Malaysian Journal of Mathematical Sciences 5(1): 125-141 (2011)

A Multistate Approach to Estimating the Net Survival Function in the Presence of Competing Risks

[1]Fang Yen Yen, [2]Noorani Ahmad and [3]Suraiya Kassim
[1, 2, 3] School of Mathematical Sciences,
Universiti Sains Malaysia,
11800 USM Minden, Penang, Malaysia
E-mail: [1]fangyenyen@hotmail.com, [2]nahmad@cs.usm.my and
[3]ksuraya@cs.usm.my

ABSTRACT

In competing risks analysis, formulation and estimation of the net survival function is usually done by the traditional Latent Failure Time approach with Kaplan Meier Estimator. However, this approach involves identifiability problems and based on unverified assumptions of independent risks and equal hazard of the crude and the net. It has been argued that even under independent risks, the equal hazard assumption may not be true in many practical problems. An extended multistate approach by Islam (1994) is proposed in estimating the net survival function without the equal hazard assumption and allows for the presence of informative eliminated risks. A comparison of the results of the proposed procedure to that of the Kaplan Meier Estimator is illustrated on an adapted dataset. The proposed method shows that when noninformative eliminated risks are assumed, the net hazard and the crude hazard are equal, as in the Kaplan Meier Estimator. The proposed procedure is shown to be useful when informative eliminated risks are present and may result in unequal hazard even under independent risks.

Keywords: Competing risks, informative eliminated risks, Kaplan Meier Estimator, Latent Failure Time approach, multistate approach, net survival function.

INTRODUCTION

Survival analysis is a statistical tool in analyzing an event time data. Survival data occurs in many areas of research such as medical, engineering (reliability data), demography (decrement data) and sociology (historical/event data) where time to occurrences of events are recorded. The event can be a failure of a component, death of a subject/individual, occurrence of a disease, recurrent of cancer etc. The occurrence of an event other than the event of interest is called a censoring event. The well-known Kaplan Meier Estimator (KME) or product limit estimator is a common nonparametric estimation procedure to estimate survival distribution function by the independent noninformative censoring assumption (Kaplan

and Meier (1958). However, Islam (1994) has proposed a nonparametric multistate approach for estimating the survival function as an alternative procedure, when the censoring is informative.

In some studies, several causes of an event can happen and subjects may fail from some causes other than the causes of interest. This is called a competing risks situation, whereby several risks compete with each other to act as a cause of failure and the occurrence of a risk preclude the occurrence of other risks (Gooley *et al*. (1999)). Competing risks is an advanced survival tool in analyzing survival or failure time data when there are more than one possible causes or types of failure that react simultaneously on a subject within a well-defined population. Historically, earlier works have been done by Daniel Bernoulli in 1760 in estimating the survival rate of a population if smallpox was eliminated as a cause of death. This is a classical competing risks problem. By this approach, the theory of competing risks was discovered *(*David and Moeschberger (1978); Dietz and Heesterbeek (2002)). The work of competing risks and elimination studies were later followed by D' Aldermbert (1761), Makeham (1874) and Farr (1875) as cited in Karn (1933), Dietz and Heesterbeek (2002) and Tsiatis (2005). Classically, competing risks has been a tool for actuarial and demography sciences. There were some studies on competing risks with cause-elimination-life-table such as Cornfield (1957), Elveback (1958), Kimball (1969), Chiang (1968), Manton and Poss (1979) and Nour (1981). Historical reviews and literature of actuarial method to competing risks problem can be found in Seal (1977). Later, the trend was moving from classical problem to a modern competing risks analysis that majority has been applied in the other fields of study such as medical, engineering or computer sciences (Ma and Krings (2008)).

According to Chiang (1968), the three important quantities to be considered in competing risks analysis are:

(i) the crude probability (observable)
(ii) the partial crude probability (unobservable)
(iii) the net probability (unobservable)

The net survival probability is generally defined as:

(i) the probability of survival (from remaining risks) after a risk has been eliminated from the population, or
(ii) the probability of survival if a specific risk is the only risk that reacts in the population.

In competing risks studies, it is of interest to estimate the nonidentifiable net survival probability in order to study the effect of one risk acting on a population. The estimation of the net survival probability is a type of "cause-elimination' analysis, whereby all other risks other than that of interest, are hypothetically eliminated (Hougaard (2000)). Traditionally, the latent failure time (LFT) approach is used to formulate competing risks problem. However, this approach involves identifiability problems. Only with additional assumptions made on the joint distribution of the LFT, such as independent risks, we are able to estimate the nonidentifiable quantities in competing risks.

However, all the assumptions being made cannot be verified or tested based on competing risks data, as the observable failure time, T is only the minimum of the m risks latent failure time, $T=min(T_1,T_2 \ldots T_m)$ (Tsiatis (1975); Crower (1994, 2001)). Two common assumptions by traditional approach to formulating the net survival function are:

(i) independent risks and
(ii) elimination of causes other than a cause, say j, is just by letting the failure rate of the eliminated causes equal zero without changing the failure rate of cause j, which means equal hazard of the crude and the net. (Noninformative eliminated risks).

Assumption (ii) means subjects survive from eliminated causes do have the same failure risks as the general population; they do not provide any information to the net survival time of interest. Therefore, it is called noninformative eliminated risks.

On the nonidentifiability issues, one of the perspectives taken in literatures is to focus only on observable quantities (modern competing risks analysis). Very few literatures are found on the nonparametric estimation of net survival function (classical competing risks problem) especially after the 1990s.

A modern framework is based on cause-specific hazard function and cumulative incident function (Prentice *et al.* (1978); Kalbfleisch and Prentice (2002)). Nevertheless, the study of the non-observable quantity of the net survival function is important since it is a significant quantity for actuarial, demography sciences, medical research or engineering study (Hoogaard (2000); Mathew (2002); Nelson (2003); Tsiatis (2005)).

Competing risks that are to be eliminated are treated as a random censoring in survival analysis (Klein and Moeschberger, 2003). With the assumptions (i) and (ii), the KME to estimate the net survival function assumed those eliminated risks as noninformative or random censoring (Kaplan and Meier (1958)). Chiang (1968)'s proportional hazard assumption also involves the basic assumption of independent risks and noninformative eliminated risk in estimating the complement of the net survival function. In the Markov formulation for the net survival probability, the underlying assumption used is the same as the two assumptions above (Aalen (1978)). However, in most practical problems, the two assumptions might be biased, especially assumption (ii). It has been argued that even in an independent process, the elimination of causes of failure might change the crude hazard rate in several ways, i.e. the crude hazard may not equal to the net hazard (Gail (1975); Prentice *et al*. (1978); Elandt-Johnson and Johnson (1980); Kalbfleisch and Prentice (2002); Lawless (2003)). It is therefore important to consider the net probability that does not need the noninformative assumption and allows for the presence of informative eliminated risks in which equal hazard assumption is not assumed.

In this paper, we show how the traditional LFT approach with KME is used in estimating the net survival probability when only a cause is acting. Due to the argument that eliminated causes do provide some information to the net survival to allow unequal hazard, we propose a nonparametric multistate approach to estimating the net survival probability when informative eliminated risks is present. We extend the idea from Islam (1994) into a new result within a competing risks framework and show how informative risks can be incorporated into the procedure. The results show a change of crude hazard after the elimination of causes of death under informative risks, even when the risks react independently.

THE LATENT FAILURE TIME APPROACH

This approach is the traditional approach to the formulation of competing risks problem. Suppose each individual and subject from a homogenous population is exposed to three potential risks of death, $j = 1, 2, 3$ with corresponding latent times T_1, T_2 and T_3. We define a time random variable $T = (T_1, T_2, T_3)$ as the latent death (failure) time that could happen to any individual under study. However, in competing risks analysis, only at most one cause of death would happen to an individual throughout the study. Thus, only the minimum time to death among the latent death time

T is observable within each individual, and it is defined as $T = \min\{T_1, T_2, T_3\}$.

Define the multivariate lifetime distribution or the joint distribution of T, by the joint survivor function as

$$S_{1,2,3}(x_1, x_2, x_3) = P(T_1 \geq x_1, T_2 \geq x_2, T_3 \geq x_3).$$

The overall survivor function is defined as

$$S_T(x) = P(T \geq x) = S_{1,2,3}(x, x, x)$$

and the corresponding overall hazard function is

$$\lambda_T(x) = -d \ \log \ S_T(x)/dx.$$

We can identify the cause-specific hazard function, CSH (crude hazard) as

$$\lambda_j(x) = -\delta \ \log \ S_{1,2,3}(x_1, x_2, x_3)\Big/ dx_j \Big|_{x_1 = x_2 = x_3 = x}, \ j = 1,2,3 \qquad (1)$$

the instantaneous rate of failure from cause j, when all three causes are operating simultaneously in the population. For simultaneous risks, we have

$$\lambda_T(x) = \sum \lambda_j(x), \ \ j = 1,2,3.$$

All functions in terms of CSH, (1) are estimable, other than that, they are inestimable without further assumptions (nonidentifiable) based on competing risks data (Gail (1975); Prentice *et. al.* (1978)).

Let the nonidentifiable marginal distribution of T with marginal survival function for $T_j, j = 1,2,3$ be defined as

$$S_j(x_j) = \mathrm{P}(T_j \geq x_j) = S(0, x_j, 0), j = 1,2,3 \qquad (2)$$

and the corresponding marginal hazard be

$$h_j(x) = -d \log S_j(x)/dx, \ j = 1,2,3.$$

To overcome the nonidentifiability problem in competing risks analysis, suppose that times to failure T_1, T_2 and T_3. are independent. The joint survival function then becomes

$$S_{1,2,3}(x_1, x_2, x_3) = P(T_1 \geq x_1)P(T_2 \geq x_2)P(T_3 \geq x_3) = S_1(x_1)S_2(x_2)S_3(x_3).$$

It can be easily shown that

$$\lambda_j(x) = h_j(x) \quad \text{for } j = 1,2,3 \tag{3}$$

Equation (3) allows for an indirect estimation of the non-observable probability of the marginal survival function.

The marginal survival function (2), often referred to as the net survival function, is of interest. In a competing risks setup, the net survival function involves the hypothetical situation that only risk *j* react to the population with all other risks being eliminated. The basic assumption is that the effect of eliminating causes is just letting the corresponding arguments in $S_{1,2,3}(x_1, x_2, x_3)$ to be nullified, with no effect on $S_{1,2,3}(x_1, x_2, x_3)$ (Chiang (1968)). It is equivalent to assuming that the elimination of causes only let the hazard of eliminated causes to be zero and the hazard of remaining cause unchanged. The net survival probability for any cause $j(j = 1,2,3)$ alone, can be formulated as

$$S_j(x) = \exp\left[-\int_0^x \lambda_j(u)du\right]. \tag{4}$$

This implies equal hazard rate of the crude (before elimination) hazard and the net (after elimination) hazard, (Crowder, 2001) where eliminated causes do not provide any information (noninformative risks) to the net survival function. If we are interested to derive the net survival probability of cause 1, the net survival probability is then defined as

$$S_1(x) = S_{1,2,3}(x, 0, 0) = \mathrm{P}(T_1 \geq x)$$

with corresponding hazard $\lambda_1(x)$. This implies elimination of cause 2 and cause 3 nullifies $\lambda_2(x)$ and $\lambda_3(x)$ without altering $\lambda_1(x)$. However, this condition is a weaker assumption than the statistically independent

assumption since even when risks react independently, $\lambda_1(x)$ might be altered in several ways when cause 2 and cause 3 are not present.

Let us define a new independent latent random variable of $T_j^*, j=1,2,3$ in a new hypothetical situation when only cause 1 is present. The cause 1 net survival function is

$$S_1^*(x) = S_{1,2,3}^*(x,0,0) = \mathrm{P}(T_1^* \geq x) = \exp\left[-\int_0^x \lambda_1^*(u)du\right],$$

with a new hazard that may be different from the crude hazard $\lambda_1(x)$, in which the estimation procedure by the multistate approach will be shown in later section.

Kaplan-Meier Estimator

The estimation of the net survival function, (Crowder (2001)) follows the Kaplan Meier Estimator (KME) method, a consistent estimator of the $\exp\left[-\int_0^x \lambda_j(u)du\right]$ (Tsiatis (2005); Kaplan and Meier (1958)). It treats eliminated causes as noninformative censoring and assumes only cause *j* is present after the other causes have been eliminated. The KME of the net survival function follows the LFT approach; it is consistent only when the net hazard function is equal to the CSH function (crude hazard). The estimator of the net survival function of cause *j* is

$$\hat{S}_j(x) = \prod_{t_i \leq x}\left[1 - d_{ij}/n_i\right], j=1,2,3 \tag{5}$$

where d_{ij} is the number of individuals that fail from cause *j* at time t_i, and n_i is the number of individuals being at risks at the beginning of time t_i.

THE PROPOSED MULTISTATE APPROACH IN ESTIMATING THE NET SURVIVAL FUNCTION

This section shows a multistate approach to estimating the net survival probability when only a cause is acting after elimination of all other causes, an extension from Islam (1994). For simplicity, assume there are

three risks, $j = 1,2,3$ that react simultaneously on any individual in a population from the beginning at time t_0. During each time interval before elimination, each individual is exposed to four possible outcomes: the individual is alive as a survivor, dies from cause 1, dies from cause 2 or dies from cause 3.

Assume a new hypothetical situation where only risk 1 is present, which also means competing causes other than cause 1 have been eliminated. After the elimination of cause 2 and cause 3, assume all individuals who died from both causes were saved and they are called *hypothetic survivors*. They are expected to die from cause 1 or remain alive in an ideal world. Therefore, there are overall six expected outcomes (six states) for each individual in this hypothetical situation (Figure 1): the individual

(i) is alive as survivor,
(ii) dies from cause 1,
(iii) saved from eliminated cause 2 and survives (hypothetic survivor),
(iv) saved from eliminated cause 3 and survives (hypothetic survivor),
(v) dies from cause 1 after being a cause 2 hypothetic survivor or
(vi) dies from cause 1 after being a cause 3 hypothetic survivor.

Outcomes (i), and (ii) are observable but (iii), (iv), (v) and (vi) are unobservable and can only be estimated indirectly with some assumptions. Figure 1 shows the six possible state outcomes and their relationship.

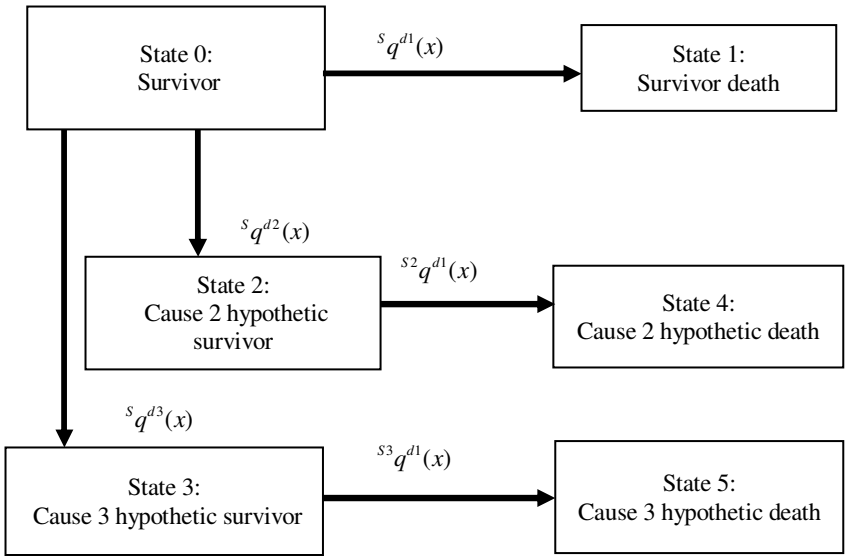

Figure 1: The multistate (six states) relationship when only cause 1 is present.

Assume $t_0 = 0$ and during a time interval $(0, x)$, let (Figure 1):

(a) ${}^{S}q^{d1}(x)$ be the probability of dying from cause 1 in the survivor state (the transition probability from state 0 to state 1 at time x).

(b) ${}^{S}q^{dj}(x), j = 2,3$ be the probability of moving from the survivor state to cause 2 and cause 3 hypothetic survivor states respectively.

(c) ${}^{S2}q^{d1}(x)$ and ${}^{S3}q^{d1}(x)$ be defined as the inestimable probability of dying from cause 1 when in cause 2 and cause 3 hypothetic survivor states respectively.

The probability of dying from cause 2 and cause 3 respectively before elimination is equal to the probability of being saved from eliminated cause 2 and cause 3 respectively and is represented in (b).

Let $P^S(x)$ be the probability of surviving in the interval $(0, x)$ in the survivor state (the probability of being in state 0 at time x). Suppose that all individuals in state 0 are free from any cause of death at time *0*, hence $P^S(0) = 1$. From Figure 1, $P^S(x)$ can be expressed as follows:

$$P^S(x) = 1 - {}^{S}q^{d1}(x) - {}^{S}q^{d2}(x) - {}^{S}q^{d3}(x) \tag{6}$$

Let $P^{S2}(x)$ and $P^{S3}(x)$ respectively be the probability of surviving in the interval $(0,x)$ when in states 2 and 3 (probability of being in state 2 and state 3 respectively at time x). Suppose there are no individuals in state 2 and state 3 at time *0* and hence $P^{S2}(0)=P^{S3}(0)=0$. Then $P^{S2}(x)$ and $P^{S3}(x)$ can be expressed as:

$$P^{S2}(x) = {}^{S}q^{d2}(x) - {}^{S2}q^{d1}(x) \tag{7}$$

and

$$P^{S3}(x) = {}^{S}q^{d3}(x) - {}^{S3}q^{d1}(x). \tag{8}$$

The net survival probability after eliminating cause 2 and cause 3 during the interval $(0,x)$ is the sum of the survivor probabilities of the three survivor states (state 0, state 2 and state 3) as given in Equation (6), (7) and (8), i.e. the probability of being in all three survivor states at time x are

$$P(x) = P^{S}(x) + P^{S2}(x) + P^{S3}(x) = 1 - {}^{S}q^{d1}(x) - {}^{S2}q^{d1}(x) - {}^{S3}q^{d1}(x). \tag{9}$$

${}^{S2}q^{d1}(x)$ and ${}^{S3}q^{d1}(x)$ are only estimable indirectly from observed data.

Consider now time intervals (t_i, t_{i+1}) of the survival time. Denote:

- ${}^{S}q_i^{d2}$ and ${}^{S}q_i^{d3}$ = the probability of transiting from state 0 to state 2 and state 3 respectively (after eliminating cause 1) during the interval,
- ${}^{S2}q_i^{d1}$ = the probability of transiting from state 2 to state 4,
- ${}^{S3}q_i^{d1}$ = the probability of transiting from state 3 to state 5, and
- ${}^{S}q_i^{d1}$ = the probability of transiting from state 0 to state 1.

After elimination of cause 2 and cause 3 deaths, individuals are saved from both causes. Assume that after the individuals enter state 2 or state 3 respectively, both hypothetic survivors can only meet death at the next interval time. By definition, it is proven (Islam (1994)) that

$${}^{S2}q^{d1}(x) = \sum_{t_i<x} {}^{S2}q_i^{d1} P^{S2}(t_i) \quad \text{and} \quad {}^{S3}q^{d1}(x) = \sum_{t_i<x} {}^{S3}q_i^{d1} P^{S3}(t_i). \tag{10}$$

The noninformative eliminated risks assumption from the LFT approach implies that eliminated risks (cause 2 and cause 3) do not provide any information to the net survival probability and the probability of dying from cause 1 is the same for remaining actual survivor and both hypothetic survivors. However there might be a bias due to the complexity of some practical or real problem. By incorporating adjustment factors to estimate the net survival probability, we take into account the unequal probability of dying for actual survivor state individuals and hypothetic (cause 2 and/or cause 3) survivor states individuals. By assuming that they are proportional at each time interval, we let a_i and b_i be adjustment factors for the i^{th} interval, and we have

$$ {}^{S2}q_i^{d1} = a_i \, {}^{S}q_i^{d1} \text{ and } {}^{S3}q_i^{d1} = b_i \, {}^{S}q_i^{d1}. \tag{11} $$

From equations (9), (10) and (11), we get

$$ P(x) = 1 - {}^{S}q^{d1}(x) - \sum_{t_i<x} a_i \, {}^{S}q_i^{d1} P^{S2}(t_i) - \sum_{t_i<x} b_i \, {}^{S}q_i^{d1} P^{S3}(t_i). \tag{12} $$

When both $a_i = 1$ and $b_i = 1$, it is reduced to the noninformative assumption used by the KME treated eliminated causes as noninformative censoring and cause 1 death as the only risk, (Cornfield (1957)). If at least one of the adjustment factors is not equal to one ($a_i \neq 1$ and/or $b_i \neq 1$) in at least one time interval, then the informative eliminated cause 2 and/or cause 3 do provide some information to the net survival probability. Equation (12) is an alternative multistate approach to replace Equation (5) of the Kaplan Meier method when informative eliminated risks are present. The value of the adjustment factors may vary in different intervals, depending on the researcher's need or background information of the process under study.

Estimation

During each time interval (t_i, t_{i+1}), an individual at the beginning of interval may live as a survivor, dies from cause 1, dies from cause 2 or dies from cause 3. The transitions of states follow a multinomial distribution. Under the assumption of independent risks without censoring, the general form of the likelihood function for a multinomial distribution is

$$ L\alpha \left({}^{S}q_i^{d1} \right)^{{}^{S}d_i^{d1}} \left({}^{S}q_i^{d2} \right)^{{}^{S}d_i^{d2}} \left({}^{S}q_i^{d3} \right)^{{}^{S}d_i^{d3}} \left(1 - {}^{S}q_i^{d1} - {}^{S}q_i^{d2} - {}^{S}q_i^{d3} \right)^{n_i - {}^{S}d_i^{d1} - {}^{S}d_i^{d2} - {}^{S}d_i^{d3}} \tag{13} $$

The estimator of ${}^{S}q_i^{dj}$ is given by

$$ {}^{S}\hat{q}_i^{dj} = {}^{S}d_i^{dj} \big/ n_i , \quad j = 1,2,3 \tag{14} $$

where ${}^{S}d_i^{dj}$ is the observable number of individuals whose death is from cause j at interval $(t_i,\ t_{i+1})$, and n_i is the number of individuals at risk at the beginning of the interval. To estimate the unobservable probability, (Gail (1975)) we can use an indirect estimator, such as

$$ {}^{S2}\hat{q}_i^{d1} = a_i\ {}^{S}\hat{q}_i^{d1} \text{ and } {}^{S3}\hat{q}_i^{d1} = b_i\ {}^{S}\hat{q}_i^{d1} \tag{15} $$

RESULTS AND DISCUSSION

To illustrate the proposed multistate approach, we use a part of Hoel and Walburg's data adapted from Klein and Moeschberger (2003). The data is a result from an experiment on the study of the effects of radiation on life lengths of mice with three causes of death: thymic lymphoma (cause 1), reticulum cell sarcoma (cause 2), or other causes (cause 3). There is no censoring because the mice all died by the end of the experiment. Suppose that we are interested in studying the survival probability of mice when only cause 1 is present, where we examine the impact on the survival rate of mice if we could eliminate the death due to cause 2 and cause 3.

By using the KME method (Equations 4 and 5) and the multistate method by assuming noninformative risks (Equations 12, 13, 14 and 15), we display the results of estimating the net survival probability if only risk 1 is acting and its corresponding net hazard in Table 1. The results show no difference in the net survival probability and its corresponding net hazard by the two different approaches when adjustment factors, $a_i = b_i = 1$ (noninformative risks) is assumed. The crude hazard simply equals the net hazard under the two methods.

TABLE 1: The net survival probabilities and corresponding hazards by the KME method and the multistate approach in three different survivor states.

Interval	KME with LFT Approach		Multistate Approach with $a_i = b_i = 1$				
	$\hat{S}_1(x)$	$\hat{\lambda}_1(x) = \hat{h}_1(x)$	$\hat{P}^{S}(x)$	$\hat{P}^{S2}(x)$	$\hat{P}^{S3}(x)$	$\hat{P}(x)$	$\hat{h}_1(x)$
1	0.93671	0.06329	0.92406	0.00000	0.01266	0.93671	0.06329
2	0.79556	0.15068	0.75949	0.00000	0.03607	0.79556	0.15068
3	0.76904	0.03333	0.72152	0.00000	0.04752	0.76904	0.03333
4	0.71507	0.07018	0.64557	0.01266	0.05685	0.71507	0.07018
5	0.68703	0.03922	0.59494	0.02482	0.06728	0.68703	0.03922
6	0.67241	0.02128	0.37975	0.11290	0.17977	0.67241	0.02128
7	0.62759	0.06667	0.22785	0.15601	0.24373	0.62759	0.06667
8	0.62759	0.00000	0.07595	0.16886	0.38297	0.62759	0.00000
9	0.62759	0.00000	0.02532	0.18132	0.42095	0.62759	0.00000

Table 2 shows the values of the net survival probability and the net hazard by using two different assumptions in the multistate approach when known informative eliminated risks are present by assuming $a_i = b_i = 0.5$ and $a_i = b_i = 1.5$. By the multistate method, the crude hazard is not simply equal to the net hazard when $a_i \neq 1$ and $b_i \neq 1$. As shown in Table 2, when a_i and b_i are smaller than 1, the net hazard by the multistate approach is less than or equal to the hazards obtained by the LFT approach. But if a_i and b_i are greater than 1, the net hazard by the multistate approach is bigger or equal compared to the LFT approach. Moreover, the table shows that the net survival probability by the multistate approach is different from the KME method which applies the LFT approach. This result allows accessing the impact of cause 1 on the survival and failure probability when different informative rates are present.

TABLE 2: Estimation of the net survival probabilities based on the assumptions $a_i = b_i = 0.5$ and $a_i = b_i = 1.5$ with its corresponding hazards by the multistate approach.

Interval	Multistate Approach				KME with LFT approach	
	Assumed $a_i = b_i = 0.5$		Assumed $a_i = b_i = 1.5$			
	$\hat{P}(x)$	$\hat{h}_1(x)$	$\hat{P}(x)$	$\hat{h}_1(x)$	$\hat{S}_1(x)$	$\hat{h}_1(x)$
1	0.93671	0.06329	0.93671	0.06329	0.93671	0.06329
2	0.79651	0.15068	0.79461	0.15068	0.79556	0.15068
3	0.77058	0.03329	0.76754	0.03337	0.76904	0.03333
4	0.71823	0.07004	0.71206	0.07031	0.71507	0.07018
5	0.69198	0.03904	0.68432	0.03938	0.68703	0.03922
6	0.67884	0.02112	0.67047	0.02136	0.67241	0.02128
7	0.65118	0.06604	0.63918	0.06686	0.62759	0.06667
8	0.65118	0.00000	0.63918	0.00000	0.62759	0.00000
9	0.65118	0.00000	0.63918	0.00000	0.62759	0.00000

CONCLUSION

The equal hazard assumption in the traditional LFT approach and KME method in estimating the net survival function might not be true in many practical problems, when informative risks (eliminated) are present. Without the equal hazard assumption, we have proposed a more flexible multistate approach in estimating the net survival probability when informative risks (eliminated) are present by incorporating adjustment factors into the procedure. When noninformative risks are assumed, the proposed procedure is reduced to the traditional approach. The results shows unequal hazard before and after elimination by different assumptions on the adjustment factors (contradictory to the LFT approach of equal hazard). In addition, the proposed multistate approach is sensitive to the changes of hazard when informative risk is present, even under independent risks assumption.

The procedure can be extended to studies that involve more than three risks, consideration of censoring, a non-complete elimination process, partial crude probability or consideration of more complex multistate elimination. Moreover, it is easy to extend the proposed method to deal with a situation where different types of informative eliminated risks are present in one sample. The adjustment factors is allowed to vary in different intervals, thus the impact of eliminated risks with different informative rate can be studied in detail by considering each interval situation, allowing results that cannot be given by the traditional approach. As a conclusion, the proposed multistate approach without the equal hazard assumption is a simple procedure to be used as an alternative approach to the traditional approach, especially when some known informative eliminated risks are provided.

ACKNOWLEDGEMENT

This work is done with the financial support of the USM Fellowship Scholarship Scheme, Universiti Sains Malaysia.

REFERENCES

Aalen, O. 1978. Nonparametric Estimation of Partial Transition Probabilities in Multiple Decrement Models. *Ann. Statist.*, **6**(3):534-545.

Chiang, C. L. 1968. *Introduction to Stochastic Processes in Biostatistics*. John Wiley.

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

Crowder, M. 2001. *Classical Competing Risks*. New York: Chapman & Hall/CRC.

Cornfield, J. 1957. The estimation of the probability of developing a disease in the presence of competing risks. *Americal Journal Public Health*, **47**:601-607.

David, H. A. and Moeschberger, M. L. 1978. *The Theory of Competing Risks*. Macmillan Publishing.

Dietz, K. and Heesterbeek, J. A. P. 2002. Daniel Bernoulli's Epidemiological Model Revisited. *Mathematical Biosciences*, **180**:1–21.

Elandt-Johnson, R.C and Johnson, N. L. 1980. *Survival models and data analysis*. New York: John Wiley and Sons.

Elveback, L. 1958. Estimation of survivorship in chronic disease: the "actuarial" method. *Journal of the American Statistical Association*, **53**:420-440.

Fang, Y.Y., Noorani Ahmad and Suraiya Kassim. 2009. On estimating the survival probability after the elimination of a cause of failure: A traditional and a multistate approach. *Proceedings of the symposium of USM fellowship holders,* **1**: 324-333

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

Gooley, T. A., Leisenring, W., Crowley, J. and Storer, B. E. 1999. Estimation of failure probabilities in the presence of competing risks: new representations of old estimators. *Statistics in Medicine*, **18**: 695–706.

Hougaard, P. 2000. *Analysis of Multivariate Survival Data*. Statistics for Biology and Health. New York: Springer.

Islam, M. A. 1994. An Alternative to the Product Limit Method: The Multistate Approach. *The Statistician,* **43**(2):237-245.

Kalbfleisch, J. D. and Prentice, R.L. 2002. *The Statistical Analysis of Failure Time Data (2nd Ed).* New York: Wiley.

Kaplan, E.L. and Meier, P. 1958. Nonparametric Estimation from Incomplete Observations. *J. Amer. Statist. Assoc.,* **53**: 457-481.

Karn, M. N. 1933. A further study of constructing life tables when certain causes of death are eliminated. *Biometrika*, **25**: 191-201.

Kimball, A.W. 1969. Models for the estimation of competing risks from grouped data. *Biometrics,* **25**: 329-337.

Klein, J. P and Moeschberger, M. L. 2003. *Survival Analysis. Techniques for Censored and Truncated Data.* Statistics for Biology and Health. Springer: New York.

Lawless, J. F. 2003. *Statistical Models and Methods for Lifetime Data*. John Wiley

Nelson, W. 2003. *Applied life time data analysis.* John Wiley and Sons, Ltd.

Ma, Z. S. and Krings, A.W. 2008. Competing Risks Analysis of Reliability, Survivability, and Prognostics and Health Management (PHM), *Proceedings of 2008 IEEE-AIAA AeroSpace Conference.* Big Sky, Montana. March 1-8, 2008.

Manton, K. G. and Poss, S.S. 1979. Effect of dependency among causes of death for cause elimination life table. *Demography*, **16**(2):313-327.

Mathew, A and Pandey M. 2002. Attributing death to cancer: cause-specific survival estimation. *Journal of Postgraduate Medicine*, **48**:322–6.

Nour, E. S. 1981. Multiple decrement analysis with dependent risks. *Biometrical Journal,* **23**:175–183.

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

Seal, H. L. 1977. Studies in the history of probability and statistics. XXXV. Multiple decrements *or* competing risks*, Biometrika*, **64**: 429-439.

Tsiatis, A. A. 1975. A nonidentifiability aspect of the problem of competing risks. *Proceedings of the National Academy of Sciences of the United States of America*, **72**(1): 20-22.

Tsiatis A. A. 2005. Competing risks. *In: Armitage P, Colton T (eds) Encyclopedia of biostatistics*. John. Wiley and Sons Ltd, Chicester England: 1025–1035.