Survival analysis: time-dependent effects and time-varying risk factors

Friedo W. Dekker[1,2], Renée de Mutsert[1], Paul C. van Dijk[2,4], Carmine Zoccali[3] and Kitty J. Jager[2]

[1]Department of Clinical Epidemiology, Leiden University Medical Centre, Leiden, The Netherlands; [2]ERA-EDTA Registry, Department of Medical Informatics, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands; [3]CNR-IBIM Clinical Epidemiology and Pathophysiology of Renal Diseases and Hypertension, Renal and Transplantation Unit, Ospedali Riuniti, Reggio Cal, Italy and [4]Department of Clinical Epidemiology, Biostatistics and Bio-informatics, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands

In traditional Kaplan–Meier or Cox regression analysis, usually a risk factor measured at baseline is related to mortality thereafter. During follow-up, however, things may change: either the effect of a fixed baseline risk factor may vary over time, resulting in a weakening or strengthening of associations over time, or the risk factor itself may vary over time. In this paper, short-term versus long-term effects (so-called time-dependent effects) of a fixed baseline risk factor are addressed. An example is presented showing that underweight is a strong risk factor for mortality in dialysis patients, especially in the short run. In contrast, overweight is a risk factor for mortality, which is stronger in the long run than in the short run. In addition, the analysis of how time-varying risk factors (so-called time-dependent risk factors) are related to mortality is demonstrated by paying attention to the pitfall of adjusting for sequelae. The proper analysis of effects over time should be driven by a clear research question. Both kinds of research questions, that is those of time-dependent effects as well those of time-dependent risk factors, can be analyzed with time-dependent Cox regression analysis. It will be shown that using time-dependent risk factors usually implies focusing on short-term effects only.

Kidney International (2008) **74,** 994–997; doi:10.1038/ki.2008.328; published online 16 July 2008

Correspondence: *Friedo W. Dekker, Leiden University Medical Centre, Department of Clinical Epidemiology, Leiden, The Netherlands. E-mail: f.w.dekker@lumc.nl*

Received 21 February 2008; revised 3 May 2008; accepted 13 May 2008; published online 16 July 2008

In two earlier papers in this series, we described the use of Kaplan–Meier curves and Cox regression techniques to analyze survival data.[1,2] In those analyses, we studied the effect of one or more risk factors assessed at a certain moment in time on subsequent survival. For instance, Tripepi *et al.*[3] investigated whether left atrial volume was associated with mortality in 249 patients on dialysis. Their study population consisted of prevalent patients who were on dialysis for 15–100 months, and inclusion in the study was considered as the baseline to analyze subsequent mortality over a period of 5 years. This analysis yielded an adjusted relative risk (RR, in Cox regression usually denoted as HR = hazard ratio) of 1.02 per milliliter higher left atrial volume. Just as in this example, an HR usually relates to the entire follow-up period in a study (Figure 1).

However, relating all future survival to a risk factor assessed at a single moment in time may not always be what one wants from a clinical point of view. Two other approaches could be relevant. First, some fixed risk factors may have a different effect on short-term survival than on long-term survival, the so-called time-dependent effects. A well-known example is mortality, which is higher directly after renal transplantation than after having survived the first 3 months. Another example where the effect will depend on the selected time window will be addressed below.

Second, a risk factor itself may change over time. For instance, left atrial volume may well increase in patients as they become older and survive longer on dialysis.[4] It seems attractive to be able to take the updated information on this risk factor into account when studying its association with mortality. To that end, an analysis would be needed that uses serial measurements of this risk factor as a determinant for subsequent survival in a model that uses time-varying or time-dependent risk factors. In the present paper, we describe (1) the interpretation of short-term and long-term effects of fixed risk factors on survival as well as (2) the effects of risk factors that vary over time in a time-dependent analysis. In the context of this paper, we use 'effects' and 'associations' interchangeably, and consistent with Rothman we consider the relative risk (RR or HR) as a measure of effect.

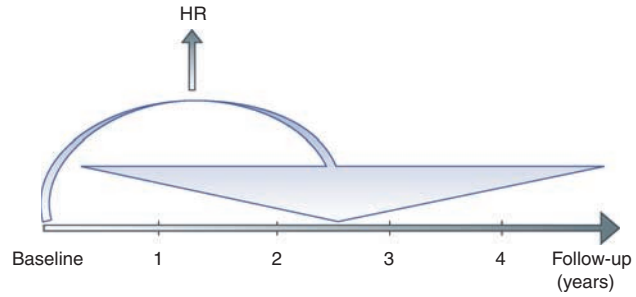

Figure 1 | Long-term effect of baseline risk factor on mortality.

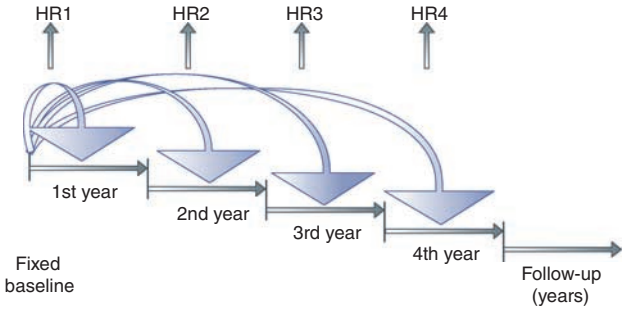

Figure 2 | Time-stratified effects of fixed baseline risk factor on mortality.

Time-dependent effects: short-term versus long-term effects

In studying whether mortality is associated with a certain risk factor, it is important to think about the length of follow-up that needs to be taken into account. For instance, obesity in the general population is a well-known risk factor for cardiovascular diseases and mortality, but only after a long term of exposure, that is after more than 10 years of follow-up. Studies in dialysis patients, however, did not find an effect of obesity on mortality, or they even found a reversed association. This difference in results in dialysis patients as compared with the general population has been labeled as 'reverse epidemiology.'[5] Some explanations for this phenomenon have been suggested, among which confounding, an important factor, is related to the difference in the length of follow-up. Although general population studies can have more than 20 years of follow-up, the survival in patients on dialysis is most often much less than 10 years, irrespective of their body mass index (BMI). Consequently, in most studies, so far the average length of follow-up is quite different between the general population and the patients on dialysis, and this may hamper a fair comparison of the effect of obesity between these groups. Indeed, when the effect was compared in populations with the same length of follow-up, the effect of obesity at the start of dialysis on mortality was similar in hemodialysis patients compared with the general population.[6]

A clear way of showing the influence of the length of follow-up is by reporting mortality rates separately for each year of follow-up.

Example 1: BMI and survival in peritoneal dialysis patients

Snyder *et al.*[7] reported yearly mortality rates in 41,197 incident peritoneal dialysis (PD) patients, broken down by BMI, which was assessed as a fixed risk factor at the start of PD. They showed that in the first year of PD, the adjusted RR of obesity at the start of PD was below 1.0, indicating a protective effect of obesity in the first year of follow-up. Conditional upon having survived the first year on PD, in the second year the adjusted RR of obesity at the start of PD was somewhat above 1.0, whereas conditional upon having survived both the first and the second years on PD, in the third year the adjusted RR of obesity at the start of PD on mortality was even higher. These conditional analyses enabled the authors to show that the effect of obesity at the start of PD on mortality increased over time.

The conditional approach used by Snyder *et al.* may be better than just comparing 1-year survival with 3-year survival, as the time windows overlap and 1-year survival is completely included in 3-year survival, which obscures the time-dependency of the effect. A graphical representation of the year-specific or time-stratified HRs is provided in Figure 2. Interestingly, using the same approach, the detrimental effect of baseline underweight on mortality decreased over time of follow-up, both in PD and in hemodialysis patients,[7,6] implying that the short-term effect of underweight is stronger than the long-term effect.[8]

In the example above, the effects of baseline risk factors became stronger or weaker over time, and these were explored in the analysis by stratifying by time. We saw that the RR of obesity at the start of PD was not the same during the years of follow-up and that in more extreme examples the survival curves could even cross each other. This phenomenon is also referred to as 'non-proportional hazards,' as explained in the previous paper in this series.[2] Instead of manually stratifying by time window, this can also be explored by running a Cox regression analysis in such a way that it gives a separate RR for each time period. For instance, the BMI at the start of dialysis can be used to predict mortality in the first year. Then, BMI at the start of dialysis is used again, but now to predict mortality in the second year, conditionally on having survived the first year. Later, BMI at the start of dialysis is used again, but now to predict mortality in the third year, conditionally on having survived the second year, and so on. This is depicted in Figure 2. This conditional or time-stratified RR is also referred to as a time-dependent RR, as the RR is dependent on the specific time window you look at. Note that the risk factor we study here is still measured only once, that is, at the baseline moment of the study, and thus in this example it can be considered a fixed risk factor. Note also that a separate RR is estimated for each time window you study.

Time-dependent risk factors for survival

In contrast to the example above, a Cox regression analysis can also be used to study the effect of a risk factor whose

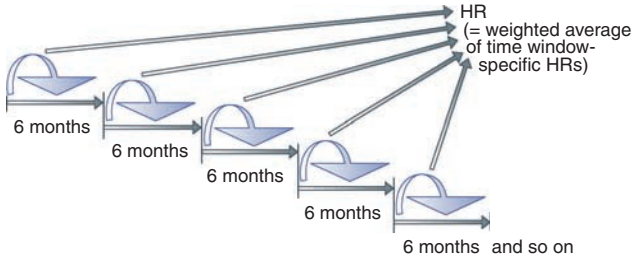

Figure 3 | Effect of time-varying risk on mortality.

value changes over time. Such risk factors are called time-varying risk factors or time-dependent covariates. In Cox regression with time-dependent risk factors, one defines a 'time-varying' factor that refers to serial measurements of that risk factor during follow-up, and includes that 'time-varying' or 'time-dependent' risk factor in a Cox regression model. Most statistical packages will easily do this analysis. The results, however, are not always easy to interpret, and it is therefore easy to make mistakes. Here, we will explain this type of analysis.

Basically, in a time-dependent analysis, the follow-up time for each patient is divided into different time windows. First, for each time –window, a separate Cox analysis is carried out using the specific value of the time-dependent variable at the beginning of that specific time window (Figure 3). Second, a weighted average of all the time window-specific results is calculated. This weighted average of a series of relatively short-term effects is presented as the result of the analysis as one RR (Figure 3). Of course, non-time-dependent variables, for example 'sex,' can be also used as fixed confounders or covariates for all the time windows included.

Example 2: Time-dependent effect of BMI on mortality

Kovesdy *et al.*[9] studied the association of BMI with mortality in 512 male patients with chronic kidney disease not yet on dialysis. Using the lowest decile of BMI as a reference category, they showed that higher BMI was associated with lower mortality, both in a 'fixed-covariate' or traditional Cox regression and in a time-dependent Cox model in which BMI was updated every 6 months. This may imply that in this study compared with the lowest BMI, a high BMI was associated with lower mortality both for the short term (6 months) and for a somewhat longer term (a median follow-up of 2.3 years).

Caveats

It may seem appropriate to always use a time-dependent Cox regression model that takes into account that risk factors may change over time, rather than using a traditional Cox model with only fixed baseline risk factors. However, a time-dependent Cox regression provides an answer to a different research question compared with a traditional Cox regression analysis. Although a traditional Cox analysis also addresses the relatively long-term effects of a risk factor on mortality (Figure 1), a time-dependent Cox analysis only addresses relatively short-term effects (Figure 3). As always, clinical reasoning and a sound research question should drive the choice for a proper analysis of the data.

Another caveat in time-dependent Cox regression analysis can emerge when confounders are also measured repeatedly during follow-up and included as time-varying or time-dependent variables in the model. We know from the earlier paper in this series addressing confounding that it is generally inappropriate to adjust for a covariate that is or can be a result of the risk factor we study.[10] In a traditional Cox model, our risk factor and potential confounders are all measured at or before baseline, and we only need to make sure that covariates are not in the causal pathway of the risk factor toward the outcome we study. With time-dependent covariates, however, there is an even greater risk that a covariate during follow-up is (partly) a result of the risk factor we study. In other words, a time-dependent covariate could be a confounder, but could also be an intermediate in the causal pathway. In an important paper, Wolfe[11] warns very clearly about what he calls 'adjusting for sequelae.' For instance, if one would compare mortality between patients starting on hemodialysis versus PD, it is not permissible to adjust for hospitalizations or newly developed comorbidities during follow-up in a time-dependent Cox model, as it is quite conceivable that hospitalizations or new comorbidities are at least affected by the type of treatment chosen.[11]

Conclusion

In addition to the traditional Cox regression analysis studying mortality associated with a baseline risk factor, two time-dependent approaches exist. In one approach, the effect of a fixed baseline risk factor on mortality in different time windows is studied (time-stratified effects). This results in separate HRs for distinct time windows. In the second approach, a risk factor that changes over time is studied in relation to subsequent mortality. This approach results in one HR that can be considered as a weighted average of short-term effects on mortality. Clinical reasoning and a sound research question should drive the choice for a proper analysis of the data.

REFERENCES

1. Jager KJ, van Dijk PC, Zoccali C *et al.* The analysis of survival data: the Kaplan–Meier method. *Kidney Int* 2008 (in press).
2. van Dijk PC, Jager KJ, Zwinderman AH *et al.* The analysis of survival data in nephrology. Basic concepts and methods of Cox regression. *Kidney Int* 2008 (in press).
3. Tripepi G, Benedetto FA, Mallamaci F *et al.* Left atrial volume in end-stage renal disease: a prospective cohort study. *J Hypertens* 2006; **24**: 1173–1180.
4. Tripepi G, Benedetto FA, Mallamaci F *et al.* Left atrial volume monitoring and cardiovascular risk in patients with end-stage renal disease: a prospective cohort study. *J Am Soc Nephrol* 2007; **18**: 1316–1322.
5. Kalantar-Zadeh K. Causes and consequences of the reverse epidemiology of body mass index in dialysis patients. *J Ren Nutr* 2005; **15**: 142–147.
6. de Mutsert R, Snijder MB, van der Sman-de Beer F *et al.* The association between body mass index and mortality is similar in the hemodialysis population and the general population at high age and equal duration of follow-up. *J Am Soc Nephrol* 2007; **18**: 967–974.
7. Snyder JJ, Foley RN, Gilbertson DT *et al.* Body size and outcomes on peritoneal dialysis in the United States. *Kidney Int* 2003; **64**: 1838–1844.

8. Kalantar-Zadeh K, Horwich TB, Oreopoulos A *et al.* Risk factor paradox in wasting diseases. *Curr Opin Clin Nutr Metab Care* 2007; **10**: 433–442.
9. Kovesdy CP, Anderson JE, Kalantar-Zadeh K. Paradoxical association between body mass index and mortality in men with CKD not yet on dialysis. *Am J Kidney Dis* 2007; **49**: 581–591.
10. Jager KJ, Zoccali C, MacLeod A *et al.* Confounding: what it is and how to deal with it. *Kidney Int* 2008; **73**: 256–260.
11. Wolfe RA, Strawderman RL. Logical and statistical fallacies in the use of Cox regression models. *Am J Kidney Dis* 1996; **27**: 124–129.