When do we need competing risks methods for survival analysis in nephrology?

Marlies Noordzij1, Karen Leffondré2, Karlijn J. van Stralen1, Carmine Zoccali3, Friedo W. Dekker4 and Kitty J. Jager1

1ERA-EDTA Registry, Department of Medical Informatics, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands,
2Université Bordeaux Segalen, ISPED, Centre de recherche INSERM U897 Épidémiologie et Biostatistique, Bordeaux, France,
3Department of Clinical Epidemiology, Leiden University Medical Centre, Leiden, The Netherlands and
4CNR–IBIM, Clinical Epidemiology and Pathophysiology of Renal Diseases and Hypertension, Renal and Transplantation Unit, Ospedali Riuniti, Reggio Calabria, Italy

Correspondence and offprint requests to: Marlies Noordzij; E-mail: m.noordzij@amc.uva.nl

Keywords: censoring, competing risks, epidemiology, statistics, survival analysis

ABSTRACT

Survival analyses are commonly applied to study death or other events of interest. In such analyses, so-called competing risks may form an important problem. A competing risk is an event that either hinders the observation of the event of interest or modifies the chance that this event occurs. For example, when studying death on dialysis, receiving a kidney transplant is an event that competes with the event of interest. Conventional methods for survival analysis ignoring the competing event(s), such as the Kaplan–Meier method and standard Cox proportional hazards regression, may be inappropriate in the presence of competing risks, and alternative methods specifically designed for analysing competing risks data should then be applied. This problem deserves more attention in nephrology research and in the current article, we therefore explain the problem of competing risks in survival analysis and how using different techniques may affect study results.

INTRODUCTION

A substantial part of the medical research papers include survival analyses. Survival analysis is the analysis of time until a certain event occurs, for example, time to renal transplantation or death. In the interpretation of results of survival analyses, competing risks can be an important problem. Competing risks occur when subjects can experience one or more events or outcomes which ‘compete’ with the outcome of interest. In those cases, the competing risk hinders the observation of the event of interest or modifies the chance that this event occurs. In the field of nephrology, there are many situations in which competing risks play a role. For example, when studying the time until a peritonitis episode occurs in peritoneal dialysis (PD) patients, death, kidney transplantation and transfer to haemodialysis can be considered as competing risks because patients who experience one of these events are no longer at risk of developing PD-related peritonitis. Another example is a study in diabetes mellitus patients investigating the time until end-stage renal disease (ESRD) occurs. In this case, death before reaching ESRD is a competing risk [1].

In oncology and cardiovascular medicine, this analytical problem of competing risks has been acknowledged for many years, whereas in nephrology, it has been acknowledged only recently in a few publications [1–5]. As it also deserves more attention in the field of nephrology, we summarize in this article the problem of competing risks and show how using different analysis techniques may impact on results and conclusions. This article is aimed at readers who would like to apply competing risk methods themselves. We will focus on an example studying patient survival on dialysis, where death
on dialysis is the event of interest and kidney transplantation is a competing risk for death on dialysis.

### Basic Principles of Survival Analysis

In survival analyses, all subjects who are at risk of experiencing an event are part of the so-called risk set. The risk set usually consists at each point in time of individuals who have been followed-up till that time and have not yet experienced the event of interest just before that time point [6]. The concept of a risk set is important in understanding the competing risks methods that are discussed in this article and to decide which method to apply.

In standard survival analysis, the survival time of subjects who do not experience the outcome of interest during the observation period is censored at the end of follow-up. In those cases, we do not know whether and when such a patient will experience the event, we only know that he or she has not done so by the end of the observation period. Censoring may occur for various reasons. A patient may be lost to follow-up during the study or may experience another event (such as recovery of renal function) which makes further follow-up impossible or useless. Finally, the observation period may end before the patient has experienced the event of interest. Censored time-to-events can therefore be considered as a form of incomplete data.

An important assumption of standard survival analytical methods such as the Kaplan–Meier method is that censoring is ‘independent’ [6]. This independent censoring assumption implies that patients who are censored at a certain time point should be representative for those still at risk (and thus in the risk set) at that point in time. This is, for example, usually the case when a patient’s survival time is censored because he or she was lost to follow-up, for instance, due to migration. In this situation, we can assume that this occurred at random and patients who are censored are likely to be at a similar risk of experiencing the event of interest as patients who are not.

For unadjusted survival analysis, generally Kaplan–Meier analyses are applied [7]. The Kaplan–Meier method estimates the probability to survive up until a certain time point (time \(t\)) in the presence of censored cases. For subjects whose data are censored, either because they left the study or because they ran into the end of the study period, all information until their time of censoring is included in the analysis. Medical papers often present the complement of the Kaplan–Meier estimate \([1−\text{KM}(t)]\), which gives the estimated probability of dying before time \(t\). By means of a log-rank test, one can statistically test whether there are significant differences in the survival between two or more groups. However, when using the Kaplan–Meier method, one cannot easily quantify an effect size. Such an effect size is therefore usually calculated as a hazard ratio (HR) using Cox proportional hazards analysis [8]. When comparing an exposed group with an unexposed group, the HR is the ratio between the hazard of the event in the exposed group and the hazard of the event in the unexposed group. The hazard of the event can vary over time in each group and can be interpreted at each time as the instantaneous risk of developing the event at that time, given that a subject is still at risk of the event at that time. The Cox model, however, assumes that the HR between the two groups is constant over time. This is the proportional hazard assumption [8]. In these Cox regression analyses, it is also possible to adjust for (potential) confounders.

### The Problem of Competing Risks

As explained earlier, a competing risk is an event that either hinders the observation of the event of interest or modifies the chance that this event occurs. For example, when performing a study with mortality on dialysis as the outcome of interest, a patient may receive a kidney transplant. This transplant is a competing risk because after the transplantation, this patient is not on dialysis anymore and therefore no longer at risk of dying while being on dialysis. In this case, the competing event, i.e. receiving a kidney transplant, hinders the occurrence of the event of interest.

It should also be mentioned that at any time before experiencing the first event, patients should be at risk of both events. This means for our example that patients who die on dialysis should be at risk of receiving a transplant at any time before dying. If another event made it impossible to receive a transplant, this event may be considered as an additional competing event. Although a competing risk analysis may include several types of competing events, for the sake of simplicity, we focus on only one type of competing event in this article.

### Prognostic Versus Aetiological Research

To decide which method for survival analysis in the presence of competing risks should be used, it is important to know what kind of research question one aims to answer. In general, there are two types of research questions which can be answered with epidemiological studies [9]. Aetiological research aims to investigate the causal relationship between risk factors or determinants and a given outcome. To this end, it uses HRs to estimate an effect size. In contrast, prognostic research aims to predict the probability of a given outcome at a given time for an individual patient.

### Prognostic Research in the Presence of Competing Risks

**Unadjusted**

To predict the unadjusted probability of a certain outcome to occur, one can use the Kaplan–Meier method. However, in the presence of competing risks, using the Kaplan–Meier method is problematic. The method can handle only one single event at a time: all other events are treated as censored observations and the complement of the Kaplan–Meier estimate \([1−\text{KM}]\) is interpreted as the probability of the event of interest in a hypothetical world in which the competing event
does not exist. This kind of interpretation is not realistic in clinical practice [10, 11]. The independent censoring assumption is violated, meaning that the patients who experience a competing event at a given time often do not have the same chance of developing the event of interest after that time as the patients who are continued to be followed-up. As a result, the Kaplan–Meier method generally overestimates the probability of the event of interest and thus yields misleading results in the presence of competing risks.

To overcome these problems that arise when using the Kaplan–Meier method in the presence of competing risks, an alternative method is available. This method, also referred to as the cumulative incidence competing risk (CICR) method, is based on the so-called cumulative incidence function. The CICR accounts for all types of events; in the case of competing events, the cumulative incidence function is estimated both for the event of interest and for all competing events, and their estimates depend on each other [12]. Unlike in the application of the Kaplan–Meier method, competing events are not handled as regular censoring events without influence on the cumulative incidence function for the event of interest. Instead, the cumulative incidence, i.e. the probability of dying before time $t$, is lowered by the occurrence of the competing event and patients experiencing the competing event are considered to be no longer at risk for the event of interest. The CICR method has been described in detail by Verduijn et al. [5].

While the log-rank test is used to test whether the survival functions are significantly different between groups when censoring is independent, this test cannot be used in the presence of competing risks [13]. Different tests based on cumulative incidence functions have therefore been developed in the context of competing risks [14, 15].

**Example: Kaplan–Meier method versus CICR method**

Using ERA-EDTA Registry data, we studied patient survival from Day 91 after the start of dialysis with death on dialysis as the event of interest. Follow-up time was censored at loss of follow-up and at the end of the observation period. In this example, kidney transplantation is the competing event because a patient who receives a transplant is no longer at risk of death on dialysis. In addition to patients who die and those who receive a kidney transplant, there are also patients who do not experience an event at all. The probability of being alive and not having received a kidney transplant at a given time $t$ is given by the event-free survival (EFS) probability. Note that at any time point, a patient has either died before receiving a transplant (event of interest) or has already received a transplant (competing event), or is still alive without transplantation.

Table 1 presents the baseline characteristics of the included patients and the number of (first) events that occurred among them during 5 years of follow-up. To study the influence of the competing event, kidney transplantation, we estimated the probabilities of dying (before receiving a transplant) before time $t$, receiving a transplant before time $t$, and being alive and not having received a transplant until time $t$ (EFS) at 1, 2 and 5 years after Day 91 of dialysis, using both the traditional Kaplan–Meier method and the CICR method. The results of these analyses are summarized in Figure 1. For 1-year patient survival on dialysis, both methods yielded similar probabilities of death and transplantation. However, at 2 years, the Kaplan–Meier method yielded probabilities that added up to a total of 104%, against 100% for the CICR method. Finally, after 5 years of follow-up, the difference between the methods was even more pronounced. The probabilities of death (60%) and transplantation (33%) were overestimated by Kaplan–Meier so that the total of probabilities amounted to 118%, while the CICR method yielded lower probabilities, still adding up to 100%. The probability of having no events (EFS) remained the same for both methods, but the Kaplan–Meier method overestimated the probabilities of both death and transplantation with percentages of almost 10% each at 5 years after the start of dialysis. The difference in findings for these two methods can be explained by the different manner of calculating the probabilities.

These results demonstrate that the Kaplan–Meier method overestimates the probabilities of both the event of interest and the competing event(s), while the estimate for EFS is unbiased. This overestimation of probabilities is increasing with follow-up time. The Kaplan–Meier method is therefore inappropriate to analyse patient survival in the presence of competing risks and using the CICR method is recommended. When there are no competing risks, the Kaplan–Meier and CICR methods yield the same results.

**Adjusted**

Although there are different methods for competing risks regression available [16–19], there is currently consensus that for prognostic studies, the so-called subdistribution hazards approach proposed by Fine and Gray [20] is the most appropriate method to use. Because there is a direct relationship

| Table 1. Baseline characteristics of dialysis patients in the ERA-EDTA Registry 91 days after the start of dialysis treatment, categorized by status after 5 years of follow-up |
|---|---|---|---|
| | Total | Deceased | Transplantation | Event-free |
| $n$ | 73 382 | 37 067 | 16 008 | 20 298 |
| Age (years)$^a$ | 62.9 (15.6) | 70.1 (11.2) | 47.4 (14.4) | 62.0 (14.3) |
| Sex (% male) | 61.3 | 61.1 | 63.2 | 59.6 |

$^a$Mean (standard deviation).
between the covariates and the cumulative incidence function, the subdistribution hazards model directly provides individual prediction based on covariates or, in other words, estimated probabilities of an event, given a patient's characteristics. An important feature of this method is that subjects who experience a competing event remain in the risk set (instead of being censored), although they are in fact no longer at risk of the event of interest. This is illustrated in Figure 2 which is adapted from Lau et al. (with permission) [6]. As a consequence, the subdistribution HR (SHR) resulting from this method cannot be interpreted as an HR [6, 19]. However, when used for prediction, the SHR is only used as part of the calculation of an individual patient's risk.

AETIOLOGICAL RESEARCH IN THE PRESENCE OF COMPETING RISKS

For aetiological research, the proportional cause-specific hazards model may be more appropriate than the subdistribution hazards method. This is because the regression parameters estimated by this method directly quantify the HRs among those individuals who are actually at risk of developing the event of interest [6, 19]. In this case, Cox regression analysis is applied for each of the specific event types. So, separate Cox regression models are used to study the event of interest, for example, death on dialysis and the competing event(s), for example, transplantation. In each of these models, the competing events are treated as censored observations. Note that one does not need the independence of competing events to obtain valid estimates from such a cause-specific approach [19]. Another advantage of this cause-specific proportional hazard model is that it is easy to fit (by simply censoring for competing events) with any type of statistical software. It is important to realize, however, that because the competing events are treated as censored observations, during follow-up, the number of patients at risk is reduced, as is illustrated in Figure 3 [6]. Therefore, HRs calculated using this approach are interpreted as 'among those patients who did not (yet) experience the event of interest or a competing event'.

In the following example of an aetiological study in a competing risks setting, we illustrate how applying the two different methods may influence the results of multivariable survival analyses, and why these results need a different interpretation.

Example multivariable survival analysis: cause-specific hazard approach versus subdistribution hazard approach

Again, we studied patient survival on dialysis using ERA-EDTA Registry data (Table 1). We were interested in the influence of the competing event of kidney transplantation when estimating the effects of sex and age on the risk of death on dialysis. We studied 5-year survival on dialysis from Day 91 after the start of dialysis and follow-up time was censored at loss to follow-up and at the end of the observation period. We compared the results for the event of interest (death) and the competing event (transplantation) using both the cause-specific approach and the subdistribution proportional hazards model. Both methods were performed using STATA version 12 because this statistical software provides the package stcrreg which fits competing risks regression models according to the subdistribution hazard method [20].

First, we investigated the association between death on dialysis and sex. To estimate cause-specific HRs for males and females for the risk of death and transplantation, we performed a standard Cox regression model for each of the events in which the other (competing) event was censored for. We found that the hazards of dying and of transplantation were
**FIGURE 2:** Overview of the calculation of the subdistribution hazard: The risk set starts with 20 individuals (grey). Over time, individuals may experience the event of interest (death, black) or the competing event (transplantation, white) and those having a competing event are maintained in the risk set. Consequently, over time, a greater proportion of the risk set becomes full of individuals who have had the competing event prior to that time. The subdistribution hazard (SDH) for death is given at the bottom of the figure along with the cause-specific hazard (CSH) for death for comparison. Note that, because individuals are maintained in the risk set, the SDH of the event of interest tends to be lower than the CSH (adapted from Lau et al. [6]).

**FIGURE 3:** Overview of the calculation of the cause-specific hazard: The risk set starts with 20 individuals (grey). Over time, individuals have either the event of interest (death, black) or the competing event (transplantation, white). As individuals have either event, they are removed from the remaining risk sets. The calculation for the cause-specific hazard for both events is given at the bottom of the figure (adapted from Lau et al. [6]).
both only slightly but significantly higher for males than for females (Table 2: HRs of 1.04 and 1.09, respectively). Subsequently, we repeated the analyses using the subdistribution proportional hazards model and found SHRs of 1.03 and 1.07. This is in agreement with the simulation results of Latouche and Porcher [21] who found that when the HR for the competing event (in our case transplantation) is close to one, so when there are only slight differences in the hazard of transplantation between males in females, the two approaches give similar results for the event of interest (in our case death).

Second, we investigated the association between death on dialysis and age at the start of dialysis treatment, comparing young (<65 years) versus old (≥65 years) patients. Again, we first calculated cause-specific HRs and repeated the analyses using the subdistribution proportional hazards model. Using the cause-specific approach, we found that the hazard of death was 2.57 times increased in older when compared with younger patients, while the hazard of receiving a kidney transplant was 90% lower for older patients than for younger patients. With the subdistribution proportional hazards model, we found an SHR for death of 3.47 for old when compared with younger patients, instead of the HR of 2.57 that we found using the cause-specific approach (Table 2).

The latter example demonstrates that both approaches for dealing with competing risk data may yield different results, which is explained by the different composition of the risk sets. For our example, it is important to keep in mind that patients who receive a kidney transplant are generally younger and thus have a lower risk of dying than those who do not receive a transplant. In the cause-specific model for death, the patients who received a transplant were censored and thus removed from the risk sets after their time of transplantation, whereas they were kept in the risk sets after transplantation in the subdistribution model. As a result, at each time point, the risk sets in the cause-specific approach comprised a higher proportion of older people than those in the subdistribution approach. The contrast between the subject who dies at a given time (who is more likely to be in the older subgroup) on the one hand and all subjects still at risk at that time on the other hand was therefore lower in the cause-specific model than in the subdistribution model. As a result, the estimate obtained with the cause-specific approach (HR = 2.57) was closer to one than that obtained with the subdistribution model (SHR = 3.47). However, as explained before, the HR and the SHR do not have the same interpretation. The HR of 2.57 means that at any time after dialysis initiation, dialysis patients older than 65 years had a hazard of dying 2.57 times higher than those younger than 65 years, among patients on dialysis who were alive and did not receive a transplant at that time. The SHR higher than one (SHR = 3.47) means that the cumulative incidence of death is higher in patients older than 65 years at the start of dialysis when compared with younger patients. However, the numerical value of 3.47 is not straightforward to interpret since it reflects the mortality rate ratio among patients who are alive or have been transplanted before. The SHR of 0.07 for transplantation reflects the transplant rate ratio of older versus younger patients among subjects who have not yet received a transplant or have already died without transplantation. So, the SHR is in fact a different quantity than an HR, representing a ratio in a non-existing population including those who experienced the competing event. This quantity is mainly of interest for prediction, and this is the reason why subdistribution hazards models are often considered less appropriate than cause-specific models for aetiological questions.

Table 2. HRs and SHRs with 95% confidence interval for all-cause mortality in male versus female (reference group) dialysis patients and in old (≥65 years) versus young (<65 years, reference group) dialysis patients

<table>
<thead>
<tr>
<th></th>
<th>Cause-specific approach HR (95% CI)</th>
<th>Subdistribution hazard approach SHR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Death</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Male&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1.04 (1.02–1.07)</td>
<td>1.03 (0.87–1.23)</td>
</tr>
<tr>
<td>Young</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Old&lt;sup&gt;b&lt;/sup&gt;</td>
<td>2.57 (2.52–2.63)</td>
<td>3.47 (3.39–3.55)</td>
</tr>
<tr>
<td><strong>Transplantation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Male&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1.09 (1.05–1.12)</td>
<td>1.07 (1.04–1.11)</td>
</tr>
<tr>
<td>Young</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Old&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.10 (0.10–0.11)</td>
<td>0.07 (0.07–0.08)</td>
</tr>
</tbody>
</table>

<sup>a</sup>Adjusted for age.  
<sup>b</sup>Adjusted for sex.
In summary, for prognostic research, applying the subdistribution proportional hazards model is recommended, and for aetiological research, the cause-specific hazards model provides quantities that are easy to interpret. Indeed, the SHR resulting from the subdistribution method cannot be interpreted as an HR, because patients who are in fact no longer at risk of the event of interest remain in the risk set. An advantage of the cause-specific approach is that the estimated HR can be interpreted as an HR among those patients who are alive and did not receive a transplant before. Another advantage of the cause-specific approach is that it is easier to handle time-dependent covariates than with the subdistribution hazards model [22]. Which method to use in the absence and presence of competing risks, and for each type of research question is summarized in Table 3.

## SOFTWARE

The cause-specific model can be estimated using any software that handles the Cox model. The user only has to fit separate Cox models for each event of interest, using adequate event and censoring times for competing events. However, not all software include the CICR method and the subdistribution hazard model. For SPSS, a macro is available to perform the CICR method [5]. Rosthoj et al. [23] published a manual on how to use SAS macros for the estimation of the cumulative incidence function based on a Cox regression model for competing risks. In addition, recently an SAS macro for the subdistribution hazard model has been developed ([http://cemsiis.meduniwien.ac.at/en/kb/science-research/software/statistical-software/pshreg](http://cemsiis.meduniwien.ac.at/en/kb/science-research/software/statistical-software/pshreg)). When using STATA, the stcrreg procedure fits subdistribution hazards models [20]. Finally, the freely available statistical software R includes different options for performing competing risk analyses such as the cmprsk package. Useful manuals for performing competing risks analyses using R were published by Scrucca et al. [24, 25]

## CONCLUSION

It is of major importance to be aware of the presence of any competing risks when performing survival analyses. The Kaplan–Meier method for unadjusted survival analysis can handle only one outcome and yields unreliable results for the estimation of survival probability in the presence of competing risks. The use of the alternative CICR method is therefore recommended.

For multivariable survival analysis, in a competing risks setting, different approaches are available. In general, the subdistribution hazard is most suitable for prediction of a survival probability, while for aetiological studies, when HRs need to be derived, the cause-specific approach is most appropriate.

## CONFLICT OF INTEREST STATEMENT

None declared.

## ACKNOWLEDGEMENTS

The research leading to these results has received funding from the European Community’s Seventh Framework Programme under grant agreement number HEALTH-F2-2009-241544 (SysKID).
REFERENCES


Received for publication: 8.2.2013; Accepted in revised form: 11.7.2013