Competing-risks regression

Roberto G. Gutierrez

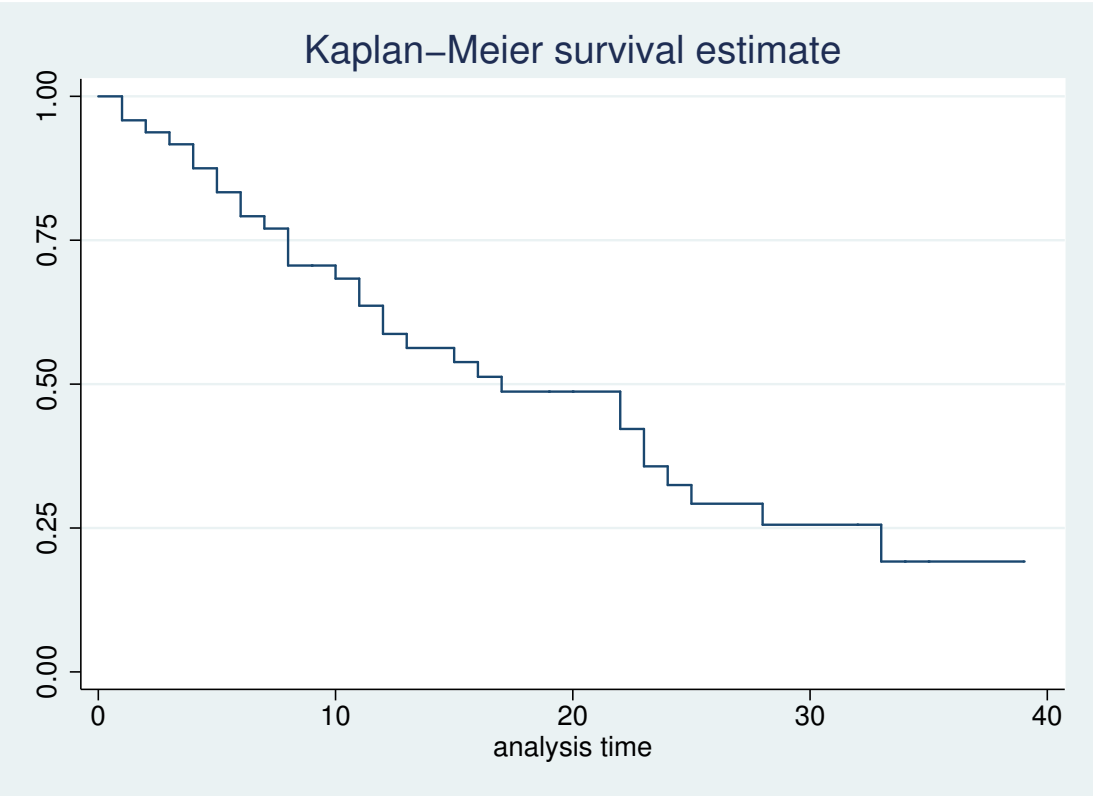
Director of Statistics
StataCorp LP

2009 Australian and New Zealand Stata Users Group Meeting

- Competing risks are events that prevent an event of interest from occurring, rather than just prevent you from seeing it happen (censoring)
- When you have competing events, you want to focus on the cumulative incidence function rather than the survival function
- You can use Cox regression with competing risks, but you have to either
  1. Modify your interpretations
  2. Go through a lot of work to assess covariate effects
- Competing-risks regression by the method of Fine and Gray (1999) is a useful alternative
- Implemented in the `stcrreg` command, new to Stata 11

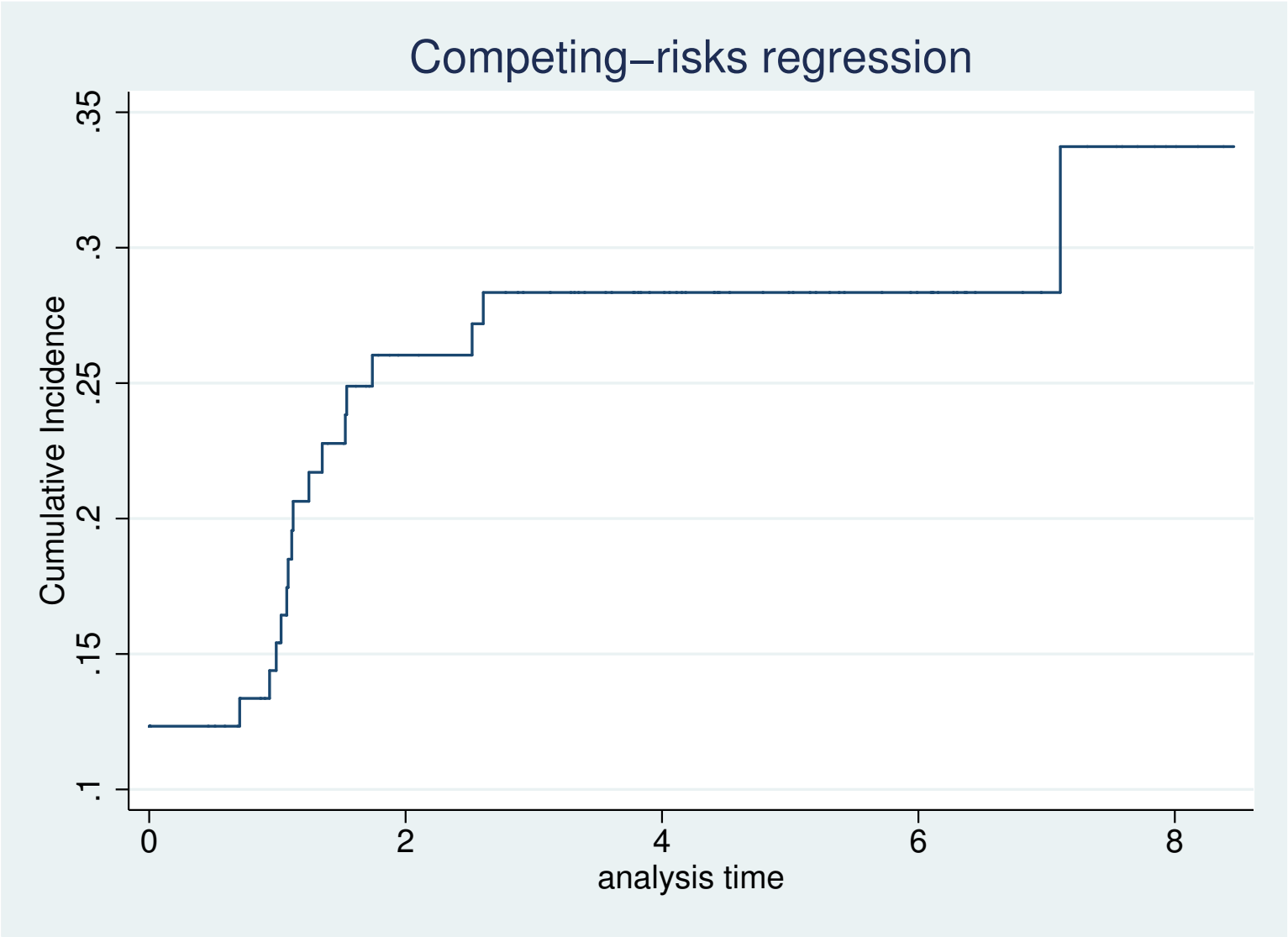
- A competing-risk event is an event that impedes what a researcher actually "wants" to see
- For example, if a researcher is interested in recurrence of breast cancer, a heart attack would be competing event
- This is not the same as censoring. With censoring, the event of interest still happens at a later time. You just don't get to see when
- What confuses the matter is that, for some purposes, it is okay to think of competing events as censorings

```
. sysuse cancer, clear
(Patient Survival in Drug Trial)
. stset studytime died
  (output omitted)
. sts graph
```



- If the event of interest is recurrent breast cancer, a Kaplan-Meier curve answers the question "What is the probability of no breast cancer before 5 months?"
- In a competing-risks setting, a Kaplan-Meier curve is inadequate for two reasons
- First, it fails to acknowledge that breast cancer may never occur, i.e. the probability of breast cancer after time zero is not equal to 1.0

- Second, with competing risks it is better to reverse the temporal ordering of the question
- It makes better sense to ask "What is the probability of breast cancer within 5 months?" than to ask "What is the probability that nothing happens for the first five months, but when something does happen I want it to be breast cancer and not a heart attack?"
- As such, you should use the cumulative incidence function (CIF) instead
- A CIF is just the probability that the event of interest occurs before a given time


Competing–risks regression
Cumulative Incidence
.1
.15
.2
.25
.3
.35
0
2
4
6
8
analysis time

Example

- Geskus (2000) analyzed data from 324 HIV-positive men from the Amsterdam Cohort Studies on HIV infection and AIDS.
- During the course of infection, the syncytium inducing (SI) phenotype appeared in some
- Appearance of SI typically means bad news
- Time to SI appearance in the absence of AIDS is the analysis of interest
- In this context, AIDS with no SI acts as the competing event
- Covariate of interest is variable `ccr5` indicating a mutation in the C-C chemokine receptor 5 gene

```
. webuse hiv_si, clear
(HIV and SI as competing risks)
. describe
Contains data from http://www.stata-press.com/data/r11/hiv_si.dta
  obs:           324                          HIV and SI as competing risks
 vars:             4                          3 Apr 2009 13:40
 size:         5,184 (99.9% of memory free)   (_dta has notes)
-------------------------------------------------------------------------------
              storage  display     value
variable name   type   format      label      variable label
-------------------------------------------------------------------------------
patnr           int    %8.0g                  ID
time            float  %9.0g                  Years from HIV infection
status          byte   %10.0g      stat       1 = AIDS, 2 = SI, 0 = event-free
ccr5            byte   %9.0g       ccr5       1 if WM (deletion in C-C
                                                chemokine receptor 5 gene)
-------------------------------------------------------------------------------
Sorted by:
```

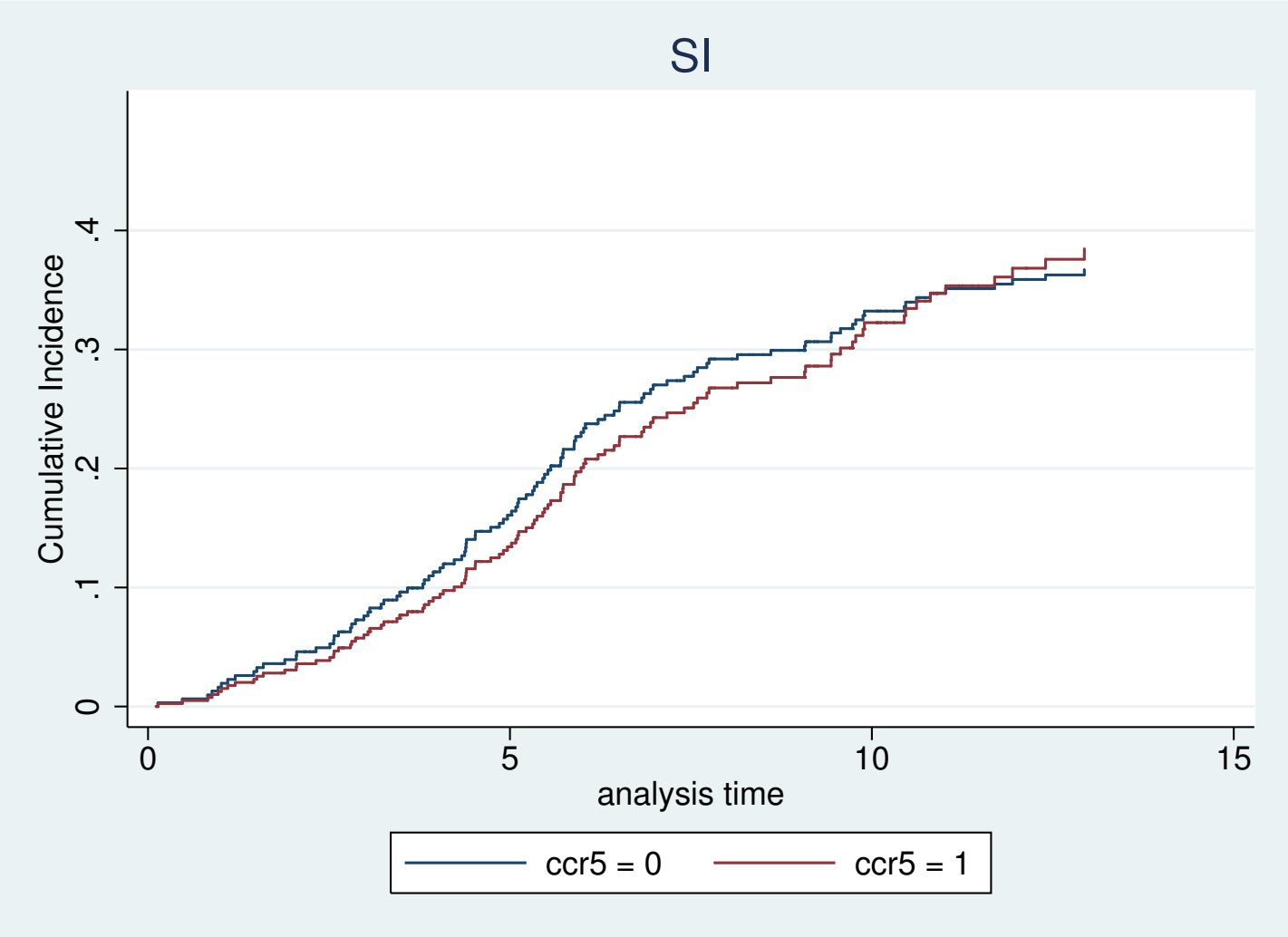
- The research question is: "How does the `ccr5` mutation affect the incidence of SI in the presence of AIDS as a competing risk?"
- Standard Cox regression does not answer that question
- That does not mean it is completely useless, however
- First, you can modify your interpretation if you wish
- Second, you can use more Cox regression to answer our research question, but it is a lot of work. But it is possible

```
. stset time, failure(status == 2)        // SI is the event of interest
  (output omitted)
. stcox ccr5
  (output omitted)
Cox regression -- no ties

No. of subjects =          324                     Number of obs    =       324
No. of failures =          107
Time at risk    =  2261.959996
                                                   LR chi2(1)       =      1.19
Log likelihood  =   -549.73443                     Prob > chi2      =    0.2748
```

_t	Haz. Ratio	Std. Err.	z	P>|z|	[95% Conf.	Interval]
ccr5	.7755334	.1846031	-1.07	0.286	.4863914	1.23656

- The previous results may be interpreted as those for a world where the competing event does not occur
- A full CIF analysis with Cox regression would require the following additional steps
  1. Predict the hazard contributions from the previous Cox regression
  2. Fit a Cox regression with AIDS instead as the event of interest
  3. Predict the hazard contributions from that model fit
  4. Calculate the estimated CIF manually for both `ccr5==0` and `ccr5==1`; see page 209 of [ST] for details
  5. Plot the results
- Moral: Assessing covariate effects on the CIF using Cox regression is a lot of work

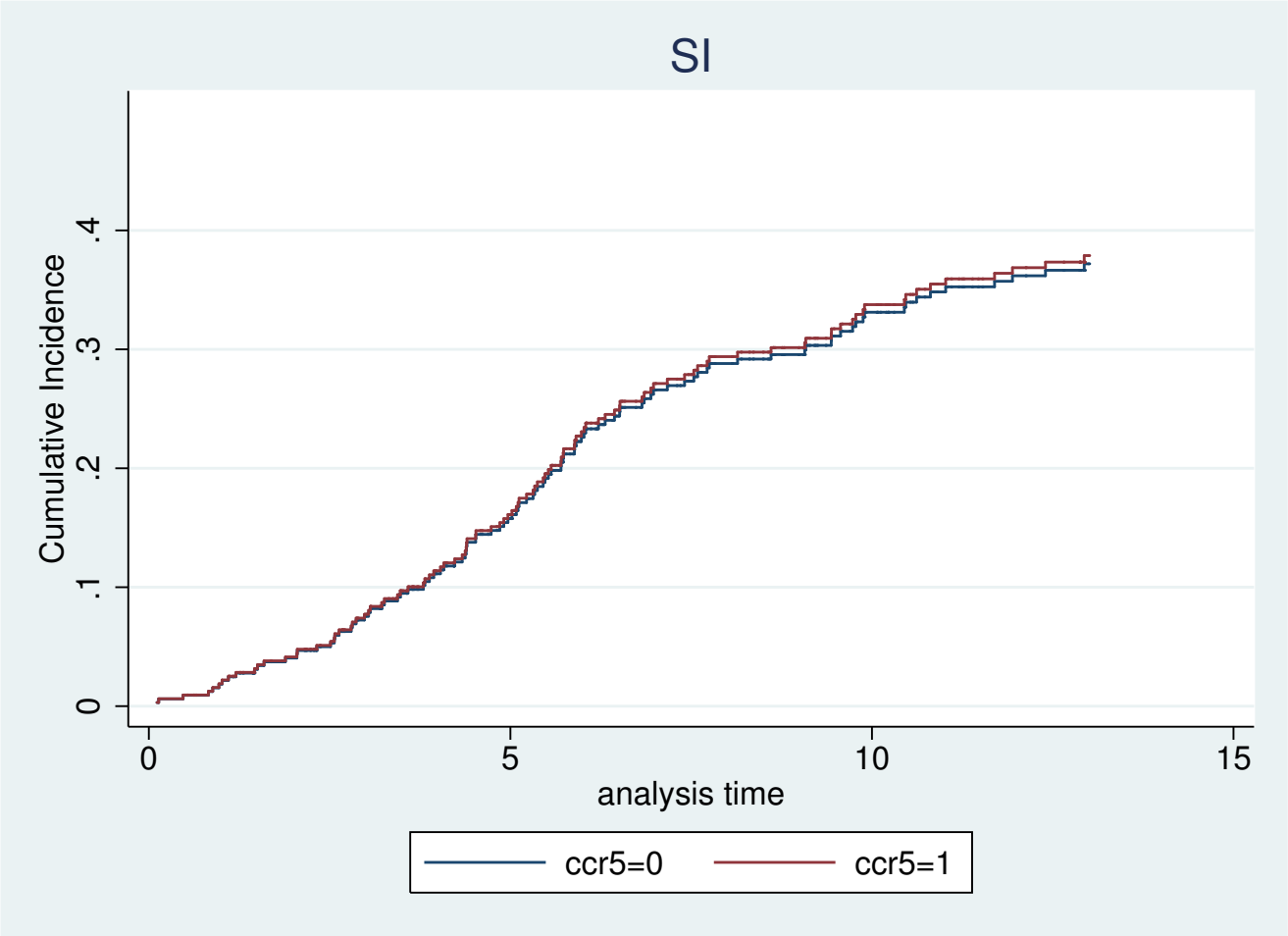

SI
Cumulative Incidence
0
.1
.2
.3
.4
0
5
10
15
analysis time
ccr5 = 0
ccr5 = 1

- An easier way to do CIF covariate analysis is with competing risks regression, according to the model of Fine and Gray (1999)
- They posit a model for the hazard of the subdistribution for the failure event of interest, known as the *subhazard*
- Put simply, they model the CIF directly knowing full well it is not a proper distribution function
- Covariates affect the subhazard proportionally, similar to Cox regression
- You do this in Stata 11 using `stcrreg`. `stcurve` after `stcrreg` will plot comparative CIFs for you

```
. stset time, failure(status == 2)        // SI is the event of interest
  (output omitted)
. stcrreg ccr5, compete(status == 1)      // AIDS is the competing event
  (output omitted)
Competing-risks regression                          No. of obs       =        324
                                                    No. of subjects  =        324
Failure event  : status == 2                        No. failed       =        107
Competing event: status == 1                        No. competing    =        113
                                                    No. censored     =        104
                                                    Wald chi2(1)     =       0.01
Log pseudolikelihood = -579.06241                   Prob > chi2      =     0.9172
```

_t	SHR	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
ccr5	1.023865	.2324119	0.10	0.917	.6561827	1.597574

```
. stcurve, cif at1(ccr5=0) at2(ccr5=1) title(SI) range(0 13) yscale(range(0 0.5))
```



- Competing risks are events that prevent an event of interest from occurring
- If you have competing risks, you want to look at CIFs instead of survival functions
- CIF analysis with Cox regression is possible, but difficult
- `stcrreg` followed by `stcurve` is the easier way to go
- Keep in mind, however, that easier does not mean correct. There are model assumptions to be made for either of the two approaches

Fine, J. and R. Gray. 1999. A proportional hazards model for the subdistribution of a competing risk. *Journal of the American Statistical Association*. **94**: 496–509.

Geskus, R. B. 2000. On the inclusion of prevalent cases in HIV/AIDS natural history studies through a marker-based estimate of time since seroconversion. *Statistics in Medicine*. **19**: 1753–1769.