ELSEVIER

CrossMark

Journal of Clinical Epidemiology 70 (2016) 101−105

Journal of Clinical Epidemiology

Competing risk bias in Kaplan−Meier risk estimates can be corrected

Carl van Walraven[a,b,c,d,*], Steven Hawken[c]

[a]*Medicine, University of Ottawa*
[b]*Epidemiology and Community Medicine, University of Ottawa*
[c]*Ottawa Hospital Research Institute*
[d]*ICES uOttawa*

Accepted 24 August 2015; Published online 26 September 2015

Abstract

Background: Kaplan−Meier (KM) analyses are frequently used to measure outcome risk over time. These analyses overestimate risk whenever competing events are present. Many published KM analyses are susceptible to such competing risk bias. This study derived and validated a model that predicted true outcome risk based on the biased KM risk.

Methods: We simulated survival data sets having a broad range of 1-year true outcome and competing event risk. Unbiased true outcome risk estimates were calculated using the cumulative incidence function (CIF). Multiple linear regression was used to determine the independent association of CIF-based true outcome risk with the biased KM risk and the proportion of all outcomes that were competing events.

Results: The final model found that both the biased KM-based risk and the proportion of all outcomes that were competing events were strongly associated with CIF-based risk. In validation populations that used a variety of distinct survival hazard functions, the model accurately predicted the CIF ($R^2 = 1$).

Conclusions: True outcome risk can be accurately predicted from KM estimates susceptible to competing risk bias.

Keywords: Kaplan-Meier estimates; Survival analysis; Product-limit; Competing risks; Bias; Cumulative incidence function

1. Introduction

The Kaplan−Meier (KM) estimator is commonly used to measure the survival function for time to event data. Observations in the KM analysis that end without a true outcome occurring are censored. True outcome risk in censored subjects is assumed to be the same as those who remain in the cohort [1]. This assumption is incorrect when subjects are censored because of a competing event. Competing events preclude the true outcome from occurring [2], so that subjects sustaining a competing event have an outcome risk of zero. When the KM analysis censors such people at the time of their competing event, their outcome risk is assumed to be the same as others in the cohort. As a result, risk estimates from KM analyses with competing events exceed true outcome risk [2].

"Competing risk bias" is very common in the medical literature. Koller et al. [3] reviewed 35 observational studies with KM estimates published in high-impact medical journals and found that 24 (67%) were susceptible to competing risk bias. We recently examined 100 studies with KM estimates that were randomly selected from high-impact medical journals and found that 46% were susceptible to competing risk bias [4]. The potential of competing risks biasing study results has recently been highlighted in general medical journals [2].

When competing events are present, risk estimates from KM analyses are overinflated, but the extent of that bias is difficult to gauge. In this study, we attempted to derive and internally validate a method to determine the extent to which KM risk estimates susceptible to competing risk bias are overinflated.

2. Methods

2.1. Study terminology and overview

In this study, an event of interest is termed a "true outcome." Events whose occurrence prohibits true outcomes from occurring are termed "competing events."

Conflict of interest: None.

Funding: This study was supported by the Department of Medicine, University of Ottawa.

* Corresponding author. ASB1-003 1053 Carling Ave, Ottawa, Ontario K1Y 4E9, Canada. Tel.: +1 613-761-4903; fax: +1 613-761-5492.

E-mail address: carlv@ohri.ca (C. van Walraven).

http://dx.doi.org/10.1016/j.jclinepi.2015.08.010
0895-4356/

What is new?

Key findings

- Many published Kaplan−Meier (KM) analyses overestimate risk because competing events are present.
- This study found that unbiased outcome risk that accounts for the presence of competing risks can be accurately predicted from using the biased KM risk estimate and the proportion of all outcomes that were competing events.

What this adds to what was known?

- This finding permits one to calculate unbiased risk estimates from published KM analyses susceptible to competing risk bias.

What is the implication and what should change now?

- The utility of biased KM risk estimates susceptible to competing risk bias is increased because this model can predict true, unbiased risk estimates.

Together, true outcomes and competing events make up ''all outcomes.'' Outcome risk based on the KM estimate was termed ''Kaplan−Meier risk'' and was calculated as 1 − KM estimate.

This was a simulation study. In step 1, we created ''base data sets'' having a broad range of event hazards. In step 2, we combined pairs of base data sets in random combinations to generate ''analytical data sets'' having a complete range of true outcome and competing event risks. In step 3, we used a random sample of analytical data sets to derive a multivariable linear model to predict the unbiased true outcome risk (ie, outcome risk not influenced by competing events). In step 4, we tested the model's performance in validation data sets using various survival functions. All analyses were done using SAS 9.3 (Cary, NC).

2.2. Step 1—creating base data sets

Here, we created 75 ''base data sets'' that simulated the survival of 1,000 patients over 1 year of observation using a modification of the SAS macro %prepdata [5]. This macro uses the RANEXP function to randomly generate survival times on the basis of an exponential hazard function. This hazard was varied, so that the 1-year event risks in the base data sets ranged between 0.6% and 100%.

2.3. Step 2—creating analytical data sets

In this step, 2 of the 75 base data sets from step 1 were randomly selected and combined to form an analytical data set. All outcomes in one base data set were classified as ''true outcomes,'' whereas those from the other were classified as ''competing events.'' We first randomly combined all base data sets in every combination to create 5,625 analytical data sets having 2,000 observations each. Creating our analytical data sets using this combination approach gave us more control of the relative number of true outcomes and competing events compared to randomly assigning outcomes when creating the base data sets to true outcomes or competing events.

These analytical data sets had true outcome risks that never exceeded 0.5 because none of the observations from the ''competing event'' base data set could have a true outcome. We therefore created another 5,625 combined data sets in which a random number of observations from the ''competing event'' base data set were used. This resulted in a total of 11,250 analytical data sets having between 1,002 and 2,000 observations and a 1-year true outcome risk that varied between 0.2% and 97.9%.

2.4. Step 3—linear model to predict true outcome risk unbiased from competing events

Our objective was to create a model that would predict the unbiased true outcome risk in the presence of competing risks. The cumulative incidence function (CIF, also known as the ''cumulative risk'' and ''cumulative incidence estimate'') estimates event risk that is insensitive to the presence of competing events [6]. In the absence of competing events, the CIF equals 1 − the KM estimate. We used the %CIF macro to calculate the CIF for each survival data set [7].

Two-thirds of data sets from step 2 were randomly selected to derive a model that predicted CIF-based true outcome risk. We hypothesized that this could be predicted from the KM risk (ie, 1 − KM estimate) and the proportion of all outcomes that were competing events (previous studies have found that the KM risk overestimates the CIF-based risk more when the number of competing events relative to true outcomes increases [2,8]). We therefore calculated for each data set KM risk (using PROC LIFETEST) and the proportion of all outcomes that were competing events.

We log-odds transformed all study variables (ie, the CIF-based outcome risk, the KM risk, and the proportion of all outcomes that were competing events). We then used multivariable linear modeling (PROC REG) to determine the association of the transformed KM risk, the proportion of all outcomes that were competing events, and the interaction of these two variables with the CIF-based risk. We used fractional polynomials [9] to determine the best nonlinear fit between explanatory variables and the outcome.

2.5. Step 4—evaluation of model performance

The final model was evaluated in the validation data sets. Deviation between predicted and observed values of the