Time-dependent covariates in the proportional subdistribution hazards model for competing risks

JAN BEYERSMANN∗
Freiburg Centre for Data Analysis and Modelling,
University of Freiburg, Freiburg, Germany and
Institute of Medical Biometry and Medical Informatics,
University Medical Centre Freiburg, Freiburg, D-79104, Germany
jan@fdm.uni-freiburg.de

MARTIN SCHUMACHER
Institute of Medical Biometry and Medical Informatics,
University Medical Centre Freiburg, Freiburg, D-79104, Germany

SUMMARY
Separate Cox analyses of all cause-specific hazards are the standard technique of choice to study the effect of a covariate in competing risks, but a synopsis of these results in terms of cumulative event probabilities is challenging. This difficulty has led to the development of the proportional subdistribution hazards model. If the covariate is known at baseline, the model allows for a summarizing assessment in terms of the cumulative incidence function. Mathematically, the model also allows for including random time-dependent covariates, but practical implementation has remained unclear due to a certain risk set peculiarity. We use the intimate relationship of discrete covariates and multistate models to naturally treat time-dependent covariates within the subdistribution hazards framework. The methodology then straightforwardly translates to real-valued time-dependent covariates. As with classical survival analysis, including time-dependent covariates does not result in a model for probability functions anymore. Nevertheless, the proposed methodology provides a useful synthesis of separate cause-specific hazards analyses. We illustrate this with hospital infection data, where time-dependent covariates and competing risks are essential to the subject research question.

Keywords: Fine and Gray model; Hospital infection; Multistate model.

1. INTRODUCTION
The proportional hazards model (Cox, 1972) is probably the single most important regression technique in survival analysis. If only time-independent covariates are considered, its results may be straightforwardly interpreted in terms of the failure time distribution due to the one-to-one relationship with the hazard (Andersen and others, 1993, Chapter VII.2.3). The proportional hazards model can also handle random

∗To whom correspondence should be addressed.
time-dependent covariates as well as competing risks (Therneau and Grambsch, 2000, Chapters 3.7.1 and 8.4), but interpretation of the results becomes more involved. Including time-dependent covariates does not result in a model for probability functions (Andersen and others, 1993, Chapter VII.2.3). In the presence of competing risks, one needs to perform separate analyses of all cause-specific hazards. We illustrate the interpretational difficulties in the data example of Section 6. As one consequence in competing risks theory, direct modeling of the cumulative incidence function, that is, the expected proportion of individuals experiencing the event of interest as time progresses, has emerged as a very active research field in the last decade (Fine and Gray, 1999; Fine, 2001; Klein and Andersen, 2005; Jeong and Fine, 2006, 2007; Scheike and Zhang, 2007).

In their pioneering work, Fine and Gray (1999) have suggested a proportional hazards model for the subdistribution hazard of the event of interest. The analysis of a baseline covariate may be straightforwardly interpreted in terms of the cumulative incidence function. The key is that while the cumulative incidence function is an involved function of all cause-specific hazards, the subdistribution hazard reestablishes a one-to-one relationship and consequently offers a summarizing analysis of separate cause-specific hazards analyses. The model has proven useful in applications, especially in medical research. A Medline search on January 8, 2008, found 201 citations, of which 155 appeared in medical journals. Mathematically, the model also allows random time-dependent covariates, but a practical implementation has remained unclear (Latouche and others, 2005; Putter and others, 2007). The subdistribution hazard is a hazard attached to a failure time with probability mass at infinity, corresponding to the events competing with the one of interest. As a technical consequence, an individual observed to fail from a competing risk is assumed to still be at risk between its real-life failure time and its potential future censoring time. The question is how to define the random time-dependent covariate on this interval.

The aim of this article was to provide for a subdistribution hazard-type analysis of a time-dependent covariate, which serves as a synthesis of separate cause-specific hazards analyses. We argue that one should look at a covariate process stopped at the real-life failure time (Jewell and Nielsen, 1993) and that this is in line with the standard cause-specific hazards modeling.

To do so, we proceed as follows. We first restrict attention to categorical and discrete time-dependent covariates. In Section 2, we recall that a proportional hazards model with a time-dependent binary covariate may equivalently be described through a multistate model with 2 transient states (Andersen, 1986; Klein and Moeschberger, 2003, Chapter 9.5). The relationship between the transition hazards into the same absorbing state is assumed to be proportional. This approach extends straightforwardly to k-level categorical and discrete covariates. In Section 3, we then consider the proportional subdistribution hazards model from a process point of view, which is nonstandard. We show that the model can be viewed in terms of a modified stochastic process, which is constructed from the original competing risks process. We combine these 2 observations in Section 4. For a competing risks process with a time-dependent binary covariate, we consider the equivalent multistate model and modify it into a subdistribution process. General real-valued time-dependent covariates are discussed in Section 5, connecting the multistate model–based approach to stopping the covariate process at the real-life failure time. We show the usefulness of the methodology for hospital infection data in Section 6 and offer a discussion in Section 7.

The hospital data example illustrates the need for a concise analysis of a time-dependent covariate in the presence of competing risks. Hospital infections constitute a severe clinical complication. An adequate analysis of the cumulative probability of infection needs to account for both the timing of the infection and its impact on the competing outcomes, discharge and death. Notably, time dependency of this risk exposure is often neglected (Irala-Estévez and others, 2001; Beyersmann, Gastmeier and others, 2006); this is, in fact, a general problem in the clinical literature, leading to biased analyses (van Walraven and others, 2004). In the example, we find that hospital pneumonia leads to more patients dying, however, not through increasing the cause-specific hazard of death (which remains essentially unaltered) but through reducing the cause-specific discharge hazard. In our practical work, we have found it most difficult to
communicate such results. We see that the new subdistribution hazard-type methodology allows for an easier interpretation and graphical presentation.

In the following, we simply write CSH for cause-specific hazard (and always use $\alpha(t)$), SH for subdistribution hazard (and always use $\lambda(t)$), and CIF for the cumulative incidence function. We also write PCSVHM for proportional CSH model and PSHM for proportional SH model.

2. Time-dependent binary covariate, proportional hazards and multistate models

For the 2-state model of Figure 1, top left, let $W(t) \in \{0, 2\}$ denote the state of an individual at time $t$, $t \in [0, \infty)$; as usual, the stochastic process $(W(t))_{t \in [0, \infty)}$ is assumed to have right-continuous sample paths. The survival time $U$ is then given as

$$U = \inf\{t \in [0, \infty) \mid W(t) = 2\}$$

with hazard

$$\alpha(t)dt = P(U \in dt \mid U \geq t),$$

where we write $dt$ for both the length of the infinitesimal interval $[t, t + dt)$ and the interval itself. Observation of $U$ may be subject to an independent right-censoring variable $C$. For a binary covariate $Z$, $Z(t) \in \{0, 1\}$, the proportional hazards assumption claims

$$\alpha(t; Z(t)) = \alpha_0(t) \cdot \exp(\beta Z(t)).$$

We require $Z$ to be predictable, that is, $Z(t)$ has already to be known at $t$—(Gill, 1984; Andersen and others, 1993, Chapter VII.1; Therneau and Grambsch, 2000, Chapter 3.7.1; Kalbfleisch and Prentice, 2002, Chapter 6.3). In our example, we look at $1(\text{Infection in } [0, t))$, $1(\cdot)$ denoting the indicator function, which is left-continuous in $t$. Estimation of $\beta$ is based on maximizing the partial likelihood (Gill, 1984).

A related multistate model is now given by the model of Figure 1, bottom left. Let $V(t) \in \{0, 1, 2\}$ denote the state an individual occupies at time $t$, $t \in [0, \infty)$. The connection to the 2-state model process $(W(t))_{t \in [0, \infty)}$ is given by

$$V(t) = 2 \iff W(t) = 2$$

and

$$\forall t \text{ with } W(t) \neq 2 : V(t) = Z(t+).$$

![Diagram of the 2-state model](image.png)

![Diagram of the 3-state model](image.png)

Fig. 1. Left: The 2-state model of survival analysis (top) extended to a 3-state model (bottom). State 2 is absorbing in both models. In the 3-state model, the transient states 0 and 1 correspond to a binary covariate. Right: The “original” competing risks process $X(t)$ with CSHs $\alpha_{02}(t)$ and $\alpha_{03}(t)$ (top) modified into the subdistribution process $\xi(t)$ with SH $\lambda(t)$ (bottom).
Because of (2.4), the definition (2.1) of the survival time $U$ remains unaffected. Definition (2.5) is due to
the fact that multistate processes have right-continuous sample paths (Andersen and Keiding, 2002). In
our example, $V(t) = 1$ will therefore have the interpretation $I($Infection in $[0, t])$. Let $a_{lm}(t)$ denote the
transition hazard of $(V(t))_{t \in [0, \infty)}$ between state $l$ and state $m$, $l \neq m \in \{0, 1, 2\}$. For time-inhomogeneous
Markov processes (Andersen and Keiding, 2002), the transition hazard is

$$a_{lm}(t)dt = P(V(t + dt) = m \mid V(t-) = l);$$

(2.6)

see Datta and Satten (2001) for a definition in the non-Markov setting. The proportional hazards assumption
now claims

$$a_{12}(t) = a_{02}(t) \cdot \exp(\beta).$$

(2.7)

The following observation is crucial. If there are no ties in the data, the partial likelihood for model (2.3)
is also a partial likelihood for the proportional hazards model (2.7), as we may exchange the $Z_j(t)$’s by
$Z_j(t+)$ as we wish without altering the likelihood. As a consequence, assuming no ties, we may for a
binary covariate in a proportional hazards model just as well consider a 3-state model with 2 transient
states corresponding to the values of the covariate and proportional transition hazards into the absorbing
state. We wish to stress that the interpretation of the states changes when moving from the model of
Figure 1, top left, to the one of Figure 1, bottom left. The event $W(t) = 0$ has the interpretation “no
terminal event by time $t$.” The event $V(t) = 0$ has the interpretation “no terminal event by time $t$,
covariate equals 0 at time $t+$.”

3. The Proportional Subdistribution Hazards Model

In this section, we reformulate the PSHM (Fine and Gray, 1999), moving from an original competing risks
process to a modified process, which forms the basis for the PSHM. Let $(X(t))_{t \in [0, \infty)}$ denote the original
competing risks process, $X(t) \in \{0, 2, 3\}$. The process starts in state 0 at time $t = 0$, and only transitions
into the competing states 2 and 3 are possible; $(X(t))_{t \in [0, \infty)}$ is assumed to have right-continuous sample
paths. Analogous to Section 2, let $a_{0m}(t)$, $m \in \{2, 3\}$, denote the CSHs of the process. The state space of
the process and its possible transitions are illustrated in Figure 1, top right. The survival or failure time $T$
is given as

$$T = \inf\{t \in [0, \infty) \mid X(t) \in \{2, 3\}\},$$

(3.1)

and the failure cause is $X_T \in \{2, 3\}$.

We are interested in analyzing the occurrence of event 2, which may be precluded by the occurrence
of the competing event 3. The CIF of event 2 is

$$\text{CIF}(t) = P(T \leq t, X_T = 2) = \int_{[0,t]} P(X(u) = 0) \cdot a_{02}(u)du,$$

(3.2)

which, through $P(X(u) = 0)$, depends on both CSHs. The subdistribution approach can be based on a
modified process $(\xi(t))_{t \in [0, \infty), \xi(t) \in \{0, 2\}}$ with

$$\xi(t) = 1(X(t) \neq 3) \cdot X(t) + 1(X(t) = 3) \cdot X(T-),$$

(3.3)

that is, $\xi$ is derived from $X$ by stopping (Andersen and others, 1993, p 62) the original process at $T-$ if
the competing event occurs at $T$. Definition (3.3) is interpreted as leaving an individual in the last transient
state occupied if the competing event occurs. As a consequence, the interpretation of the states changes.
The event $X(t) = 0$ has the interpretation “no terminal event by time $t$.” The event $\xi(t) = 0$ has the
interpretation “no terminal type 2 event by time \( t \).” This is meaningful, because we may now define a time \( \vartheta \) until the occurrence of event 2,
\[
\vartheta = \inf\{t \in [0, \infty) \mid \xi(t) = 2\}.
\] (3.4)
As usual, the infimum of the empty set is defined as infinity. Thus, \( \vartheta = \infty \) iff \( X_T = 3 \). For finite times \( t \), the distribution function of \( \vartheta \) now equals the CIF (3.2) of event 2.

Gray (1988) introduced the SH \( \lambda(t) \) attached to \( \vartheta \),
\[
\lambda(t)dt = P(\vartheta \in dt \mid \vartheta \geq t),
\] (3.5)
such that for finite \( t \),
\[
\text{CIF}(t) = P(\vartheta \leq t) = 1 - \exp\left(-\int_0^t \lambda(u)\,du\right).
\] (3.6)
The relationship between SH and CSH is easily seen to be (Beyersmann and Schumacher, 2007)
\[
a_{02}(t) = \left(1 + \frac{P(X(t) = 3)}{P(X(t) = 0)}\right) \cdot \lambda(t).
\] (3.7)
Equation (3.7) implies that \( \lambda(t) \leq a_{02}(t) \) and that proportional CSHs models in general imply nonproportional SHs (Latouche and others, 2007); we return to these issues in the data example.

Fine and Gray analyze the SH \( \lambda(t) \), suggesting a PSHM. Here a technical difficulty in applying standard partial likelihood arguments arises in that we will in general not know the minimum of \( \vartheta \) and \( C \) if \( \vartheta = \infty \). This problem has been solved by Fine and Gray using inverse probability of censoring weighting techniques (Robins and Rotnitzky, 1992). The partial likelihood considerations of Fine and Gray allow for a random time-dependent covariate, but a practical implementation has remained unclear (Latouche and others, 2005; Putter and others, 2007). The problem is in part an interpretational one. Say, individual \( i \) is observed to fail from the competing event 3 at time \( T_i \). Thus, \( \vartheta_i = \infty \). The partial likelihood of the PSHM will still consider individual \( i \) at risk between the “real-life failure time” \( T_i \) and the potential future censoring time \( C_i \). The question is how to define a time-dependent covariate on \( [T_i, C_i) \). In Section 4, we see how this can be done for the multistate model formulation of a time-dependent covariate, combining the latter with the modification principle toward the subdistribution process (3.3) of the present section.

4. Time-dependent binary covariates in the PSHM
We extend the original competing risks process of Section 3 by an additional transient state 1. The transient states 0 and 1 correspond to a time-dependent binary covariate as in Section 2. Slightly abusing notation, we call this process \((X(t))_{t \in [0, \infty)}\) with transition hazards \( \alpha_{lm}(t) \), \( l \neq m, l \in \{0, 1\}, m \in \{0, 1, 2, 3\} \); the state space together with the possible transitions is illustrated in Figure 2, left. The failure time \( T \) is still given by (3.1) with failure cause \( X_T \in \{2, 3\} \). Also, the CIF of event 2 is still given by the left hand side of (3.2). We may now move on to a subdistribution-type process \((\xi(t))_{t \in [0, \infty)}\), \( \xi(t) \in \{0, 1, 2\} \) as before by means of (3.3), that is, by leaving an individual in the last transient state occupied if the competing event occurs; see Figure 2, right. The time \( \vartheta \) until occurrence of event 2 is then still given by (3.4). Let \( \lambda_{lm}(t) \), \( l \neq m, l \in \{0, 1\}, m \in \{0, 1, 2\} \), denote the transition hazards of the modified process \((\tilde{\xi}(t))_{t \in [0, \infty)}\). As in (2.7), we may now consider the proportional model \( \lambda_{12}(t) = \lambda_{02}(t) \cdot \exp(b) \). Or, as explained in Section 2, we consider the time-dependent covariate \( Z(t) \) with \( Z(t) = \xi(t-) \), \( t \in [0, \vartheta] \), and the PSHM \( \lambda(t; Z(t)) = \lambda_0(t) \cdot \exp(bZ(t)) \) with \( \lambda(t) \) as in the definition (3.5). Thus, the time-dependent covariate has been naturally introduced by considering the corresponding multistate model (Section 2) and
Fig. 2. The original competing risks process \( X(t) \) with additional transient state 1 (left) modified into a subdistribution process \( \xi(t) \) (right).

modifying it into a subdistribution process (Section 3). The extension to a \( k \)-level categorical and discrete covariate is straightforward. We discuss general, real-valued time-dependent covariates in Section 5.

5. Real-valued time-dependent covariates in the PSHM

The case of a time-dependent binary covariate immediately translates to the practically important cases of a \( k \)-level categorical covariate, introducing \( k \) transient states to the multistate model, and of a discrete covariate with finite range, assuming a linearity assumption between the hazards into the absorbing state. For such data, we have argued, based on multistate models and the fact that suitably stopping the original competing risks process at time \( T - \) can be seen to be at the core of the PSHM, to analyze the time-dependent covariate

\[
Z(t) = \tilde{Z}(t \wedge T),
\]

where \( \tilde{Z}(t) \) is the original time-dependent covariate to be included in standard CSH analyses, that is, \( Z \) is derived by stopping the original covariate at \( T \), which is tantamount to stopping the related multistate process at \( T - \). By doing so, the problem discussed at the end of Section 3 disappears, as \( Z(t) \) is well defined between real-life failure time and potential future censoring time.

Before illustrating in the data example that this approach actually achieves the desired synthesis of separate CSH analyses for \( \tilde{Z}(t) \), we now argue that stopping is perfectly in accordance with the general concept of a hazard conditioned on a time-dependent covariate; in fact, this can be seen to be the reason why the proposed methodology offers the desired synthesis. Consider again the original CSHs \( \alpha_{0m}(t), \ m = 2, 3 \), conditional on a time-dependent covariate, that is,

\[
a_{0m}(t; \tilde{Z}(t) = z) = P(T \in dt, X_T = m | T \geq t, \tilde{Z}(t) = z) = P(T \in dt, X_T = m | T \geq t, \tilde{Z}(t \wedge T) = z).
\]

Note that (5.2) is not trivial. First, it reflects that looking at survival conditional on some time-dependent information is only of interest or meaningful if one is still alive. Second and more formally, there is no reason why we should not be interested in \( \{T < t, \tilde{Z}(t) = z\} \), say, instead of \( \{T \geq t, \tilde{Z}(t) = z\} \). However, only \( \{T < t, \tilde{Z}(t \wedge T) = z\} \) will in general be a legitimate expression. (Think of \( \tilde{Z}(t) \) as blood pressure, for instance, and of \( T \) as time to death. There is no blood pressure after death.) Note that (5.2) is in accordance with the notion of a hazard conditioned on a time-dependent covariate as defined by Jewell and Nielsen (1993, (2.2)) and that in the sense of (5.2), a CSH analysis of \( \tilde{Z}(t) \) is indistinguishable from analyzing \( Z(t) = \tilde{Z}(t \wedge T) \).

Let us now move to the SH for event 2, cf. (3.5), that is,

\[
\lambda(t)dt = P( T \in dt, X_T = 2 | T \geq t \text{ or } T < t, X_T = 3 ).
\]
Note that the conditioning in (5.3) should not be interpreted in the way that an individual who has failed from the competing cause 3 is considered to still be able to fail from the cause of interest 2. Rather, the additional conditioning on \( (T < t, X_T = 3) \) should be understood as a way to inflate the risk set, slowing down the SH, such that the distribution of the attached failure time \( \tilde{\vartheta} \) of (3.4) equals the CIF on the real line, cf. (3.7). Now, introduce conditioning on a time-dependent covariate in the sense of (5.2) and (5.1), that is,
\[
\lambda(t; Z(t) = z) = \lambda(t; \tilde{Z}(t \wedge T) = z) = P(T \in dt, X_T = 2 \mid T \geq t \text{ or } (T < t, X_T = 3), \tilde{Z}(t \wedge T) = z).
\]
Quantity (5.4) is valid for general, real-valued time-dependent covariates and coincides with the quantity we have derived from a multistate setup for \( k \)-level categorical and discrete covariates. We now show the usefulness of the proposed methodology in the example of Section 6.

6. Example

We consider the SIR 3 cohort study at the Charité University Hospital in Berlin, Germany, with prospective assessment of data to examine the effect of hospital-acquired infections in intensive care (Beyersmann, Gastmeier and others, 2006). Hospital-acquired infections are a major medical problem leading to an increase in both morbidity and mortality. We focus on pneumonia, one of the most frequent and severe infections. The aim of the present analysis was to study the impact of pneumonia on intensive care unit mortality, with discharge alive acting as a competing event. A total of 1876 intensive care patients admitted between February 2000 and July 2001 were included in the study cohort. Censoring was purely administrative. Overall, 214 (11.4%) patients died. In total, 30 (1.6%) observations were censored, that is, still in hospital as the study was finished. For 220 (11.7%) patients, pneumonia was diagnosed on admission. Of these, 48 (21.8%) died. In total, 158 (8.4%) patients acquired pneumonia in hospital. Of these, 33 (20.9%) died. We find that both pneumonia on admission, analyzed in Section 6.1, and hospital-acquired pneumonia, analyzed in Section 6.2, increase mortality, as indicated by the crude rates. However, the effect of these pneumonias on the cumulative incidence of mortality will be difficult to assess directly using CSH analyses but straightforward from an SH analysis. The time-independent covariate “pneumonia on admission” is included here to illustrate the usefulness of the PSHM and to better explain results for the time-dependent covariate “hospital-acquired pneumonia.” In this section, we contrast results from PCSHMs and PSHMs without further ado, although proportional CSHs will in general imply nonproportional SHs, as we have stated following (3.7). In case of a misspecified PSHM, the analysis can still be interpreted as a time-averaged hazard ratio (Struthers and Kalbfleisch, 1986; Latouche and others, 2007). We return to this issue in Section 7.

Practical implementation of SH analyses is discussed in Sections 3 and 4 of Fine and Gray (1999) and in Section 3.3.2 of Putter and others (2007). We note that the administrative censoring of our data example, where an individual’s potential censoring time is known at the start of observation, puts us in the situation of Section 3.2 of Fine and Gray (1999).

6.1 Pneumonia on admission

The aim of the present subsection was to analyze the impact of the time-independent covariate \( Z = 1(\text{Pneumonia diagnosed on admission}) \) on mortality. Using the R-package survival, we entered pneumonia status on admission into PCSHMs, allowing for different baseline CSHs for discharge and death, respectively (Therneau and Grambsch, 2000, Chapter 8.4). Baseline hazards correspond to no pneumonia on admission. We estimated the discharge CSH ratio as 0.42 with 95% confidence interval [0.37, 0.48].
The respective result for the death CSH ratio was 0.94 ([0.68, 1.30]). The results are in agreement with clinical experience. Pneumonia prolongs hospital stay, that is, reduces the combined discharge–death hazard, and since the magnitude of the effect on the discharge hazard is greater than the one on the death hazard, pneumonia increases mortality (e.g. Menendez and others, 2001; Wilson, 2001). However, the latter aspect is often hard to understand for practitioners. Looking only at the point estimates of the CSH ratios, pneumonia appears even to reduce the death CSH slightly. Still, pneumonia is not protective, because the discharge CSH is reduced much more. Pneumonia increases mortality, but death in hospital may be a bit delayed. This is well illustrated by the plots of Figure 3. The upper plots display the Nelson–Aalen estimates (Andersen and others, 1993, Chapter IV.1) of the cumulative CSHs \( \int_0^t \alpha_0(u)du \), \( i = 2, 3 \), stratified for pneumonia diagnosis on admission. In agreement with the PCSHM analyses, pneumonia status on admission leaves the CSH for death essentially unchanged but reduces the CSH for discharge.

A synopsis of these 2 separate CSH analyses is now offered by the SH of death. In the case of the time-independent covariate “pneumonia status on admission,” we may plot cumulative SHs \( \int_0^t \lambda(u)du \), stratified for pneumonia diagnosis on admission or, equivalently (see (3.6)), we may consider stratified plots of the CIF for death. We have chosen to present stratified plots of the Aalen–Johansen estimator (Andersen and others, 1993, Chapter IV.4) of the CIF in Figure 3, bottom, where we have also plotted the respective CIFs for discharge. Here we see, again in agreement with the PCSHM analyses, that the death CIF for patients with pneumonia on admission eventually runs above the one for those without, but

---

**Fig. 3.** Top: Cumulative CSHs for death and discharge, respectively, and stratified for pneumonia diagnosis on admission. Bottom: CIFs for death and discharge, respectively, and stratified for pneumonia diagnosis on admission. The solid lines are for patients without pneumonia on admission. The dashed lines are for patients with pneumonia on admission.
its increase comes a bit later. Using Robert Gray’s R-package cmprsk, we entered pneumonia status on admission into a PSHM for the SH of death. Again, no pneumonia corresponded to the baseline SH. We estimated the death SH ratio as 2.27 ([1.66, 3.10]). The SH analysis directly displays that pneumonia on admission increases intensive care unit mortality, which was obscured by the CSH analyses.

6.2 Hospital-acquired pneumonia

The aim of the present subsection was to analyze the impact of the time-dependent covariate \( Z(t) = 1(\text{Hospital-acquired pneumonia in } [0, t)) \) on mortality. We do not model recovery from pneumonia, and hence backward transitions from state 1 (hospital-acquired pneumonia) into state 0 (initial, free of hospital-acquired pneumonia) of the multistate models of Figure 2 are not possible. Including a time-dependent covariate in a hazard model does not result in a probability model, as we are only modeling transitions into the absorbing states of Figure 2 but not changes in the covariate value. Because the transition probabilities of a multistate model also depend on these changes, there is no simple probability plot like Figure 3, bottom, available for illustrating the effect of a time-dependent covariate (e.g. Beyersmann, Gerds and others, 2006). However, plots of the cumulative hazards from the multistate models of Figure 2 are still available (Andersen and others, 1993, Chapter IV.1). Figure 4, top, shows the Nelson–Aalen estimates of the cumulative CSHs \( \int_0^t g_{lm}(u)du, \ l = 0, 1, \ m = 2, 3 \) (Figure 2, left), and Figure 4, bottom, shows

Fig. 4. Top: Nelson–Aalen estimates of the cumulative CSHs of the original competing risks process into the competing absorbing states “death” and “discharge.” Bottom: Nelson–Aalen estimates of the cumulative SHs of the subdistribution process into the single absorbing state death. The left and right plots are identical except for different scales of the y-axis. The left plot uses the same scale as the CSH plots. The smaller scale of the right plot better depicts the difference between the SHs. The solid lines are cumulative hazards out of the initial state. The dashed lines are cumulative hazards out of the infectious state.
the Nelson–Aalen estimates of the cumulative SHs $\int_0^t \lambda_1(u) \, du$, $l = 0, 1$ (Figure 2, right). Figure 4, top, shows that hospital-acquired pneumonia leaves the CSH for death essentially unchanged but reduces the CSH for discharge. Like in the time-independent case, this will lead to more patients dying after hospital-acquired pneumonia, but the picture is again complex and challenging to interpret and communicate. The advantage of the new methodology is illustrated in Figure 4, bottom. The cumulative SH for death out of the infectious state clearly runs above the cumulative SH out of the initial state. Note that Figure 4, bottom, displays the same plot twice but with different scales for the y-axis. The y-axis of the left SH plot is the same as in the CSH plots. A comparison of these plots illustrates that the SH evolves at a lower magnitude than the CSH, as we have discussed following (5.3).

Formal analyses confirm the impression gathered from the plots. Using the R-package survival, we analyzed hospital-acquired pneumonia status as a time-dependent covariate (Therneau and Grambsch, 2000, Chapter 3.7.1) in PCSHMs and in a PSHM, respectively. For the CSH analyses, we estimated the death CSH ratio as 0.91 ([0.62, 1.35]) and the discharge CSH ratio as 0.59 ([0.51, 0.69]). These results are comparable to the time-independent case. In terms of the point estimates of the hazard ratios, we find that hospital-acquired pneumonia reduces the death CSH. But as the discharge CSH is reduced much more, hospital-acquired pneumonia leads to increased mortality. In addition, the effect is not very pronounced and not significant on the death CSH (in accordance with Figure 4, top left) and both pronounced and significant on the discharge CSH (in accordance with Figure 4, top right). A synthesis of the 2 separate CSH analyses is offered by the new SH analysis. We estimated the SH ratio as 3.35 ([2.29, 4.89], directly displaying the increase in mortality.

7. Discussion

The PSHM by Fine and Gray (1999) has proven useful in applications, but, so far, inclusion of random time-dependent covariates has been unclear in practice. Giving the model a natural stochastic process formulation, we have shown how time-dependent covariates can easily be treated within the SH framework. Unlike the time-independent case, the methodology does not result in a model for the CIF anymore; this is analogous to a standard survival Cox model with time-dependent covariates, which is no longer a model for the survival function (Andersen and others, 1993, Chapter VII.2.5). Nevertheless, the new methodology is helpful in that it offers a synopsis of separate CSH analyses, which will be more straightforward to interpret and communicate. This is analogous to an SH analysis with baseline covariates, with the difference that in the time-independent case such a synopsis can additionally be interpreted in terms of the CIF. This has been illustrated in our data analysis, which now allows the clinician to judge the effect of hospital-acquired pneumonia on death directly without having to take an analysis of the competing event into account.

We have not discussed goodness of fit in the analysis although, as previously mentioned, proportional CSHs will in general imply nonproportional SHs and proportional SHs will in general imply nonproportional CSHs of interest. While the choice between modeling CSHs and modeling SHs is essentially determined by the subject research question, some preliminary goodness-of-fit analysis suggested a better fit of the PCSHMs than of the PSHM. (Results are not shown here.) If one opts for the SH as the target quantity, this would call for refining the PSHM analysis; for example, our goodness-of-fit analysis and Figure 4 suggest modeling piecewise constant SH ratios, which would result in a less pronounced effect of hospital-acquired pneumonia than in the PSHM analysis. We also note that even in case of a misspecified PSHM, the SH analysis would be still meaningful (and helpful) in terms of the so-called least false parameter (Hjort, 1992). The idea here is that even if the model chosen for the analysis is misspecified, the estimator should be asymptotically consistent although not for the regression parameter of the misspecified model but for the so-called least false parameter, which can be interpreted as a time-averaged hazard ratio (Struthers and Kalbfleisch, 1986; Latouche and others, 2007).
ACKNOWLEDGMENTS

We thank an associate editor for helpful comments. Conflict of Interest: None declared.

FUNDING

Deutsche Forschungsgemeinschaft (FOR 534).

REFERENCES


[Received June 15, 2007; revised February 26, 2008; accepted for publication March 21, 2008]