Sains Malaysiana 42(5)(2013): 673–683

A Non-parametric Survival Estimate After Elimination of a Cause of Failure

(Penganggaran Kemandirian Tak-Berparameter Selepas Penghapusan Punca Risiko)

FANG YEN YEN* & SURAIYA KASSIM

ABSTRACT

In competing risks analysis, the primary interest of researchers is the estimation of the net survival probability (NSP) if a cause of failure could be eliminated from a population. The Kaplan-Meier product-limit estimator under the assumption that the eliminated risk is non-informative to the other remaining risks, has been widely used in the estimation of the NSP. The assumption implies that the hazard of the remaining risks before and after the elimination are equal and it could be biased. This paper addressed this possible bias by proposing a non-parametric multistate approach that accounts for an informative eliminated risk in the estimation procedure, whereby the hazard probabilities of the remaining risks before and after the elimination of a risk are not assumed to be equal. When a non-informative eliminated risk was assumed, it was shown that the proposed multistate estimator reduces to the Kaplan-Meier estimator. For illustration purposes, the proposed procedure was implemented on a published dataset and the change in hazard after elimination of a cause is investigated. Comparing the results to those obtained from using the Kaplan-Meier method, it was found that in the presence of (both constant and non-constant) informative eliminated risk, the proposed multistate approach was more sensitive and flexible.

Keywords: Competing risks; Kaplan-Meier estimator; latent-failure-time approach; multistate approach; net survival probability

ABSTRAK

Dalam analisis risiko bersaing, minat utama penyelidik ialah penganggaran kebarangkalian kemandirian bersih (NSP) sekiranya punca risiko boleh dihapuskan daripada satu populasi. Penganggar had-hasil darab Kaplan-Meier, dengan andaian bahawa punca risiko yang dihapuskan adalah tidak bermaklumat kepada punca risiko yang lain, telah digunakan secara meluas dalam penganggaran NSP. Andaian ini membawa implikasi bahawa kadaran bahaya baki risiko sebelum dan selepas penghapusan adalah sama dan ia mungkin tak saksama. Kertas ini menangani kemungkinan ketaksamaan ini dengan mencadangkan suatu pendekatan multi-keadaan tak-berparameter yang mengambil kira risiko dihapus yang bermaklumat dalam prosedur penganggaran, dengan kebarangkalian bahaya bagi risiko lain sebelum dan selepas penghapusan suatu risiko tidak diandaikan sama. Apabila risiko dihapus diandaikan tak bermaklumat, ditunjukkan bahawa penganggar multi-keadaan yang dicadangkan menurun kepada penganggar Kaplan-Meier. Bagi tujuan illustrasi, prosedur yang dicadangkan dilaksanakan pada satu set data yang telah diterbitkan dan perubahan kadar bahaya selepas penghapusan suatu risiko disiasat. Membandingkan keputusan yang diperoleh dengan keputusan daripada kaedah Kaplan-Meier, didapati bahawa dengan kehadiran risiko dihapus yang bermaklumat (malar dan bukan malar), pendekatan multi-keadaan yang dicadangkan adalah lebih peka dan lebih lentur.

Kata kunci: Kebarangkalian kemandirian bersih; pendekatan masa-risiko-terpendam; pendekatan multi-keadaan; penganggar Kaplan-Meier; risiko bersaing

INTRODUCTION

Competing risks survival analysis is a statistical tool for analyzing time-to-event (failure/death) data in the presence of more than one possible cause of failure (multiple events) and each cause is known as a competing risk. In the setup, the risks react simultaneously on subjects of a population and the occurrence of a cause of failure precludes the occurrence of other risks of failure (Gooley et al. 1999). Competing risks situations occur in many areas of research such as epidemiology, engineering, medical research and actuarial science.

In competing risks analysis, a crude probability is the probability of failure given that all risks are acting together. It arises directly from observations of the actual situation and can be estimated directly from the original competing risks data. However, a net probability is the probability of failure from the other remaining causes given that a cause is eliminated. It is not observable since this quantity arises only in an assumed hypothetical situation. (Chiang 1968). Thus, a further assumption is needed to get an indirect estimate.

One of the main interests in competing risks study is the estimation of the net survival probability (NSP) from

the remaining risks, after a cause of failure is eliminated from a population study (Agrawal et al. 2009; Dhar et al. 2008; Farley et al. 2001; Klein 2010; Ma & Krings 2008; Mathew & Pandey 2002). It helps to answer the important research question of what would happen to the overall survival rate of a defined population if a cause of failure in the population is eliminated, thus allowing a researcher to predict the impact of an elimination of a cause on the survival and hazard rate. Historically, the estimation of the survival rate of a population if smallpox was eliminated as a cause was done by Bernoulli in 1760 (David & Moeschberger 1978; Dietz & Heesterbeek 2002).

In competing risks survival analysis, formulation of the NSP can be done by the traditional latent-failure-times (LFT) approach or the conventional multistate Markov method. However, the NSP cannot be expressed in terms of the cause-specific hazard function (crude hazard) and is therefore, non-identifiable (Kalbfleisch & Prentice 2002). As an elimination involves hypothetically, the non-observable NSP, it is thus unestimable without additional assumptions (Crowder 1994; Tsiatis 1975). Moreover, since the observable failure time, T, is only the minimum of the failure times of m causes, i.e. $T = \min\{T_1, T_2, \ldots, T_m\}$ for each individual, any assumptions being made cannot be verified or tested experimentally based on competing risks data.

The estimation of the NSP has thus, become one of the main problems in competing risks analysis (Kalbfleisch & Prentice 2002; Prentice et al. 1978) and to date, few literatures on the problem can be found as more focus is being given to observable (crude) quantities. Discussions on the non-identifiability issues and estimation of non-observable probabilities problems can be found in Crowder (2001), Gail (1975), Kalbfleisch and Prentice (2002), Prentice et al. (1978) and Tsiatis (1975).

The issue of non-identifiability can be resolved and the NSP made identifiable only with some additional unverifiable assumptions. The most common additional assumption is that the eliminated cause is independent of the remaining causes or equivalently, that the elimination of a cause of failure does not provide any information (non-informative) to the NSP. The well-known Kaplan-Meier product limit estimator (KMPLE) (Kaplan & Meier 1958) which involves non-informative censoring (eliminated risk), works under the assumption of equal hazard before and after elimination. However, it has been argued by many researchers (Elandt-Johnson & Johnson 1980; Gail 1975, 1982; Kalbfleisch & Prentice 2002; Lawless 2003; Prentice et al. 1978) that even under an independent risks mechanism, there might be some information provided by the eliminated risk to the process to cause a change in the hazards after the elimination.

The objective of this paper was to propose a procedure that takes informative eliminated risk (IER) into account for estimating the NSP. The next section shows how the traditional LFT approach and the existing multistate Markov method are used in the formulation of the NSP, followed by the estimation procedure of the NSP with the KMPLE. Section three gives an alternative estimation procedure of the NSP using the proposed multistate approach, based on the argument of unequal hazards after the elimination of a cause of failure. The procedure shows how IER is being incorporated into the estimation process. It is an extension of the approach by Islam (1994) on estimating the survival probability in the presence of informative censoring. In section four, an investigation in the change in the hazard probabilities under different constant and non-constant IER situations is carried out. The estimated results obtained from using the proposed multistate approach are summarized in the last section.

Formulation and Estimation of the Net Survival Probability

Generally, the two approaches in the formulation of the net survival probability (NSP): the latent-failure-times (LFT) approach and the multistate Markov approach are different in their formulation, but similar in their interpretations. By the LFT approach, the m latent failure times of all existing risks, denoted as $T_1, T_2, \ldots, T_m$, are modeled and a joint distribution of the survival function, $S(t_1, \ldots, t_m)$ is assumed. By the multistate method, competing risks are modeled as a simple Markov process without consideration of the latent death times. However, without the modeling of $S(t_1, \ldots, t_m)$, the assumption used in the conventional multistate Markov method is no different than those used in the independent LFT model. Both the independent LFT model and the conventional multistate Markov model assume that the elimination of a cause reduces the hazard of only the cause to zero while leaving the hazards of the other causes unchanged. This implies that the hazard of each uneliminated risk are equal in two different set of conditions (before and after elimination of a cause).

The Latent-Failure-Time Approach

For simplicity, only three risks are considered. Assumed that each identically and independently distributed individual in a population is exposed to potential risks of failure j (or cause j), denoted as C_j, $j = 1, 2, 3$. Let random variable T_j correspond to each possible failure time due to C_j. The observable quantity T is the minimum time to failure among the LFT, that is, $T = \min\{T_1, T_2, T_3\}$ and throughout the study, only one cause of failure would happen to a subject, at the most.

We define the multivariate lifetime distribution or the joint distribution of the LFT by the joint survival function, that is:

$$S_{1,2,3}(t_1, t_2, t_3) = \mathrm{P}(T_1 > t_1, T_2 > t_2, T_3 > t_3). \quad (1)$$

The overall survival function of T in the presence of all causes is then defined as $S(t) = \mathrm{P}(T > t) = S_{1,2,3}(t,t,t)$, with the corresponding overall hazard function given by $h(t) = (-dS(t)/dt)/S(t)$. From (1), the crude cause-specific hazard (CSH) of C_j is derived as:

$$h^j(t) = \frac{-\delta S_{1,2,3}(t_1,t_2,t_3)/\delta t_j\big|_{t_1=t_2=t}}{S_{1,2,3}(t_1,t_2,t_3)_{t_1=t_2=t_3=t}}, \quad j = 1,2,3, \tag{2}$$

where $h(t) = \Sigma_j h^j(t)$, $j = 1,2,3$. The CSH is the instantaneous rate of failure from C_j when all three causes are operating simultaneously in the population. It is a basic estimable crude quantity in a competing risks framework and functions in terms of CSH are also estimable (Kalbfleisch & Prentice 2002). For example, the overall survival probability of T in terms of CSH, is given by:

$$S(t) = \exp\left(-\int_0^t h(u)du\right) = \exp\left(-\int_0^t \sum_{j=1,2,3} h^j(u)du\right), \tag{3}$$

and the crude cause-specific cumulative incidence function (CSCIF) in terms of CSH, is:

$$F^j(t) = \int_0^t h^j(u)\exp\left(-\int_0^u h(s)ds\right)du, \quad j = 1,2,3. \tag{4}$$

The CSCIF is the probability of failure from cause C_j by time t. Functions which are not in terms of CSH are inestimable (non-identifiable) without additional assumptions; the NSP is an example of such a function.

Assume a hypothetical situation where a cause of failure say, cause 1 (C_1), was eliminated. By the LFT approach, the formulation of the NSP will assume that the elimination of C_1 is equivalent to letting the crude hazard of C_1 equals to zero (i.e. $h^1(t) = 0$), without altering the hazard rate of the remaining uneliminated causes (Elandt-Johnson & Johnson 1980; Gail 1975; Hougaard 2000; Kalbfleisch & Prentice 2002; Lawless 2003; Prentice et al. 1978). To derive the NSP under the LFT model, the following three assumptions were made (Gail 1982):

AI: A structure for joint survival distribution $S_{1,2,3}(t_1, t_2, t_3)$ is assumed. AII: The effect of eliminating a risk can be expressed in terms of $S_{1,2,3}(t_1, t_2, t_3)$ in which it is to nullify the corresponding argument of $S_{1,2,3}(t_1, t_2, t_3)$. AIII: Elimination mechanism will only have effect as in AII without otherwise altering $S_{1,2,3}(t_1, t_2, t_3)$.

Assumptions AII and AIII are equivalent to assuming that the effect of eliminating C_j is just to nullify the corresponding crude hazard, $h^j(t)$ without altering the other hazards. The inappropriateness of assumptions AI-AIII in practical problems has been argued in many literatures (Cornfield 1957; Crowder 1994; Elandt-Johnson & Johnson 1980; Gail 1975, 1982; Hougaard 2000; Kalbfleisch & Prentice 2002; Klein & Moeschberger 1987; Lawless 2003; Prentice et al. 1978; Putter et al. 2007; Slud & Byar 1988).

Define $T_{2,3\bullet1} = \min\{T_2, T_3\}$ as the time to failure in the hypothetical situation that C_1 has been eliminated. Under AI-AIII, define the NSP from C_2 and C_3,

$$S_{2,3\bullet1}(t) = S_{1,2,3}(0,t2,t3)\,\big|\,t_2 = t_3 = t. \tag{5}$$

The net CSH of risk j ($j = 2, 3$) is then derived as:

$$h^j_{2,3\bullet1}(t) = \frac{-\delta S_{1,2,3}(0,t_2,t_3)/\delta t_j\big|_{t_2=t_3=t}}{S_{1,2,3}(0,t_2,t_3)\big|_{t_2=t_3=t}}, \quad j = 2,3. \tag{6}$$

In the case when the LFT of C_2 and C_3 are independent of C_1, it follows from (2) and (6) that the net CSH of risk j ($j = 2,3$) is equal to its corresponding crude CSH (Tsiatis 2005), that is:

$$h^j_{2,3\bullet1}(t) = h^j(t). \tag{7}$$

From (7), the overall net hazard after the elimination of C_1, is then obtained as $h^j_{2,3\bullet1}(t) = \Sigma_{j=2,3} h^j_{2,3\bullet1}(t) = \Sigma_j h^j(t)$. In terms of the crude CSH, the NSP as given in (5), can be written as:

$$\begin{aligned} S_{2,3\bullet1}(t) &= \exp\left(-\int_0^t \sum_j h^j_{2,3\bullet1}(u)du\right) \\ &= \exp\left(-\int_0^t \sum_j h^j(u)du\right), \quad j = 2,3. \end{aligned} \tag{8}$$

Comparing formulations (3) and (8), it can be seen that the overall net hazard is simply the sum of the crude hazards of the remaining uneliminated causes. The elimination of C_1 just nullifies hazard $h^1(t)$ without altering hazards $h^2(t)$ and $h^3(t)$. However, the condition is a stronger assumption than the statistically independent assumption since even when risks react independently, $h^2(t)$ and $h^3(t)$ might be altered in several other ways.

THE CONVENTIONAL MULTISTATE MARKOV FORMULATION

A multistate Markov model is an alternative method to the formulation of a competing risks problem (Aalen et al. 2008; Andersen et al. 2002; Chiang 1968; Kalbfleisch & Prentice 2002). It does not involve potential failure times to each cause of death and a hypothetical failure time is not assumed for each cause of death of each individual. Figure 1 shows a competing risks multistate model consisting of four possible states in a simple Markov chain. There are one transient (survivor or alive) state and three absorbing (death) states corresponding to three possible causes of death for each $C_j, j \in J = \{1,2,3\}$. Transitions are possible only from the survivor state (state 0) to the death states (state j).

Define the transition intensity from state 0 to a state $j \in J$ at time t by:

$${}^0h^j(t) = \lim_{\Delta\to\infty} \frac{\mathrm{P}\left(t \le T < t+\Delta \text{ at state } j \,\middle|\, T \ge t \text{ at state } 0\right)}{\Delta}, \quad j \in J, \tag{9}$$

which is the crude CSH function, $h^j(t)$. Define the transition probability ${}^hP^j(s, t)$ as the probability of being in state j at time t, given that the process was in state h at time s. Therefore, ${}^0P^0(0, t)$, the state occupation (survival) probability at time t, is exactly $S(t)$, i.e.

$${}^0P^0(0,t) = \exp\left(-\int_0^t \sum_j {}^0h^j(u)du\right), \quad j \in J. \tag{10}$$

The transition probability from state 0 to state j at time t is equivalent to the CSCIF, $F^j(t)$, defined in (4). It is the survival probability at a time prior to t, multiply by the conditional failure probability from C_j in a small interval $(t, t + \Delta)$, $(\approx {}^0h^j(t)\Delta)$:

$$ {}^0P^j(0, t) = \int_0^t {}^0P^0(0, u)\ {}^0h^j(u)du, \qquad j \in J. \tag{11}$$

In a competing risks multistate model, the estimation of the survival probability after an absorbing state is eliminated (the NSP), is of interest. Hoem (1969) studied the process after some possible transitions were eliminated and introduced the terminology of partial transition probability in a partial Markov process. An example of a partial Markov process is the competing risks model where a cause has been eliminated and its cause-specific transition intensity is replaced by '0' (Andersen et al. 1993).

By the Markov formulation, the transition intensity of uneliminated state j is assumed unchanged after the elimination of an absorbing state. For example, elimination of a state, say state 1 in a three competing risks situation, where only states 2 and 3 are present, we have ${}^0h^j_{2,3\bullet1}(t) = {}^0h^j(t)$ for $j = 2, 3$. The NSP after elimination of state 1 (C_1) is given by:

$$ {}^0P^0_{2,3\bullet1}(0, t) = \exp\left(-\int_0^t \sum_j {}^0h^j(u)du\right), \quad j = 2,3, \tag{12}$$

where the transition intensity of eliminated state 1 is replaced by zero (${}^0h^1(t) = 0$) and the transition intensities of uneliminated states $j = 2,3$ remain unchanged (Andersen et al. 1993; Gail 1982). Formulation (12) implies that the net hazard is equal to the crude hazard in the modeling of the NSP, similar to what is implied by (8). Thus, the conventional Markov formulation does not solve the classical competing risks problem as it is equivalent to the independent LFT model (Gail 1982; Hougaard 2000).

THE PRODUCT LIMIT ESTIMATOR

The non-parametric estimation of the NSP (8) using the Kaplan-Meier product limit estimator (KMPLE) is carried out under the assumption that the eliminated cause C_1 is a non-informative censoring observation (non-informative eliminated risk) at all the observed times, the remaining causes C_2 and C_3 are grouped together as a single cause with a similar failure pattern (Andersen et al. 1993; Gail 1982; Kaplan & Meier 1958; Tsiatis 2005). Suppose the observed failure times t_i are ordered as $t_1 < t_2 < \ldots$. The KMPLE of the NSP at times beyond time t, after the elimination of C_1 is given by (Kaplan & Meier 1958):

$$\begin{aligned} \hat{S}_{2,3\bullet1}(t) &= \prod_{i:t_i \le t}\left(1 - \sum_j \hat{h}^j(t_i)\right) \\ &= \prod_{i:t_i \le t}\left(1 - \sum_j \left(d_{ji}/n_i\right)\right), \qquad j = 2,3, \end{aligned} \tag{13}$$

where d_{ji} is the number of individuals who fail from cause C_j at time t_i and n_i is the number of individuals at risk at a time just prior to time t_i.

Alternatively, we can estimate the NSP by using the following Nelson-Aalen estimator for the cumulative CSH:

$$ \hat{H}^j(t) = \int_0^t \hat{h}^j(u)du = \sum_{i:t_i \le t}\left(d_{ji}/n_i\right), \tag{14}$$

where its increment, $d\hat{H}^j(t)$, is the empirical hazard estimate of $\hat{h}^j(t_i) = d_{ji}/n_i$ at time t_i (Davis & Lawrance 1989; Lawless 2003). From the Nelson-Aalen estimator , the estimated NSP of is given by:

$$\begin{aligned} \hat{S}_{2,3\bullet1}(t) &= \exp\left(-\sum_j \sum_{i:t_i \le t} \frac{d_{ji}}{n_i}\right) \\ &= \prod_{i:t_i \le t} \exp\left(-\sum_j \frac{d_{ji}}{n_i}\right) \\ &\approx \prod_{i:t_i \le 1}\left(1 - \sum_j \frac{d_{ji}}{n_i}\right), \qquad j = 2,3. \end{aligned} \tag{15}$$

Since $\exp\left(-\sum_j d_{ji}/n_i\right) \approx 1 - \left(\sum_j d_{ji}/n_i\right)$ for small $\Sigma_j d_{ji}/n_i$, estimator (15) is an approximate of the KMPLE (13). Thus, as the sample size increases, the KMPLE approaches a continuous distribution.

THE PROPOSED MULTISTATE APPROACH

This section shows the proposed multistate approach to estimating the net survival probability (NSP) after the elimination of a cause. The procedure is an extension of the approach by Islam (1994), whereby a competing risk that is to be eliminated is treated as a random censoring observation. Consider n number of subjects under study at the beginning time t_o. For simplicity sake, assume there are three risks, $j = 1, 2, 3$ that react simultaneously on any individual under study, in the absence of censoring. The procedure can be extended to cases with m causes with or without censoring, without loss of generality. Given a hypothetical situation that cause C_1 has been eliminated, we are interested in estimating the NSP from the remaining causes.

Consider a time interval $[0, t]$ where at each time point $t_i \in [0,t]$, before the elimination of causes, an individual (alive prior to the time point) is in one of the following four possible outcomes (or states) (Figure 1): alive as survivor; fails from cause 1 (C_1); fails from cause 2 (C_2) or fails from cause 3 (C_3).

After the elimination of cause C_1, those individuals who are saved from C_1 failure are called *cause 1-survivors* and are expected to fail from either cause C_2 or C_3 or remain alive. Overall, each individual in this hypothetical situation has six expected outcomes (or can be in six possible states) (Figure 2): alive as survivor; fails from cause 2 (C_2); fails from cause 3 (C_3); saved from eliminated cause 1 (C_1) and survives as survivor (cause 1-survivor); cause 1-survivor fails from cause 2 (C_2) or cause 1-survivor fails from cause 3 (C_3).

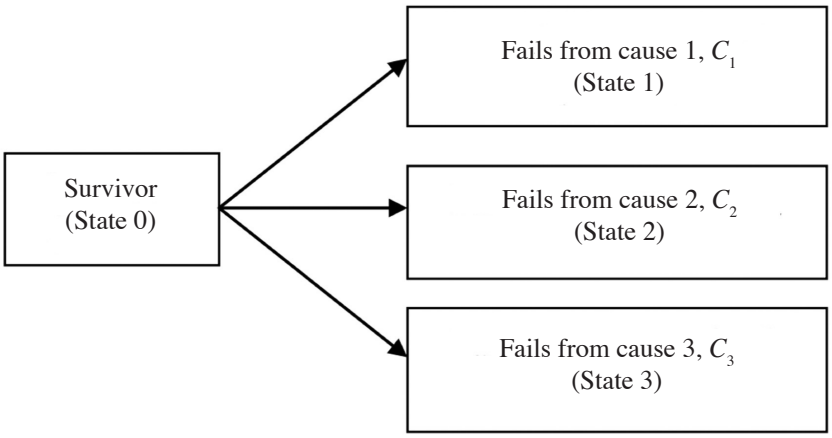

FIGURE 1. A competing risks four states model

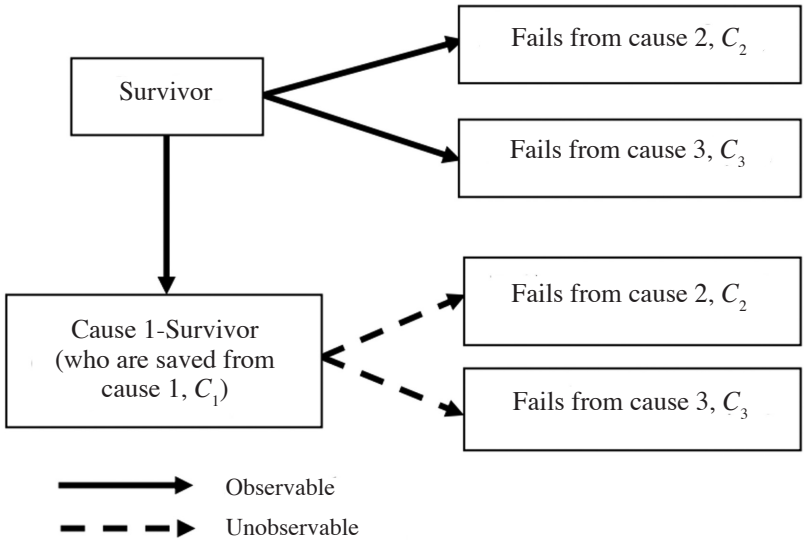

FIGURE 2. The multistate (six states) relationship situation after elimination of C_1

Outcomes (i), (ii) and (iii) are observable but outcomes (iv), (v) and (vi) are unobservable. However, an indirect estimation is possible with some assumptions.

Consider an observed survival time point t_i, given that an individual has survived up to a time prior to t_i. Denote:

${}^{S}h^{C2}(t_i)$ and ${}^{S}h^{C3}(t_i)$ as the conditional probability of transiting from the survivor state to C_2 and C_3 failure states, respectively;

${}^{S}h^{S1}(t_i)$ as the conditional probability of transiting from the survivor state to cause 1-survivor state after the elimination of C_1 (the hazard probability), i.e. ${}^{S}h^{S1}(t_i) = {}^{S}h^{C1}(t_i)$ since all individuals saved from C_1 is assumed to be the cause 1-survivor;

${}^{S1}h^{C2}(t_i)$ and ${}^{S1}h^{C3}(t_i)$ as the conditional probability of transiting from cause 1-survivor state to C_2 and C_3 failure states, respectively.

Let $S^{S}(t)$ and $S^{S1}(t)$ be the survival probability in the survivor state and cause 1-survivor state, respectively. Assuming $t_0 = 0$, we have $S^{S}(0) = 1$ and $S^{S1}(0) = 0$. Assume that after the elimination of C_1, a cause 1-survivor can only fail at the next observed failure time point when the individual enter the cause 1-survivor state. For the time interval $[0,t]$, let ${}^{S}F^{Cj}(t)$, j = 1,2,3 be the probability of failure from C_j at a survivor state by time t (discrete type CSCIF as defined in and):

$$ {}^{S}F^{Cj}(t) = \sum_{i:t_i \le t} {}^{S}h^{Cj}(t_i) S^{S}(t_i^-) \quad j = 1,2,3, \tag{16} $$

where $S^{S1}(t_i^-)$ is the probability of surviving at a survivor state at time prior to t_i. Let ${}^{S1}F^{C2}(t)$ and ${}^{S1}F^{C3}(t)$ be defined as the hypothetical probability of failure from C_2 and C_3, respectively, at cause 1-survivor state by time t:

$$ {}^{S1}F^{Cj}(t) = \sum_{i:t_i \le t} {}^{S1}h^{Cj}(t_i) S^{S1}(t_i^-) \quad j = 2,3, \tag{17} $$

where $S^{S1}(t_i^-)$ is the probability of surviving at cause 1-survivor state at time prior to t_i.

From Figure 1, it can be seen that the (crude) probability of surviving in the interval [0,*t*] at the survivor state is,

$$S^{S}(t) = 1 - {}^{S}F^{C1}(t) - {}^{S}F^{C2}(t) - {}^{S}F^{C3}(t). \quad (18)$$

The probability of surviving in the interval [0,*t*] at cause 1-survivor state is:

$$S^{S1}(t) = {}^{S}F^{C1}(t) - {}^{S1}F^{C2}(t) - {}^{S1}F^{C3}(t). \quad (19)$$

In (18), ${}^{S}F^{C1}(t)$ represents the probability of failure from C_1 before its elimination, while in (19), ${}^{S}F^{C1}(t)$ represents the probability of being saved from eliminated C_1. The NSP after the elimination of C_1 during the interval [0, *t*] is therefore, the sum of two types of survivor probabilities given in (18) and (19), i.e.,

$$\begin{aligned} S_{2,3\bullet1}(t) &= S^{S}(t) + S^{S1}(t) \\ &= 1 - {}^{S}F^{C2}(t) - {}^{S}F^{C3}(t) - {}^{S1}F^{C2}(t) - {}^{S1}F^{C3}(t). \end{aligned} \quad (20)$$

From (16) and (17), the NSP of is therefore:

$$\begin{aligned} S_{2,3\bullet1}(t) = 1 &- \sum_{j}\sum_{i:t_i\le t} {}^{S}h^{Cj}(t_i)\, S^{S}(t_i^-) \\ &- \sum_{j}\sum_{i:t_i\le t} {}^{S1}h^{Cj}(t_i)\, S^{S1}(t_i^-), \quad j = 2,3 \end{aligned} \quad (21)$$

The ${}^{S1}h^{Cj}(t_i)$ in (21) are non-observable and thus, only indirectly estimate from observed quantities, i.e. ${}^{S}h^{Cj}(t_i)$.

The assumption of non-informative eliminated risk from the KMPLE implies that the eliminated risk do not provide any information to the survival probability and that the conditional probability of failure (from C_2 and C_3) is the same in both the actual survivor and cause 1-survivor states (i.e. ${}^{S1}h^{Cj}(t_i) = {}^{S}h^{Cj}(t_i), j = 2,3$. The results from making such an assumption might, however, be biased due to the complexity of some practical real problem. Thus, there is a need to account for the possibility of unequal probability of failure at the actual survivor state and cause 1-survivor state, after the elimination of C_1.

The possibility of a bias is addressed in the proposed procedure by the assumption that the conditional probability of failure (from C_2 and C_3) at the actual survivor state and cause 1-survivor state, are proportional to each other at each observed time point, i.e. ${}^{S1}h^{Cj}(t_i) \propto {}^{S}h^{Cj}(t_i)$, $j = 2,3$ through an adjustment factor. The idea, which is an extension of Islam's (1994) work, makes the procedure of estimating the NSP, after the elimination of C_1, more flexible.

Letting a_i and b_i be adjustment factors at the i^{th} time point, we have:

$${}^{S1}h^{C2}(t_i) = a_i\, {}^{S}h^{C2}(t_i) \text{ and } {}^{S1}h^{C3}(t_i) = b_i\, {}^{S}h^{C3}(t_i). \quad (22)$$

From (21) and (22), we obtain:

$$\begin{aligned} S_{2,3\bullet1}(t) = 1 &- \sum_{j=2,3}\sum_{i:t_i\le t} {}^{S1}h^{Cj}(t_i)\, S^{S}(t_i^-) \\ &- \sum_{i:t_i\le t} a_i\, {}^{S}h^{C2}(t_i) S^{S1}(t_i^-) \\ &- \sum_{i:t_i\le t} b_i\, {}^{S}h^{C3}(t_i)\, S^{S1}(t_i^-). \end{aligned} \quad (23)$$

The adjustment factor measures the extent of the effect of a specific elimination on deaths among the hypothetical survivors at each lifetime. Its value may be varied at different time points and/or for each eliminated risk, depending on the researcher's need or background information of the process under study. Thus, it can take on any positive value or it can be zero.

ESTIMATION OF THE NET SURVIVAL PROBABILITY

The transitions of states at each time t_i in the survivor state, before the elimination of any causes, follow a multinomial distribution. In the absence of censoring, the general form of the likelihood function for a multinomial distribution is:

$$L \,\alpha\, \Pi_j\, {}^{S}h^{Cj}(t_i)^{d_{ji}}(1 - \Sigma_j\, {}^{S}h^{Cj}(t_i)^{n_i - \Sigma_j d_{ji}}, \quad j = 1,2,3.$$

The estimates of ${}^{S}h^{Cj}(t_i)$ are:

$${}^{S}h^{Cj}(t_i) = d_{ji}/n_i, \quad j = 1,2,3, \quad (24)$$

where d_{ji} is the observable number of individuals whose failure is from C_j at time t_i, and n_i is the number of individuals at risk, just prior to time t_i. To estimate the unobservable conditional probabilities, the following indirect estimators form are used:

$${}^{S1}\hat{h}^{C2}(t_i) = a_i\left(d_{2i}/n_i\right) \text{ and}$$

$${}^{S1}\hat{h}^{C3}(t_i) = b_i\left(d_{3i}/n_i\right). \quad (25)$$

Substituting the estimators in (25) into (23), the estimated net survival probability becomes:

$$\begin{aligned} \hat{S}_{2,3\bullet1}(t) = 1 &- \sum_{j=1,2,3}\sum_{i:t_i\le t}\left(d_{ji}/n_i\right)\hat{S}^{S}\left(t_i^-\right) \\ &- \sum_{i:t_i\le t} a_i\left(d_{2i}/n_i\right)\hat{S}^{S1}\left(t_i^-\right) \\ &- \sum_{i:t_i\le t} b_i\left(d_{3i}/n_i\right)\hat{S}^{S1}\left(t_i^-\right). \end{aligned} \quad (26)$$

where:

$$\begin{aligned} \hat{S}^{S}(t) &= 1 - \sum_{j=1,2,3}\sum_{i:t_i\le t}\left(d_{ji}/n_i\right)\hat{S}^{S}\left(t_i^-\right), \\ \hat{S}^{S1}(t) &= \sum_{i:t_i\le t}\left(d_{1i}/n_i\right)\hat{S}^{S}\left(t_i^-\right) \\ &\quad - \sum_{i:t_i\le t}\left(\left(a_i d_{2i} + b_i d_{3i}\right)/n_i\right)\hat{S}^{S1}\left(t_i^-\right). \end{aligned}$$

When $a_i = b_i = 1$ for all t_i, i.e. ${}^{S1}h^{C2}(t_i)$ and ${}^{S1}h^{C3}(t_i) = {}^{S}h^{C3}(t_i)$, the above condition reduces to the assumption used by the KMPLE and is reduced to (Appendix I). If at least one of the adjustment factors is not equal to at least one of the t_i's ($a_i \neq 1$ and/or $b_i \neq 1$), then ${}^{S1}h^{C2}(t_i) \neq {}^{S}h^{C2}(t_i)$ and/or ${}^{S1}h^{C3}(t_i) \neq {}^{S}h^{C3}(t_i)$. This implies that eliminated C_1 do provide some information to the NSP through the informative rates a_i and b_i, and the condition is known as the *informative eliminated risks* (IER) situation.

If the adjustment factors are equal for different risks (i.e. $a_i = b_i \neq 1$) and is a constant throughout the whole lifetime, the condition is known as a constant IER situation. However, it is a non-constant situation if the adjustment factors have different values over different risks (i.e. $a_i \neq b_i$) for at least one time point and/or if any of the adjustment factor has a non-fixed value over time. Thus, for both constant and non-constant IER situations, the adjustment factor makes the proposed method a useful informative tool in estimating the NSP.

Illustrative Example

To illustrate the proposed multistate approach, Hoel's (1972) published data set is used. It was a result from a laboratory experiment on two groups (germ-free group and control group) of male mice which had been subjected to a radiation dose of 300r at the age 5 to 6 weeks. The number of days until death and the corresponding cause of death were recorded for each mice. There were two major causes (types of cancer tumour) of death called thymic lymphoma (C_1) and reticulum cell sarcoma (C_2), while other causes of death were combined into a single group and referred as the third cause (C_3). For illustration purpose, only the data from germ-free environmental group is used. There is no censoring involved because all of the mice died by the end of the experiment.

Applying the KMPLE method and the proposed multistate formulation on the data set, respectively, we obtain estimates of the NSP after the elimination of C_3 and its corresponding estimated overall net hazard probability at times 100, 200, … 1100 days. Let a_i and b_i be a known informative rate provided from eliminated C_3 to C_1 and C_2, respectively. Table 1 compares the result from the KMPLE method against the proposed multistate method with adjustment factors $a_i = b_i = 1$. The results showed that there was no difference in the estimated NSP and its corresponding estimated overall net hazard from using the two different approaches. It can also be seen from Table 1 that the overall net hazard is just simply the sum of the crude hazards of the remaining uneliminated risks, $\hat{h}^1(t_i)$ and $\hat{h}^2(t_i)$ for both methods, i.e. $\hat{h}_{1,2\bullet 3}(t_i) = \hat{h}^1(t_i) + \hat{h}^2(t_i)$.

Under the multistate method, if the adjustment factors $a_i \neq 1$ and $b_i \neq 1$, then it is implied that IER is present and the overall net hazard is not just simply the sum of the crude CSHs. This fact is illustrated in Table 2 where the estimated NSP and the overall net hazards for which $a_i \neq 1$ and $b_i \neq 1$, are displayed. It can be seen that the estimated NSP and overall net hazard by the multistate approach are different from those of the KMPLE method. When the values of a_i and b_i are smaller than one ($a_i < 1$ and $b_i < 1$), the estimated overall net hazard by the multistate approach is less than or equal to those of the KMPLE, while the estimated NSP is greater.

In addition, when a_i and b_i are greater than one ($a_i > 1$ and $b_i > 1$), the estimated overall net hazard by the multistate approach is greater than or equal to those from the KMPLE, while the estimated NSP is smaller. The net CSH of C_1 and C_2 are different from its corresponding crude CSH (implying the change of hazard) after the elimination of C_3, i.e. $\hat{h}^j_{1,2\bullet 3}(t_i) \neq \hat{h}^j(t_i), j = 1,2$ (for mathematical proof, see Appendix II) and the overall net hazard is not just simply the sum of the crude CSH of uneliminated risks

TABLE 1. Estimated net survival probabilities, $\hat{S}_{1,2\bullet 3}(t)$ and net overall hazard, $\hat{h}_{1,2\bullet 3}(t)$ after elimination of cause 3: The proposed multistate approach and the product limit method

	KMPLE †		Multistate approach with a_i=b_i=1			
Time, t (days)	$\hat{S}_{1,2\bullet 3}(t)$	$\hat{h}_{1,2\bullet 3}(t) = \hat{h}^1(t) + \hat{h}^2(t)$‡	$\hat{S}^S(t)$ §	$\hat{S}^{S3}(t)$ ¶	$\hat{S}_{1,2\bullet 3}(t)$	$\hat{h}_{1,2\bullet 3}(t)$
≤100	1.00000	0.00000	1.00000	0.00000	1.00000	0.00000
≤200	0.93902	0.06098	0.92683	0.01220	0.93902	0.06098
≤300	0.80311	0.14474	0.76829	0.03482	0.80311	0.14474
≤400	0.76487	0.04762	0.71951	0.04536	0.76487	0.04762
≤500	0.68709	0.10169	0.63415	0.05294	0.68709	0.10169
≤600	0.64745	0.05769	0.58537	0.06208	0.64745	0.05769
≤700	0.53954	0.16667	0.37805	0.16149	0.53954	0.16667
≤800	0.43511	0.19355	0.21951	0.21560	0.43511	0.19355
≤900	0.41094	0.05556	0.07317	0.33777	0.41094	0.05556
≤1000	0.34245	0.16667	0.02439	0.31806	0.34245	0.16667
≤1100	0.34245	0.00000	0.00000	0.34245	0.34245	0.00000

† Kaplan-Meier product limit estimator
‡ estimated crude cause-specific hazard of cause 1 and cause 2
§ estimated survival probability in survivor state
¶ estimated survival probability in cause 3-survivor state

TABLE 2. Estimated net survival probability, $\hat{S}_{1,2\bullet 3}(t)$ and the estimated overall net hazard, $\hat{h}_{1,2\bullet 3}(t)$ based on the assumptions $a_i=b_i=0.5$ and $a_i=b_i=1.5$ using the proposed multistate method and its comparison to the product limit estimator

time, t (days)	The Multistate Approach				KMPLE†	
	Assumed $a_i=b_i=0.5$		Assumed $a_i=b_i=1.5$			
	$\hat{S}_{1,2\bullet 3}(t)$	$\hat{h}_{1,2\bullet 3}(t)$	$\hat{S}_{1,2\bullet 3}(t)$	$\hat{h}_{1,2\bullet 3}(t)$	$\hat{S}_{1,2\bullet 3}(t)$	$\hat{h}_{1,2\bullet 3}(t)$
≤100	1.00000	0.00000	1.00000	0.00000	1.00000	0.00000
≤200	0.93902	0.06098	0.93902	0.06098	0.93902	0.06098
≤300	0.80400	0.14380	0.80223	0.14568	0.80311	0.14474
≤400	0.76656	0.04656	0.76322	0.04863	0.76487	0.04762
≤500	0.69100	0.09857	0.68338	0.10461	0.68709	0.10169
≤600	0.65277	0.05532	0.64254	0.05977	0.64745	0.05769
≤700	0.54959	0.15806	0.53068	0.17408	0.53954	0.16667
≤800	0.45982	0.16334	0.41320	0.22138	0.43511	0.19355
≤900	0.44095	0.04104	0.38486	0.06858	0.41094	0.05556
≤1000	0.39811	0.09716	0.29475	0.23416	0.34245	0.16667
≤1100	0.39811	0.00000	0.29475	0.00000	0.34245	0.00000

† Kaplan-Meier product limit estimator

(i.e. $\hat{h}^{j}_{1,2\bullet 3}(t_i) \neq \hat{h}^{1}(t_i)+\hat{h}^{2}(t_i)$). These implied that the KMPLE might under-estimate the survival probability and over-estimate the hazard rate in the presence of IER C_3, i.e. $a_i = b_i < 1$. However, the KMPLE might over-estimate the survival probability and under-estimate the hazard rate when $a_i = b_i > 1$.

To demonstrate the flexibility of the proposed multistate procedure, a non-constant adjustment factor is incorporated into it, to take into account a non-constant IER situation. For this purpose, consider two assumptions: (i) $a_i = b_i = 0.5$ for i:$t_i \leq 700$ and $a_i = b_i = 1.5$ for i:$t_i > 700$, (ii) $a_i = 0.5$, $\forall i$ and $b_i = 1$, $\forall i$. Table 3 shows the estimated NSP and its corresponding estimated net overall hazard based on the non-constant IER assumptions (i) and (ii) above. It can be seen that the estimated overall hazard probability is not equal to the sum of the crude hazard, implying a change in hazard after the elimination of a cause and that the estimated NSP by the proposed procedure is different from the KMPLE. Thus, even in the presence of non-constant IER situation, the proposed procedure is a more flexible method as compared with the product limit method. It is a more advanced method that it allows the investigation of the change in hazard after the elimination of a cause.

DISCUSSION

The proposed multistate method is shown to be more flexible and comprehensive in estimating the NSP as compared with the product limit method as it allows for various types of IER situation to be considered, including constant and non-constant IER situations. The proposed procedure produces estimates of the NSP that are more appropriate since the change of hazard is considered in the estimation procedure.

The adjustment factor that is incorporated in the proposed multistate formulation is a flexible parameter that defines how deaths among saved individuals (from eliminated risks) can happen, after elimination of a cause. As it can be varied over time and/or over different risk (a non-constant IER) situations, the assumption about the deaths among saved individuals can thus, be specified accordingly. This makes the proposed method more advanced than the product limit method in that, uneliminated risks are neither combined nor treated as a group of risks with the same behavior over different time. In addition, the proposed multistate method allows saved individuals to die from not only within the next time interval, but also at any time interval after the elimination of a cause. Therefore, any underlying elimination effect at each time point is not hidden. This facility is an advantageous feature of the proposed method over the conventional ones.

Furthermore, the proposed multistate method is simpler in form and therefore, easier to handle than the product limit method. This is because, unlike the product limit method, the formulation of the proposed method does not involve the multiplicative of the conditional sub-survival function (crude cause-specific survival probability), which is an improper function. Also, it does not assume a fixed structure of the multivariate distribution, unlike the conventional LFT approach. Hence, we can avoid making strong unverifiable assumption about the nature of the failure mechanism.

However, when applying the proposed procedure, care has to be taken in determining the value of the adjustment factor to obtain optimum results. Good background information and prior knowledge of the process under study is necessary for this task. Currently, the optimal value of the adjustment factor is determined subjectively by different experts or researchers who have different opinions and information about a process. In view of the inconsistencies, a standard procedure on how to choose

TABLE 3. Estimated net survival probability, $\hat{S}_{1,2\bullet 3}(t)$ and overall net hazard, $\hat{h}_{1,2\bullet 3}(t)$ based on the non-constant informative eliminated risk assumption using the proposed multistate method and its comparison to the product limit estimator

time, t (days)	The multistate approach: $a_i = b_i = 1.5$ for $t_i > 700$; $a_i = b_i = 0.5$ for $t_i \le 700$;		The multistate approach: $a_i = 0.5$; $b_i = 1$		KMPLE†	
	$\hat{S}_{1,2\bullet 3}(t)$	$\hat{h}_{1,2\bullet 3}(t)$	$\hat{S}_{1,2\bullet 3}(t)$	$\hat{h}_{1,2\bullet 3}(t)$	$\hat{S}_{1,2\bullet 3}(t)$	$\hat{h}_{1,2\bullet 3}(t)$
≤100	1.00000	0.00000	1.00000	0.00000	1.00000	0.00000
≤200	0.93902	0.06098	0.93902	0.06098	0.93902	0.06098
≤300	0.80400	0.14380	0.80400	0.14380	0.80311	0.14474
≤400	0.76656	0.04656	0.76656	0.04656	0.76487	0.04762
≤500	0.69100	0.09857	0.69060	0.09909	0.68709	0.10169
≤600	0.65277	0.05532	0.65184	0.05612	0.64745	0.05769
≤700	0.54959	0.15806	0.54389	0.16560	0.53954	0.16667
≤800	0.42662	0.22375	0.44397	0.18371	0.43511	0.19355
≤900	0.39717	0.06904	0.41931	0.05556	0.41094	0.05556
≤1000	0.30397	0.23465	0.34942	0.16667	0.34245	0.16667
≤1100	0.30397	0.00000	0.34942	0.00000	0.34245	0.00000

† Kaplan-Meier product limit estimator

the value of the adjustment factor need to be developed to ensure the robustness of the results. Alternatively, one can place a possible range of values to the adjustment factor to study the bounds of estimated NSP.

Another limitation of the proposed method is that it does not cater for the presence of all possible effects of covariates. A procedure that allows for adjustments of covariates effects is important since competing risks data are always accompanied by several covariates. Possible research can be done on extensions of the proposed multistate approach to the regression problem in the examination of the effects of covariate on the net survival probability.

CONCLUSION

In this paper we proposed a multistate approach to estimate the net survival probability (NSP) in a competing risks problem. Adjustment factors are incorporated into the procedure to account for the presence of informative eliminated risks (IER), for both constant and non-constant cases. The conventional product limit formulation with non-informative eliminated risks assumption is a special case of the proposed methodology. Results from applying the procedure allow evaluating the change of hazard before and after the elimination of a cause, which is contradictory to the results from the KMPLE which showed equal hazard. In addition, the proposed multistate approach was more sensitive to the presence of IER since the independence assumption was not made. The underlying impact of elimination can be studied in detail by considering each interval or time point situation, presenting results that otherwise, cannot be given by traditional approaches.

The proposed procedure can be readily extended to studies that involve more than three risks, with or without the consideration of censoring (both informative and non-informative), a non-complete (partial) elimination process or a more complex multistate elimination study. It can be used to study the gain in life after the elimination of a cause of death, as well as to study the loss in life if a new risk of death should exist. In conclusion, the proposed multistate approach is a simple yet advanced estimation procedure that can be used as an alternative approach to the traditional approaches in estimating the net survival probability, especially when a known IER is provided.

APPENDIX I

From (26), when $a_i = b_i = 1$, we have:

$$\begin{aligned}
\hat{S}_{2,3\bullet 1}(t) &= 1 - \sum_{j=2,3}\sum_{i:t_i \le t} (d_{ji}/n_i)\hat{S}^{S}(t_i^-) - \sum_{i:t_i \le t} (d_{2i}/n_i)\hat{S}^{S1}(t_i^-) \\
&\quad - \sum_{i:t_i \le t} (d_{3i}/n_i)\hat{S}^{S1}(t_i^-) \\
&= 1 - \sum_{i:t_i \le t} (d_{2i}/n_i)\left(\hat{S}^{S1}(t_i^-) + \hat{S}^{S}(t_i^-)\right) \\
&\quad - \sum_{i:t_i \le t} (d_{3i}/n_i)\left(\hat{S}^{S1}(t_i^-) + \hat{S}^{S}(t_i^-)\right) \\
&= 1 - \sum_{i:t_i \le t} \left(\sum_j d_{ji}/n_i\right)\left(\hat{S}^{S1}(t_i^-) + \hat{S}^{S}(t_i^-)\right) \\
&= 1 - \sum_{i:t_i \le t} \left(\sum_j d_{ji}/n_i\right)\left(\hat{S}_{2,3\bullet 1}(t_i^-)\right) \\
&= -\sum_j d_{j1}/n_1 \left(\hat{S}_{2,3\bullet 1}(t_0)\right) - \sum_j d_{j2}/n_2 \left(\hat{S}_{2,3\bullet 1}(t_1)\right) \\
&\quad - \sum_j d_{j3}/n_3 \left(\hat{S}_{2,3\bullet 1}(t_2)\right) - \ldots \\
&= 1 - \sum_j d_{j1}/n_1 - \sum_j d_{j2}/n_2 \left(1 - \sum_j d_{j1}/n_1\right) \\
&\quad - \sum_j d_{j3}/n_3 \left(\hat{S}_{2,3\bullet 1}(t_2)\right) - \ldots
\end{aligned}$$

$$
\begin{aligned}
&= \left(1-\sum_j d_{j1}/n_1\right)-\sum_j d_{j2}/n_2\left(1-\sum_j d_{j1}/n_1\right)\\
&\quad -\sum_j d_{j3}/n_3\left(\left(1-\sum_j d_{j1}/n_1\right)\right.\\
&\quad \left.-\sum_j d_{j2}/n_2\left(1-\sum_j d_{j1}/n_1\right)\right)-\ldots\\
&= \left(1-\sum_j d_{j1}/n_1\right)\left(1-\sum_j d_{j2}/n_2\right)\\
&\quad \left(1-\sum_j d_{j3}/n_3\right)\ldots\\
&= \prod_{i:t_i\le t}\left(1-\sum\nolimits_{j=2,3} d_{j1}/n_i\right), \qquad j=2,3
\end{aligned}
$$

Thus, it has been shown that (26) from the proposed multistate formulation is reduced to (13) of the Kaplan-Meier product limit estimator when non-informative eliminated risk is assumed, i.e. $a_i = b_i = 1$.

APPENDIX II

From (26), we prove that there is change of hazard after the elimination of a cause of failure, C_3 in the presence of informative eliminated risk. After elimination of C_3, the estimated net survival probability using the proposed formulation is given by:

$$
\begin{aligned}
\hat{S}_{1,2\bullet 3}(t) &= 1-\sum_{j=1,2}\sum_{i:t_i\le t}\left(d_{ji}/n_i\right)\hat{S}^S\left(t_i^-\right)-\sum_{i:t_i\le t}\left(a_i d_{2i}/n_i\right)\hat{S}^{S3}\left(t_i^-\right)\\
&\quad -\sum_{i:t_i\le t}\left(b_i d_{2i}/n_i\right)\hat{S}^{S3}\left(t_i^-\right)\\
&= 1-\sum_{i:t_i\le t}\left(d_{1i}/n_i\right)\hat{S}^S\left(t_i^-\right)-\sum_{i:t_i\le t}\left(d_{1i}/n_i\right)a_i\hat{S}^{S3}\left(t_i^-\right)\\
&\quad -\sum_{i:t_i\le t}\left(d_{2i}/n_i\right)\hat{S}^S\left(t_i^-\right)-\sum_{i:t_i\le t}\left(d_{2i}/n_i\right)b_i\hat{S}^{S3}\left(t_i^-\right)\\
&= 1-\sum_{i:t_i\le t}\left(d_{1i}/n_i\right)\left(\hat{S}^S(t_i^-)+a_i\hat{S}^{S3}(t_i^-)\right)\\
&\quad -\sum_{i:t_i\le t}\left(d_{2i}/n_i\right)\left(\hat{S}^S(t_i^-)+b_i\hat{S}^{S3}(t_i^-)\right)\\
&= 1-\sum_{i:t_i\le t}\frac{\left(\hat{S}^S\left(t_i^-\right)+a_i\hat{S}^{S3}\left(t_i^-\right)\right)}{\hat{S}_{1,2\bullet 3}\left(t_i^-\right)}\left(d_{1i}/n_i\right)\hat{S}_{1.2\bullet 3}\left(t_i^-\right)\\
&\quad -\sum_{i:t_i\le t}\frac{\left(\hat{S}^S\left(t_i^-\right)+b_i\hat{S}^{S3}\left(t_i^-\right)\right)}{S_{1,2\bullet 3}\left(t_i^-\right)}\left(d_{2i}/n_i\right)\hat{S}_{1.2\bullet 3}\left(t_i^-\right)
\end{aligned}
$$

and thus, the estimated net cause-specific hazard for each C_1 and C_2 are:

$$
\hat{h}^1_{1,2\bullet 3}(t_i) = \frac{\left(\hat{S}^S\left(t_i^-\right)+a_i\hat{S}^{S3}\left(t_i^-\right)\right)}{\hat{S}_{1,2\bullet 3}\left(t_i^-\right)}\left(d_{1i}/n_i\right) \text{ and}
$$

$$
\hat{h}^2_{1,2\bullet 3}(t_i) = \frac{\left(\hat{S}^S\left(t_i^-\right)+b_i\hat{S}^{S3}\left(t_i^-\right)\right)}{\hat{S}_{1,2\bullet 3}\left(t_i^-\right)}\left(d_{2i}/n_i\right),
$$

where the estimated overall net hazard,

$$
\begin{aligned}
\hat{h}^1_{1,2\bullet 3}(t_i) &= \hat{h}^1_{1,2\bullet 3}(t_i)+h^2_{1,2\bullet 3}(t_i)\\
&\quad \frac{\left(\hat{S}^S\left(t_i^-\right)+a_i\hat{S}^{S3}\left(t_i^-\right)\right)}{\hat{S}_{1,2\bullet 3}\left(t_i^-\right)}\left(d_{1i}/n_i\right)+\frac{\left(\hat{S}^S\left(t_i^-\right)+b_i\hat{S}^{S3}\left(t_i^-\right)\right)}{\hat{S}_{1,2\bullet 3}\left(t_i^-\right)}\left(d_{2i}/n_i\right)\\
&\neq \hat{h}^1(t_i)+\hat{h}^2(t_i),
\end{aligned}
$$

which is not equal to the sum of the estimated crude cause-specific hazard of C_1 and C_2.

ACKNOWLEDGEMENT

This research work has been done with the support of USM Fellowship Scholarship Scheme, Universiti Sains Malaysia. We would like to thank Prof. M. Ataharul Islam for his valuable comments on the initial idea of the proposed method.

REFERENCES

Aalen, O., Borgan, Ø. & Gjessing, H.K. 2008. *Survival and Event History Analysis: A Process Point of View*. New York: Springer.

Agrawal, V., Marinescu, V., Agarwal, M. & McCullough, P.A. 2009. Cardiovascular implications of proteinuria: An indicator of chronic kidney disease. *Nat. Rev. Cardiol.* 6(4): 301-311.

Andersen, P.K., Abildstrom, S.Z. & Rosthoj, S. 2002. Competing risks as a multi-state model. *Stat. Methods Med. Res.* 11(2): 203-215.

Andersen, P.K., Borgan, O., Gill, R.D. & Keiding, N. 1993. *Statistical Models Based on Counting Processes*. New York: Springer Verlag.

Chiang, C.L. 1968. *Competing Risks: Introduction to Stochastic Processes in Biostatistics*. New York: John Wiley & Sons.

Cornfield, J. 1957. The estimation of the probability of developing a disease in the presence of competing risks. *Am. J. Public Health Nations Health* 47(5): 601-607.

Crowder, M. 1994. Identifiability crises in competing risks. *International Statistical Review* 62(3): 379-391.

Crowder, M. 2001. *Classical Competing Risks*. Boca Raton: Chapman & Hall/CRC.

David, H.A. & Moeschberger, M.L. 1978. *The Theory of Competing Risks*. London: Griffin.

Davis, T.P. & Lawrance, A.J. 1989. The likelihood for competing risk survival analysis. *Scand. J. Statist.* 16: 23-28.

Dhar, M., Lahiri, S., Vijayasimha, R., Ashok, N. & Murthy, N. 2008. Study of social loss or gain from a particular cause: Scope for the developing countries - A review with reference to cancer. *Asian Pacific Journal of Cancer Prevention* 9(4): 797-804.

Dietz, K. & Heesterbeek, J.A.P. 2002. Daniel Bernoulli's epidemiological model revisited. *Mathematical Biosciences* 180(1-2): 1-21.

Elandt-Johnson, R.C. & Johnson, N.L. 1980. *Survival Models and Data Analysis*. New York: John Wiley & Sons.

Farley, T.M., Ali, M.M. & Slaymaker, E. 2001. Competing approaches to analysis of failure times with competing risks. *Stat. Med.* 20(23): 3601-3610.

Gail, M. 1975. A review and critique of some models used in competing risk analysis. *Biometrics* 31(1): 209-222.

Gail, M. 1982. Competing risks. In *Encyclopedia of Statistical Sciences* edited by Kotz S. & Johnson, N.L. New York: Wiley.

Gooley, T.A., Leisenring, W., Crowley, J. & Storer, B.E. 1999. Estimation of failure probabilities in the presence of competing risks: New representations of old estimators. *Stat. Med*. 18(6): 695-706.

Hoel, D.G. 1972. A representation of mortality data by competing risks. *Biometrics* 28(2): 475-488.

Hoem, J.M. 1969. Purged and partial markov chains. *Scand. Aktuarietidskrift* (52): 147-155.

Hougaard, P. 2000. Competing risks model. *Analysis of Multivariate Survival Data*. New York: Springer.

Islam, M.A. 1994. An alternative to the product limit method: The multistate approach. *Journal of the Royal Statistical Society. Series D (The Statistician)* 43(2): 237-245.

Kalbfleisch, J.D. & Prentice, R.L. 2002. *The Statistical Analysis of Failure Time Data*. 2nd ed. New York: Wiley.

Kaplan, E.L. & Meier, P. 1958. Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association* 53(282): 457-481.

Klein, J.P. 2010. Competing risks. *Wiley Interdisciplinary Reviews: Computational Statistics* 2(3): 333-339.

Klein, J.P. & Moeschberger, M.L. 1987. Independent or dependent competing risks: Does it make a difference. *Communications in Statistics - Simulation and Computation* 16(2): 507-533.

Lawless, J.F. 2003. *Statistical Models and Methods for Lifetime Data*. 2nd ed. New York: John Wiley.

Ma, Z.S. & Krings, A.W. 2008. Competing risks analysis of reliability, survivability, and prognostics and health management (PHM). *Proc. IEEE-AIAA AeroSpace Conf., March 1-8, 2008. Big Sky, MT.*: 20.

Mathew, A. & Pandey, M. 2002. Attributing death to cancer: Cause-specific survival estimation. *Journal of Postgraduate Medicine* 48(4): 322-326.

Prentice, R.L., Kalbfleisch, J.D., Peterson, A.V. Jr., Flournoy, N., Farewell, V.T. & Breslow, N.E. 1978. The analysis of failure times in the presence of competing risks. *Biometrics* 34(4): 541-554.

Putter, H., Fiocco, M. & Geskus, R.B. 2007. Tutorial in biostatistics: Competing risks and multi-state models. *Stat. Med*. 26(11): 2389-2430.

Slud, E. & Byar, D. 1988. How dependent causes of death can make risk factors appear protective. *Biometrics* 44(1): 265-269.

Tsiatis, A.A. 1975. A nonidentifiability aspect of the problem of competing risks. *Proc. Natl. Acad. Sci. USA* 72(1): 20-22.

Tsiatis, A.A. 2005. Competing risks. In *Encyclopedia of Biostatistics*. 2nd ed. edited by Armitage, P. & Colton, T. Chichester, West Sussex: John Wiley & Son.

School of Mathematical Sciences
Universiti Sains Malaysia
11800 USM, Penang
Malaysia

*Corresponding author; email: fangyenyen@hotmail.com

Received: 9 May 2012
Accepted: 17 September 2012