Estimating cumulative incidence functions in competing risks data with dependent left-truncation

Regina Stegherr1 | Arthur Allignol2 | Reinhard Meister3 | Christof Schaefer4 | Jan Beyersmann1

Both delayed study entry (left-truncation) and competing risks are common phenomena in observational time-to-event studies. For example, in studies conducted by Teratology Information Services (TIS) on adverse drug reactions during pregnancy, the natural time scale is gestational age, but women enter the study after time origin and upon contact with the service. Competing risks are present, because an elective termination may be precluded by a spontaneous abortion. If left-truncation is entirely random, the Aalen-Johansen estimator is the canonical estimator of the cumulative incidence functions of the competing events. If the assumption of random left-truncation is in doubt, we propose a new semiparametric estimator of the cumulative incidence function. The dependence between entry time and time-to-event is modeled using a cause-specific Cox proportional hazards model and the marginal (unconditional) estimates are derived via inverse probability weighting arguments. We apply the new estimator to data about coumarin usage during pregnancy. Here, the concern is that the cause-specific hazard of experiencing an induced abortion may depend on the time when seeking advice by a TIS, which also is the time of left-truncation or study entry. While the aims of counseling by a TIS are to reduce the rate of elective terminations based on irrational overestimation of drug risks and to lead to better and safer medical treatment of maternal disease, it is conceivable that women considering an induced abortion are more likely to seek counseling. The new estimator is also evaluated in extensive simulation studies and found preferable compared to the Aalen-Johansen estimator in non-misspecified scenarios and to at least provide for a sensitivity analysis otherwise.

1 | INTRODUCTION

This work is motivated by observational studies of Teratology Information Services (TIS) on drug-exposed pregnancies. Pregnant women or their doctors contact TIS after the pregnancy has been recognized and if a drug risk assessment is...
needed. In addition to the individual counseling, TIS will then prospectively follow up pregnancies if consent is provided. One statistical consequence is that women have a delayed study entry. In addition, the analysis will have to account for competing risks, because, eg, a spontaneous abortion may preclude an induced abortion/elective termination.

The standard method to estimate cumulative event probabilities in such situations is the Aalen-Johansen estimator, generalizing the Kaplan-Meier estimator to multiple event types under the assumption of random left-truncation (and right-censoring). However, the analysis of left-truncated data comes with additional aspects when compared to only right-censored data.

One issue is that left-truncation may lead to early overly small risk sets. Events observed from such overly small risk sets may lead to instability in both the Kaplan-Meier and the Aalen-Johansen estimator that propagates over the whole time span. Lai and Ying for the Kaplan-Meier estimator and, more recently, Friedrich et al and Rousson et al for the Aalen-Johansen estimator, therefore, stabilized these estimators by discarding events in too small risk sets. This was, however, not a concern in our data example below.

Another aspect is that, just like right-censoring, delayed study entry may be related to the distribution of the event times. However, unlike the assumption of random right-censoring, the assumption of random left-truncation may be tested. The aim of this work will be to model possibly dependent left-truncation in order to estimate cumulative event probabilities if the assumption of random left-truncation is in doubt. Keiding and Moeschberger described scenarios where the left-truncation and the time-to-event may not be independent. This challenge occurs when the populations are heterogeneous in a way not accounted for in the modeling. Keiding and Moeschberger further gave two examples of situations with dependent left-truncation; one including a frailty and another about cross-sectional samples of prevalent cohorts, where patients with a longer disease duration are more likely to enter a study. The problem is that those who are selected to enter the study do not adequately represent the population base since only an individual whose event time is greater than its left-truncation time enters the study. A positive association between the left-truncation time and the time-to-event results in an underestimation of the risk of the event and a negative association results in an overestimation.

A recent simulation study investigating the impact of dependent left-truncation in the context of competing risks has been given by Bakoyannis and Touloumi.

For standard survival data, there are several methods to test the assumption of random left-truncation on the so-called observable region, where the left-truncation time is smaller than the time-to-event. An early method using Kendall’s tau was introduced by Tsai. Efron and Petrosian derived a permutation test closely related to the Kaplan-Meier estimator. Furthermore, Martin and Betensky suggested testing quasi-independence of the left-truncation time and the event time, ie, independence on the observable region only. In this work, we will follow Keiding and Gill who suggested to test independence by including the left-truncation time as a covariate into a Cox proportional hazards model. Our choice has two reasons. Firstly, the approach immediately applies to competing risks data by including the left-truncation time in a Cox proportional hazards model for the cause-specific hazards of the competing risks, see the work of Anzures-Cabrera and Hutton for a related approach in a parametric competing risks framework. Secondly, the approach may be used to model dependent left-truncation. For standard survival data, this was taken up by Mackenzie who used inverse probability weighting arguments and a Cox proportional hazards model to adapt the Kaplan-Meier estimator to dependent left-truncation.

The aim of this work is to extend Mackenzie’s inverse probability of left-truncation weight (IPLW) approach to competing risks and to estimate cumulative event probabilities of a competing risk in the presence of dependent left-truncation. This is complicated by the presence of multiple cause-specific hazards, where we have as many hazards as there are competing event types. The presence of multiple hazards also requires an additional subtlety in the modeling of dependent left-truncation: In our data example, there will be concerns that one cause-specific hazard (induced abortion) is affected by the study entry time, but not the others (spontaneous abortion, live birth), and this will need to be accounted for in the modeling. Furthermore, and in contrast to Mackenzie, we also evaluate our new estimator in simulation studies where the Cox proportional hazards model is misspecified.

Our motivating example is from a study about the usage of coumarin derivatives in pregnancy where the aim was to evaluate the effect of medication intake on the risk of spontaneous abortion, see the work of Meister and Schaefer for details. In short, the data were collected by TIS Berlin, Germany, and contain three competing events: spontaneous abortion, induced abortion, and live birth. There is a concern that women who are contemplating an induced abortion are more likely to contact the TIS although, in general, TIS does not advise induced abortions. Nevertheless, a former analysis using cause-specific Cox models found that the left-truncation time may be associated with an increased cause-specific hazard of induced abortion. The aim of our applied analysis will be to provide for a sensitivity analysis of the cumulative pregnancy outcome probabilities, allowing for possibly dependent left-truncation.
The remainder of this paper is organized as follows: Section 2 briefly revisits the case of survival data to clarify the target (or estimand) of the proposed inferential procedures. Both the competing risks model and our new estimator of the cumulative event probabilities in the presence of dependent left-truncation are in Section 3. Section 4 reports simulation results and a reanalysis of data on drug-exposed pregnancies accounting for possibly dependent left-truncation is in Section 5. A summary concludes this paper in Section 6. Supporting information, in particular reporting additional simulation results, is provided online.

2 | THE ESTIMAND OF MACKENZIE

The basic model of survival data considers i.i.d. replicates of \((T, C)\), where \(T\) is the time-to-event and \(C\) is a right-censoring time. The common assumption is that \(T\) and \(C\) are independent, possibly conditional on baseline covariates. Analogously, left-truncated survival data are typically viewed as copies of the pair \((L, T)\), where \(L\) now is the left-truncation time. While right-censoring switches observation “off,” left-truncation can be viewed as switching observation “on,” such that an individual is under study if \(L < T\), while individuals with \(T \leq L\) are not sampled. Again, the common assumption is that \(T\) and \(L\) are independent, and, already, Kaplan and Meier\(^{15}\) suggested to use their estimator for estimation of \(P(T > t) = \exp(- \int_0^t a(u)du)\) in the presence of left-truncation, where \(a(t)\) is the survival hazard. Common to both models also is that both \(T, L,\) and \(C\) are viewed as latent variables. A difference between right-censoring and left-truncation is that a possible dependency between \(L\) and \(T\) can be investigated by including \(L\) as a covariate in a regression model for the hazard of \(T\), say \(a(t | L = \ell)\). One option to account for such dependence is to only assume conditional independence of \(L\) and \(T\) given baseline covariates. In the example of pregnancy data, such covariates are difficult to identify as the reasons for induced abortions are various. The characteristics of women which influence induced abortions are hard to detect and interpret. A different approach is IPLW. To this end, Mackenzie\(^{14}\) suggested to consider the conditional survival functions

\[
P(T > t | L = \ell) = \exp \left( - \int_0^t a(u | L = \ell)du \right)
\]  

(1)

and to subsequently marginalize w.r.t. the distribution \(F_L\) of \(L\)

\[
P(T > t) = \int P(T > t | L = \ell) dF_L(\ell).
\]

(2)

Section 3 extends Mackenzie’s suggestion to the situation where the target of estimation is the unconditional cumulative event probability or cumulative incidence function of a competing risk.

However, before we proceed, three remarks are in place. Firstly, Equations (1) and (2) demonstrate that the left-truncation time is truly viewed as a latent variable, but not as the time-to-occurrence of an internal time-dependent covariate.\(^{16}\) An internal time-dependent covariate is conceptually tied to existence or survival of the individual, and hence, the quantity \(a(u | L = \ell)\) is without any useful meaning for \(u \leq \ell\) and \(P(T > t | L = \ell)\) must be one for \(t < \ell\). The consequence is that the present approach does not apply to common examples of so-called prevalent sampling where individuals are included conditional on an initial event such as disease onset, and where this initial event is modeled as a left-truncation event, see the work of Cheng et al\(^{17}\) for a recent example. In our motivating data example, the interpretation is that a drug-exposed pregnancy comes with a latent time-to-TIS-contact. However, if, say, the pregnancy ends in a spontaneous abortion beforehand, the contact to TIS will not be realized.

Secondly, while our target of estimation will be an unconditional probability, a conditional probability such as \(P(T > t | L = \ell)\) may very well be preferred for individual counseling or prediction. In fact, there is a relation to recent calls for dynamic prediction.\(^{18}\) If, as in our motivating data example, \(L = \ell\) is the time of TIS contact for seeking advice, individual counseling should be based on \(P(T > t | T \geq \ell)\) or, in the case of dependence between \(T\) and \(L\), \(P(T > t | T \geq \ell, L = \ell)\). Such predictions may very well be obtained by using \(L = \ell\) in a hazard regression, see (7) in Section 3. However, as we will also see below, our practical aim will be to provide for a sensitivity analysis of the standard Aalen-Johansen estimator that assumes random left-truncation.

Thirdly, the relation \(P(T > t) = \int P(T > t | L = \ell) dF_L(\ell)\) uses the law of total probability. The interpretation is that \(P(T > t)\) is the survival probability of an individual picked randomly from the population at hand. Conditional on \(L\), however, the survival probabilities differ between individuals. We will exploit this in the simulations of Section 4, where
we determine the target probability by using the Aalen-Johansen estimator on complete data sets without truncation. We also note that, because of the law of total probability, the target quantity is not the (cumulative) hazard \( a(t \mid L = \ell) \) marginalized w.r.t. \( F_\ell \). In other words, the hazard belonging to \( P(T > t) \) is not \( \int a(t \mid L = \ell) dF_\ell(\ell) \) but \( a(t) = \int a(t \mid L = \ell) P(T > t \mid L = \ell) dF_\ell(\ell) / P(T > t) \).

We will briefly revisit these aspects in Section 4 using simulations. In particular, the simulation algorithm in Section 4.1 will illustrate the points made above about latent times and marginal vs conditional probabilities. A brief simulation for survival data without competing risks in Section 4.2 further illustrates the question of the estimand addressed above.

### 3. Competing Risks Model and Estimation of Cumulative Event Probabilities

In the following, we consider a competing risks process \((X_t)_{t \geq 0}\) denoting the state of an individual for every point in time \( t, X_t \in \{0, 1, \ldots, k\} \). As in the data example, we consider the case with \( k = 3 \) absorbing states, but the model can be easily written for two or more than three competing events. Every individual starts in the initial state 0 at study entry, \( P(X_0 = 0) = 1 \). As long as no event occurs the individual stays in state 0. If an event of type \( j, j \in \{1, 2, 3\} \), occurs the individual moves into the corresponding state. The event time \( T = \inf \{t > 0 \mid X_t \neq 0\} \) is defined as the time the competing risk process moves out of the initial state 0. The process \((X_t)_{t \geq 0}\) is assumed to be right-continuous, i.e., the competing risks process is in one of the states 1, 2, and 3 at time \( T \).

The cause-specific hazards \( a_0(t) \) can be defined via

\[
a_0(t)dt := P(T \in [t, t + dt), X_T = j \mid T \geq t), j = 1, 2, 3.
\]

Also, write \( A_0(t) := \int_0^t a_0(u)du, j = 1, 2, 3 \), for the cumulative cause-specific hazards. The sum of the \( a_0 \)'s is the all-cause hazard \( A_0, a_0(t)dt = P(T \in [t, t + dt) \mid T \geq t) \), with cumulative all-cause hazard \( A_0(t) \). Hence, all cause-specific hazards enter the survival function \( P(T > t) = \exp(-\int_0^t a_0(u)du) \).

In a competing risks setting, the interest often is in the cumulative incidence function of state \( j = 1, 2, 3 \)

\[
P(T \leq t, X_T = j) := \int_0^t P(T \geq u)a_0(j)du.
\]

Note that \( P(T \geq u) = P(T > u) \), but estimation will use an estimator of \( P(T \geq u) \) rather than \( P(T > u) \).

For ease of presentation, we assume \((X_t)_{t \geq 0}\) to only be subject to left-truncation and not to right-censoring but including independent right-censoring is straightforwardly incorporated as usual. This corresponds to the typical situation in pregnancy data. Let \( L \) denote the left-truncation time, i.e., the time of study entry. An individual’s event time and event type \((T, X_T)\) can only be observed if \( L < T \). Otherwise, the individual experiences an event before study entry and does not enter the study, i.e., only \( n \) out of \( m \) individuals enter the study, \( n \leq m \).

#### 3.1 Random left-truncation

In the standard random left-truncation model, the event time \( T \) and the left-truncation time \( L \) are assumed to be independent. For estimation, we introduce the following counting process notation. Let \( n \) be the number of individuals under study as above. Therefore, we have \( n \) i.i.d. replicates of \((L_i, T_i, X_{Ti})\).\(^{19,20}\) The aggregated at-risk process denoting the number of individuals observed in state 0 just before time \( t \) is given by \( Y(t) = \sum_{i=1}^n I(L_i < t \leq T_i) \) where \( I(.) \) denotes the indicator function. Furthermore, \( N_0(t) = \sum_{i=1}^n I(L_i < T_i \leq t, X_{Ti} = j) \) is the number of observed \( 0 \rightarrow j \)-transitions in the time interval \([0, t]\), \( j = 1, 2, 3 \). We also define the total number of transitions out of state 0 in the interval \([0, t]\) as \( N_0(t) = N_01(t) + N_02(t) + N_03(t) \). The increment of the counting process, i.e., the number of observed events of type \( j \) at \( t \), is written as \( \Delta N_0(j) = N_0(j) - N_0(-) \). Moreover, \( \Delta N_0(t) = \Delta N_01(t) + \Delta N_02(t) + \Delta N_03(t) \) is the total number of events observed at time \( t \). Then, the Aalen-Johansen estimator of the cumulative incidence function is given by

\[
\hat{P}(T \leq t, X_T = j) = \sum_{s \leq t} \hat{P}(T \geq s) \frac{\Delta N_0(j)(s)}{Y(s)}, j = 1, 2, 3,
\]
Dependent left-truncation

Note that Equations (6) and (7), which define and use region \( T \) model. This assumption is also only testable on the observable region influences the risk of experiencing an event.2 In a single endpoint survival situation, Keiding and Gill12 therefore Unlike the assumption of random right-censoring, it may be investigated whether or not the information of left-truncation 3.2

3.2  Dependent left-truncation

Unlike the assumption of random right-censoring, it may be investigated whether or not the information of left-truncation influences the risk of experiencing an event.2 In a single endpoint survival situation, Keiding and Gill12 therefore suggested to include the left-truncation time as a covariate in a Cox proportional hazards model. The approach can be analogously applied to the competing risk setting by including the left-truncation time in the Cox models for all cause-specific hazards.2 To model dependent left-truncation, we assume the cause-specific hazards to follow a Cox proportional hazards model. This assumption is also only testable on the observable region \( T > L \) and remains untestable on the remaining region \( T \leq L \).14 Thus,

\[
\alpha_0(t | \ell = \ell)dt := \Pr(T \in [t, t + dt), X_T = j | T \geq t, L = \ell)
\]

with \( \alpha_0(t | L = \ell) = \alpha_{0;0}(t) \exp(\beta_0 \ell) \) for \( j = 1, 2, 3 \) and with cumulative cause-specific hazards \( \alpha_0(t | \ell) = \int_0^t \alpha_0(t | L = \ell)du \).

Next, the cumulative incidence function conditional on \( L = \ell \) is defined as

\[
\Pr(T \leq t, X_T = j | L = \ell) := \int_0^t \Pr(T > u - |L = \ell| \alpha_0(u | L = \ell)du, j = 1, 2, 3.
\]

Note that Equations (6) and (7), which define and use \( \alpha_0(u | L = \ell) \) for all times \( u \), rely on the fact that \( L = \ell \) is not viewed as an internal time-dependent event happening at \( \ell \), but \( L \) is viewed as a latent variable, see Section 2. As also explained in Section 2, the conditional probability (7) would be useful for individual prediction, ideally updated for the past event-free time. However, in what follows, we will be interested in a comparator for the Aalen-Johansen estimator, which can be based on integration of (7) with respect to the left-truncation distribution \( F_L \)

\[
\Pr(T \leq t, X_T = j) = \int_0^\infty \Pr(T \leq t, X_T = j | L = \ell) \cdot dF_L(\ell), j = 1, 2, 3.
\]

For estimation of this function the parameters \( \beta_0 \) need to be estimated first. As in the standard Cox proportional hazards model, this works by maximizing the log-likelihood of the model where \( L \) is included as a covariate. The predicted cumulative cause-specific hazards for individual \( i \) with entry time \( \ell_i \) are

\[
\hat{A}_{0j}(t | L = \ell_i) = \hat{A}_{0j;0}(t) \exp(\hat{\beta}_0 \ell_i), j = 1, 2, 3,
\]

where the cumulative baseline hazard can be estimated using the Breslow estimator \( \hat{A}_{0j;0}(t) = \sum_{s \leq t} \frac{\Delta N_{0j}(s)}{Y(s) \exp(\beta_0 \ell_i)} \). After that, the predicted cumulative all-cause hazard is given by \( \hat{A}_0(t | l = \ell_i) = \sum_j \hat{A}_{0j}(t | L = \ell_i) \). Furthermore, the predicted conditional survival probability is denoted by

\[
\hat{S}(t | L = \ell_i) = \hat{P}(T > t | L = \ell_i) = \prod_{u \leq t} (1 - \Delta \hat{A}_0(u | L = \ell_i)).
\]

With Equation (10) the predicted conditional cumulative incidence functions can be derived

\[
\hat{P}(T \leq t, X_T = j | L = \ell_i) = \sum_{u \leq t} \hat{S}(u - |L = \ell_i) \cdot \Delta \hat{A}_0(u | L = \ell_i).
\]

In a scenario with dependent left-truncation, we need to account for not observing all events due to left-truncation, ie, only \( n \) out of \( m \) events can be observed. Therefore, \( m \) is unknown and an estimator of it is needed. Inverse probability weighting arguments are used for its derivation. The main idea is that, if the (marginal) inclusion probability is
Q = P(T > L) ≈ \frac{n}{m}$, one individual under study represents 1/$Q$ individuals of the total population. So, estimating $m$ with $\hat{m} = \sum_{i=1}^{n} \frac{1}{P(T > \ell_i | L = \ell_i)}$ leads to an estimated probability of being included in the study of $\hat{Q} = \frac{n}{\hat{m}}$, and the left-truncation distribution $F_L(\ell)$ can be estimated by

$$\hat{F}_L(\ell) = \frac{\hat{Q}}{n} \sum_{i=1}^{n} \frac{1(\ell_i \leq \ell)}{P(T_i > \ell_i | L_i = \ell_i)}.$$

Finally, an estimator the of marginal cumulative incidence function in a scenario with dependent left-truncation is

$$\hat{P}(T \leq t, X_T = j) = \frac{\hat{Q}}{n} \sum_{i} \sum_{i=1}^{n} \hat{P}(T \leq t, X_T = j | L_i = \ell_i) \cdot \frac{1(\ell_k = \ell_i)}{P(T > \ell_i | L_i = \ell_i)}.$$

### 4 | SIMULATION STUDY

Section 4.1 describes the simulation algorithm. The algorithm further illustrates points made about latent times and marginal vs conditional probabilities made in Section 2. Section 4.2 revisits the standard survival case illustrating the estimand as discussed in Section 2. Competing risks results are in Section 4.3.

### 4.1 | Setup

Similar to our real data example, we simulated competing risks data with $k = 2$ competing states as the two primary competing events are induced abortion and spontaneous abortion since live births are usually observed much later. A total number of $m = 500$ individuals were simulated. To begin, the left-truncation times are simulated following an exponential distribution with parameter chosen such that the marginal probability of being included in the study is $Q = 60\%$. In a second step, the event times are generated conditional on the first step and from an exponential distribution with parameter corresponding to the all-cause hazard $\alpha_0(t | L = \ell_i)$, $i = 1, \ldots, m$. The event type is determined in a third step: Given an event occurs at time $t_i$, it is of type 1 with conditional probability $\frac{\alpha_0(t_i | L = \ell_i)}{\sum_{k=0}^{\ell_i} \alpha_0(t_i | L = \ell_i)}$. Finally, the left-truncated data set for analysis consists of the $n \leq m$ individuals with $L_i < T_i$, $i = 1, \ldots, n$. We note two facts about the distribution of the simulated data: Firstly, the algorithm is run identically for all initial $m$ units. Hence, the $m$ units are marginally i.i.d. This is why the standard Aalen-Johansen estimator for the marginal cumulative event probability (or Kaplan-Meier for the marginal survival function of $T_i$) computed in the untruncated data set with all $m$ units are (asymptotically) unbiased estimators. However, conditional on $L_i = \ell_i$, the units differ. Secondly, we reiterate that $(L_i, T_i)$ are viewed as latent variables, see Section 2. Phrased in the context of our data example, the interpretation is that a drug-exposed pregnancy $i$ comes with a latent time-to-TIS-contact $L_i$. If, say, the pregnancy ends in a spontaneous abortion at $T_i = t_i < \ell_i = L_i$, the contact to TIS will not be made anymore. This is not unlike the latent times model for censored survival data, $(T_i, C_i)$, where the right-censoring time $C_i$ is not observed if $T_i$ happens first, but $C_i$ is still assumed to conceptually exist. This point of view is put to further use here by modeling the dependence between $L_i$ and $T_i$ via conditional hazards (6).

### 4.2 | Survival data without competing risks

The aim of this section is to illustrate the estimand discussed in Section 2 for standard survival data. To this end, assume left-truncation times to follow an exponential distribution with parameter 0.17 and survival data to follow a conditional distribution with hazard $\alpha(t | L = \ell_i) = 0.06 \cdot \exp(-0.3 \cdot \ell_i)$; we will reuse these specifications in the context of competing risks. The marginal survival distribution then is

$$P(T > t) = \int_0^\infty P(T > t | L = \ell) dF_L(\ell) = \int_0^\infty \exp(-0.06 \cdot \exp(-0.3 \cdot \ell) \cdot t) \cdot \exp(-0.17 \cdot \ell) \cdot 0.17 \, d\ell,$$

where the integral on the right hand side of the previous display may be evaluated numerically.

We simulated 1000 data sets, each with an initial number of $m$ individuals. Figure 1 shows the true curve based on the above formula, and the average across all data sets of Mackenzie’s IPLW estimator of the survival function, which is a special case of (one minus) our estimator (13), the Kaplan-Meier estimator evaluated in the complete (untruncated)
data sets and the Kaplan-Meier estimator computed based on the left-truncated data only. The former three curves are indistinguishable, while the latter can be seen to be biased.

### 4.3 Competing risks data

Four simulation scenarios were conducted. In accordance with the coumarin data example, only the cause-specific hazard of one event type depends on the left-truncation time. Here, in all scenarios, it is the cause-specific hazard of event type 2. The cause-specific hazards of event type 1 were chosen constant as $a_{01}(t|L = \ell_i) = 0.1$ in scenarios (A), (B), and (D). For event type 2, three different options were chosen; in the scenario (A) and (B), it follows a Cox model with $\beta_{02} = 0.3$ in scenario (A) and $\beta_{02} = -0.3$ in (B), ie, $a_{02}(t|L = \ell_i) = 0.06 \cdot \exp(\beta_{02} \cdot \ell_i)$. In scenarios (C) and (D), model misspecifications are taken into account. Therefore, in scenario (C), a gamma frailty, $Z \sim \Gamma(2, 2)$, was included, ie, $a_{01}(t|L = \ell_i) = 0.1 \cdot z_i$ and $a_{02}(t|L = \ell_i) = 0.06 \cdot z_i \cdot \exp(0.3 \cdot \ell_i)$. Scenario (D) investigates parameter estimation when there is an additive structure of the cause-specific hazard of events of type 2, ie, $a_{02}(t|L = \ell_i) = 0.06 + 0.3 \cdot \ell_i$. This additivity is ignored while estimating the cumulative incidence functions in the sense that $a_{02}$ is analyzed using a Cox model.

The true values were numerically computed by averaging the cumulative incidence functions from 10 simulated complete, ie, all $m$ individuals are observed, studies with $m = 100,000$.

Figure 2 displays the true cumulative incidence functions, their Aalen-Johansen estimates, and the estimates of our new IPLW estimator, each with empirical pointwise 95% confidence interval. The results of scenario (A) and (B) are in
TABLE 1  Mean squared error of the estimated cumulative incidence functions at time 40 using the new IPLW estimator or the Aalen-Johansen estimator

| Estimator        | (A)     | (B)     | (C)     | (D)     |
|------------------|---------|---------|---------|---------|
| CIF 1 IPLW estimator | 0.0086  | 0.0022  | 0.0068  | 0.0256  |
| CIF 1 Aalen-Johansen | 0.0088  | 0.0108  | 0.0101  | 0.0083  |
| CIF 2 IPLW estimator | 0.0086  | 0.0021  | 0.0078  | 0.0256  |
| CIF 2 Aalen-Johansen | 0.0088  | 0.0111  | 0.0116  | 0.0083  |

Abbreviations: CIF, cumulative incidence function; IPLW, inverse probability of left-truncation weighting.

TABLE 2  Estimated mean and empirical 95% confidence interval (CI) of the parameters that models the association between the left-truncation time and the hazard of an event in the Cox proportional hazards models including the left-truncation time as a covariate for all four simulation scenarios

| (A)     | (B)     | (C)     | (D)     |
|---------|---------|---------|---------|
| mean $\hat{\beta}_{01}$ (95% CI) | $-0.001$ ($-0.055, 0.053$) | $0.000$ ($-0.014, 0.013$) | $-0.092$ ($-0.140, -0.044$) | $-0.013$ ($-0.184, 0.158$) |
| mean $\hat{\beta}_{02}$ (95% CI) | $0.303$ ($0.254, 0.350$) | $-0.306$ ($-0.352, -0.253$) | $0.199$ ($0.158, 0.236$) | $1.051$ ($0.949, 1.150$) |

accordance with the remarks in the work of Keiding and Moeschberger,7 ie, a negative association results in an overestimation of the cumulative incidence function using the Aalen-Johansen estimator and a positive association results in an underestimation. In scenarios (A), (B), and (C), the new estimator is virtually unbiased at the price of an increased variance. The empirical confidence interval also includes the mean Aalen-Johansen estimator. On the other hand, the empirical 95% confidence interval of the Aalen-Johansen estimator does not cover the true value. In scenario (C), our approach to modeling dependent left-truncation is misspecified, as it is in scenario (D). In contrast to (C), however, in scenario (D), our new estimator is no longer unbiased and has a strongly increased variance, the mean squared error is the highest in this setting.

To further investigate this, we considered the test of random left-truncation via Cox modeling. Table 2 displays the estimated mean and empirical 95% confidence intervals of the parameters that model the association between the left-truncation time and the hazard of an event in the Cox model.

On average, results were in line with the model specifications for scenarios (A) and (B). In scenario (C), both cause-specific hazards models are misspecified, while in scenario (D), only the cause-specific hazard for event type 2 is misspecified. Hence, the average result for the latter is also in line with the model specifications. On the other hand, the average results reported in Table 2 for scenario (C) and for the cause-specific hazard for event type 2 in scenario (D) reflect model misspecification. However, and this is a notable difference between (C) and (D), results for (C) are of a similar magnitude as compared to (A) and (B), while the magnitude of the effect reported for the type 2 hazard in (D) clearly stands out. In other words, and as measured by Table 2, the degree of misspecification appears to be larger in (D), which helps to explain the results seen in Figure 2.

We also note that, in some simulated data sets (1 in (A), 9 in (B), 8 in (C), and 48 in (D)), too heavy parametric extrapolation occurred such that, in (10) negative values were produced. These studies were not included in the averages reported in Figure 2. Furthermore, the problem was, again, most pronounced in scenario (D).

In the statistic software R, the Aalen-Johansen estimator can be calculated with the function `etm` of the `etm` package.21 The conditional cumulative incidence function can be calculated with the help of the `mstate` package.22

5 | COUMARIN DATA

We reanalyze the abortion data set, which we have made publicly available in the R package `etm`,1,2,21 where for the present analysis, we have broken ties by adding a small uniform random noise to facilitate the semiparametric modeling approach from above. The data set is originally from a prospective cohort study of pregnant women exposed to coumarin...
derivatives.\textsuperscript{1} It was collected by the Teratology Information Service (TIS) of Berlin, Germany. The TIS counsels women on the toxicity of medications during pregnancy. Often, the first time contacting the TIS is in the first trimester of pregnancy.

The data consist of 1186 women. Of these, 173 were exposed to coumarin, an anticoagulant, for medical reasons, and 1013 served as controls as they were not exposed to any potential teratogens. There are three competing events in the data: induced abortion, live birth, and spontaneous abortion. Of the control group, 20 women had an induced abortion, 69 a spontaneous abortion, and 924 children were born alive. In the coumarin group, on the other hand, 38 women had an induced abortion, 43 a spontaneous abortion, and 92 experienced a live birth. We note that the pregnancy outcome of every woman is observed, meaning there is no right-censoring in the data. Figure S7 of the Supporting Information shows the time under study and the event type of the women divided by group. The figure also shows that the real data analysis must put attention to possible extrapolation. Especially in the coumarin group, there are few late entries. Most women entering after week 22 experience a live birth.

As typical in pregnancy data, the women enter the study several weeks after conception, so the data are left-truncated. Since the data were collected as a consequence of women or clinicians seeking advice by the TIS, there is the concern that women who take an induced abortion into consideration may be more likely to contact the TIS. However, TIS does usually not advise induced abortions. Furthermore, more women consider an induced abortion for social reason than for medical ones.\textsuperscript{23} Still, there may be the suspicion of dependence between left-truncation and event time. Since life birth happens at a much later time and spontaneous abortions happen unexpectedly, these events are assumed to be independent of the left-truncation time in the sense that the study entry time does not have an effect on the respective cause-specific hazards. Figure 3 displays the competing risks setting with dependent left-truncation in the abortion data.

To begin, we tested for possible dependence between the induced-abortion hazard and the left-truncation time using a Cox model including the study entry time. Table 3 shows the estimated hazard ratio with corresponding 95% confidence interval, where our first interest here is whether or not the value one is contained in the confidence interval. The subtlety of the competing risks setting is illustrated by the fact that no significant effect was found on the all-cause hazard, but when distinguishing the cause-specific hazards, a significant effect of the study entry time for induced abortion was found in the coumarin group. Possible explanations why a later study entry time increases the risk of induced abortion in the coumarin group could be, on the one hand, that women taking coumarin are more likely to be pregnant unplanned since otherwise the coumarin intake may have been stopped before conception. Unplanned pregnancies are detected later and women considering an induced abortion may be in a decision dilemma and take their time to contact the TIS. On the other hand, the time induced abortions are allowed is limited.

To proceed with the modeling approach, we chose to use as a covariate $\ell_i = \ell_i \cdot I(\ell_i \leq 15)$, because an elective termination after gestational week 15 is typically only possible for severe medical reasons. Figure 4 displays the marginal cumulative incidence functions with corresponding 95% confidence intervals of the two estimators together with the simple proportions of the pregnancy outcomes. The pointwise confidence intervals of the new estimator were obtained with a bootstrap by drawing $n$ times with replacement from the study population. In comparison to the simple proportions, eg, the ones for induced abortion are $\frac{38}{173} \approx 22\%$ in the coumarin group and $\frac{20}{1013} \approx 2\%$ in the control group.

![FIGURE 3](image-url) Competing risks setting in the coumarin data

|                     | Control        | Coumarin       |
|---------------------|----------------|----------------|
| All-cause           | 0.994 (0.986, 1.002) | 1.025 (0.99, 1.061) |
| Induced abortion    | 0.968 (0.722, 1.299) | 1.191 (1.022, 1.388) |
| Live birth          | 0.994 (0.986, 1.002) | 1.016 (0.977, 1.056) |
| Spontaneous abortion| 1.03 (0.926, 1.145)  | 1.044 (0.905, 1.203) |

TABLE 3 Estimated hazard ratio and 95% confidence interval of the Cox proportional hazard models including the left-truncation time as a covariate.
the Aalen-Johansen and the new IPLW estimator take the time-dependency of the risk set into account. Therefore, the simple proportions lead to an underestimation of the risk for induced abortion and spontaneous abortion in both groups and an overestimation of the probability of live birth. For the control group, there is no visible difference between the Aalen-Johansen estimator and the IPLW estimator but the confidence intervals of the IPLW estimator are wider. The same goes for the estimates of the cumulative incidence function of spontaneous abortion in the coumarin group. Since in the coumarin group, the left-truncation time had an influence on the time-to-induced abortion, we see a difference between the Aalen-Johansen estimate and the estimate of our new IPLW estimator. The Aalen-Johansen estimator potentially underestimated the risk of experiencing an induced abortion. However, this potential underestimation must be contrasted with two aspects. One is that simple proportions are still commonly used in the field, and compared with these, both the standard Aalen-Johansen estimator and the new estimator avoid underestimation of the absolute induced abortion risk as presented by the crude proportions. The second aspect is that the confidence intervals of the new IPLW estimator are now much wider compared to the Aalen-Johansen estimator as a consequence of the small sample size of only 173 women in the coumarin group and of the inverse probability weighting arguments. In Section 3 of the Supporting information, we conducted a simple sensitivity analysis w.r.t. sample size. Artificially, increasing the size of the coumarin group resulted in a decrease of the width of the confidence interval. Nevertheless, the confidence interval of the Aalen-Johansen estimator was still narrower. This is in line with our simulation results which found comparable MSEs in scenarios (A) to (C). Also reported in the Supporting Information is a simple stratified IPLW analysis with smaller confidence intervals and point estimates closer to the standard Aalen-Johansen estimator.

6 | SUMMARY

Estimating pregnancy outcome probabilities must account for both left-truncation and competing risks. In general, left-truncation must be taken into account in time-to-event studies whenever the natural time origin potentially lies before study entry, and competing risks are present whenever there is more than one event type. Because the time of delayed study entry is an element of the past w.r.t. the time scale of analysis, we may test and model associations between the left-truncation time and the time-to-event. In this work, we have extended a recent proposal by Mackenzie for estimating survival functions in the presence of dependent left-truncation. Using inverse probability of left-truncation weights obtained from Cox modeling of the cause-specific hazards, we obtained an estimator of the cumulative event
probabilities when the random left-truncation assumption is in doubt. The new estimator was found to be unbiased in simulation studies with correctly specified model for dependent left-truncation. Complementing Mackenzie's investigations, we also studied our new estimator under misspecification of the left-truncation model with negligible bias in the presence of a frailty term, but more pronounced bias when the true truncation model was additive rather than proportional. In all settings, the new IPLW estimator was much more variable than the original Aalen-Johansen estimator, but with comparable MSEs both in the correctly specified models and in the misspecified frailty case.

Both the simulation results and the real data analysis illustrated that our new estimator may be used as a sensitivity analysis of a standard Aalen-Johansen analysis, if the assumption of random left-truncation is in doubt. In the real data analysis, we found a signal that the Aalen-Johansen estimator potentially underestimated the absolute risk of induced abortion in the coumarin group, but the potential bias was moderate when compared to the still commonly used incidence proportions. The approach is, in principle, extendable to general multistate models, although, arguably, the problem of increased variation will even be more pronounced in such general models.

Accounting for dependent left-truncation relies on modeling the dependency. Some remarks are in place. To begin, an alternative approach for the estimation of the survival function under dependent left-truncation uses copulas to model the joint distribution of the left-truncation and the event time, but the extension to competing risks is less straightforward. An alternative popular method of analysis in the context of competing risks is based on the subdistribution hazard, ie, the hazard attached to the cumulative incidence function interpreting the latter as a distribution function. Originally developed in the presence of right-censoring only, left-truncation has been incorporated by several authors; we refer to the work of Zhang et al who also allow for dependence of the left-truncation time on baseline covariates. The subdistribution hazard has been criticized, questioning whether it has any useful interpretation. In this work, we have deliberately chosen to use the cause-specific hazards rather than the subdistribution hazard. The reason was the ability to specifically model an impact on the hazard of induced abortion only.

In this work, as repeatedly emphasized above, we have also relied on a latent times model \((L, T)\) similar to the common latent times model \((T, C)\) for censored data. An important consequence is that our approach does not apply to common examples of prevalent sampling where left-truncation is due to an internal time-dependent event. For such categorical time-dependent covariates, Cortese and Andersen discuss multistate modeling as a joint model for both the intermediate event and competing risks, see also the book by Beyersmann et al. In a progressive multistate model, a possible approach to estimate a “marginal” cumulative incidence function could be based on estimating the sum of state occupation probabilities corresponding to a competing risk, with absorbing states either reached directly from an initial state or after having passed through the intermediate state. Here, the practical problem will be to estimate the transition intensity from the initial to the intermediate state based on left-truncated data, but such a line of research merits future work, possibly under simplifying assumptions on this intensity.

The real data analysis also illustrated that some care is needed when modeling dependent left-truncation. Here, we only modeled an impact of study entry times within the first 15 weeks, because a later elective termination is typically only possible for severe medical reasons. In the Supporting Information online, we also provide a simple stratified IPLW analysis, which resulted in both less variation and estimates closer to the Aalen-Johansen estimator than the IPLW estimator based on Cox regression. These analyses suggest that investigating modeling alternatives and model choice for accounting for dependent left-truncation merits future research. Compared to the Cox modeling approach in this paper and in the work of Mackenzie, possible directions of hazards modeling include fully parametric modeling as in the work of Anzuces-Cabrera and Hutton who employ a Weibull proportional hazards model including a stratified left-truncation indicator as covariate and extensions of the Cox model as in the work of Matsuura and Eguchi who express the hazard as product of a hazard conditional on covariates (but not the left-truncation time) and a “nuisance function” that depends on the left-truncation time. Interestingly, Matsuura and Eguchi also use a stratification approach in their real data analysis. IPLW is not addressed in these papers. Of future interest could also be investigating the IPLW estimators of Mackenzie and of the present work under an additive hazards model for the truncation hazard, see, eg, Section 4.2.7 in the work of Aalen et al for a discussion of using the additive hazards model in inverse probability weighting. Another aspect is that, in the real data analysis, we have chosen to break ties in order to facilitate the semiparametric modeling approach. Another possibility and topic of future research would be time-discrete modeling.

Finally, in terms of subject matter considerations, we note that selection bias may be present in the TIS data collection as only women who consent are followed-up. On the one hand, it has been discussed that the risk profile of women who enroll in TIS studies does not substantially differ from women who do not enroll. On the other hand, many TIS experienced that women with a higher education may be more likely to seek contact. Schaefer et al give a thorough discussion of such biases. These authors also note that a possible selection effect of socio-economic status or education
will likely be present in both drug-exposed groups and in control groups and is not expected to be linked to the potentially harmful drug under investigation. They also stress as a major advantage of TIS data its prospective character. One consequence is that recall bias should be negligible, because very recent drug intake leads to consulting TIS.

ACKNOWLEDGEMENT

This work was partially supported through grants BE 4500/3-1, ME 910/2-1, and SCHA 1749/3-1 from the German Research Foundation (Deutsche Forschungsgemeinschaft).

CONFLICT OF INTEREST

The authors declare no potential conflict of interests.

FINANCIAL DISCLOSURE

None reported.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in the abortion data set, which we included in the R package etm.1,2,21

ORCID

Regina Stegherr https://orcid.org/0000-0002-0528-348X

REFERENCES

1. Meister R, Schaefer C. Statistical methods for estimating the probability of spontaneous abortion in observational studies—analyzing pregnancies exposed to coumarin derivatives. Reproductive Toxicology. 2008;26(1):31-35.
2. Allignol A, Schumacher M, Beyersmann J. A note on variance estimation of the Aalen-Johansen estimator of the cumulative incidence function in competing risks, with a view towards left-truncated data. Biometrical Journal. 2010;52(1):126-137.
3. Aalen O, Johansen S. An empirical transition matrix for non-homogeneous Markov chains based on censored observations. Scand Stat Theory Appl. 1978;5(3):141-150.
4. Lai TL, Ying Z. Estimating a distribution function with truncated and censored data. Ann Stat. 1991;19(1):417-442.
5. Friedrich S, Beyersmann J, Winterfeld U, Schumacher M, Allignol A. Nonparametric estimation of pregnancy outcome probabilities. Ann Appl Stat. 2017;11(2):840-867.
6. Rousson V, Allignol A, Aurousseau A, Winterfeld U, Beyersmann J. Stabilizing cumulative incidence estimation of pregnancy outcome with delayed entries. Biometrical Journal. 2019;61(5):1290-1302.
7. Keiding N, Moeschberger M. Independent delayed entry. In: Klein JP, Goel PK, eds. Survival Analysis: State of the Art. Dordrecht, The Netherlands: Springer; 1992:309-326.
8. Bakoyannis G, Touloumi G. Impact of dependent left truncation in semiparametric competing risks methods: a simulation study. Comm Statist Theory Methods. 2017;46(3):2025-2042.
9. Tsai W-Y. Testing the assumption of independence of truncation time and failure time. Biometrika. 1990;77(1):169-177.
10. Efron B, Petrosian V. A simple test of independence for truncated data with applications to redshift surveys. Astrophysical Journal. 1992;399:345-352.
11. Martin E, Betensky R. Testing quasi-independence of failure and truncation times via conditional Kendall's tau. J Am Stat Assoc. 2005;100(470):484-492.
12. Keiding N, Gill RD. Random truncation models and Markov processes. Ann Stat. 1990;18(2):582-602.
13. Anzures-Cabrera J, Hutton JL. Competing risks, left truncation and late entry effect in A-bomb survivors cohort. J Appl Stat. 2010;37(5):821-831.
14. Mackenzie T. Survival curve estimation with dependent left truncated data using Cox's model. Int J Biostat. 2012;8(1).
15. Kaplan E, Meier P. Nonparametric estimation from incomplete observations. J Am Stat Assoc. 1958;53:457-481.
16. Kalbfleisch J, Prentice R. The Statistical Analysis of Failure Time Data. 2nd ed. Hoboken, NJ: Wiley; 2002.
17. Cheng Y-J, Wang M-C, Tsai C-Y. Estimations of the joint distribution of failure time and failure type with dependent truncation. Biometrics. 2019;75(2):428-438.
18. Schumacher M, Hieke S, Ihorst G, Engelhardt M. Dynamic prediction: a challenge for biostatisticians, but greatly needed by patients, physicians and the public. Biometrical Journal. 2019. Early view.
19. Andersen PK, Gill RD, Borgan Ø, Keiding N. Statistical Models Based on Counting Processes. New York, NY: Springer; 1993.
20. Aalen O, Borgan O, Gjessing H. Survival and Event History Analysis: A Process Point of View. New York, NY: Springer Science & Business Media; 2008.
21. Allignol A, Schumacher M, Beyersmann J. Empirical transition matrix of multi-state models: the etm package. J Stat Softw. 2011;38(4):1-15.
22. de Wreede LC, Fiocco M, Putter H. mstate: an R package for the analysis of competing risks and multi-state models. J Stat Softw. 2011;38(7):1-30.
23. Willand I. Statistisches Jahrbuch Deutschland 2014. Wiesbaden, Germany: Statistisches Bundesamt; 2014.
24. Chaieb L, Rivest L-P, Abdous B. Estimating survival under a dependent truncation. Biometrika. 2006;93(3):655-669.
25. Beaudoin D, Lakhal-Chaieb L. Archimedean copula model selection under dependent truncation. Statist Med. 2008;27(22):4440-4454.
26. Zhang X, Zhang M-J, Fine J. A proportional hazards regression model for the subdistribution with right-censored and left-truncated competing risks data. Statist Med. 2011;30(16):1933-1951.
27. Andersen PK, Keiding N. Interpretability and importance of functionals in competing risks and multistate models. Statist Med. 2012;31(11-12):1074-1088.
28. Cortese G, Andersen PK. Competing risks and time-dependent covariates. Biometrical Journal. 2010;52(1):138-158.
29. Beyersmann J, Allignol A, Schumacher M. Competing Risks and Multistate Models With R. New York, NY: Springer Science & Business Media; 2012.
30. Matsuura M, Eguchi S. Modeling late entry bias in survival analysis. Biometrics. 2005;61(2):559-566.
31. Tutz G, Schmid M. Modeling Discrete Time-to-Event Data. New York, NY: Springer; 2016.
32. Johnson KA, Weber PA, Jones KL, Chambers CD. Selection bias in teratology information service pregnancy outcome studies. Teratology. 2001;64(2):79-82.
33. Schaefer C, Ornoy A, Clementi M, Meister R, Weber-Schoendorfer C. Using observational cohort data for studying drug effects on pregnancy outcome—methodological considerations. Reproductive Toxicology. 2008;26(1):36-41.

SUPPORTING INFORMATION
Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Stegherr R, Allignol A, Meister R, Schaefer C, Beyersmann J. Estimating cumulative incidence functions in competing risks data with dependent left-truncation. Statistics in Medicine. 2020;39:481–493.
https://doi.org/10.1002/sim.8421