November 2, 2023

# The state learner
# – a super learner for right-censored data

Anders Munch and Thomas A. Gerds

**Abstract**

In survival analysis, prediction models are needed as stand-alone tools and in applications of causal inference to estimate nuisance parameters. The super learner is a machine learning algorithm which combines a library of prediction models into a meta learner based on cross-validated loss. Unfortunately, the commonly used partial likelihood loss is not suited for super learning, and inverse probability of censoring weighted loss functions require a pre-specified estimator of the censoring distribution. To relax this, we introduce the state learner, a new super learner for survival analysis, which evaluates the loss based on the observed data simultaneously using libraries of predictions models for the event(s) of interest and the censoring distribution. We establish an oracle inequality for the state learner and investigate its performance through numerical experiments. We illustrate how the state learner allows us to estimate causal effects in a competing risks setting without having to pre-specify models for neither the cause-specific hazards of interest nor the censoring distribution.

## 1 Introduction

A super learner is a machine learning algorithm that combines a finite set of learners into a meta learner by estimating prediction performance in hold-out samples using a pre-specified loss function [van der Laan et al., 2007]. When the aim is to make a prediction model, super learners typically combine strong learners, such as Cox regression models and random survival forests [Gerds and Kattan, 2021]. While the general idea of combining strong learners based on cross-validation data stems from earlier work [Wolpert, 1992, Breiman, 1996], the name super learner is justified by an oracle inequality [van der Laan and Dudoit, 2003, van der Vaart et al., 2006]: The super learner is guaranteed to perform almost as well as the model which minimizes the expected performance, i.e., the model we would select if we could evaluate the prediction performance in an infinite hold-out sample.

We are concerned with the choice of the loss function for super learning in survival analysis. Existing super learner algorithms for right-censored data use partial log-likelihood loss or inverse probability of censoring weighted loss [Polley and van der Laan, 2011, Keles et al., 2004, Golmakani and Polley, 2020, Westling et al., 2021]. The use of the partial log-likelihood loss restricts the class of learners and excludes for example simple Kaplan-Meier based learners and also more complex random survival forest algorithms. For this reason Golmakani and Polley [2020] restrict their learners to Cox proportional hazard models. A lesser known fact is that a super learner constructed with the negative partial log-likelihood loss implicitly depends on the censoring distribution [Hjort, 1992, Whitney et al., 2019]. A disadvantage of inverse probability of censoring weighted loss functions is that they requires a pre-specified model for the censoring distribution. Westling et al. [2021] tackle this challenge by iterating between super learning of the censoring distribution and the event time distribution.

In this article we define the state learner, a new super learner for right-censored data, which simultaneously evaluates the loss for learners of the event time distribution and the censoring distribution. The loss function which is used to define the state learner is only based on observable quantities. The state learner can be applied to all types of survival estimators, works in the presence of competing risks, and does not require a single pre-specified estimator of the conditional censoring distribution. To analyze the theoretical properties of the state learner we focus on the so-called discrete super learner which 'combines' the library of learners by picking the one that minimizes the cross-validated loss. The state learner uses separate libraries to model each competing event and the censoring distribution. We show that the oracle selector of the state learner is consistent if all libraries contain a consistent learner and prove a finite sample oracle inequality.

The state learner can be used to select a model which predicts the probability of an event based on covariates in the presence of competing risks. Another application is in targeted learning where conditional event probabilities occur as high-dimensional nuisance parameters which need to be estimated at a certain rate [van der Laan and Rose, 2011, Rytgaard et al., 2021, Rytgaard and van der Laan, 2022]. We show how a targeted estimator can be obtained from the state learner, and that a second order product structure for the asymptotic bias term of the targeted estimator is retained when the state learner is used to estimate nuisance parameters.

The article is organized as follows. We introduce our notation and framework in Section 2. In Section 3 we define super learning in general with right-censored data, and in Section 4 we introduce the state learner. Section 5 provides theoretical guarantees for the state learner. In Section 6 we discuss the use of the state learner in the context of targeted learning. We report a numerical study in Section 7, and analyze a prostate cancer data set in Section 8. Finally, we relate the state learner to existing approaches in Section 9 and discuss some limitations of our proposal. Appendices A and B contain proofs.

## 2 Notation and framework

In a competing risk framework [Andersen et al., 2012], let $T$ be a time to event variable, $D \in \{1, 2\}$ the cause of the event, and $X \in \mathcal{X}$ a vector of baseline covariates taking values in a bounded subset $\mathcal{X} \subset \mathbb{R}^p$, $p \in \mathbb{N}$. Let $\tau < \infty$ be the maximal length of follow-up. We use $\mathcal{Q}$ to denote the collection of all probability measures on $[0, \tau] \times \{1, 2\} \times \mathcal{X}$ such that $(T, D, X) \sim Q$ for some unknown $Q \in \mathcal{Q}$. For $j \in \{1, 2\}$, the cause-specific conditional cumulative hazard functions are defined by $\Lambda_j \colon [0, \tau] \times \mathcal{X} \to \mathbb{R}_+$ such that

$$\Lambda_j(t \mid x) = \int_0^t \frac{Q(T \in \mathrm{d}s, D = j \mid X = x)}{Q(T \geq s \mid X = x)}.$$

For ease of notation we assume throughout that $\Lambda_j(\cdot \mid x)$ is continuous for all $x$ and $j$. We denote by $S$ the conditional event-free survival function,

$$S(t \mid x) = \exp\{-\Lambda_1(t \mid x) - \Lambda_2(t \mid x)\}. \tag{1}$$

Let $\mathcal{M}$ denote the space of all conditional cumulative hazard functions on $[0, \tau] \times \mathcal{X}$. Any distribution $Q \in \mathcal{Q}$ can be characterized by

$$\begin{aligned} Q(\mathrm{d}t, j, \mathrm{d}x) = \{&S(t- \mid x)\Lambda_1(\mathrm{d}t \mid x)H(\mathrm{d}x)\}^{1_{\{j=1\}}} \\ &\{S(t- \mid x)\Lambda_2(\mathrm{d}t \mid x)H(\mathrm{d}x)\}^{1_{\{j=2\}}}, \end{aligned}$$

where $\Lambda_j \in \mathcal{M}$ for $j = 1, 2$ and $H$ is the marginal distribution of the covariates.

We consider the usual right-censored setting in which we observe data $O = (\tilde{T}, \tilde{D}, X)$, where $\tilde{T} = \min(T, C)$ for a right-censoring time $C$, $\Delta = \mathbb{1}\{T \leq C\}$, and $\tilde{D} = \Delta D$. Let $\mathcal{P}$ denote a set of probability measures on the sample space $\mathcal{O} = [0, \tau] \times \{0, 1, 2\} \times \mathcal{X}$ such that $O \sim P$ for some unknown $P \in \mathcal{P}$. We assume that the event time and the censoring time are conditionally independent given covariates, $T \perp\!\!\!\perp C \mid X$. This implies that any distribution $P \in \mathcal{P}$ is characterized by a distribution $Q \in \mathcal{Q}$ and a conditional cumulative hazard function for $C$ given $X$ [c.f., Begun et al., 1983, Gill et al., 1997]. We use $\Gamma \in \mathcal{M}$ to denote the conditional cumulative hazard function for censoring. We assume that $\Gamma(\cdot \mid x)$ is continuous for all $x$, and let $G(t \mid x) = \exp\{-\Gamma(t \mid x)\}$ denote the survival function of the conditional censoring distribution. In our setting with competing risks, this yields

$$
\begin{aligned}
P(\mathrm{d}t, j, \mathrm{d}x) &= \{G(t- \mid x)S(t- \mid x)\Lambda_1(\mathrm{d}t \mid x)H(\mathrm{d}x)\}^{\mathbb{1}\{j=1\}} \\
&\quad \{G(t- \mid x)S(t- \mid x)\Lambda_2(\mathrm{d}t \mid x)H(\mathrm{d}x)\}^{\mathbb{1}\{j=2\}} \\
&\quad \{G(t- \mid x)S(t- \mid x)\Gamma(\mathrm{d}t \mid x)H(\mathrm{d}x)\}^{\mathbb{1}\{j=0\}} \\
&= \{G(t- \mid x)Q(\mathrm{d}t, j, \mathrm{d}x)\}^{\mathbb{1}\{j\neq 0\}} \\
&\quad \{G(t- \mid x)S(t- \mid x)\Gamma(\mathrm{d}t \mid x)H(\mathrm{d}x)\}^{\mathbb{1}\{j=0\}}.
\end{aligned} \tag{2}
$$

Hence, we may write $\mathcal{P} = \{P_{Q,\Gamma} : Q \in \mathcal{Q}, \Gamma \in \mathcal{G}\}$ for some $\mathcal{G} \subset \mathcal{M}$. We also have

$$
P(\tilde{T} > t \mid X = x) = S(t \mid x)G(t \mid x) = \exp\{-\Lambda_1(t \mid x) - \Lambda_2(t \mid x) - \Gamma(t \mid x)\}.
$$

We further assume that there exists $\kappa < \infty$ such that $\Lambda_j(\tau- \mid x) < \kappa$, for $j \in \{1, 2\}$, and $\Gamma(\tau- \mid x) < \kappa$ for almost all $x \in \mathcal{X}$. Note that this implies that $G(\tau- \mid x)$ is bounded away from zero for almost all $x \in \mathcal{X}$. Under these assumptions, the conditional cumulative hazard functions $\Lambda_j$ and $\Gamma$ can be identified from $P$ by

$$
\Lambda_j(t \mid x) = \int_0^t \frac{P(\tilde{T} \in \mathrm{d}s, \tilde{D} = j \mid X = x)}{P(\tilde{T} \geq s \mid X = x)}, \tag{3}
$$

$$
\Gamma(t \mid x) = \int_0^t \frac{P(\tilde{T} \in \mathrm{d}s, \tilde{D} = 0 \mid X = x)}{P(\tilde{T} \geq s \mid X = x)}. \tag{4}
$$

Thus, we can consider $\Lambda_j$ and $\Gamma$ as operators which map from $\mathcal{P}$ to $\mathcal{M}$.

# 3 Super learning with right-censored survival data

A super learner estimates a parameter $\Psi$ which can be identified from the observed data distribution $P \in \mathcal{P}$. In this section, to introduce the concept of super learning, we simply consider estimation of the function $\Lambda_j$. The parameter $\Psi : \mathcal{P} \to \mathcal{M}$ is then identified via equation (3) by $\Psi(P) = \Lambda_j$.

As input to the super learner we need a sample $\mathcal{D}_n = \{O_i\}_{i=1}^n$ of i.i.d. observations from some unknown $P \in \mathcal{P}$ and a finite collection of candidate learners $\mathcal{A}$. Each learner $a \in \mathcal{A}$ is a map $a \colon \mathcal{O}^n \to \mathcal{M}$ which takes a data set as input and returns an estimate $a(\mathcal{D}_n) \in \mathcal{M}$ of $\Lambda_j$. In what follows, we use the short-hand notation $P[f] = \int f(o)P(\mathrm{d}o)$. A super learner evaluates the performance of $a \in \mathcal{A}$ using a loss function $L \colon \mathcal{M} \times \mathcal{O} \to \mathbb{R}_+$ by estimating the expected loss $P[L(a(\mathcal{D}_n), \cdot)]$ using cross-validation. Specifically, the expected loss of $a \in \mathcal{A}$ is estimated by splitting the data set $\mathcal{D}_n$ into $K$ disjoint approximately equally sized subsets $\mathcal{D}_n^1, \mathcal{D}_n^2, \dots, \mathcal{D}_n^K$ and then calculating the cross-validated loss

$$
\hat{R}_n(a; L) = \frac{1}{K} \sum_{k=1}^{K} \frac{1}{|\mathcal{D}_n^k|} \sum_{O_i \in \mathcal{D}_n^k} L\big(a(\mathcal{D}_n^{-k}), O_i\big), \quad \text{with} \quad \mathcal{D}_n^{-k} = \mathcal{D}_n \setminus \mathcal{D}_n^k.
$$

The subset $\mathcal{D}_n^{-k}$ is referred to as the $k$'th training sample, while $\mathcal{D}_n^k$ is referred to as the $k$'th test or hold-out sample. The discrete super learner is defined as

$$
\hat{a}_n = \operatorname*{argmin}_{a \in \mathcal{A}} \hat{R}_n(a; L).
$$

The final estimator of $\Psi(P) = \Lambda_j$ is then the selected learner applied to the full data set, i.e., $\hat{a}_n(\mathcal{D}_n)$. The oracle learner is defined as the learner that minimizes the average loss according to the data-generating distribution $P$, i.e.,

$$\tilde{a}_n = \operatorname*{argmin}_{a \in \mathcal{A}} \tilde{R}_n(a; L), \quad \text{with} \quad \tilde{R}_n(a; L) = \frac{1}{K} \sum_{k=1}^{K} P\big[L\big(a(\mathcal{D}_n^{-k}), \cdot\big)\big].$$

Note that both the discrete super learner and the oracle learner depend on the library of learners and on the number of folds $K$, and that the oracle learner is a function of the data and the unknown data-generating distribution. These dependencies are suppressed in the notation.

# 4 The state learner

The problem with most existing super learners for right-censored data is that they depend on a pre-specified estimator of the censoring distribution. The main idea of the state learner is to jointly use learners of $\Lambda_1$, $\Lambda_2$, and $\Gamma$, and the relations in equation (2), to learn a feature of the observed data distribution $P$. The discrete state learner ranks a tuple of learners of $(\Lambda_1, \Lambda_2, \Gamma)$ based on how well they jointly model the observed data. To formally introduce the state learner, we define the multi-state process

$$\eta(t) = \mathbb{1}\{\tilde{T} \leq t, \tilde{D} = 1\} + 2\,\mathbb{1}\{\tilde{T} \leq t, \tilde{D} = 2\} + 3\,\mathbb{1}\{\tilde{T} \leq t, \tilde{D} = 0\}, \quad \text{for} \quad t \in [0, \tau].$$

Note that at time $t$, we observe that each observation is in one of four mutually exclusive states (Figure 1). The conditional distribution of $\eta(t)$ given $X = x$ is determined by the

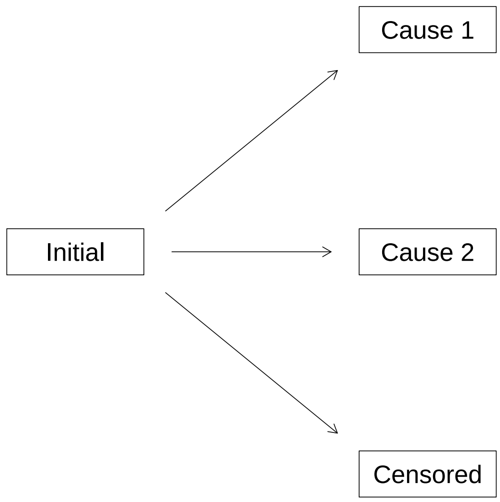


Figure 1: Illustration of the multi-state process $\eta$ used by the state learner. Note that 'censored' is a state, hence the process is always observed at any time.

function

$$F(t, k, x) = P(\eta(t) = k \mid X = x), \quad \text{for all} \quad t \in [0, \tau], k \in \{0, 1, 2, 3\}, x \in \mathcal{X}. \tag{5}$$

The function $F$ describes the conditional state occupation probabilities corresponding to the observed multi-state process $\eta$.

We propose to construct a super learner for $F$, i.e., the target of this super learner is $\Psi(P) = F$ where the parameter is identified through equation (5). Because each quadruple $(\Lambda_1, \Lambda_2, \Gamma, H)$ characterizes a $P \in \mathcal{P}$ which in turn determines $(F, H)$, a learner for $F$ can be constructed from learners of $\Lambda_1$, $\Lambda_2$, and $\Gamma$ as follows:

$$\begin{aligned}
F(t,1,x) &= P(\tilde{T} \leq t, \Delta = 1 \mid X = x) = \int_0^t e^{\{-\Lambda_1(s|x) - \Lambda_2(s|x) - \Gamma(s|x)\}} \Lambda_1(\mathrm{d}s \mid x), \\
F(t,2,x) &= P(\tilde{T} \leq t, \Delta = 2 \mid X = x) = \int_0^t e^{\{-\Lambda_1(s|x) - \Lambda_2(s|x) - \Gamma(s|x)\}} \Lambda_2(\mathrm{d}s \mid x), \\
F(t,3,x) &= P(\tilde{T} \leq t, \Delta = 0 \mid X = x) = \int_0^t e^{\{-\Lambda_1(s|x) - \Lambda_2(s|x) - \Gamma(s|x)\}} \Gamma(\mathrm{d}s \mid x), \\
F(t,0,x) &= P(\tilde{T} > t \mid X = x) = 1 - F(t,1,x) - F(t,2,x) - F(t,3,x).
\end{aligned} \tag{6}$$

The state learner requires three libraries of learners, $\mathcal{A}_1$, $\mathcal{A}_2$, and $\mathcal{B}$, where $\mathcal{A}_1$ and $\mathcal{A}_2$ contain learners of the conditional cause-specific cumulative hazard functions of the event time distribution $\Lambda_1$ and $\Lambda_2$, respectively, and $\mathcal{B}$ contains learners of the conditional cumulative hazard function of the censoring distribution. Based on the Cartesian product of libraries of learners for $(\Lambda_1, \Lambda_2, \Gamma)$ we construct a library $\mathcal{F}(\mathcal{A}_1, \mathcal{A}_2, \mathcal{B})$ of learners for $F$:

$$\mathcal{F}(\mathcal{A}_1, \mathcal{A}_2, \mathcal{B}) = \{\varphi_{a_1,a_2,b} : a_1 \in \mathcal{A}_1, a_2 \in \mathcal{A}_2, b \in \mathcal{B}\},$$

where in correspondance with the relations in equation (6),

$$\begin{aligned}
\varphi_{a_1,a_2,b}(\mathcal{D}_n)(t,1,x) &= \int_0^t e^{\{-a_1(\mathcal{D}_n)(s|x) - a_2(\mathcal{D}_n)(s|x) - b(\mathcal{D}_n)(s|x)\}} a_1(\mathcal{D}_n)(\mathrm{d}s \mid x), \\
\varphi_{a_1,a_2,b}(\mathcal{D}_n)(t,2,x) &= \int_0^t e^{\{-a_1(\mathcal{D}_n)(s|x) - a_2(\mathcal{D}_n)(s|x) - b(\mathcal{D}_n)(s|x)\}} a_2(\mathcal{D}_n)(\mathrm{d}s \mid x), \\
\varphi_{a_1,a_2,b}(\mathcal{D}_n)(t,3,x) &= \int_0^t e^{\{-a_1(\mathcal{D}_n)(s|x) - a_2(\mathcal{D}_n)(s|x) - b(\mathcal{D}_n)(s|x)\}} b(\mathcal{D}_n)(\mathrm{d}s \mid x), \\
\varphi_{a_1,a_2,b}(\mathcal{D}_n)(t,0,x) &= 1 - \sum_{j=1}^{3} \varphi_{a_1,a_2,b}(\mathcal{D}_n)(t,j,x).
\end{aligned}$$

To evaluate how well a function $F$ predicts the observed multi-state process we use the integrated Brier score $\bar{B}_\tau(F,O) = \int_0^\tau B_t(F,O)\,\mathrm{d}t$, where $B_t$ is the Brier score [Brier et al., 1950] at time $t \in [0,\tau]$,

$$B_t(F,O) = \sum_{j=0}^{3} \left(F(t,j,X) - \mathbb{1}\{\eta(t) = j\}\right)^2.$$

As described in Section 3, each learner $\varphi_{a_1,a_2,b}$ in the library $\mathcal{F}(\mathcal{A}_1, \mathcal{A}_2, \mathcal{B})$ is evaluated using the cross-validated loss,

$$\hat{R}_n(\varphi_{a_1,a_2,b}; \bar{B}_\tau) = \frac{1}{K} \sum_{k=1}^{K} \frac{1}{|\mathcal{D}_n^k|} \sum_{O_i \in \mathcal{D}_n^k} \bar{B}_\tau\left(\varphi_{a_1,a_2,b}(\mathcal{D}_n^{-k}), O_i\right),$$

and the discrete state learner is

$$\hat{\varphi}_n = \underset{(a_1,a_2,b) \in \mathcal{A}_1 \times \mathcal{A}_2 \times \mathcal{B}}{\operatorname{argmin}} \hat{R}_n(\varphi_{a_1,a_2,b}; \bar{B}_\tau).$$

## 5 Theoretical results for the state learner

In this section we establish theoretical guarantees for the state learner. Proposition 5.1 can be derived from the fact that the integrated Brier score (also called the continuous ranked probability score) is a strictly proper scoring rule [Gneiting and Raftery, 2007]. This implies that if we minimize the average loss of the integrated Brier score, we recover the parameters of the data-generating distribution. In particular, the oracle of a state learner will be consistent if the library of learners contains at least one learner that is consistent for estimation of $F$. Recall that the function $F$ implicitly depends on the data generating probability measure $P \in \mathcal{P}$ but that this was suppressed in the notation. We now make this dependence explicit by writing $F_0$ for the function which is obtained by substituting a specific $P_0 \in \mathcal{P}$ for $P$ in equation (6).

**Proposition 5.1.** *For $P_0 \in \mathcal{P}$ define*

$$F^* = \operatorname*{argmin}_F P_0[\bar{B}_\tau(F, \cdot)],$$

*where the minimum is taken over all $F$, such that $F$ is a conditional state occupation probability function for some measure $P$ as defined in equation (5). Then $F^*(t, j, \cdot) = F_0(t, j, \cdot)$ $H$-almost surely for any $j \in \{0, 1, 2, 3\}$ and almost any $t \in [0, \tau]$.*

*Proof.* See Appendix A. □

We establish a finite sample oracle result for the state learner. Our Corollary 5.2 is in essence a special case of a general cross-validation result by van der Vaart et al. [2006]. We assume that we split the data into equally sized folds, and for simplicity of presentation we take $n$ to be such that $|\mathcal{D}_n^{-k}| = n/K$ with $K$ fixed. We will allow the number of learners to grow with $n$ and write $\mathcal{F}_n = \mathcal{F}(\mathcal{A}_{1,n}, \mathcal{A}_{2,n}, \mathcal{B}_n)$ as short-hand notation and to emphasize the dependence on $n$. In the following we let $\|\cdot\|_{P_0}$ denote the norm

$$\|F\|_{P_0} = \left\{ \sum_{j=0}^{3} \int_{\mathcal{X}} \int_0^\tau F(t, j, x)^2 \,\mathrm{d}t H_0(\mathrm{d}x) \right\}^{1/2}. \tag{7}$$

**Corollary 5.2.** *For all $P_0 \in \mathcal{P}$, $n \in \mathbb{N}$, $k \in \{1, \dots, K\}$, and $\delta > 0$,*

$$\begin{aligned}\mathbb{E}_{P_0}\left[\|\hat{\varphi}_n(\mathcal{D}_n^{-k}) - F_0\|_{P_0}^2\right] \leq (1+2\delta)\,\mathbb{E}_{P_0}\left[\|\tilde{\varphi}_n(\mathcal{D}_n^{-k}) - F_0\|_{P_0}^2\right]& \\ + (1+\delta) 16 K\tau \left(13 + \frac{12}{\delta}\right) \frac{\log(1 + |\mathcal{F}_n|)}{n}&.\end{aligned}$$

*Proof.* See Appendix A. □

Corollary 5.2 has the following asymptotic consequences.

**Corollary 5.3.** *Assume that $|\mathcal{F}_n| = O(n^q)$, for some $q \in \mathbb{N}$ and that there exists a sequence $\varphi_n \in \mathcal{F}_n$, $n \in \mathbb{N}$, such that $\mathbb{E}_{P_0}\left[\|\varphi_n(\mathcal{D}_n^{-k}) - F_0\|_{P_0}^2\right] = O(n^{-\alpha})$, for some $\alpha \leq 1$.*

*(i) If $\alpha = 1$ then $\mathbb{E}_{P_0}\left[\|\hat{\varphi}_n(\mathcal{D}_n^{-k}) - F_0\|_{P_0}^2\right] = O(\log(n) n^{-1})$.*

*(ii) If $\alpha < 1$ then $\mathbb{E}_{P_0}\left[\|\hat{\varphi}_n(\mathcal{D}_n^{-k}) - F_0\|_{P_0}^2\right] = O(n^{-\alpha})$.*

*Proof.* See Appendix A. □

In Section 6 we demonstrate how the state learner can be used for targeted learning. A targeted estimator is typically obtained from estimators of the nuisance parameters $\Lambda_1$, $\Lambda_2$, and $\Gamma$. By equations (3) and (4) and the definition of $F$, we have

$$\Gamma(t,x) = \int_0^t \frac{F(\mathrm{d}s,3,x)}{F(s-,0,x)}, \quad \text{and} \quad \Lambda_j(t,x) = \int_0^t \frac{F(\mathrm{d}s,j,x)}{F(s-,0,x)}, \quad j \in \{1,2\}, \tag{8}$$

and thus a targeted estimator can also be obtained from an estimator of $F$ using equation (8). The key feature of a targeted estimator is that it is asymptotically equivalent to a sum of i.i.d. random variables plus a second order remainder term [Bickel et al., 1993, Fisher and Kennedy, 2021]. For our setting of competing risks, the remainder term is typically dominated by terms of the form

$$P\left[\int_0^\tau w_n(s,\cdot)\hat{M}_{1,n}(s \mid \cdot)\hat{M}_{2,n}(\mathrm{d}s \mid \cdot)\right], \tag{9}$$

where $(\hat{M}_{1,n}, \hat{M}_{2,n})$ is any of the nine combinations of $\hat{M}_{1,n} \in \{[\Gamma - \hat{\Gamma}_n], [\Lambda_1 - \hat{\Lambda}_{1,n}], [\Lambda_2 - \hat{\Lambda}_{2,n}]\}$ and $\hat{M}_{2,n} \in \{[\Gamma - \hat{\Gamma}_n], [\Lambda_1 - \hat{\Lambda}_{1,n}], [\Lambda_2 - \hat{\Lambda}_{2,n}]\}$, and $w_n$ is some data-dependent function with domain $[0,\tau] \times \mathcal{X}$ [van der Laan and Robins, 2003]. In particular, a targeted estimator will be asymptotically linear if the 'products' of the estimation errors $\hat{M}_{1,n}$ and $\hat{M}_{2,n}$ in equation (9) are $o_P(n^{-1/2})$. Proposition 5.4 states that if equation (9) holds for the a targeted estimator based on estimator $\hat{\Lambda}_{1,n}$, $\hat{\Lambda}_{2,n}$, and $\hat{\Gamma}_n$, then a similar product structure holds for a targeted estimator based on $\hat{F}_n$. We state the result for the special case that $\hat{M}_{1,n} = \Gamma - \hat{\Gamma}_n$ and $\hat{M}_{2,n} = \Lambda_1 - \hat{\Lambda}_{1,n}$, but similar results holds for any combinations of $\Gamma - \hat{\Gamma}_n$, $\Lambda_1 - \hat{\Lambda}_{1,n}$, and $\Lambda_2 - \hat{\Lambda}_{2,n}$.

**Proposition 5.4.** *Assume that $w(s,x) \leq c$, $F(s,0,x) \geq 1/c$ and $\hat{F}_n(s,0,x) \geq 1/c$ for some $c > 0$ for all $s \in [0,\tau]$ and $x \in \mathcal{X}$. Then there are real-valued uniformly bounded functions $w_n^a$, $w_n^b$, $w_n^c$, and $w_n^d$ with domain $[0,\tau]^2 \times \mathcal{X}$ such that*

$$\begin{aligned}
&P_0\left[\int_0^\tau w(s,\cdot)\left\{\Gamma_0(s,\cdot) - \hat{\Gamma}_n(s,\cdot)\right\}[\Lambda_0 - \hat{\Lambda}_n](\mathrm{d}s,\cdot)\right] \\
&= P_0\left[\int_0^\tau \int_0^s w_n^a(s,u,\cdot)[F_0 - \hat{F}_n](u-,0,\cdot)[F_0 - \hat{F}_n](s-,0,\cdot)F_0(\mathrm{d}u,2,\cdot)F_0(\mathrm{d}s,1,\cdot)\right] \\
&\quad + P_0\left[\int_0^\tau \int_0^s w_n^b(s,u,\cdot)[F_0 - \hat{F}_n](u-,0,\cdot)F_0(\mathrm{d}u,2,\cdot)[F_0 - \hat{F}_n](\mathrm{d}s,1,\cdot)\right] \\
&\quad + P_0\left[\int_0^\tau \int_0^s w_n^c(s,u,\cdot)[F_0 - \hat{F}_n](\mathrm{d}u,2,\cdot)[F_0 - \hat{F}_n](s-,0,\cdot)F_0(\mathrm{d}s,1,\cdot)\right] \\
&\quad + P_0\left[\int_0^\tau \int_0^s w_n^d(s,u,\cdot)[F_0 - \hat{F}_n](\mathrm{d}u,2,\cdot)[F_0 - \hat{F}_n](\mathrm{d}s,1,\cdot)\right].
\end{aligned}$$

*Proof.* See Appendix B. □

# 6 Targeted learning

Features of the observed data distribution $P \in \mathcal{P}$ are rarely of interest. We are instead interested in a parameter $\theta\colon \mathcal{Q} \to \Theta$ that expresses a property of the uncensored population governed by the measure $Q \in \mathcal{Q}$. The parameter space $\Theta$ can be a subset of $\mathbb{R}^d$ or a subset of a function space, such as $\mathcal{M}$. In subsection 6.1 we give a concrete example from causal inference where $\theta$ is the average treatment effect and $\Theta = [-1,1]$. Under the assumption of conditional independent censoring and positivity, $\theta$ is identifiable from $\mathcal{P}$ which means that there exists an operator $\Psi\colon \mathcal{P} \to \Theta$ such that $\theta(Q) = \Psi(P_{Q,\Gamma})$ for all $\Gamma \in \mathcal{M}$. By equation (2) this imply that we may write

$$\theta(Q) = \Psi(P) = \tilde{\Psi}^0(\Lambda_1, \Lambda_2, H),$$

for some operator $\tilde{\Psi}^0$. The state learner provides a ranking of tuples $(a_1, a_2, b) \in \mathcal{A}_1 \times \mathcal{A}_2 \times \mathcal{B}$, and we use $\hat{a}_{1,n}$, $\hat{a}_{2,n}$, and $\hat{b}_n$ to denote the learners corresponding to the discrete state learner $\hat{\varphi}_n$. Letting $H(\mathcal{D}_n)$ denote the empirical measure of $\{X_1, \dots, X_n\}$, we can obtain a simple plug-in an estimator of $\theta$ as

$$\hat{\Psi}^0(\mathcal{D}_n) = \tilde{\Psi}^0(\hat{a}_{1,n}(\mathcal{D}_n), \hat{a}_{2,n}(\mathcal{D}_n), H(\mathcal{D}_n)). \tag{10}$$

The asymptotic distribution of $\hat{\Psi}^0_{\tau,j}$ is difficult to analyze due to the model selection step involved when estimating the nuisance parameters $\Lambda_1$ and $\Lambda_2$. In addition, the estimator will typically have an asymptotic bias that vanishes at a sub-optimal rate. Using tools from semi-parametric efficiency theory, it is possible to construct a so-called targeted or debiased estimator with smaller asymptotic bias and an asymptotic distribution which we know how to estimate [Bickel et al., 1993, van der Laan and Rose, 2011, Chernozhukov et al., 2018]. A targeted estimator is based on the efficient influence function for the parameter $\tilde{\Psi}^0$ and relies on an estimator of $\Gamma$ in addition to estimators of $\Lambda_1$ and $\Lambda_2$. The efficient influence function is a $P$-zero mean and square integrable function indexed by the nuisance parameters $(\Lambda_1, \Lambda_2, \Gamma)$, which we denote by $\psi(\cdot; \Lambda_1, \Lambda_2, \Gamma)$. The name is justified because any regular asymptotically linear estimator that has $\psi$ as its influence function is efficient, meaning that it has smallest asymptotic variance among all regular asymptotically linear estimators.

An example of a targeted estimator is the one-step estimator, defined as

also TMLE, depending on implementation

$$\hat{\Psi}_{\mathrm{OS}}(\mathcal{D}_n) = \tilde{\Psi}^0(\hat{a}_{1,n}(\mathcal{D}_n), \hat{a}_{2,n}(\mathcal{D}_n), H(\mathcal{D}_n)) + \mathbb{P}_n[\psi(\cdot; \hat{a}_{1,n}(\mathcal{D}_n), \hat{a}_{2,n}(\mathcal{D}_n), \hat{b}_n(\mathcal{D}_n))], \tag{11}$$

where $\mathbb{P}_n$ is the empirical measure of a sample $\{O_i\}_{i=1}^n$. We can make the following asymptotic expansion of the one-step estimator [Pfanzagl and Wefelmeyer, 1982, Fisher and Kennedy, 2021, Kennedy, 2022],

$$\hat{\Psi}_{\mathrm{OS}}(\mathcal{D}_n) - \Psi(P) = \mathbb{P}_n[\psi(\cdot; \Lambda_1, \Lambda_2, \Gamma)] + \mathrm{Rem}(\hat{\Lambda}_{1,n}, \hat{\Lambda}_{2,n}, \hat{\Gamma}_n, P) + o_P(n^{-1/2}),$$

where the remainder term is typically such that [van der Laan and Robins, 2003, Rytgaard et al., 2022]

$$\mathrm{Rem}(\hat{\Lambda}_{1,n}, \hat{\Lambda}_{2,n}, \hat{\Gamma}_n, P) = \mathcal{O}_P\Big\{(\Lambda_1 - \hat{\Lambda}_{1,n})^2 + (\Lambda_2 - \hat{\Lambda}_{2,n})^2 + (\Gamma - \hat{\Gamma}_n)^2\Big\}. \tag{12}$$

Hence, under a suitable Donsker class regularity condition [Bickel et al., 1993, Kennedy, 2016], when equation (12) holds and the nuisance parameters $\Lambda_1$, $\Lambda_2$, and $\Gamma$ are consistently estimated at rate $o_P(n^{-1/4})$ then

$$\sqrt{n}(\hat{\Psi}_{\mathrm{OS}}(\mathcal{D}_n) - \Psi(P)) \rightsquigarrow \mathcal{N}(0, P[\psi(\cdot; \Lambda_1, \Lambda_2, \Gamma)^2]). \tag{13}$$

In particular, equation (13) and Slutsky's lemma imply that we can obtain asymptotically valid $(1-\alpha) \cdot 100\%$ confidence intervals by calculating

$$\left[\hat{\Psi}_{\mathrm{OS}}(\mathcal{D}_n) - q_{\alpha/2}\hat{\sigma}(\mathcal{D}_n),\ \hat{\Psi}_{\mathrm{OS}}(\mathcal{D}_n) + q_{\alpha/2}\hat{\sigma}(\mathcal{D}_n)\right],$$

where $q_\alpha$ is the $(1-\alpha)$-quantile or the standard normal distribution, and

$$\hat{\sigma}(\mathcal{D}_n) = \left(\mathbb{P}_n\left[\psi(\cdot; \hat{a}_{1,n}(\mathcal{D}_n), \hat{a}_{2,n}(\mathcal{D}_n), \hat{b}_n(\mathcal{D}_n))^2\right]\right)^{1/2}.$$

## 6.1 Cause-specific average treatment effect

We now demonstrate how the state learner and the general estimation strategy outlined above can be used to estimate a concrete parameter of interest in the competing risk setting. We assume that the covariate vector $X \in \mathbb{R}^p$ can be separated into a binary treatment indicator $A$ and a vector of potential confounders, $W \in \mathcal{W} \subset \mathbb{R}^{p-1}$. We abuse notation slightly by writing $\Lambda_j(t \mid x) = \Lambda_j(t \mid a, w)$ and $S(t \mid x) = S(t \mid a, w)$ when $x = (a, w)$. We use $\mu$ to denote the marginal distribution of $W$ and $\pi$ to denote the conditional probability of treatment,

$$\pi(w) = P(A = 1 \mid W = w).$$

We assume throughout that $\pi$ is uniformly bounded away from 0 and 1 on $\mathcal{W}$. As both $A$ and $W$ are fully observed for all individuals we can use a super learner to estimate $\pi$ [ref], and we denote this estimator by $\hat{\pi}_n$. We use the empirical measure of $\{W_1, \ldots, W_n\}$ to estimate $\mu$, and denote this estimator by the $\hat{\mu}_n$. As parameter of interest we consider the standardized difference in absolute risk of cause 1 before time $\tau$, add

what is this called?

$$\theta_\tau(Q) = \int_{\mathcal{W}} \left\{ \int_0^\tau S(s- \mid w, 1)\Lambda_1(\mathrm{d}s \mid w, 1) - \int_0^\tau S(s- \mid w, 0)\Lambda_1(\mathrm{d}s \mid w, 0) \right\} \mu(\mathrm{d}w).$$

Under a set of additional assumptions, $\theta_\tau$ can be given the causal interpretation

$$\theta_\tau(Q) = \mathbb{E}\left[T^{A=1} \leq \tau, D^{A=1} = 1\right] - \mathbb{E}\left[T^{A=0} \leq \tau, D^{A=0} = 1\right],$$

where $(T^A, D^A)$ denotes potential outcomes [Hernán and Robins, 2020]. In this case, the interpretation of $\theta_\tau$ is the difference in the average risk of cause 1 occurring before time $\tau$ in the population if everyone had been given treatment $(A = 1)$ compared to if no one had been given treatment.

Using equation (1), we may write $\theta_\tau(Q) = \tilde{\Psi}_t^0(\Lambda_1, \Lambda_2, \mu)$, where

$$\begin{aligned}\tilde{\Psi}_t^0(\Lambda_1, \Lambda_2, \mu) = &\int_{\mathcal{W}} \int_0^\tau e^{-\Lambda_1(s-|w,1)-\Lambda_2(s-|w,1)} \Lambda_1(\mathrm{d}s \mid w, 1)\mu(\mathrm{d}w) \\ &- \int_{\mathcal{W}} \int_0^\tau e^{-\Lambda_1(s-|w,0)-\Lambda_2(s-|w,0)} \Lambda_1(\mathrm{d}s \mid w, 0)\mu(\mathrm{d}w).\end{aligned} \tag{14}$$

The efficient influence function for the parameter $\tilde{\Psi}_\tau$ depends on the set $(\Lambda_1, \Lambda_2, \Gamma, \pi)$ of nuisance parameters, and is given by [Rytgaard and van der Laan, 2022] other refs?

$$\begin{aligned}&\psi_\tau(O; \Lambda_1, \Lambda_2, \Gamma, \pi) \\ &= \sum_{l=1}^2 \int_0^\tau \left[h_{\tau,l,1}(t, O; \Lambda_1, \Lambda_2, \Gamma, \pi) - h_{\tau,l,0}(t, O; \Lambda_1, \Lambda_2, \Gamma, \pi)\right] \\ &\qquad\qquad \times \Big(N_l(\mathrm{d}t) - \mathbb{1}\{\tilde{T} \geq t\}\Lambda_l(\mathrm{d}t \mid A, W)\Big) \\ &\quad - \sum_{a=0}^1 \left\{ (-1)^a \int_0^\tau e^{-\Lambda_1(s-|W,a)-\Lambda_2(s-|W,a)} \Lambda_1(\mathrm{d}s \mid W, a) \right\} - \tilde{\Psi}_t^0(\Lambda_1, \Lambda_2, \mu),\end{aligned} \tag{15}$$

where

$$\begin{aligned}h_{\tau,l,a}(t, O; \Lambda_1, \Lambda_2, \Gamma, \pi) = &\frac{\mathbb{1}\{A = a\} e^{\Gamma(t-|A,W)}}{\pi(W)^a (1 - \pi(W))^{1-a}} \\ &\times \left(1 - \frac{\int_t^\tau e^{-\Lambda_1(s-|A,W)-\Lambda_2(s-|A,W)} \Lambda_1(\mathrm{d}s \mid A, W)}{e^{-\Lambda_1(t|A,W)-\Lambda_2(t|A,W)}}\right)^{\mathbb{1}\{l=1\}} \\ &\times \left(- \frac{\int_t^\tau e^{-\Lambda_1(s-|A,W)-\Lambda_2(s-|A,W)} \Lambda_1(\mathrm{d}s \mid A, W)}{e^{-\Lambda_1(t|A,W)-\Lambda_2(t|A,W)}}\right)^{\mathbb{1}\{l=2\}}\end{aligned}$$

Equation (14) and (15) then allow us to construct a one-step estimator by using the definition given in equation (11), which gives the estimator

$$\tilde{\Psi}^0_{t,j}(\hat{a}_1(\mathcal{D}_n), \hat{a}_2(\mathcal{D}_n), \hat{\mu}_n(\mathcal{D}_n)) \tag{16}$$

## 7 Numerical experiments

We compare the state learner with two other discrete super learners and an oracle selector in a simulation study without a competing event. The two other super learners are based on inverse probability of censoring weighted (IPCW) Brier scores [Graf et al., 1999, Gerds and Schumacher, 2006], and we refer to these as IPCW super learners. These super learners depend on an estimator of the censoring distribution, and we consider IPCW super learners that use either the Kaplan-Meier estimator (IPCW(KM)) or a correctly specified Cox model (IPCW(Cox)) to estimate the censoring distribution. Both IPCW super learners are fitted using the R-package `riskRegression` [Gerds et al., 2023]. Each discrete super learner provides a learner for the cumulative hazard function for the outcome of interest, and from this a risk prediction model can be obtained. We measure performance of each super learner in terms of the Brier score of the provided risk prediction model at a specific time horizon. The Brier score is approximated using a large ($n = 20,000$) independent data set of uncensored data. The oracle selector uses the large data set of uncensored event times to select the learner with the lowest expected Brier score. The expected Brier score of the oracle selector serves as a lower benchmark value. For all super learners we split the data into five folds for training and testing.

Note that given a learner for the cumulative hazard function of the outcome event, we can typically use the same method to construct a learner of the cumulative hazard function of the censoring distribution. This would typically work by training the learner on the data set $\mathcal{D}^c_n$, where $\mathcal{D}^c_n = \{O^c_i\}^n_{i=1}$ with $O^c_i = (\tilde{T}_i, 1 - \Delta, X_i)$. When we say that we use a learner for the cumulative hazard function of the outcome to learn the cumulative hazard function of the censoring time, we mean that the learner is trained on $\mathcal{D}^c_n$.

**Scenario 1** We first generate a simple dataset to demonstrate that an IPCW super learner can perform poorly when the censoring model is misspecified. We start by generating a binomial baseline covariate $A$ with success probability 30%. We then generate outcome and censoring variables according to a Cox-Weibull distribution with hazard rates of approximately 0.5 and 2.5, respectively. We use a library $\mathcal{A}$ consisting of two learners, the (marginal) Kaplan-Meier estimator and the Kaplan-Meier estimator stratified on the binary covariate. For the state learner we use the same learners to construct the library $\mathcal{B}$.

The results are shown in Figure 2. We see that the IPCW super learner based on a misspecified censoring model (IPCW(KM)) has a larger Brier score than the other super learners and that its performance does not improve much with sample size. The performance of the state learner is comparable to or slightly better than the IPCW super learner based on a correctly pre-specified censoring model. The performance of both the IPCW(Cox) super learner and the state learner are close to that of the oracle for most sample sizes.

**Scenario 2** We next generate data according to a more complex distribution motivated from a real dataset in which censoring depends on the baseline covariates. We simulate data based on a prostate cancer study described in [Kattan et al., 2000]. The outcome of interest was the time to tumor recurrence, and five baseline covariates were used to predict

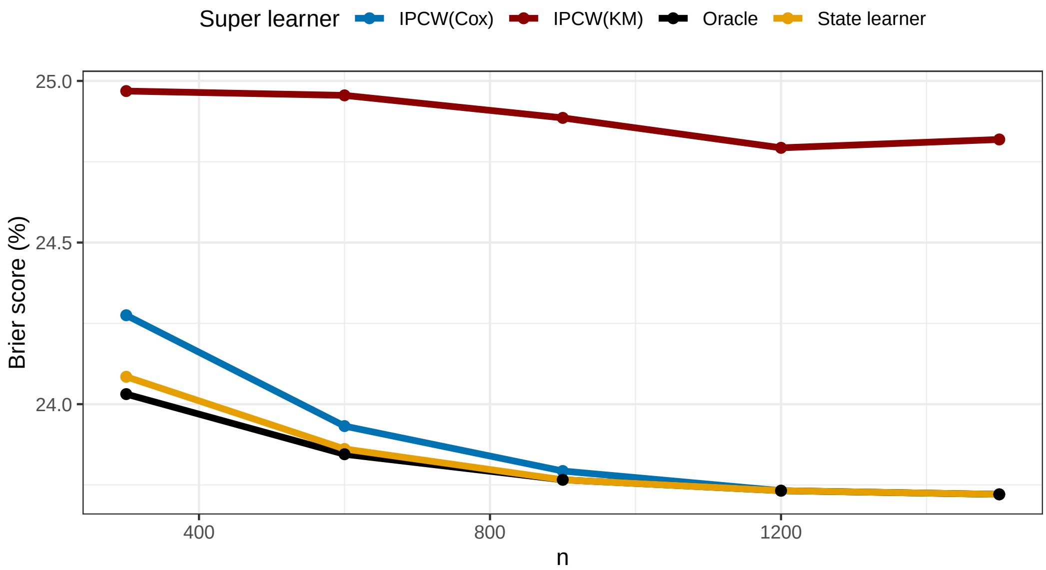


Figure 2: Results for scenario 1 of the simulation study. For the learner selected by each of the four discrete super learners, the Brier score calculated in a large independent data set without censoring is plotted against sample size. The results are based on 200 repetitions.

outcome: prostate-specific antigen (PSA, ng/mL), Gleason score sum (GSS, values between 6 and 10), radiation dose (RD), hormone therapy (HT, yes/no) and clinical stage (CS, six values). The study was designed such that a patient's radiation dose depended on when the patient entered the study [Gerds et al., 2013]. This in turn implied that the time of censoring depended on the radiation dose. The data were re-analyzed in [Gerds et al., 2013] where a sensitivity analysis was conducted based on simulated data. We use the same simulation setup, where event and censoring times are generated according to parametric Cox-Weibull models estimated from the original data, and the covariates are generated according to either marginal Gaussian normal or binomial distributions estimated from the original data [c.f., Gerds et al., 2013, Section 4.6]. We use the library consisting of the nine learners described in Table 1. For the state learner we use the same library to learn the censoring distribution.

The results are shown in Figure 3. We see that the Brier score of the IPCW super learner based on a misspecified censoring model (IPCW(KM)) decreased with sample size, but is larger than that of the other super learners for all sample sizes. The state learner and the IPCW super learner based on a correctly pre-specified censoring model demonstrate similar performance for all sample sizes. The performance of both the IPCW(Cox) super learner and the state learner approaches the benchmark provided by the oracle selector for large sample sizes.

# 8 Real data application

The original prostate cancer data analyzed by Kattan et al. [2000], which we introduced in Section 7, include a competing event in the form of death without tumor recurrence. To illustrate our method we fit the state learner to the original data set consisting of 1,042 patients. We consider death without tumor recurrence and recurrence of tumor as two competing events of interest. We include the fives learner `KM`, `Cox strata CS`, `Lasso`, `Elastic`,

| Family | Model | Description |
|---|---|---|
| Marginal | `KM` | The Kaplan-Meier estimator |
| Cox | `Cox` | All five covariates included with additive effects |
| | `Cox strata CS` | Cox model stratified on CS |
| | `Cox strata HT` | Cox model stratified on HT |
| | `Cox spline` | PSA and RD modeled with splines |
| Penalized Cox | `Lasso` | Cox model with $L_1$-norm penalty |
| | `Ridge` | Cox model with $L_2$-norm penalty |
| | `Elastic` | Cox model with $L_1$- and $L_2$-norm penalty |
| Random forest | `RF` | Random forest with 50 trees and default settings |

Table 1: Overview of the nine learners used in scenario 2 of the simulation study. The Kaplan-Meier estimator was fitted using the package `prodlim` [Gerds, 2019]. All Cox models included all five covariates in the model and were fitted using the package `survival` [Therneau, 2022]. All penalized Cox models included all five covariates as linear predictors and were fitted using the package `glmnet` [Simon et al., 2011, Friedman et al., 2010]. The random forest was fitted with the package `randomForestSRC` [Ishwaran and Kogalur, 2023].

and `RF` which are described in Table 1. We use the same library of learners to learn $\Lambda_1$, $\Lambda_2$, and $\Gamma$. In this case, $\Lambda_1$ denotes the cause-specific cumulative hazard function of tumor recurrence, and $\Lambda_2$ denotes the cause-specific cumulative hazard function of death without tumor recurrence.

This gives a library consisting of $5^3 = 125$ learners for the conditional state occupation probability function $F$ defined in equation (5). We use five folds for training and testing the models, and we repeat training and evaluation five times with different splits. The integrated Brier score for all learners are shown in Figure 4, and the top 10 combinations of learners are displayed in Table 2. We see that the prediction performance is mostly affected by the choice of learner for the censoring distribution. Several combinations of learners give similar performance as measured by the integrated Brier score, as long as a random forest is used to model the censoring distribution.

| Tumor learner | Death learner | Censoring learner | Integrated Brier score |
|---|---|---|---|
| `Elastic` | `Elastic` | `RF` | $7.03 \pm 0.02$ |
| `Elastic` | `KM` | `RF` | $7.03 \pm 0.02$ |
| `Lasso` | `Elastic` | `RF` | $7.04 \pm 0.02$ |
| `Lasso` | `KM` | `RF` | $7.04 \pm 0.02$ |
| `Elastic` | `Lasso` | `RF` | $7.04 \pm 0.02$ |
| `Cox strata CT` | `Elastic` | `RF` | $7.04 \pm 0.03$ |
| `Lasso` | `Lasso` | `RF` | $7.04 \pm 0.02$ |
| `Cox strata CT` | `Lasso` | `RF` | $7.04 \pm 0.03$ |
| `Cox strata CT` | `KM` | `RF` | $7.04 \pm 0.03$ |
| `Elastic` | `RF` | `RF` | $7.04 \pm 0.02$ |

Table 2: The 10 best performing models in terms of integrated Brier score. The reported standard errors are based on five repetitions using different splits. The models are described in Table 1. We refer to learners of $\Lambda_1$, $\Lambda_2$, and $\Gamma$ as 'Tumor learner', 'Death leaner', and 'Censoring learner', respectively.

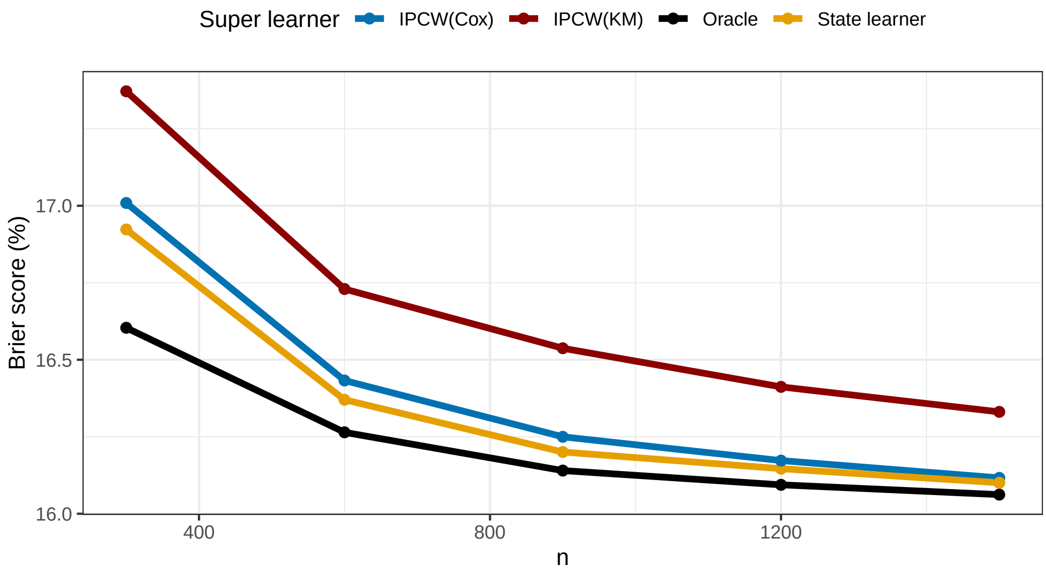


Figure 3: Results for scenario 2 of the simulation study. For the learner selected by each of the four discrete super learners, the Brier score calculated in a large independent data set without censoring is plotted against sample size. The results are based on 200 repetitions.

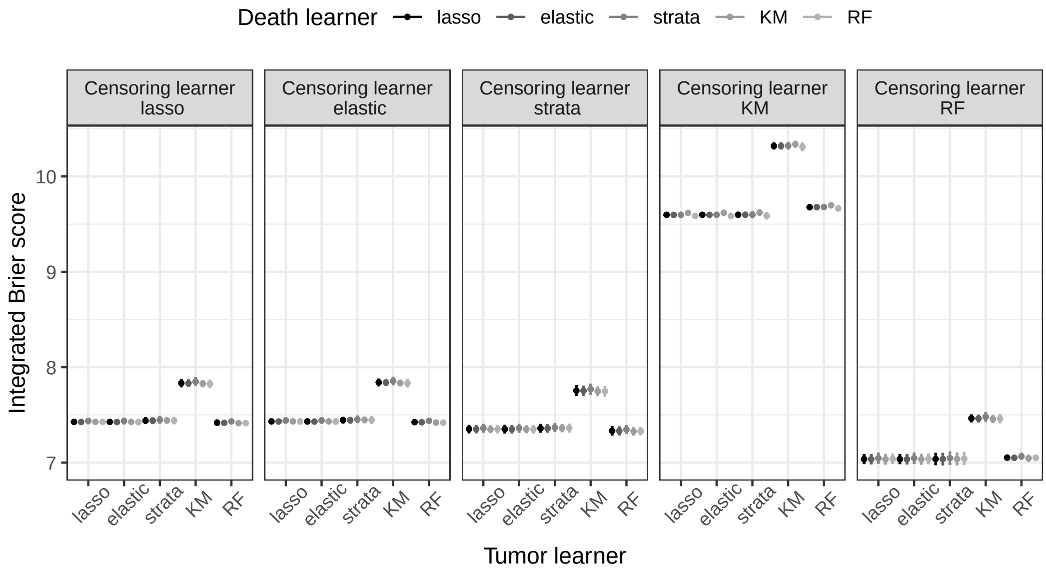


Figure 4: The results of applying the 125 combinations of learners to the prostate cancer data set. The learners are `KM` (KM), `Cox strata CS` (strata), `Lasso` (lasso), `Elastic` (elastic), and `RF` (RF) as described Table 1. The error bars are based on five repetitions using different splits. We refer to learners of $\Lambda_1$, $\Lambda_2$, and $\Gamma$ as 'Tumor learner', 'Death learner', and 'Censoring learner', respectively.

# 9 Discussion

We have proposed a new super learner that can be used with right-censored data and competing events. In this section, we compare our proposal to existing super learners and discuss avenues for further research.

## 9.1 Existing super learners for right-censored data

Machine learning based on right-censored data commonly use the negative partial log-likelihood as loss function [e.g., Li et al., 2016, Yao et al., 2017, Lee et al., 2018, Katzman et al., 2018, Gensheimer and Narasimhan, 2019, Lee et al., 2021, Kvamme and Borgan, 2021]. However, this loss function is unsuited for super learning, because many canonical survival learners (e.g., the Kaplan-Meier estimator, random survival forest, and semi-parametric Cox models) provide cumulative hazard functions that are piece-wise constant in the time argument, and hence assign zero probability to event times not observed in the training data. This implies that when data are observed in continuous time, any of these learners will almost surely have infinite loss in any independent hold-out sample according to the negative partial log-likelihood loss. When a proportional hazards model is assumed, the baseline hazard function can be profiled out of the likelihood to give a new partial log-likelihood loss [Cox, 1972], which has been suggested as a loss function for super learning [Golmakani and Polley, 2020, Verweij and van Houwelingen, 1993]. While this allows the library of learners to include Cox' proportional hazard models, the drawback is that the library is in fact *only* allowed to include these models. The advantage of the state learner is that it does not require evaluation of the density of $F(\cdot, j, x)$ and does not assume a particular semi-parametric structure for $\Lambda_j$ but can be used with any library of learners.

Another approach for super learning with right-censored data is to use an inverse probability of censoring weighted (IPCW) loss function [Graf et al., 1999, van der Laan and Dudoit, 2003, Molinaro et al., 2004, Keles et al., 2004, Hothorn et al., 2006, Gerds and Schumacher, 2006, Gonzalez Ginestet et al., 2021]. An IPCW loss function is attractive because the associated risk does not depend on the censoring distribution but describes a feature of the population of interested governed by the measure $Q \in \mathcal{Q}$. Similar results can be obtained using censoring unbiased transformations [Fan and Gijbels, 1996, Steingrimsson et al., 2019] or pseudo-values [Andersen et al., 2003, Mogensen and Gerds, 2013, Sachs et al., 2019]. All these methods rely on an estimator of the censoring distribution, and their drawback is that this estimator has to be pre-specified. When the data-generating mechanism is complex and not well-understood, pre-specification of the censoring distribution is a challenge. The advantage of using the state learner is that a censoring distribution need not be pre-specified but is selected automatically based on the provided library $\mathcal{B}$.

To the best of our knowledge, the only existing attempt at avoiding the need to pre-specify a censoring model is a recent proposal suggested independently by Han et al. [2021] and Westling et al. [2021]. The authors do not consider competing risks but suggest to iterate between learning $\Lambda$ and $\Gamma$ using IPCW loss functions and select the final learner when the iterative procedure has converged. No general theoretical guarantees exist for this procedure, but it would be interesting to compare its performance to that of the state learner in a simulation study.

### 9.2 A performance measure of interest

A major advantage of the state learner is that performance of each combination of learners is measured in terms of observable quantities. This means that no additional nuisance parameters need to be estimated to evaluate the loss. The drawback with this approach is that we are rarely interested in features of the observed data distribution when the data are right-censored. The finite sample oracle inequality in Corollary 5.2 concerns the function $F$, which is a feature of $P \in \mathcal{P}$, while what we are typically interested in is $\Lambda_j$ or $S$, which are features of $Q \in \mathcal{Q}$. We emphasize that while the state learner provides us with estimates of $\Lambda_j$ and $\Gamma$ based on libraries $\mathcal{A}_j$ and $\mathcal{B}$, performance are not assessed directly for these parameters, but only jointly for estimation of the parameter $F$. For settings without a competing risk, our numerical studies suggest that measuring performance with respect to estimation of $F$ also leads to good performance for estimation of $S$. Further research on this topic, both numerical and theoretical, is warranted. mention this?

[Double robustness maybe lacking]

### 9.3 Implementation

Our proposed super learner can be implemented with a broad library of learners using existing software, for instance the R-package `riskRegression` [Gerds et al., 2023]. Furthermore, while the library $\mathcal{F}(\mathcal{A}_1, \mathcal{A}_2, \mathcal{B})$ consists of $|\mathcal{A}_1||\mathcal{A}_2||\mathcal{B}|$ learners, as long as we have sufficient memory we need only fit $|\mathcal{A}_1|+|\mathcal{A}_2|+|\mathcal{B}|$ learners in each fold. To evaluate the performance of each learner we need to perform $|\mathcal{A}_1||\mathcal{A}_2||\mathcal{B}|$ operations to calculate the integrated Brier score in each hold-out sample, one for each combination of the fitted models, but these operations are often negligible compared to fitting the models. Hence the state learner is essentially no more computationally demanding than any procedure that uses super learning to learn $\Lambda_1$, $\Lambda_2$, and $\Gamma$ separately. While our proposal is based on constructing the library $\mathcal{F}$ from libraries for learning $\Lambda_1$, $\Lambda_2$, and $\Gamma$, it would also be of interest to consider learners that estimate $F$ directly.

In our numerical studies, we only considered learners of $\Lambda_j$ and $\Gamma$ that provide cumulative hazard functions which are piece-wise constant in the time argument. This simplifies the calculation of $F$ as the integrals in equation (6) reduce to sums. When $\Lambda_j$ or $\Gamma$ are absolutely continuous in the time argument, calculating $F$ is more involved, but we expect that a good approximation can be achieved by discretization. In the future, we intend to investigate the performance of the state learner when using a broader library of learners in more comprehensive simulation studies.

## A Theoretical guarantees for the state learner

In this section we provide proofs of the results stated in Section 5.

Define $\bar{B}_{\tau,0}(F, o) = \bar{B}_\tau(F, o) - \bar{B}_\tau(F_0, o)$ and $R_0(F) = P_0[\bar{B}_{\tau,0}(F, \cdot)]$.

**Lemma A.1.** *$R_0(F) = \|F - F_0\|^2_{P_0}$, where $\|\cdot\|_{P_0}$ is defined in equation (7).*

*Proof.* For any $t \in [0, \tau]$ and $k \in \{0, 1, 2\}$ we have

$$\begin{aligned}
&\mathbb{E}_{P_0}\left[(F(t,k,X) - \mathbb{1}\{\eta(t) = k\})^2\right] \\
&= \mathbb{E}_{P_0}\left[(F(t,k,X) - F_0(t,k,X) + F_0(t,k,X) - \mathbb{1}\{\eta(t) = k\})^2\right] \\
&= \mathbb{E}_{P_0}\left[(F(t,k,X) - F_0(t,k,X))^2\right] + \mathbb{E}_{P_0}\left[(F_0(t,k,X) - \mathbb{1}\{\eta(t) = k\})^2\right] \\
&\quad + 2\,\mathbb{E}_{P_0}\left[(F(t,k,X) - F_0(t,k,X))(F_0(t,k,X) - \mathbb{1}\{\eta(t) = k\})\right] \\
&= \mathbb{E}_{P_0}\left[(F(t,k,X) - F_0(t,k,X))^2\right] + \mathbb{E}_{P_0}\left[(F_0(t,k,X) - \mathbb{1}\{\eta(t) = k\})^2\right],
\end{aligned}$$

where the last equality follows from the tower property. Hence, using Fubini, we have

$$P[\bar{B}_\tau(F, \cdot)] = \|F - F_0\|_{P_0}^2 + P_0[\bar{B}_\tau(F_0, \cdot)]. \qquad \square$$

*Proof of Proposition 5.1.* The result follows from Lemma A.1. $\square$

In the following, let $\Theta$ denote the function space consisting of all conditional state occupation probability functions for some measure $P$ as defined in equation (5).

*Proof of Corollary 5.2.* First note that minimising the loss $\bar{B}_\tau$ is equivalent to minimising the loss $\bar{B}_{\tau,0}$, so the discrete super learner and oracle according to $\bar{B}_\tau$ and $\bar{B}_{\tau,0}$ are identical. By Lemma A.1, $R_0(F) \geq 0$ for any $F \in \Theta$, and so using Theorem 2.3 from [van der Vaart et al., 2006] with $p = 1$, we have that for all $\delta > 0$,

$$\begin{aligned}
\mathbb{E}_{P_0}\left[R_0(\hat{\varphi}_n(\mathcal{D}_n^{-k}))\right] \leq &(1 + 2\delta)\,\mathbb{E}_{P_0}\left[R_0(\tilde{\varphi}_n(\mathcal{D}_n^{-k}))\right] \\
&+ (1 + \delta)\frac{16K}{n}\log(1 + |\mathcal{F}_n|) \sup_{F \in \Theta}\left\{M(F) + \frac{v(F)}{R_0(F)}\left(\frac{1}{\delta} + 1\right)\right\}
\end{aligned}$$

where for each $F \in \Theta$, $(M(F), v(F))$ is some Bernstein pair for the function $o \mapsto \bar{B}_{\tau,0}(F, o)$. As $\bar{B}_{\tau,0}(F, \cdot)$ is uniformly bounded by $\tau$ for any $F \in \Theta$, it follows from section 8.1 in [van der Vaart et al., 2006] that $(\tau, 1.5P_0[\bar{B}_{\tau,0}(F, \cdot)^2])$ is a Bernstein pair for $\bar{B}_{\tau,0}(F, \cdot)$. Now, for any $a, b, c \in \mathbb{R}$ we have

$$\begin{aligned}
(a - c)^2 - (b - c)^2 &= (a - b + b - c)^2 - (b - c)^2 \\
&= (a - b)^2 + (b - c)^2 + 2(b - c)(a - b) - (b - c)^2 \\
&= (a - b)\left\{(a - b) + 2(b - c)\right\} \\
&= (a - b)\left\{a + b - 2c\right\},
\end{aligned}$$

so using this with $a = F(t, k, x)$, $b = F_0(t, k, x)$, and $c = \mathbb{1}\{\eta(t) = k\}$, we have by Jensen's inequality

$$\begin{aligned}
&P_0[\bar{B}_{\tau,0}(F, \cdot)^2] \\
&\leq 2\tau\,\mathbb{E}_{P_0}\left[\sum_{k=0}^{3}\int_0^\tau \left\{(F(t,k,X) - \mathbb{1}\{\eta(t) = k\})^2 - (F_0(t,k,X) - \mathbb{1}\{\eta(t) = k\})^2\right\}^2 \mathrm{d}t\right] \\
&= 2\tau\,\mathbb{E}_{P_0}\left[\sum_{k=0}^{3}\int_0^\tau (F(t,k,X) - F_0(t,k,X))^2 \right. \\
&\qquad\qquad \left. \times \left\{F(t,k,X) + F_0(t,k,X) - 2\mathbb{1}\{\eta(t) = k\}\right\}^2 \mathrm{d}t\right] \\
&\leq 8\tau\,\mathbb{E}_{P_0}\left[\sum_{k=0}^{3}\int_0^\tau (F(t,k,X) - F_0(t,k,X))^2\, \mathrm{d}t\right]. \\
&= 8\tau\|F - F_0\|_{P_0}^2.
\end{aligned}$$

Thus when $v(F) = 1.5 P_0[\bar{B}_{\tau,0}(F,\cdot)^2]$ we have by Lemma A.1

$$\frac{v(F)}{R_0(F)} = 1.5 \frac{P_0[\bar{B}_{\tau,0}(F,\cdot)^2]}{P_0[\bar{B}_{\tau,0}(F,\cdot)]} \leq 12\tau,$$

and so using the Bernstein pairs $(\tau, 1.5P_0[\bar{B}_{\tau,0}(F,\cdot)^2])$ we have

$$\sup_{F\in\Theta} \left\{ M(F) + \frac{v(F)}{R_0(F)} \left( \frac{1}{\delta} + 1 \right) \right\} \leq \tau \left( 13 + \frac{12}{\delta} \right),$$

For all $\delta > 0$ we thus have

$$\begin{aligned} \mathbb{E}_{P_0}\left[ R_0(\hat{\varphi}_n(\mathcal{D}_n^{-k})) \right] \leq & (1+2\delta)\, \mathbb{E}_{P_0}\left[ R_0(\tilde{\varphi}_n(\mathcal{D}_n^{-k})) \right] \\ & + (1+\delta) \log(1+|\mathcal{F}_n|) \tau \frac{16K}{n} \left( 13 + \frac{12}{\delta} \right), \end{aligned}$$

and then the final result follows from Lemma A.1. □

*Proof of Corollary 5.3.* By definition of the oracle and Lemma A.1, $\mathbb{E}_{P_0}\left[ \|\tilde{\varphi}_n(\mathcal{D}_n^{-k}) - F_0\|_{P_0}^2 \right] \leq \mathbb{E}_{P_0}\left[ \|\varphi_n(\mathcal{D}_n^{-k}) - F_0\|_{P_0}^2 \right]$ for all $n \in \mathbb{N}$. The results then follows from Corollary 5.2. □

## B The state learner with targeted learning

In this section show that a product structure is preserved when the estimator $\bar{\Psi}(\hat{F}_n, \hat{H}_n)$ is used instead of $\tilde{\Psi}(\hat{\Lambda}_n, \hat{\Gamma}_n, \hat{H}_n)$.

*Proof of Proposition 5.4.* For notational convenience we suppress $X$ in the following. The final result can be obtained by adding the argument $X$ to all functions and averaging. We use the relations from equation (8) to write

$$\begin{aligned}
&\int_0^\tau w(s) \left\{ \Gamma(s) - \hat{\Gamma}_n(s) \right\} [\Lambda - \hat{\Lambda}_n](\mathrm{d}s) \\
&= \int_0^\tau w(s) \left\{ \int_0^s \frac{F(\mathrm{d}u,2)}{F(u-,0)} - \int_0^s \frac{\hat{F}_n(\mathrm{d}u,2)}{\hat{F}_n(u-,0)} - \right\} \left[ \frac{F(\mathrm{d}s,1)}{F(s-,0)} - \frac{\hat{F}_n(\mathrm{d}s,1)}{\hat{F}_n(s-,0)} \right] \\
&= \int_0^\tau w(s) \Bigg\{ \int_0^s \left( \frac{1}{F(u-,0)} - \frac{1}{\hat{F}_n(u-,0)} \right) F(\mathrm{d}u,2) \\
&\qquad\qquad + \int_0^s \frac{1}{\hat{F}_n(u-,0)} \left[ F(\mathrm{d}u,2) - \hat{F}_n(\mathrm{d}u,2) \right] \Bigg\} \\
&\qquad \times \left[ \left( \frac{1}{F(s-,0)} - \frac{1}{\hat{F}_n(s-,0)} \right) F(\mathrm{d}s,1) + \frac{1}{\hat{F}_n(s-,0)} \Big( F(\mathrm{d}s,1) - \hat{F}_n(\mathrm{d}s,1) \Big) \right] \\
&= \int_0^\tau \int_0^s w(s) \left( \frac{1}{F(u-,0)} - \frac{1}{\hat{F}_n(u-,0)} \right) \left( \frac{1}{F(s-,0)} - \frac{1}{\hat{F}_n(s-,0)} \right) F(\mathrm{d}u,2) F(\mathrm{d}s,1) \\
&\quad + \int_0^\tau \int_0^s w(s) \left( \frac{1}{F(u-,0)} - \frac{1}{\hat{F}_n(u-,0)} \right) \frac{F(\mathrm{d}u,2)}{\hat{F}_n(u-,0)} \Big( F(\mathrm{d}s,1) - \hat{F}_n(\mathrm{d}s,1) \Big) \\
&\quad + \int_0^\tau \int_0^s \frac{w(s)}{\hat{F}_n(u-,0)} \left[ F(\mathrm{d}u,2) - \hat{F}_n(\mathrm{d}u,2) \right] \left( \frac{1}{F(s-,0)} - \frac{1}{\hat{F}_n(s-,0)} \right) F(\mathrm{d}s,1) \\
&\quad + \int_0^\tau \int_0^s \frac{w(s)}{\hat{F}_n(u-,0)} \left[ F(\mathrm{d}u,2) - \hat{F}_n(\mathrm{d}u,2) \right] \frac{1}{\hat{F}_n(s-,0)} \Big( F(\mathrm{d}s,1) - \hat{F}_n(\mathrm{d}s,1) \Big).
\end{aligned}$$

Consider the first term on the right hand side. Defining

$$w_n^*(t) = \Big(F(t-,0) - \hat{F}_n(t-,0)\Big)\left(\frac{1}{F(t-,0)} - \frac{1}{\hat{F}_n(t-,0)}\right),$$

we can write

$$\begin{aligned}
&\int_0^\tau \int_0^s w(s)\left(\frac{1}{F(u-,0)} - \frac{1}{\hat{F}_n(u-,0)}\right)\left(\frac{1}{F(s-,0)} - \frac{1}{\hat{F}_n(s-,0)}\right) F(\mathrm{d}u,2)F(\mathrm{d}s,1)\\
&= \int_0^\tau \int_0^s w(s)w_n^*(u)\Big(F(u-,0) - \hat{F}_n(u-,0)\Big)\\
&\qquad\qquad \times w_n^*(s)\Big(F(s-,0) - \hat{F}_n(s-,0)\Big) F(\mathrm{d}u,2)F(\mathrm{d}s,1)\\
&= \int_0^\tau \int_0^s w_n^a(s,u)\Big(F(u-,0) - \hat{F}_n(u-,0)\Big)\Big(F(s-,0) - \hat{F}_n(s-,0)\Big) F(\mathrm{d}u,2)F(\mathrm{d}s,1),
\end{aligned}$$

where we have defined $w_n^a(s,u) = w(s)w_n^*(s)w_n^*(u)$. By assumption, $w_n^a(s,u)$ is uniformly bounded. The same approach can be applied to the three remaining terms which gives the result. □

# References


P. K. Andersen, J. P. Klein, and S. Rosthøj. Generalised linear models for correlated pseudo-observations, with applications to multi-state models. *Biometrika*, 90(1):15–27, 2003.

P. K. Andersen, O. Borgan, R. D. Gill, and N. Keiding. *Statistical models based on counting processes*. Springer Science & Business Media, 2012.

J. M. Begun, W. J. Hall, W.-M. Huang, and J. A. Wellner. Information and asymptotic efficiency in parametric-nonparametric models. *The Annals of Statistics*, 11(2):432–452, 1983.

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

L. Breiman. Stacked regressions. *Machine learning*, 24(1):49–64, 1996.

G. W. Brier et al. Verification of forecasts expressed in terms of probability. *Monthly weather review*, 78(1):1–3, 1950.

V. Chernozhukov, D. Chetverikov, M. Demirer, E. Duflo, C. Hansen, W. Newey, and J. Robins. Double/debiased machine learning for treatment and structural parameters, 2018.

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

J. Fan and I. Gijbels. *Local polynomial modelling and its applications*. Routledge, 1996.

A. Fisher and E. H. Kennedy. Visually communicating and teaching intuition for influence functions. *The American Statistician*, 75(2):162–172, 2021.

J. Friedman, T. Hastie, and R. Tibshirani. Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software*, 33(1):1–22, 2010. doi: 10.18637/jss.v033.i01. URL https://www.jstatsoft.org/v33/i01/.

M. F. Gensheimer and B. Narasimhan. A scalable discrete-time survival model for neural networks. *PeerJ*, 7:e6257, 2019.

T. A. Gerds. *prodlim: Product-Limit Estimation for Censored Event History Analysis*, 2019. URL https://CRAN.R-project.org/package=prodlim. R package version 2019.11.13.

T. A. Gerds and M. W. Kattan. *Medical risk prediction models: with ties to machine learning*. CRC Press, 2021.

T. A. Gerds and M. Schumacher. Consistent estimation of the expected Brier score in general survival models with right-censored event times. *Biometrical Journal*, 48(6):1029–1040, 2006.

T. A. Gerds, M. W. Kattan, M. Schumacher, and C. Yu. Estimating a time-dependent concordance index for survival prediction models with covariate dependent censoring. *Statistics in medicine*, 32(13):2173–2184, 2013.

T. A. Gerds, J. S. Ohlendorff, and B. Ozenne. *riskRegression: Risk Regression Models and Prediction Scores for Survival Analysis with Competing Risks*, 2023. URL https://CRAN.R-project.org/package=riskRegression. R package version 2023.03.22.

R. D. Gill, M. J. van der Laan, and J. M. Robins. Coarsening at random: Characterizations, conjectures, counter-examples. In *Proceedings of the First Seattle Symposium in Biostatistics*, pages 255–294. Springer, 1997.

T. Gneiting and A. E. Raftery. Strictly proper scoring rules, prediction, and estimation. *Journal of the American statistical Association*, 102(477):359–378, 2007.

M. K. Golmakani and E. C. Polley. Super learner for survival data prediction. *The International Journal of Biostatistics*, 16(2):20190065, 2020.

P. Gonzalez Ginestet, A. Kotalik, D. M. Vock, J. Wolfson, and E. E. Gabriel. Stacked inverse probability of censoring weighted bagging: A case study in the infcarehiv register. *Journal of the Royal Statistical Society Series C: Applied Statistics*, 70(1):51–65, 2021.

E. Graf, C. Schmoor, W. Sauerbrei, and M. Schumacher. Assessment and comparison of prognostic classification schemes for survival data. *Statistics in medicine*, 18(17-18):2529–2545, 1999.

X. Han, M. Goldstein, A. Puli, T. Wies, A. Perotte, and R. Ranganath. Inverse-weighted survival games. *Advances in Neural Information Processing Systems*, 34, 2021.

M. Hernán and J. Robins. *Causal Inference: What If*. Boca Raton: Chapman & Hall/CRC, 2020.

N. L. Hjort. On inference in parametric survival data models. *International Statistical Review/Revue Internationale de Statistique*, pages 355–387, 1992.

T. Hothorn, P. Bühlmann, S. Dudoit, A. Molinaro, and M. J. van der Laan. Survival ensembles. *Biostatistics*, 7(3):355–373, 2006.

H. Ishwaran and U. Kogalur. *Fast Unified Random Forests for Survival, Regression, and Classification (RF-SRC)*, 2023. URL https://cran.r-project.org/package=randomForestSRC. R package version 3.2.2.

M. W. Kattan, M. J. Zelefsky, P. A. Kupelian, P. T. Scardino, Z. Fuks, and S. A. Leibel. Pretreatment nomogram for predicting the outcome of three-dimensional conformal radiotherapy in prostate cancer. *Journal of clinical oncology*, 18(19):3352–3359, 2000.

J. L. Katzman, U. Shaham, A. Cloninger, J. Bates, T. Jiang, and Y. Kluger. Deepsurv: personalized treatment recommender system using a Cox proportional hazards deep neural network. *BMC medical research methodology*, 18(1):1–12, 2018.

S. Keles, M. van der Laan, and S. Dudoit. Asymptotically optimal model selection method with right censored outcomes. *Bernoulli*, 10(6):1011–1037, 2004.

E. H. Kennedy. Semiparametric theory and empirical processes in causal inference. In *Statistical causal inferences and their applications in public health research*, pages 141–167. Springer, 2016.

E. H. Kennedy. Semiparametric doubly robust targeted double machine learning: a review. *arXiv preprint arXiv:2203.06469*, 2022.

H. Kvamme and Ø. Borgan. Continuous and discrete-time survival prediction with neural networks. *Lifetime Data Analysis*, 27(4):710–736, 2021.

C. Lee, W. Zame, J. Yoon, and M. van der Schaar. Deephit: A deep learning approach to survival analysis with competing risks. In *Proceedings of the AAAI conference on artificial intelligence*, volume 32, 2018.

D. K. Lee, N. Chen, and H. Ishwaran. Boosted nonparametric hazards with time-dependent covariates. *Annals of Statistics*, 49(4):2101, 2021.

Y. Li, K. S. Xu, and C. K. Reddy. Regularized parametric regression for high-dimensional survival analysis. In *Proceedings of the 2016 SIAM International Conference on Data Mining*, pages 765–773. SIAM, 2016.

U. B. Mogensen and T. A. Gerds. A random forest approach for competing risks based on pseudo-values. *Statistics in medicine*, 32(18):3102–3114, 2013.

A. M. Molinaro, S. Dudoit, and M. J. van der Laan. Tree-based multivariate regression and density estimation with right-censored data. *Journal of Multivariate Analysis*, 90(1): 154–177, 2004.

J. Pfanzagl and W. Wefelmeyer. *Contributions to a general asymptotic statistical theory.* Springer, 1982.

E. C. Polley and M. J. van der Laan. Super learning for right-censored data. In M. J. van der Laan and S. Rose, editors, *Targeted Learning: Causal Inference for Observational and Experimental Data*, pages 249–258. Springer, 2011.

H. C. Rytgaard and M. J. van der Laan. Targeted maximum likelihood estimation for causal inference in survival and competing risks analysis. *Lifetime Data Analysis*, pages 1–30, 2022.

H. C. Rytgaard, F. Eriksson, and M. J. van der Laan. Estimation of time-specific intervention effects on continuously distributed time-to-event outcomes by targeted maximum likelihood estimation. *Biometrics*, 2021.

H. C. Rytgaard, T. A. Gerds, and M. J. van der Laan. Continuous-time targeted minimum loss-based estimation of intervention-specific mean outcomes. *The Annals of Statistics*, 50(5):2469–2491, 2022.

M. C. Sachs, A. Discacciati, Å. H. Everhov, O. Olén, and E. E. Gabriel. Ensemble prediction of time-to-event outcomes with competing risks: A case-study of surgical complications in Crohn's disease. *Journal of the Royal Statistical Society Series C: Applied Statistics*, 68(5):1431–1446, 2019.

N. Simon, J. Friedman, T. Hastie, and R. Tibshirani. Regularization paths for Cox's proportional hazards model via coordinate descent. *Journal of Statistical Software*, 39(5): 1–13, 2011. doi: 10.18637/jss.v039.i05. URL https://www.jstatsoft.org/v39/i05/.

J. A. Steingrimsson, L. Diao, and R. L. Strawderman. Censoring unbiased regression trees and ensembles. *Journal of the American Statistical Association*, 114(525):370–383, 2019.

T. M. Therneau. *A Package for Survival Analysis in R*, 2022. URL https://CRAN.R-project.org/package=survival. R package version 3.4-0.

M. J. van der Laan and S. Dudoit. Unified cross-validation methodology for selection among estimators and a general cross-validated adaptive epsilon-net estimator: Finite sample oracle inequalities and examples. Technical report, Division of Biostatistics, University of California, 2003.

M. J. van der Laan and J. M. Robins. *Unified methods for censored longitudinal data and causality*. Springer Science & Business Media, 2003.

M. J. van der Laan and S. Rose. *Targeted learning: causal inference for observational and experimental data*. Springer Science & Business Media, 2011.

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

A. W. van der Vaart, S. Dudoit, and M. J. van der Laan. Oracle inequalities for multi-fold cross validation. *Statistics & Decisions*, 24(3):351–371, 2006.

P. J. Verweij and H. C. van Houwelingen. Cross-validation in survival analysis. *Statistics in medicine*, 12(24):2305–2314, 1993.

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

D. Whitney, A. Shojaie, and M. Carone. Comment: Models as (deliberate) approximations. *Statistical science: a review journal of the Institute of Mathematical Statistics*, 34(4):591, 2019.

D. H. Wolpert. Stacked generalization. *Neural networks*, 5(2):241–259, 1992.

J. Yao, X. Zhu, F. Zhu, and J. Huang. Deep correlational learning for survival prediction from multi-modality data. In *International conference on medical image computing and computer-assisted intervention*, pages 406–414. Springer, 2017.