s) - d\hat{\Lambda}(s|A_i, \mathbf{L}_i)\right\}$ is chosen to be piecewise constant with jumps at the observed event times, equal to the change in cumulative hazard at those times. In the special case of variable selection under the Cox model (5), this reduces to

$$\hat{\beta} = \tilde{\beta} + \frac{\sum_{i=1}^{n} \{A_i - \hat{p}(\mathbf{L}_i)\} \int_{\tau_0}^{\tau_0+\tau} \frac{1}{\hat{\Lambda}(t|A_i,\mathbf{L}_i)} \int_0^t \frac{d\hat{M}_i(s|A_i,\mathbf{L}_i)}{\hat{S}(s|A_i,\mathbf{L}_i)\hat{S}_c(s|A_i,\mathbf{L}_i)} dt}{\tau \sum_{i=1}^{n} \{A_i - \hat{p}(\mathbf{L}_i)\}^2},$$

where the second term debiases the post-lasso estimator $\tilde{\beta}$.

It follows from the general theory on nonparametric estimation (Pfanzagl, 1990; Bickel et al., 1993) and the results in the Appendix that when the nuisance parameters are estimated from a separate sample, independent from the one which is used for the calculation of $\hat{\beta}$, then $\hat{\beta}$'s asymptotic distribution is governed by its influence function in the sense that

$$\sqrt{n}\left(\hat{\beta} - \beta\right) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \phi(\mathbf{L}_i, A_i, \tilde{T}_i, \Delta_i) + o_p(1). \qquad (7)$$

This implies that $\hat{\beta}$ is asymptotically normally distributed with bias that shrinks to zero faster than the standard error, and with a variance that can be estimated as 1 over n times the sample variance of the influence functions (where population expectations, nuisance parameters and the value of β can be substituted by consistent estimates). That is,

$$\widehat{\text{Var}}\left(\hat{\beta}\right) = \frac{1}{n^2}\sum_{i=1}^{n}\left\{\hat{\phi}(\mathbf{L}_i, A_i, \tilde{T}_i, \Delta_i)\right\}^2,$$

where $\hat{\phi}(.)$ is defined like $\phi(.)$, but with all nuisance parameters substituted by consistent estimates. This is quite remarkable as it implies that the uncertainty in these nuisance parameter estimators, which is typically poorly understood when data-adaptive estimators are used, does not affect inference based on $\hat{\beta}$.

In practice, the requirement that the nuisance parameters are estimated from a separate sample, independent from the one which is used for the calculation of $\hat{\beta}$, can be ensured using sample splitting (Zheng and van der Laan, 2011), also referred to as cross-fitting (Chernozhukov et al., 2018). In particular, the data can be split into K (e.g., 10) non-overlapping folds. In the expression for β, when individual i belongs to fold $k = 1, ..., K$, then the nuisance parameter estimators are obtained by analysing all $K-1$ remaining folds. Such cross-fitting is needed to ensure asymptotic validity of the proposal, though in small to moderate sample sizes, better results may sometimes be obtained by avoiding sample splitting as it may deliver better nuisance parameter estimates. The following theorem details the additional conditions required to ensure validity of the above asymptotic result.

Theorem 2. *Estimator $\hat{\beta}$ defined in (6) is asymptotically linear (in the sense of (7)) when*

1. *all nuisance parameter estimators are consistent,*
2. *all nuisance parameters are estimated on a sample independent to the one on which $\hat{\beta}$ is evaluated,*

3. all of the following terms are $o_p(n^{-1/2})$*:*

$$E\left[\int_0^{\tau_0+\tau}\left(\int_{\max(\tau_0,t)}^{\tau_0+\tau}\hat{\Lambda}^{-1}(s|A,\mathbf{L})ds\right)\left(\frac{S_c(t|A,\mathbf{L})-\hat{S}_c(t|A,\mathbf{L})}{\hat{S}_c(t|A,\mathbf{L})}\right)d\left\{\frac{S(t|A,\mathbf{L})-\hat{S}(t|A,\mathbf{L})}{\hat{S}(t|A,\mathbf{L})}\right\}\right]$$

$$\sup_{t\in[\tau_0,\tau_0+\tau]}\left\{\hat{\Lambda}(t|A,\mathbf{L})-\Lambda(t|A,\mathbf{L})\right\}^2$$

$$E\left[\{\hat{p}(\mathbf{L})-p(\mathbf{L})\}^2\right]^{1/2}\sup_{t\in[\tau_0,\tau_0+\tau]}E\left[\{\hat{q}(t|\mathbf{L})-q(t|\mathbf{L})\}^2\right]^{1/2}$$

$$E\left[\{\hat{p}(\mathbf{L})-p(\mathbf{L})\}^2\right],$$

4. $\inf_{t\in[\tau_0,\tau_0+\tau]}\hat{\Lambda}(t|A,\mathbf{L})>\sigma$ *w.p.1 for some* $\sigma>0$*, and*

5. $|A-\hat{p}(\mathbf{L})|\leq B$ *w.p.1. for some finite upper bound B.*

The condition $\inf_{t\in[\tau_0,\tau_0+\tau]}\hat{\Lambda}(t|A,\mathbf{L})>\sigma$ w.p.1 is strong. While ideally τ_0 is close to zero, this condition requires τ_0 to be sufficiently large so that events have taken place prior to τ_0. Similar requirements as in Theorem 2 are seen in Westling et al. (2021). They suggest that slow convergence of the censoring probabilities can be compensated by fast convergence of the survival probabilities. These requirements hold when parametric models or certain semiparametric models (e.g. Cox proportional hazard models) are used for the nuisance functions, even when the fitting procedure involves l_1-penalisation (provided that typical ultra-sparsity conditions hold). However, these conditions hold much more generally. In practice, we recommend the use of ensemble learners (e.g., SuperLearner) based on a combination of parametric model-based estimators, which respect the required rates under correct model specification, as well as data-adaptive (e.g. machine learning) estimators whose convergence rate is generally less well understood. Use of the highly-adaptive lasso (Benkeser and van der Laan, 2016) has also been suggested as it is guaranteed to deliver faster than $n^{1/4}$ rates, provided that the nuisance functionals lie in a class of functions with bounded variation norm. There are currently limited options to proceed when the rate conditions of Theorem 2 are not fulfilled, so that the remaining bias is too large for root-n inference. Whilst there is already some theory (e.g., based on higher order influence functions (Robins et al., 2008)), these methods are currently computationally prohibiting even for simpler estimands. We refer to Appendix B in the Supplementary Materials for further discussion of these conditions.

4 Simulation experiments

We illustrate the proposed procedure using two sets of simulation experiments, once using variable selection under the standard Cox model (5) and once without relating to a specific survival model (using survival random forests instead). In the first set of simulation experiments, we let $\mathbf{L}$ be a 10-dimensional mean zero multivariate normal variate with unit variance and Toeplitz covariance matrix with correlations 1, 0.9, 0.8, ..., 0.1. Further, the exposure A is normally distributed with mean $\sum_{j=1}^{10} L_j/j$ (where L_j is the jth element of the vector $\mathbf{L}$) and unit variance, T is exponentially distributed with hazard $\exp\left(-2.5 + 0.5A - 0.5\sum_{j=1}^{10} L_j/j\right)$ and C is the minimum of 30 and an exponentially distributed variable with hazard $\exp(-3+0.5A+0.1\sum_{j=1}^{5} L_j/j+0.1\sum_{j=1}^{5} L_{j+5}/j)$. The analysis assumes that the hazard obeys the (correctly specified) Cox model (5), which is then fitted via post-lasso (with penalty chosen via 20-fold cross-validation to be the largest value of the penalty such that the prediction error was within 1 standard error of the minimum). In the second set of simulation experiments, we generate $\mathbf{L}$ as before and let the exposure A be Bernoulli distributed with mean $\text{expit}(-2\sqrt{|L_1 L_2|}+2\sqrt{|L_{10}|}-2\cos L_5+2\cos L_5 \cos L_6)$ and unit variance, T is Weibull distributed with shape parameter 1.5 and scale parameter

$100^{2/3}\exp\left(-A-0.5\sqrt{|L_1L_2|}+0.5\sqrt{|L_{10}|}-0.5\cos L_5+0.5\cos L_5\cos L_6\right)^{2/3}$ and C is the minimum of 12 and a Weibull distributed variable with shape parameter 1.5 and scale parameter $25\exp\left(-A-0.5\sqrt{|L_1L_{10}|}+0.5\sqrt{|L_9|}-0.5\cos L_5+0.5\cos L_7\cos L_6\right)^{2/3}$. The latter data-generating mechanisms are inspired by that in Díaz (2019) (with some modifications, e.g., we did not consider truncated normal covariate distributions, which may have been chosen in Díaz (2019) to prevent extreme inverse probability weights, we considered a continuous-time setup, among others). In the first set of simulation results, roughly 55% of subjects experienced an event, 40% got censored during the study and the remaining 5% was administratively censored. In the second set of simulation results, roughly 40% of subjects experienced an event, 30% got censored during the study and the remaining 30% was administratively censored.

The tables below show results for 1000 simulation runs for the oracle estimator (Oracle) obtained by fitting a standard Cox model with exposure A and a scalar covariate given by the corresponding part of the linear predictor that includes all covariate terms, and for plug-in estimator (4). In the first set of simulations, plug-in estimator (4) equals the post-lasso estimator; in the second set of simulations, we also provide results for the partial likelihood estimator (Part Lik) obtained under a misspecified Cox model of the form (5) (with main effects of exposure and the covariates $\mathbf{L}$) to give some appreciation for the extent of bias due to model misspecification, and for Cox models with natural splines with 3 or 4 degrees of freedom for all covariates (Spline3, Spline4). We finally also report results for the proposed debiased machine learning estimator based on survival random forests without cross-fitting (DML) and with 5-fold cross-fitting (DML-CF) (where the use of 5 folds aligns with other papers on debiased machine learning for time-to-event endpoints (Westling et al., 2021; Wen et al., 2021)). In the debiased machine learning estimators,

τ_0 was chosen as the time at which the largest estimated survival probability was upper bounded by $\exp(-10/n) \approx 0.99$ (in view of the inverse of the cumulative hazard taking extreme values at lower time points) and where $\tau_0 + \tau$ was chosen as the time at which the minimum of the product of the estimated survival and censoring probabilities exceeded $10/n$. This data-adaptive choice of the interval $[\tau_0, \tau_0 + \tau]$ may possibly induce some residual degree of plug-in bias (see later). In the tables below, we report Monte Carlo bias as the sample average of the estimates minus the true value of the estimand (0.5 and 1 in the first and second set of results), Monte Carlo standard deviation as the sample standard deviation of the estimates, the sample average of the standard errors estimated as 1 over root n times the sample standard deviation of the estimated influence functions, and finally the percentage of times the 95% Wald confidence intervals based on these estimated standard errors cover the true value of the estimand.

The simulations on variable selection (see Table 1) demonstrate the poor performance of the post-lasso estimator, which suffers large bias and excess variability as a result of variability in the selected covariate set. It moreover comes with a poorly estimated, overly optimistic standard error as a consequence of ignoring variable selection uncertainty. It therefore produces confidence intervals with poor finite sample performance. Of the two proposed estimators, the debiased machine learning estimator without cross-fitting generally performs best in terms of bias, likely as a result of estimating the nuisance parameters on larger sample sizes. The use of cross-fitting comes with a slight reduction in variability and, as expected, estimated standard errors that more accurately approximate the Monte Carlo standard deviation. While the proposed estimators are unsurprisingly more variable than the oracle estimator, the results at $n = 400$ show a much reduced variability relative to the post-lasso estimator.

Table 1: Simulation results for variable selection in the Cox proportional hazards model: sample size n, Monte Carlo bias, scaled Monte Carlo bias, Monte standard deviation (SD), average of the influence function based standard errors (Average SE) and coverage of 95% Wald confidence intervals (Cov).

Estimator	n	Bias	$n^{1/2}$Bias	SD	Average SE	95% Cov
Oracle	200	0.012	0.17	0.10	0.11	96.5
Plug-in		-0.31	-4.32	0.21	0.058	20.2
DML		0.023	0.24	0.21	0.19	90.1
DML-CF		-0.049	-0.70	0.20	0.19	94.1
Oracle	400	0.0026	0.037	0.074	0.074	95.3
Plug-in		-0.14	-2.05	0.20	0.056	52.9
DML		0.0083	0.12	0.14	0.13	92.4
DML-CF		-0.025	-0.35	0.13	0.12	93.7

The simulations on survival random forests reveal a small bias in the debiased machine learning estimator without cross-fitting, likely by ignoring the data-adaptive choice of the integration limits. Indeed, when letting the lower limit τ_0 equal the first event time at which all estimated survival probabilities differ from 1, and the upper bound $\tau_0 + \tau$ equal the administrative censoring time of 12, the bias shrunk to 0.015 at $n = 500$ and 0.017 at $n = 1000$ at the expense of more variability: Monte Carlo SD 0.27 (Mean SE 0.25, Coverage 93.7) at $n = 500$ and 0.18 (Mean SE 0.18, Coverage 94.8) at $n = 1000$. We found the small bias seen in Table 2 not to be worrisome, as the table shows that the bias shrinks at faster than 1 over root-n rate. This is no longer the case when cross-fitting is used, which results in an unacceptable bias, though much less variability (of the same magnitude as seen in the Oracle estimator) that is accurately predicted by the estimated standard errors. This larger bias is likely the result of the data-adaptive choice of τ_0, which may lead to a different value of τ_0 in each fold.

As before, we found the plug-in machine learning estimator to suffer large bias. The spline-based estimators perform generally well, but the increase in root-n times the bias with sample size, as seen when 3 degrees of freedom are used, is worrisome. Moreover, note that comparison with these estimators is somewhat unfair as they rely on the proportional hazards structure, unlike the proposed estimators, and because their degrees of freedom are not chosen data-adaptively. Like for the plug-in estimator, we indeed expect to see some bias when variables and/or degrees of freedom are chosen data-adaptively, as is common in real applications. In the Appendix, we show that similar (slightly better) results are seen at lower censoring rates, despite the increased difficulty of estimating the censoring probabilities.

Table 2: Simulation results for survival random forests: sample size n, Monte Carlo bias, scaled Monte Carlo bias, Monte standard deviation (SD), average of the influence function based standard errors (Average SE) and coverage of 95% Wald confidence intervals (Cov).

Estimator	n	Bias	$n^{1/2}$Bias	SD	Average SE	95% Cov
Oracle	500	0.0026	0.058	0.17	0.17	95.2
Part Lik		-0.34	-7.69	0.15	0.15	36.7
Spline3		-0.056	-1.26	0.18	0.17	93.7
Spline4		0.057	1.27	0.19	0.18	93.1
Plug-in		-0.65	-14.6	0.19	0.0052	0.1
DML		0.052	1.17	0.23	0.24	94.7
DML-CF		-0.16	-3.61	0.16	0.16	84.0
Oracle	1000	0.00077	0.024	0.12	0.12	94.7
Part Lik		-0.36	-11.2	0.10	0.10	7.1
Spline3		-0.091	-2.88	0.12	0.12	87.2
Spline4		0.015	0.47	0.13	0.12	94.9
Plug-in		-0.44	-13.8	0.17	0.0057	0.3
DML		0.024	0.75	0.15	0.17	97.2
DML-CF		-0.14	-4.44	0.12	0.12	79.6

5 Empirical illustration

To illustrate the proposed methods, we consider the analysis of an observational study on the predictive relationship between the initial concentration of serum monoclonal protein and death in 1338 patients with monoclonal gammopathy of undetermined significance (MGUS), residing in southeastern Minnesota (Kyle et al., 2002). MGUS is a condition in which the blood contains an abnormal protein (monoclonal protein). It affects up to 2% of persons aged 50 years or more (primarily older men). It usually causes no severe health problems, but can progress to some forms of blood cancer. Cai and van der Laan (2020) analysed these data using a one-step TMLE, but had to dichotomise the exposure, namely the monoclonal spike on serum protein electrophoresis, which is discrete but taking on 29 different levels from 0 to 3 (average 1.16, standard deviation 0.57). The methods proposed in the current paper do not require such dichotomisation and we will therefore assess the conditional association between exposure and the time to death, conditional on the baseline covariates age, gender, haemoglobin, creatinine and time of diagnosis. With access to complete data on 1338 individuals, a standard Cox analysis delivers a log hazard ratio of 0.028 (SE 0.0618) corresponding to a unit increase in exposure, after adjusting for all covariates (using main effects), and of 0.031 (SE 0.0594) after adjustment for only those covariates that are selected by the lasso. A score test of the proportional hazard assumption showed evidence of violations with respect to age (P 2 10^{-7}), haemoglobin (P 0.017), creatinine (P 0.011) and time of diagnosis (P 0.030), suggesting that the above analysis may be invalid; there was no such evidence with respect to the exposure (P 0.072) and neither was there sufficient evidence of a violation of non-linearity with respect to the exposure (P 0.061 based on a likelihood ratio test in a Cox model with 4-degree of freedom natural splines for all covariates). We thus considered the debiased machine learning proposal with survival random forests for the survival and censoring probabilities, and SuperLearner predictions for the exposure and for the log cumulative hazard at each time (using the algorithms and integration bounds considered in the simulation study). This resulted in a plug-in estimator of the log hazard ratio corresponding to a unit increase in exposure, equal to 0.0063 (SE 0.133), a debiased machine learning estimator equal to -0.063 (SE 0.175) without cross-fitting and -0.042 (SE 0.138) with 5-fold cross-fitting.

6 Discussion

Our focus in this paper has been on reporting a scalar summary of the conditional association between an exposure and a time-to-event endpoint, given covariates. For this, we were inspired by the estimand proposed in Vansteelandt and Dukes (2022) with complementary log log link and outcome given by the at risk indicator at a given time t, but we faced challenges due to censoring and the fact that there is not a unique time t of primary interest. The resulting proposal may in particular be used to test for a conditional association between an exposure and a time-to-event endpoint, conditional on covariates, without relying on the proportional hazards assumption. When this association is highly non-linear in the exposure (on the log hazard scale), then any attempt to summarise it in terms of a single number may end up being unsuccessful, unless linearity can be imposed via transformations of the exposure. In such cases, one may rather need to visualise the cumulative hazard's dependence on the exposure, which is beyond the scope of this work. The proposed estimand is likewise not designed for studies which have specific interest in effect heterogeneity. Such studies call for the development of specific effect modification parameters, e.g., along the lines of Vansteelandt and Dukes (2022), or for a parallel development that focuses on the variability (rather than central location) of $\beta(t, \mathbf{L})$ across t or $\mathbf{L}$. These developments need further work.

While our primary aim was to develop an inference that is not explicitly relying on the validity of an assumed Cox model, we have indicated that the proposed results are also useful to acknowledge variable selection uncertainty when a Cox model is assumed. In that case, the considered estimand can be interpreted as a standard hazard ratio, so that our proposal may be viewed as delivering valid data-adaptive inference for hazard ratios. Relative to competing methods for post-selection inference in survival analysis (Fang et al., 2017; Van Lancker et al., 2021), the advantage of the proposed strategy is that it makes a first order bias correction towards the interpretable estimand (3), even when the analysis relies on the Cox model, by invoking a nonparametrically estimated propensity score model. In future work, it will be of interest to compare these approaches with the proposed approach in settings with and without correct model specification.

As in Vansteelandt and Dukes (2022), we have deliberately focused on estimands that deliver an influence function that does not involve inverse weighting by the conditional exposure density, given covariates. The reason is that such weighting may render the approach unstable and sensitive to minor estimation errors in the tails of the density, especially when a continuous exposure is considered. The resulting estimands moreover downweigh subjects in whom little exposure variability is seen; this is desirable when this low variability points towards subjects who are either ineligible to other exposure levels or unwilling to change their exposure. A key limitation of the proposal, however, is that it delivers an influence function that involves inverse weighting by the cumulative hazard. We have remedied this by choosing τ_0 in the definition of the estimand not too close to 0. In the Appendix, we discuss a number of alternative estimands that were considered in preparation of this work, which also reduce to the hazard ratio under a Cox proportional hazard regression model when that model is correctly specified. Future work will examine more optimal ways of averaging $\beta(t, \mathbf{L})$ over time, e.g. by considering the weighted average that delivers the estimand with the most favourable nonparametric efficiency bound.

References

Benkeser, D. and van der Laan, M. "The highly adaptive lasso estimator." In *2016 IEEE international conference on data science and advanced analytics (DSAA)*, 689–696. IEEE (2016).

Bickel, P. J., Klaassen, C. A., Bickel, P. J., Ritov, Y., Klaassen, J., Wellner, J. A., and Ritov, Y. *Efficient and adaptive estimation for semiparametric models*, volume 4. Johns Hopkins University Press Baltimore (1993).

Buja, A., Brown, L., Kuchibhotla, A. K., Berk, R., George, E., Zhao, L., et al. "Models as approximations ii: A model-free theory of parametric regression." *Statistical Science*, 34(4):545–565 (2019).

Cai, W. and van der Laan, M. J. "One-step targeted maximum likelihood estimation for time-to-event outcomes." *Biometrics*, 76(3):722–733 (2020).

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., and Robins, J. "Double/debiased machine learning for treatment and structural parameters." *The Econometrics Journal*, 21(1):C1–C68 (2018).

Cox, D. R. "Regression models and life-tables." *Journal of the Royal Statistical Society: Series B*, 34(2):187–202 (1972).

Crump, R. K., Hotz, V. J., Imbens, G. W., and Mitnik, O. A. "Moving the goalposts: Addressing limited overlap in the estimation of average treatment effects by changing the estimand." Technical report, National Bureau of Economic Research (2006).

Dabrowska, D. M. "Smoothed cox regression." *The Annals of Statistics*, 1510–1540 (1997).

Daniel, R. M., De Stavola, B. L., and Vansteelandt, S. "Commentary: The formal approach to quantitative causal inference in epidemiology: misguided or misrepresented?" *International journal of epidemiology*, 45(6):1817–1829 (2016).

Díaz, I. "Statistical inference for data-adaptive doubly robust estimators with survival outcomes." *Statistics in Medicine*, 38(15):2735–2748 (2019).

Díaz, I., Colantuoni, E., Hanley, D. F., and Rosenblum, M. "Improved precision in the analysis of randomized trials with survival outcomes, without assuming proportional hazards." *Lifetime data analysis*, 25(3):439–468 (2019).

Fang, E. X., Ning, Y., and Liu, H. "Testing and confidence intervals for high dimensional proportional hazards models." *Journal of the Royal Statistical Society: Series B*, 79(5):1415–1437 (2017).

Hernán, M. A. "The hazards of hazard ratios." *Epidemiology*, 21(1):13 (2010).

Hernán, M. A. and Robins, J. M. *Causal inference*. Boca Raton: Chapman & Hall/CRC (2021).

Hines, O., Dukes, O., Diaz-Ordaz, K., and Vansteelandt, S. "Demystifying statistical learning based on efficient influence functions." *The American Statistician*, 1–13 (2022).

Huang, J. "Efficient estimation of the partly linear additive Cox model." *The annals of Statistics*, 27(5):1536–1563 (1999).

Huang, T.-J., Luedtke, A., and McKeague, I. W. "Efficient Estimation of the Maximal Association between Multiple Predictors and a Survival Outcome." *arXiv preprint arXiv:2112.10996* (2021).

Ishwaran, H., Kogalur, U. B., Blackstone, E. H., and Lauer, M. S. "Random survival forests." *The annals of applied statistics*, 2(3):841–860 (2008).

Kallus, N. "More efficient policy learning via optimal retargeting." *Journal of the American Statistical Association*, 116(534):646–658 (2021).

Kyle, R. A., Therneau, T. M., Rajkumar, S. V., Offord, J. R., Larson, D. R., Plevak, M. F., and Melton III, L. J. "A long-term study of prognosis in monoclonal gammopathy of undetermined significance." *New England Journal of Medicine*, 346(8):564–569 (2002).

Li, F., Morgan, K. L., and Zaslavsky, A. M. "Balancing covariates via propensity score weighting." *Journal of the American Statistical Association*, 113(521):390–400 (2018).

Lu, X., Chen, G., Singh, R. S., and K. Song, P. X. "A class of partially linear single-index survival models." *Canadian Journal of Statistics*, 34(1):97–112 (2006).

Martinussen, T. and Scheike, T. H. *Dynamic regression models for survival data*, volume 1. Springer (2006).

Martinussen, T., Vansteelandt, S., and Andersen, P. K. “Subtleties in the interpretation of hazard contrasts.” *Lifetime Data Analysis*, 26(4):833–855 (2020).

Naimi, A. I. and Kennedy, E. H. “Nonparametric double robustness.” *arXiv preprint arXiv:1711.07137* (2017).

Pearl, J. and Bareinboim, E. “External validity: From do-calculus to transportability across populations.” *Statistical Science*, 579–595 (2014).

Pfanzagl, J. “Estimation in semiparametric models.” In *Estimation in Semiparametric Models*, 17–22. Springer (1990).

Robins, J., Li, L., Tchetgen, E., van der Vaart, A., et al. “Higher order influence functions and minimax estimation of nonlinear functionals.” In *Probability and statistics: essays in honor of David A. Freedman*, 335–421. Institute of Mathematical Statistics (2008).

Robins, J. M. “Association, causation, and marginal structural models.” *Synthese*, 151–179 (1999).

Sasieni, P. “Non-orthogonal projections and their application to calculating the information in a partly linear Cox model.” *Scandinavian Journal of Statistics*, 215–233 (1992).

Scheike, T. H. and Zhang, M.-J. “An additive–multiplicative Cox–Aalen regression model.” *Scandinavian Journal of Statistics*, 29(1):75–88 (2002).

Stensrud, M. J. and Hernán, M. A. “Why Test for Proportional Hazards?” *Journal of the American Medical Association*, 323(14):1401–1402 (2020).

Stitelman, O. M., De Gruttola, V., and van der Laan, M. J. "A general implementation of TMLE for longitudinal data applied to causal inference in survival analysis." *The international journal of biostatistics*, 8(1) (2012).

Stitelman, O. M., Wester, C. W., De Gruttola, V., and van der Laan, M. J. "Targeted maximum likelihood estimation of effect modification parameters in survival analysis." *The international journal of biostatistics*, 7(1) (2011).

Struthers, C. A. and Kalbfleisch, J. D. "Misspecified proportional hazard models." *Biometrika*, 73(2):363–369 (1986).

van der Laan, M. J., Polley, E. C., and Hubbard, A. E. "Super learner." *Statistical applications in genetics and molecular biology*, 6(1) (2007).

van der Laan, M. J. and Rose, S. *Targeted Learning*. Springer Series in Statistics. New York, NY: Springer New York (2011).

Van Lancker, K., Dukes, O., and Vansteelandt, S. "Principled selection of baseline covariates to account for censoring in randomized trials with a survival endpoint." *Statistics in Medicine*, 40(18):4108–4121 (2021).

Vansteelandt, S. and Daniel, R. M. "On regression adjustment for the propensity score." *Statistics in medicine*, 33(23):4053–4072 (2014).

Vansteelandt, S. and Dukes, O. "Discussion of 'More Efficient Policy Learning via Optimal Retargeting'." *Journal of the American Statistical Association*, 675–679 (2021).

—. "Assumption-lean inference for generalised linear model parameters (with discussion)." *Journal of the Royal Statistical Society - Series B*, 84:657–739 (2022).

Wen, L., Marcus, J., and Young, J. "Intervention treatment distributions that depend on the observed treatment process and model double robustness in causal survival analysis." *arXiv preprint arXiv:2112.00807* (2021).

Westling, T., Luedtke, A., Gilbert, P., and Carone, M. "Inference for treatment-specific survival curves using machine learning." *arXiv preprint arXiv:2106.06602* (2021).

Whitney, D., Shojaie, A., Carone, M., et al. "Comment: Models as (Deliberate) Approximations." *Statistical Science*, 34(4):591–598 (2019).

Zheng, W. and van der Laan, M. J. "Cross-Validated Targeted Minimum-Loss-Based Estimation." In *Targeted Learning*, 459–474. New York, NY: Springer New York (2011).

Zhong, Q., Mueller, J., and Wang, J.-L. "Deep learning for the partially linear cox model." *The Annals of Statistics*, 50(3):1348–1375 (2022).