Cumulative incidence estimation in the presence of competing risks

Vincenzo Coviello
Department of Prevention
Azienda U.S.L. BA/1, Italy
coviello@mythnet.it

May Boggess
StataCorp
College Station TX 77840
mboggess@stata.com

Abstract. When competing risks are present, the appropriate estimate of the failure probabilities is the cumulative incidence. `stcompet` creates new variables containing the estimate of this function, its standard error, and $\ln(-\ln)$ transformed confidence bounds. Two examples are presented to illustrate the use of the new command and some key features of the cumulative incidence.

Keywords: st0059, stcompet, survival analysis, competing risks, cumulative incidence

1 Introduction

Several authors have stressed the methodological problems arising in the analysis of cohort studies or clinical trials when competing risks are present. A competing risk is defined as an event whose occurrence precludes or alters the probability of occurrence of a main event under examination. In this setting, the appropriate estimate of the probability of failure is described by the so-called cumulative incidence. This function is not calculated by many statistical software packages so that, as Gooley et al. (1999) advocate, the complement of a Kaplan–Meier estimate (`1-KM`) is frequently misused instead.

Indeed, to estimate (`1-KM`), the failures from a competing event are treated as censored at the time this event occurs. This way, we assume that the patients failing from a competing risk are no more or less likely to fail from the cause of interest than the patients still at risk beyond this time. When the aim is to estimate the failure probabilities, this censoring is inappropriate because, after a competing event has occurred, failure from the cause of interest is no longer possible.

The purpose of this paper is to make available for Stata users a new command, `stcompet`, allowing users to estimate with one instruction, quickly and reliably, the cumulative incidence, its standard error, and its confidence interval.

This paper presents methods implemented and the syntax of the new command. Then, we provide two examples to illustrate its use and some key features of the cumulative incidence.

2   Methods

The cumulative incidence for the failure type k is estimated as

$$\widehat{I}_k(t) = \sum_{j|t_j \leq t} \widehat{S}(t_{j-1}) \frac{d_{kj}}{n_j}$$

where $\widehat{S}(t_j - 1)$ is the Kaplan–Meier estimate of the overall survival function, that is, considering failures of any kind, and the second factor is an estimate of the hazard of failure type k (Marubini and Valsecchi 1995, 338, equation 10.11).

This definition implies that the cumulative incidence is a function of the hazards of all the competing events and not solely of the hazard of the event to which it refers. Furthermore, the sum of all cumulative incidences has the nice feature that it equals $1 - \widehat{S}(t)$, the complement of the overall Kaplan–Meier estimate of survival considering failures of any kind.

The variance estimator of the cumulative incidence implemented in `stcompet` is as follows:

$$\begin{aligned}
\mathrm{var}\{\widehat{I}_k(t_j)\} &= \sum_{\alpha=1}^{j} \left[\left\{ \widehat{I}_k(t_j) - \widehat{I}_k(t_\alpha) \right\}^2 \frac{d_\alpha}{n_\alpha(n_\alpha - d_\alpha)} \right] \\
&+ \sum_{\alpha=1}^{j} \left\{ \widehat{S}(t_{\alpha-1}) \right\}^2 \left(\frac{n_\alpha - d_{k\alpha}}{n_\alpha} \right) \left(\frac{d_{k\alpha}}{n_\alpha^2} \right) \\
&- 2 \sum_{\alpha=1}^{j} \left\{ \widehat{I}_k(t_j) - \widehat{I}_k(t_\alpha) \right\} \left\{ \widehat{S}(t_{\alpha-1}) \right\} \left(\frac{d_{k\alpha}}{n_\alpha^2} \right)
\end{aligned}$$

where $d_j = \sum_{k=1}^{C} d_{kj}$ and C is the number of causes of failure (Marubini and Valsecchi 1995, 341, equation 10.12).

A general version of this formula was derived by Dinse and Larson (1986) using the delta method and was used by Gaynor et al. (1993) in a study of oncological data.

Choudhury recently issued the S-Plus code to estimate the cumulative incidence and its standard error (Choudhury 2002, appendices B and C). His code has been used to validate the results achieved with the new Stata command comparing them with the S-Plus output. In our checks, the variance of the cumulative incidence computed with `stcompet` is exactly the same as that achieved with S-Plus.

The confidence interval calculated by `stcompet` is formed using $\ln(-\ln)$ transformed bounds:

$$\widehat{I}_k(t)^{\exp\left[\frac{\pm c_{\alpha/2} \widehat{\sigma}_k(t)}{\widehat{I}_k(t) \ln\{\widehat{I}_k(t)\}} \right]}$$

where $c_{\alpha/2}$ is the upper $\alpha/2$ percentile of the standard normal distribution and $\widehat{\sigma}_k(t)$ is the square root of $\mathrm{var}\{\widehat{I}_k(t_j)\}$ (Choudhury 2002, equation 4).

Choudhury demonstrates using simulations, that the $\ln(-\ln)$ transform has better coverage probability than both the linear and arcsine transforms (Choudhury 2002, section 4).

3   Syntax

stcompet *newvarname*= { ci | se | hi | lo } $\big[\,[$*newvarname*$]$= ...$\big]$ $\big[$...$\big]\big]$ [if *exp*]
$\big[$in *range*$\big]$, compet1(*numlist*) $\big[$compet2(*numlist*) compet3(*numlist*) ...
compet6(*numlist*) by(*varname*) <u>l</u>evel(#) $\big]$

As with all st commands, we must use stset before using stcompet. Furthermore, in the previous stset, we must specify the failure event as

$$\underline{\mathtt{f}}\mathtt{ailure}(\mathit{failvar}\big[\mathtt{==}\mathit{numlist}\big])$$

where the values in *numlist* are assumed to correspond to the occurrence of the event of interest.

The syntax of stcompet is analogous to that of sts gen. It creates new variables containing the estimated cumulative incidence and its standard error and bounds of confidence interval:

> *newvarname* specifies the name of the new variable where the estimate will be recorded:
>
>> ci produces the cumulative incidence function
>>
>> se produces the standard error of the cumulative incidence
>>
>> hi produces the higher bound of the confidence interval based on $\ln\{-\ln(\mathtt{ci})\}$
>>
>> lo produces the lower bound of the confidence interval based on $\ln\{-\ln(\mathtt{ci})\}$

4   Options

compet1(*numlist*) is not optional because at least one event must compete with the event of interest. A failure of a competing event occurs whenever *failvar*, specified in the previous stset, takes on any of the values of this *numlist*. The function calculated will be estimated with this competing event and the event of interest.

compet2(*numlist*) ... compet6(*numlist*) refer to failures for other competing events.

by(*varname*) produces separate functions by making separate calculations for each value of *varname*. It may take on integer or string values.

level(#) specifies the confidence level, as a percentage, for the pointwise confidence interval around the cumulative incidence functions. The default is is level(95), which is the system parameter c(level).

5 Examples

5.1 Example 1: Simulated data

Gooley et al. (1999) impressively illustrate the difference between the cumulative incidence and the complement of a Kaplan–Meier estimate (1-KM). Their example using a simulated dataset can be reproduced easily using stcompet.

To this aim, two types of failures are assumed, and a time to each type of failure is generated for 5,000 subjects with a constant hazard being 0.25 for the first type of failure and 0.99 for the second. A subject is assumed to fail from the event that occurs earliest if it occurs before time equals to two units:

```
. clear
. set obs 5000
obs was 0, now 5000
. set seed 1234
. gen t25 = -1/.25 * log(1-uniform())
. gen t99 = -1/.99 * log(1-uniform())
. gen time = min(t99,t25)
. gen byte fail = (t25<t99) + 2*(t25>=t99)
. replace fail = 0 if time>=2
(414 real changes made)
. replace time = 2 if time>2
(414 real changes made)
. stset time, f(fail==1) noshow
     failure event:  fail == 1
obs. time interval:  (0, time]
 exit on or before:  failure
  (output omitted )
. sts gen KM = s
. gen Complement = 1- KM
. stcompet CumInc=ci, compet1(2)
```

In the previous lines, the occurrence of the first type of failure is declared in the stset statement as the main cause of failure. Then, using stcompet the occurrence of the second type of failure is specified as a competing risk in the compet1() option. The command creates a new variable, named CumInc, where the estimate of the cumulative incidence is recorded for both types of failure at each time when a corresponding failure occurs. To achieve a plot (figure 1) comparing 1-KM with the cumulative incidence of the first type of failure, a new variable must be created to contain only the estimate pertaining to it.

```
. gen CI=CumInc if fail==1
(4042 missing values generated)
. twoway line Complement CI time, sort xlabel(0(.5)2)
>        ylabel(0(.1).5) ytitle("Probability") legend(off)
>        text(.31 1.2 "1-KM", place(e))
>        text(.15 1.6 "CI", place(e))
```
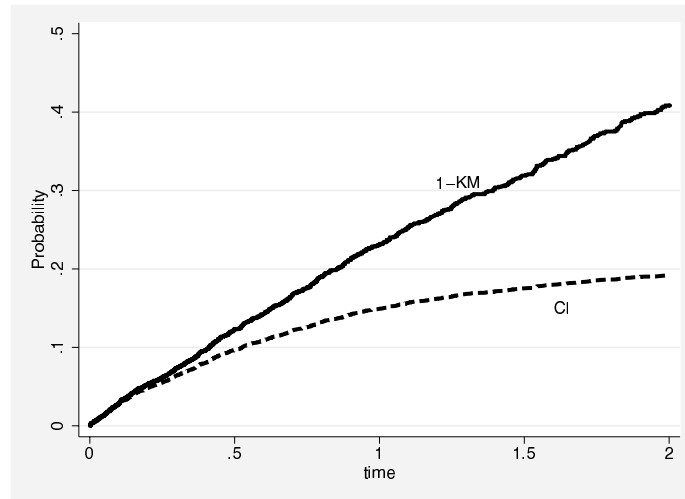
Figure 1: Complement of the KM estimate and cumulative incidence of the first type of failure.

Gooley et al. (1999) stressed in this example that, in a competing risk setting, the complement of the Kaplan–Meier overestimates the true failure probability, whereas the cumulative incidence is the appropriate quantity to use.

Extending the dataset above, the example (Kalbfleisch and Prentice 2002, 252, example 8.1) is now reproduced to illustrate a further point related to the cumulative incidence. As the cumulative incidence depends on all competing hazards, some unexpected results may be produced when two groups are compared.

We suppose that a new treatment halves the type 2 hazard, whereas it does not affect the type 1 hazard. Therefore, to the previous "untreated" subjects we add another 5,000 "treated" subjects whose hazard for the second type of failure is $0.99/2 = 0.495$:

```
. set obs 10000
obs was 5000, now 10000
. gen treat = _n>5000
. replace t25 = -1/.25 * log(1-uniform())
(10000 real changes made)
. replace t99 = -1/.495 * log(1-uniform()) in 5000/l
(5001 real changes made)
. replace time = min(t99,t25)
(6851 real changes made)
. replace fail = (t25<t99) + 2*(t25>=t99)
(6542 real changes made)
. replace fail = 0 if time>=2
(1567 real changes made)
```

Using the option by(*varname*), stcompet can estimate the cumulative incidence separately in the treated and untreated groups:

```
. stset time, f(fail==1) noshow
  (output omitted)
. stcompet Newci=ci,compet1(2) by(treat)
```

To graph this function for each failure type and treatment arm, two new variables are created starting from the estimates stored in the variable Newci:

```
. gen CI_treat1=Newci if treat==1
(6140 missing values generated)
. gen CI_treat0=Newci if treat==0
(5427 missing values generated)
. label var CI_treat1 "Treated"
. label var CI_treat0 "Untreated"
. label define fail 1 "fail 1" 2 "fail 2"
. label val fail fail
. twoway line CI_* time if fail,
>       sort yti("Cumulative Incidence")
>       xti("time") xla(0(.5)2) yla( 5)
>       by(fail, style(compact) rows(1)
>       yrescale legend(pos(12)) note(""))
>       subtitle(, ring(0) pos(11) nobexpand)
>       clp("-##" solid) clw(thick thick)
```
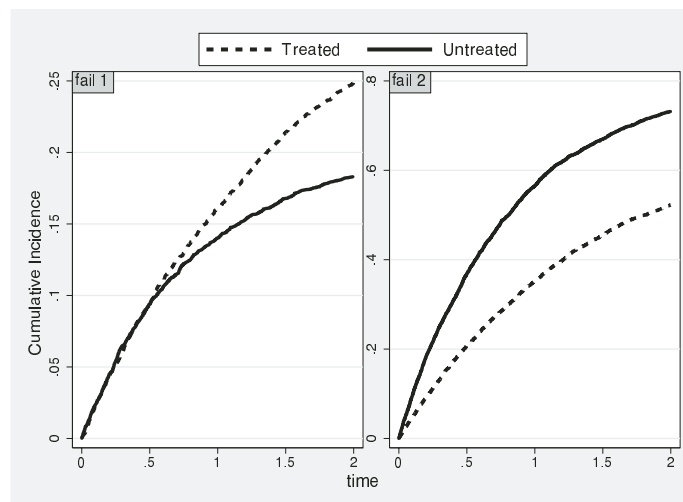

Figure 2: Cumulative incidence of type 1 and type 2 failure in treated and untreated groups.

Figure 2 shows the reduction of the cumulative incidence of the type 2 failure in treated patients but also an apparent effect of the treatment on the incidence of the type 1 failure. As noted by Kalbfleisch and Prentice in their example, the cumulative

incidence of type 1 failure increases in treated patients only as a result of the change in the type 2 rates.

5.2 Example 2: prostate cancer

As an example in a real context, consider the survival times of 506 patients with prostate cancer who are randomly allocated to a treatment with diethylstilbestrol. These data were considered by several authors, first by Byar and Green (1980) and more recently by Kay (1986) and Lunn and McNeil (1995), who use them to fit two different Cox models.[1]

In this dataset, the value of the variable status classifies the cause of death as $1 =$ cancer (the event of interest), $2 =$ cardiovascular disease, and $3 =$ other. The patients are considered treated if they received at least 1 mg of diethylstilbestrol daily.

In this situation, there are two events competing with the event of interest. Furthermore, the probability of occurrence of each of them has to be separately computed in the two treatment arms. Once Stata has been told with `stset` about the basic structure of the survival data at hand, `stcompet` can generate with just one instruction the cumulative incidence and its confidence bounds for each event and two treatment arms:

```
. use prostatecancer, clear
. stset time, f(status==1)
       failure event:  status == 1
 obs. time interval:  (0, time]
  exit on or before:  failure
    (output omitted)
. stcompet CI = ci hilim = hi lowlim = lo, compet1(2) compet2(3) by(treatment)
```

Three new variables have been created: `CI`, `hilim`, and `lowlim` for the cumulative incidence and its high and low confidence bound. The estimates relating to a specific event are recorded at each time when this failure type occurs. So analogously to the previous simulated example, two new variables are generated to distinguish the estimates between treatment arms:

```
. gen CI_tr = CI if treat
. label var CI_tr "Treated"
. gen CI_untr = CI if !treat
. label var CI_untr "Untreated"
. gen hi_tr = hilim if treat
. gen hi_untr = hilim if !treat
. gen lo_tr = lowlim if treat
. gen lo_untr = lowlim if !treat
```

---

[1]In the do-file accompanying this insert, we include the commands for fitting a Cox model by augmenting the data as described in Lunn and McNeil (1995) and Kalbfleisch and Prentice (2002, formula 8.16).

To assess simultaneously the changes in the risk of each failure type induced by treatment, we can easily achieve a plot such as figure 3:
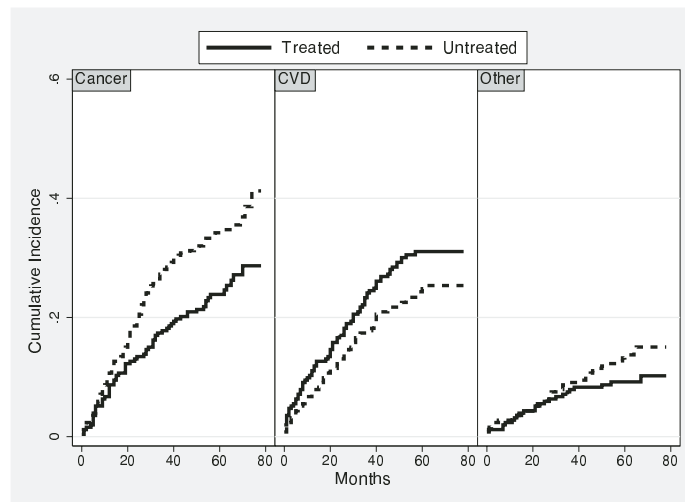

Figure 3: Cumulative incidence of cancer, cardiovascular disease and other causes of death in treated and untreated patients with prostate cancer.

Looking at the graph, we immediately notice that the treatment reduces the probability of cancer death, but simultaneously it increases the risk of cardiovascular death. Nevertheless, remembering the previous simulated example, it might be unwise to conclude that the rate of cardiovascular death is increased too.

Tests for difference between cumulative incidence have been proposed by Gray (1988). Note that, as Gooley et al. (1999) emphasize, this difference depends on the hazards of all competing risks, so testing it is not the same as testing the difference of the hazards for a specific cause. Though stcompet does not compute a test, it allows us to take a glance at this subject graphing (see command below) the confidence bounds of the cumulative incidence for a specific failure in the compared groups.

```
. twoway line CI_* hi_* lo_* time if cause==1,
>       sort c(J J J J J J) clp(l -## l -## l -##)
>       clw(medthick medthick medthick medthick medthick medthick)
>       clc(black black black black black black)
>       yla(0(.2).6) xla(0(20)80) yti("Cancer Cumulative Incidence")
>       ysca(range(0.6)) legend(order(1 2) ring(0) pos(11) cols(1))
```

In figure 4, cancer cumulative incidence for treated and untreated patients is compared. The confidence bounds overlap, so a significant difference between the cumulative incidence of these groups cannot be expected.
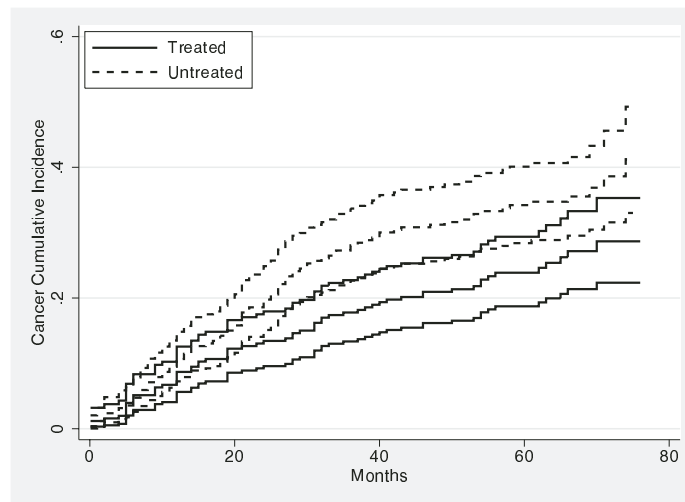
Figure 4: Cumulative incidence and corresponding confidence bounds of cancer death in treated and untreated patients with prostate cancer.

6   Acknowledgment

Enzo Coviello thanks John Moran for his support and encouragement to start this work.

7   References

Byar, D. P. and S. B. Green. 1980. Prognostic variables for survival in a randomized comparison of treatments for prostate cancers. *Bulletin du Cancer* 67: 477–490.

Choudhury, J. B. 2002. Nonparametric confidence interval estimation for competing risks analysis: application to contraceptive data. *Statistics in Medicine* 21: 1129–1140.

Dinse, G. E. and M. G. Larson. 1986. A note on semi-Markov models for partially censored data. *Biometrika* 73: 379–386.

Gaynor, J. J., F. J. Fener, C. C. Tan, D. H. Wu, C. R. Little, D. J. Straus, B. D. Clarkson, and M. F. Brennan. 1993. On the use of cause-specific failure and conditional failure probabilities: examples from clinical oncology data. *Journal of the American Statistical Association* 18: 400–409.

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

Gray, R. J. 1988. A class of k-sample tests for comparing the cumulative incidence of a competing risk. *Annals of Statistics* 16: 1141–1154.

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

Kay, R. 1986. Treatment effects in competing-risks analysis of prostate cancer data. *Biometrics* 42: 203–211.

Lunn, M. and D. McNeil. 1995. Applying Cox regression to competing risks. *Biometrics* 51: 524–532.

Marubini, E. and M. G. Valsecchi. 1995. *Analysing Survival Data from Clinical Trials and Observational Studies*. Chichester, UK: John Wiley & Sons.

About the Authors

Vincenzo Coviello is a Medical Doctor at the Department of Prevention ASL Ba/1V Andria (Italy). His main focus is on cohort studies in occupational epidemiology.

May Boggess is a Statistician for the Technical Services Department at StataCorp.