Estimating and modelling cumulative incidence functions using time-dependent weights

Paul C Lambert[1,2]

[1]Department of Health Sciences, University of Leicester, Leicester, UK
[2]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden

UK Stata Users Group, London, September 2013

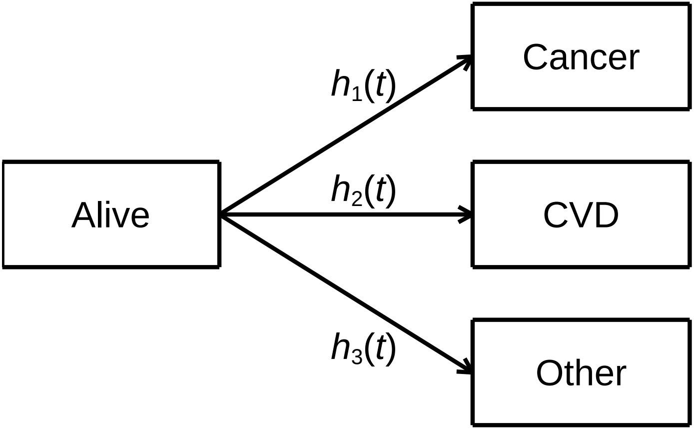
Competing Risks

- In survival analysis individuals are often at risk of more than one event.
- For example, individuals diagnosed with breast cancer are,
  - at risk of death from their cancer
  - at risk of death from other causes

Competing Risks

- In survival analysis individuals are often at risk of more than one event.
- For example, individuals diagnosed with breast cancer are,
  - at risk of death from their cancer
  - at risk of death from other causes
- The probability of dying from cancer will depend upon the mortality rate due to cancer and the mortality rate due to other causes.
- This is a classic competing risks situation.

Competing risks schematic



Cause specific hazard function

- For cause k,

$$h_k(t) = \lim_{\delta \to 0} \frac{P\left(t \leq T < t + \delta, \text{event} = k | T > t\right)}{\delta}$$

- To still be at risk at time t a subject can not have died of cause k or any of the $K - 1$ other causes.
- Total hazard (mortality) rate

$$h(t) = \sum_{k=1}^{K} h_k(t)$$

- All cause survival

$$S(t) = \exp\left(-\int_0^t h(u)du\right) = \exp\left(-\int_0^t \sum_{k=1}^{K} h_k(u)du\right)$$

Cause specific cumulative incidence function

- We want the probability of dying of cause k accounting for the competing risks.
- For cause k.

$$CIF_k(t) = P\left(T \leq t, \text{event} = k\right)$$

$$CIF_k(t) = \int_0^t S(u)h_k(u)du$$

Cause specific cumulative incidence function

- We want the probability of dying of cause k accounting for the competing risks.
- For cause k.

$$CIF_k(t) = P\left(T \leq t, \text{event} = k\right)$$

$$CIF_k(t) = \int_0^t S(u)h_k(u)du$$

$$CIF(t) = \sum_{k=1}^{K} CIF_k(t)$$

- Note: CIF does not require independence between causes.
- For further details on competing risks see references [1, 2, 3]
- Post estimation command `stpm2cif` will estimate CIFs and related measures after using `stpm2` to model cause-specific hazards [4, 5]

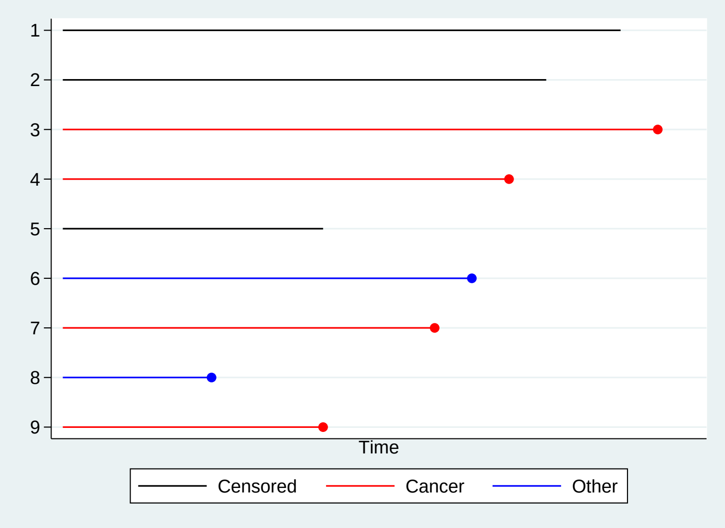
Key Paper: Geskus 2011[8]

- Geskus showed estimation and modelling of the CIF can use weighted versions of standard estimators.
- `crprep` function in R to restructure data and calculate weights[6].
- I will describe a new command `stcrprep` that has similar functionality to `crprep`, but also some extensions to enable parametric models for the CIF to be easily fitted.
- After expansion and weighting of the data,
  - `sts graph, failure` will plot CIF.
  - `sts test` will perform test for differences in CIFs[7].
  - `stcox` will fit a Fine and Gray model (same as `stcrreg`).
  - `estat phtest` can be used to assess proportional subhazards.
  - `streg`, `stpm2` can be used to fit parametric models for CIF.

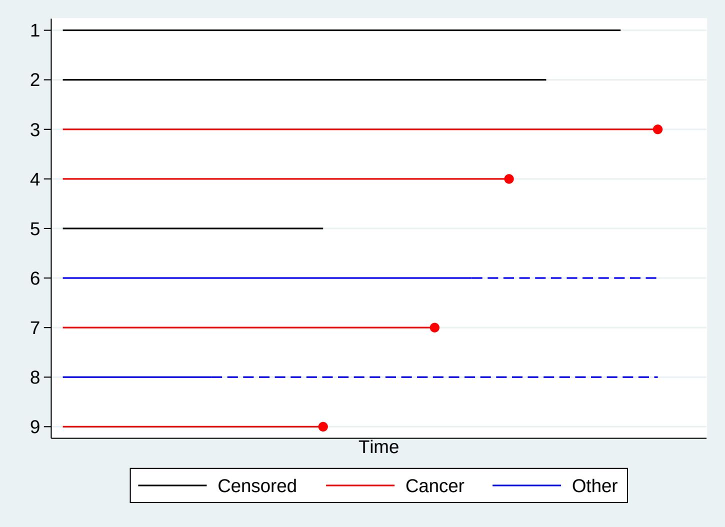
Data expansion and weighting

- Define event of interest.
- Subjects that have a competing event are kept in the risk set to the end of follow-up.
- However, there is a a chance that they would be censored after their competing event.
- Estimate censoring distribution.
- Weights depend on conditional probability of not being censored after competing event.

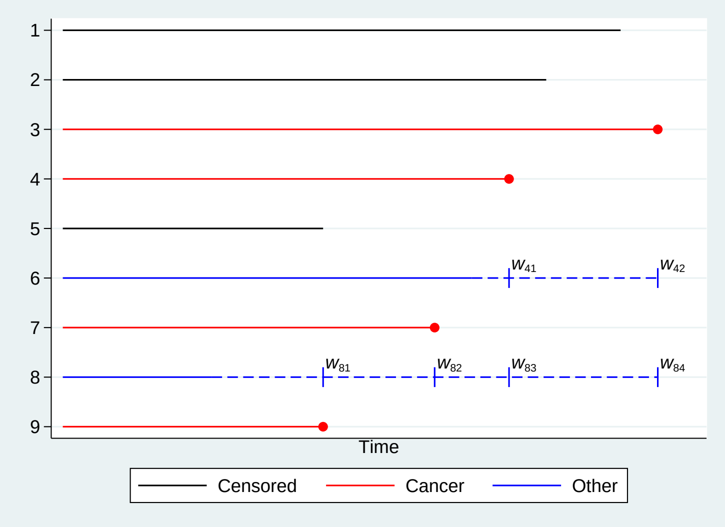
Data Expansion for Competing Events



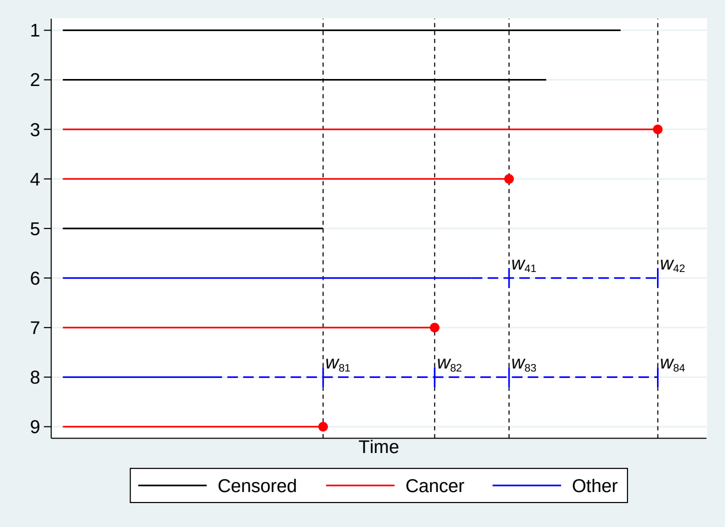
Data Expansion for Competing Events



Data Expansion for Competing Events



Data Expansion for Competing Events



Initial data

Competing event: d == 2

```
. stset t, failure(d==1,2) id(id)
  (output omitted)
. list id d _t0 _t _d, noobs sep(0)
```

id	d	_t0	_t	_d
1	1	0	3.5	1
2	2	0	2	1
3	1	0	5	1
4	2	0	5.5	1
5	0	0	3.5	0
6	1	0	6	1
7	1	0	8	1
8	0	0	6.5	0
9	0	0	7.5	0

Using stcrprep

Competing event: d == 2

```
. stcrprep, events(d) trans(1) noshorten
. gen event = d == failcode
. stset tstop [iw = weight_c], failure(event) enter(tstart) id(id)
  (output omitted)
. list id d _t0 _t _d weight_c, noobs sep(0)
```

id	d	_t0	_t	_d	weight_c
1	1	0	3.5	1	1
2	2	0	2	0	1
2	2	2	3.5	0	1
2	2	3.5	5	0	.85714286
2	2	5	6	0	.85714286
2	2	6	8	0	.28571429
3	1	0	5	1	1
4	2	0	5.5	0	1
4	2	5.5	6	0	1
4	2	6	8	0	.33333333
5	0	0	3.5	0	1
6	1	0	6	1	1
7	1	0	8	1	1
8	0	0	6.5	0	1
9	0	0	7.5	0	1

European Blood and Marrow Transplantation Data

- 1977 patients from the European Blood and Marrow Transplantation (EBMT) registry who received an allogeneic bone marrow transplantation[6].
- Events are death and relapse
  - 836 censored
  - 456 relapse
  - 685 died
- One covariate of interest, the EBMT risk score, which has been categorized into 3 groups (low, medium and high risk).

Using stcrprep

stcrprep

```
. stset time, failure(status==1,2) scale(365.25) id(patid)
  (output omitted)
. stcrprep, events(status) keep(score) trans(1 2) byg(score)
. gen event = status == failcode
. stset tstop [iw=weight_c], failure(event=1) enter(tstart) noshow
  (output omitted)
```

Using stcrprep

stcrprep

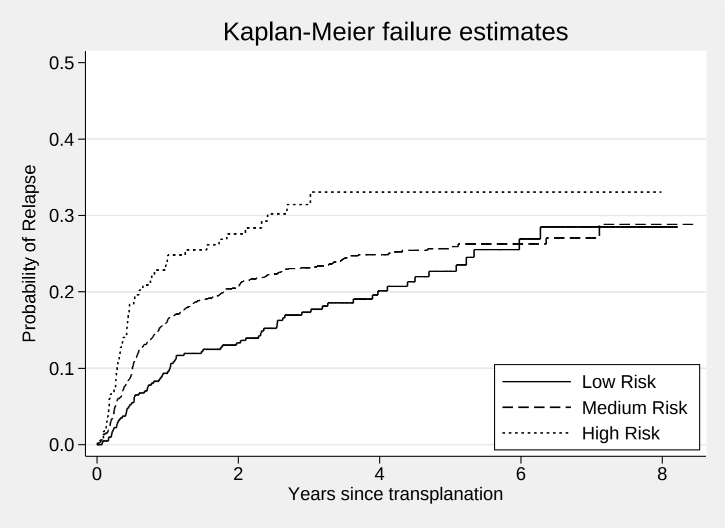
```
. stset time, failure(status==1,2) scale(365.25) id(patid)
  (output omitted)
. stcrprep, events(status) keep(score) trans(1 2) byg(score)
. gen event = status == failcode
. stset tstop [iw=weight_c], failure(event=1) enter(tstart) noshow
   (output omitted)
```

- We can now estimate the CIF using sts graph.

sts graph

```
. sts graph if failcode == 1, by(score) failure // relapse
. sts graph if failcode == 2, by(score) failure // death
```

Using sts graph to estimate cause-specific CIF



Testing for difference between cause-specific CIFs

- Use sts test

sts test

```
. sts test score if failcode == 1

Log-rank test for equality of survivor functions

             |   Events         Events
score        |  observed       expected
-------------+-------------------------
Low risk     |        79          98.77
Medium risk  |       328         322.61
High risk    |        49          34.62
-------------+-------------------------
Total        |       456         456.00

                   chi2(2) =      10.03
                   Pr>chi2 =     0.0067
```

Testing for difference between cause-specific CIFs

- Use `sts test`

sts test

```
. sts test score if failcode == 1

Log-rank test for equality of survivor functions

             |  Events         Events
score        |  observed       expected
-------------+-------------------------
Low risk     |       79           98.77
Medium risk  |      328          322.61
High risk    |       49           34.62
-------------+-------------------------
Total        |      456          456.00

                 chi2(2) =      10.03
                 Pr>chi2 =     0.0067
```

- Similar to Gray's test [7] since the number at risk is modified when compared to the standard log-rank test.

Using stcox to fit Fine and Gray Model[9]

- Use stcrprep without byg() option since Fine and Gray model assumes common censoring distribution.

```
. stcrprep, events(status) keep(score) trans(1 2)
. stset tstop [iw=weight_c], failure(event) enter(tstart)
```

Using stcox to fit Fine and Gray Model[9]

- Use stcrprep without byg() option since Fine and Gray model assumes common censoring distribution.

```
. stcrprep, events(status) keep(score) trans(1 2)
. stset tstop [iw=weight_c], failure(event) enter(tstart)
```

stcox

```
. stcox i.score if failcode == 1, nolog
Cox regression -- Breslow method for ties
No. of subjects =  72880.46857                   Number of obs   =      72880
No. of failures =          456
Time at risk    =   6026.27434
                                                 LR chi2(2)      =       9.63
Log likelihood  =   -3333.3112                   Prob > chi2     =     0.0081
```

_t	Haz. Ratio	Std. Err.	z	P>|z|	[95% Conf.	Interval]
score						
Medium risk	1.271235	.1593392	1.91	0.056	.9943389	1.625238
High risk	1.769899	.3219273	3.14	0.002	1.239148	2.52798

Comparison with stcrreg

Comparison of Estimates

```
. estimates table stcrreg stcox*, eq(1) b(%6.5f) se(%6.5f) modelwidth(12)
```

Variable	stcrreg	stcox	stcox_robust
score			
Medium risk	0.23998	0.23999	0.23999
	0.12227	0.11861	0.12225
High risk	0.57090	0.57092	0.57092
	0.18298	0.16941	0.18297

legend: b/se

- Use pweights and vce(cluster id) for robust standard errors.
- However, Geskus (2011) showed that robust standard errors are less efficient[8].

Comparison with stcrreg

Comparison of Estimates

```
. estimates table stcrreg stcox*, eq(1) b(%6.5f) se(%6.5f) modelwidth(12)
```

Variable	stcrreg	stcox	stcox_robust
score			
Medium risk	0.23998	0.23999	0.23999
	0.12227	0.11861	0.12225
High risk	0.57090	0.57092	0.57092
	0.18298	0.16941	0.18297

legend: b/se

- Use pweights and vce(cluster id) for robust standard errors.
- However, Geskus (2011) showed that robust standard errors are less efficient[8].
- Perhaps stcrreg should have a 'norobust' option.

Time Improvements (seconds)

EBMT data (1977 subjects)

`stcrreg`	-	18.2
`stcrprep`	-	14.3
`stcox`	-	1.5

`stcrprep` only needs to be run once!

Time Improvements (seconds)

EBMT data (1977 subjects)

`stcrreg`	-	18.2
`stcrprep`	-	14.3
`stcox`	-	1.5

`stcrprep` only needs to be run once!

EBMT data $\times 10$ (19770 subjects): no ties

`stcrreg`	-	2814
`stcrprep`	-	922
`stcox`	-	49

`stcrprep` only needs to be run once!

Proportional subhazards (`estat phtest`)

- Assess proportional subhazards using Schoenfeld residuals.

Proportional subhazards (estat phtest)

- Assess proportional subhazards using Schoenfeld residuals.

```
. estat phtest,
      Test of proportional-hazards assumption
      Time:  Time
```

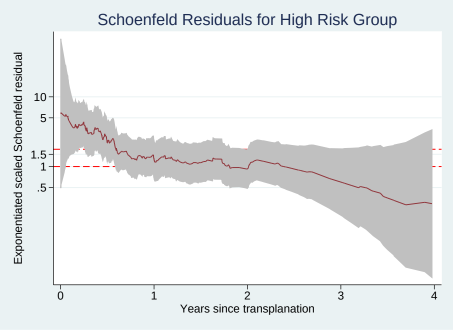
	chi2	df	Prob>chi2
global test	23.24	2	0.0000

Proportional subhazards (estat phtest)

- Assess proportional subhazards using Schoenfeld residuals.

```
. estat phtest,
      Test of proportional-hazards assumption
      Time:  Time
```

	chi2	df	Prob>chi2
global test	23.24	2	0.0000



Summary (part 1)

- Non-parametric estimates of CIF using `sts graph`.
- Other exploratory analysis (`stphplot`, `stcoxkm`)
- `stcrprep` allows fitting of Fine and Gray models with substantial speed improvements.
- A number of extensions to what is available in `stcrreg`.
  - Schoenfeld like residuals (`estat phtest`) [10]
  - Stratified models (`strata()`) [11].
  - 'Stacked models' share parameters over different events.
- Tests need a more in depth study of their properties.

Parametric approach

- Previous parametric models of the CIF required modelling of all K causes [12, 13].
- After using `stcrprep` we can fit a parametric equivalent of the Fine and Gray model

Parametric approach

- Previous parametric models of the CIF required modelling of all K causes [12, 13].
- After using `stcrprep` we can fit a parametric equivalent of the Fine and Gray model
    - Only need to model cause of interest.
- Useful for predictions, quantifying differences and non-proportional subhazards.
- Faster than Fine and Gray model as fewer splits (uses an approximation).

Parametric approach

- For those with competing events, allow to be at risk to end of potential follow-up.
- Split follow-up after competing event into (small) time-intervals.
- Apply weights to each interval.

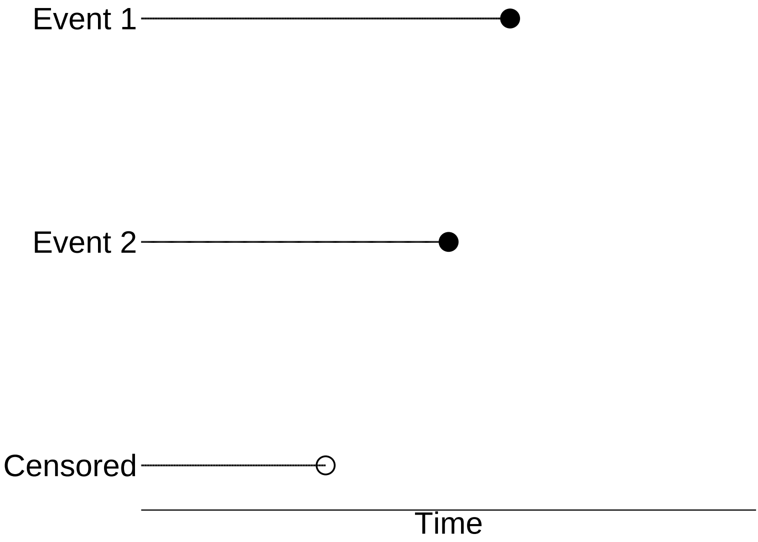
Likelihood

$$\ln L_i = d_{1i} \ln \left[h_1(t_i)\right] + (1 - d_{2i}) \ln \left[S(t_i)\right] +$$

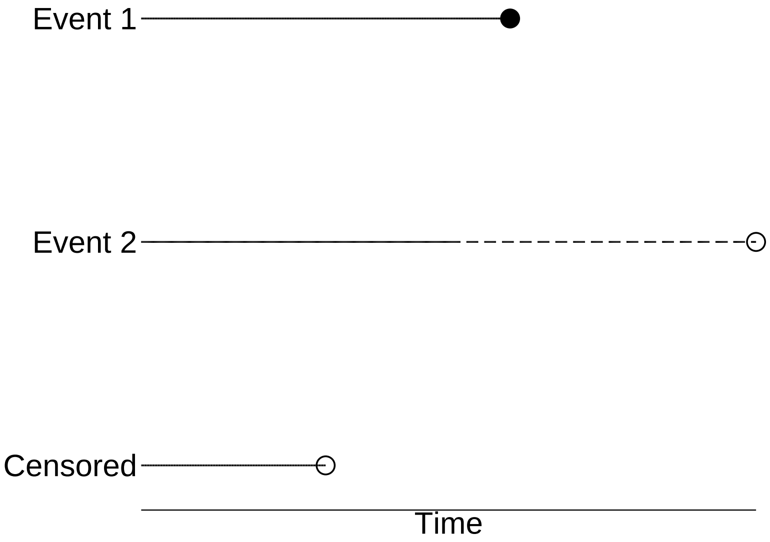
$$d_{2i} \sum_{j=1}^{J_i} w_{ij} \left(\ln \left[S(t_{ij})\right] - \ln \left[S(t_{i(j-1)})\right]\right)$$

- Need to specify parametric form of CIF for event of interest, **but** not for competing events.
- Also need weighting function. Obtained by modelling censoring distribution.

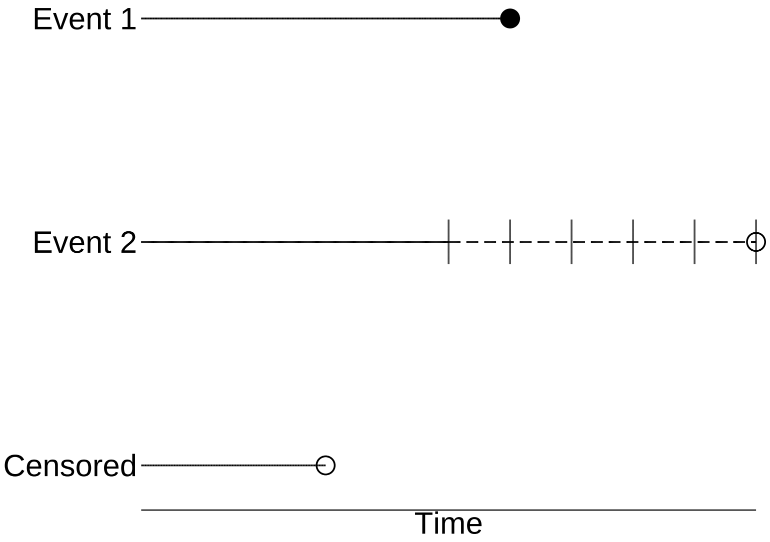
Splitting



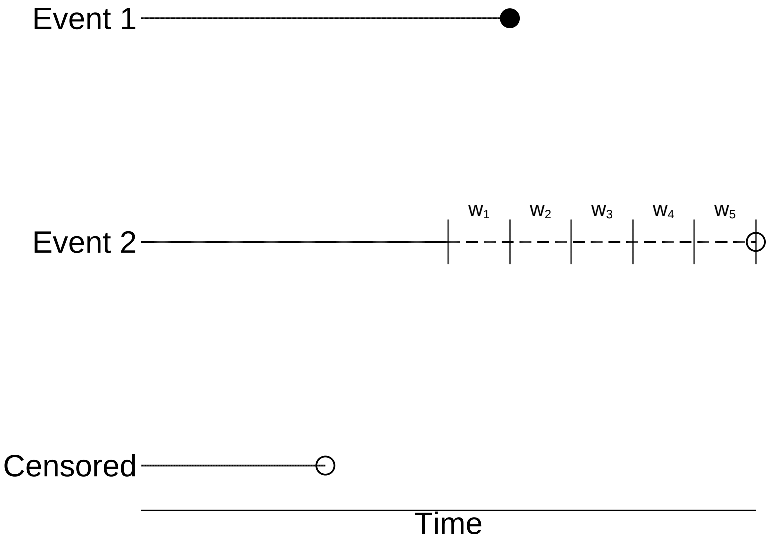
Splitting



Splitting



Splitting



The censoring distribution

- Fit a parametric model `stpm2`,
  - Option to include a variety of covariates.
  - Also to model time-dependent effects.
- `stcrreg` assumes common censoring distribution.
- Need to decide where to evaluate censoring distribution (number of split points) for weighted likelihood.

Flexible parametric models

- Possible to use any parametric approach that allows for delayed entry and weights.
- We use flexible parametric survival models that uses restricted splines to model the baseline using `stpm2` in Stata. [14, 15].

$$g[S(t|\mathbf{x}_i)] = \eta_i = s\left(\ln(t)|\gamma, \mathbf{k}_0\right) + \mathbf{x}_i\beta$$

- where $s\left(\ln(t)|\gamma, \mathbf{k}_0\right)$ is a restricted cubic spline function of $ln(t)$ with knots, $\mathbf{k}_0$.
- $g()$ is a link function.

Link Functions

- When using weights with expanded data

proportional sub hazards

$$\log(-\log(1 - CIF_k(t|\mathbf{x}_i))) = s(\ln(t)|\gamma, \mathbf{k}_0) + \mathbf{x}_i\boldsymbol{\beta}$$

proportional odds

$$\log\left(\frac{CIF_k(t|\mathbf{x}_i)}{1 - CIF_k(t|\mathbf{x}_i)}\right) = s(\ln(t)|\gamma, \mathbf{k}_0) + \mathbf{x}_i\boldsymbol{\beta}$$

relative absolute risk

$$\log(CIF_k(t|\mathbf{x}_i)) = s(\ln(t)|\gamma, \mathbf{k}_0) + \mathbf{x}_i\boldsymbol{\beta}$$

Link Functions

- When using weights with expanded data

proportional sub hazards

$$\log(-\log\left(1 - CIF_k(t|\mathbf{x}_i)\right)) = s\left(\ln(t)|\gamma, \mathbf{k}_0\right) + \mathbf{x}_i\boldsymbol{\beta}$$

proportional odds

$$\log\left(\frac{CIF_k(t|\mathbf{x}_i)}{1 - CIF_k(t|\mathbf{x}_i)}\right) = s\left(\ln(t)|\gamma, \mathbf{k}_0\right) + \mathbf{x}_i\boldsymbol{\beta}$$

relative absolute risk

$$\log\left(CIF_k(t|\mathbf{x}_i)\right) = s\left(\ln(t)|\gamma, \mathbf{k}_0\right) + \mathbf{x}_i\boldsymbol{\beta}$$

- Time-dependent effects can be fitted for any of these link functions.

Parametric proportional subhazards models 1

stcrprep

```
. stset time, failure(status==1,2) scale(365.25) id(patid)
  (output omitted)
. stcrprep, events(status) keep(score) trans(1 2) censstpm2 every(0.2)
. gen event = status == failcode
. stset tstop [iw=weight_c], failure(event) enter(tstart) noshow
     failure event:  event != 0 & event < .
obs. time interval:  (0, tstop]
 enter on or after:  time tstart
 exit on or before:  failure
            weight:  [iweight=weight_c]
------------------------------------------------------------------------------
    48116  total observations
        0  exclusions
------------------------------------------------------------------------------
    48116  observations remaining, representing
     1141  failures in single-record/single-failure data
 16367.15  total analysis time at risk and under observation
                                                at risk from t =         0
                                     earliest observed entry t =         0
                                          last observed exit t =  8.454483
```

Parametric proportional subhazards models 2

stpm2

```
. stpm2 i.score if failcode == 1, scale(hazard) df(4) eform nolog
note: delayed entry models are being fitted
Log likelihood = -1678.7162                     Number of obs   =      29147
```

	exp(b)	Std. Err.	z	P>|z|	[95% Conf.	Interval]
xb						
score						
Medium risk	1.270615	.1592552	1.91	0.056	.9938639	1.62443
High risk	1.770563	.3220405	3.14	0.002	1.239624	2.528908
_rcs1	1.431289	.0284143	18.06	0.000	1.376667	1.488077
_rcs2	1.124393	.0149958	8.79	0.000	1.095382	1.154172
_rcs3	1.037582	.0130522	2.93	0.003	1.012313	1.063481
_rcs4	.9688918	.0078559	-3.90	0.000	.9536162	.9844121
_cons	.2087425	.0235126	-13.91	0.000	.167391	.2603092

Parametric proportional subhazards models 2

stpm2

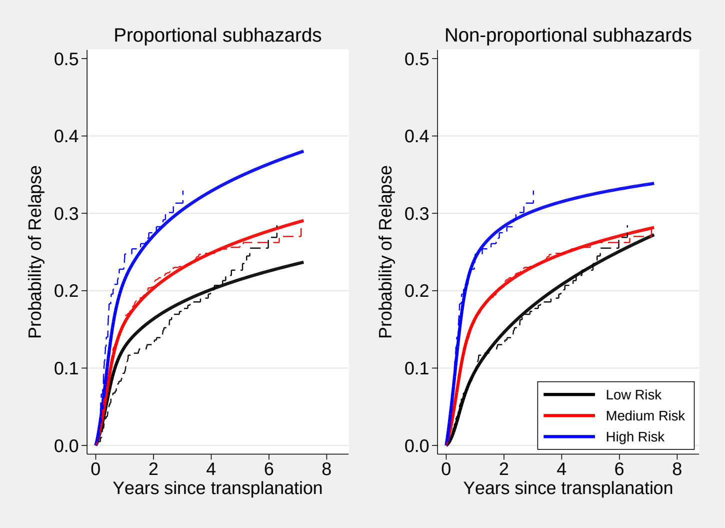
```
. stpm2 i.score if failcode == 1, scale(hazard) df(4) eform nolog
note: delayed entry models are being fitted
Log likelihood = -1678.7162                     Number of obs   =      29147
```

	exp(b)	Std. Err.	z	P>|z|	[95% Conf.	Interval]
xb						
score						
Medium risk	1.270615	.1592552	1.91	0.056	.9938639	1.62443
High risk	1.770563	.3220405	3.14	0.002	1.239624	2.528908
_rcs1	1.431289	.0284143	18.06	0.000	1.376667	1.488077
_rcs2	1.124393	.0149958	8.79	0.000	1.095382	1.154172
_rcs3	1.037582	.0130522	2.93	0.003	1.012313	1.063481
_rcs4	.9688918	.0078559	-3.90	0.000	.9536162	.9844121
_cons	.2087425	.0235126	-13.91	0.000	.167391	.2603092

- Sub-hazard ratios very similar to semi-parametric estimates.

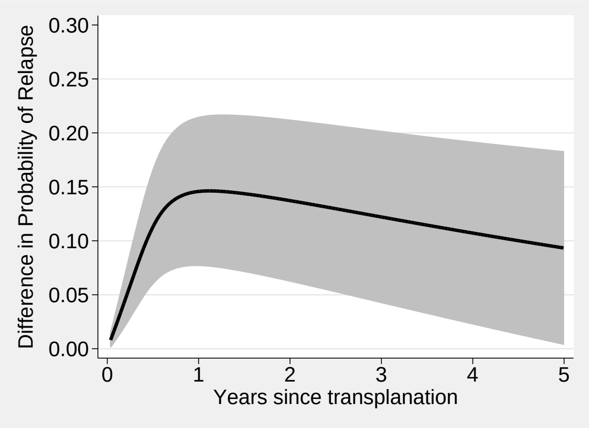
Predictions of CIF

```
predict cif, failure
```



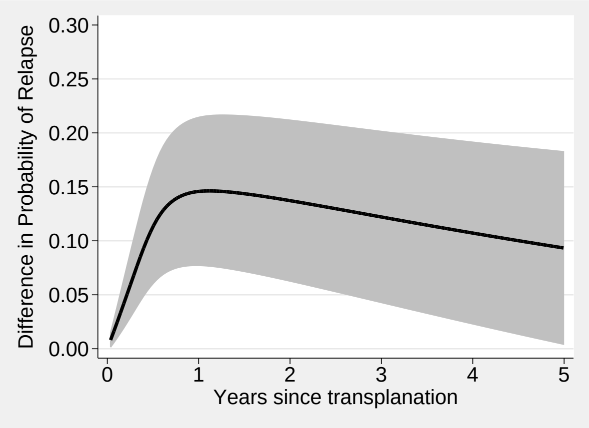
Difference in CIFs

```
. predict CIF_diff, sdiff1(score2 0 score3 0) sdiff2(score3 1) ci
```



Difference in CIFs

```
. predict CIF_diff, sdiff1(score2 0 score3 0) sdiff2(score3 1) ci
```



- Take reciprocal to estimate Number Needed to Treat (NNT) accounting for competing risks[16]

Relative absolute risks

stpm2

```
. stpm2 i.score if failcode == 1, scale(log) df(4) eform nolog
note: delayed entry models are being fitted
Log likelihood = -1680.1742                        Number of obs   =      29147
```

	exp(b)	Std. Err.	z	P>|z|	[95% Conf.	Interval]
xb						
score						
Medium risk	1.19893	.1332627	1.63	0.103	.9642325	1.490755
High risk	1.543556	.2392695	2.80	0.005	1.139137	2.091553
_rcs1	1.38459	.0247152	18.23	0.000	1.336987	1.433889
_rcs2	1.126424	.0141052	9.51	0.000	1.099115	1.154412
_rcs3	1.034958	.0127994	2.78	0.005	1.010174	1.060351
_rcs4	.9702326	.0072774	-4.03	0.000	.9560736	.9846014
_cons	.1922568	.0193746	-16.36	0.000	.1577982	.2342401

Relative absolute risks

stpm2

```
. stpm2 i.score if failcode == 1, scale(log) df(4) eform nolog
note: delayed entry models are being fitted
Log likelihood = -1680.1742                    Number of obs   =      29147
```

	exp(b)	Std. Err.	z	P>|z|	[95% Conf.	Interval]
xb						
score						
Medium risk	1.19893	.1332627	1.63	0.103	.9642325	1.490755
High risk	1.543556	.2392695	2.80	0.005	1.139137	2.091553
_rcs1	1.38459	.0247152	18.23	0.000	1.336987	1.433889
_rcs2	1.126424	.0141052	9.51	0.000	1.099115	1.154412
_rcs3	1.034958	.0127994	2.78	0.005	1.010174	1.060351
_rcs4	.9702326	.0072774	-4.03	0.000	.9560736	.9846014
_cons	.1922568	.0193746	-16.36	0.000	.1577982	.2342401

- Effect sizes are now relative risks rather than subhazard ratios.
- Assumed constant over time, but this can be relaxed.

Summary (part 2)

- Parametric version of Fine and Gray model.
- Only need to model event of interest to estimate CIF.
- Models on a variety of scales.
- Can relax the proportionality assumption.

Summary (part 2)

- Parametric version of Fine and Gray model.
- Only need to model event of interest to estimate CIF.
- Models on a variety of scales.
- Can relax the proportionality assumption.
- Need to choose split times, but can be fairly crude.
- When modelling competing risks, still useful to model cause-specific hazards.
    - See `stpm2cif`[5]

References

[1] Andersen PK, Abildstrom SZ, Rosthøj S. Competing risks as a multi-state model. *Stat Methods Med Res* 2002;**11**:203–215.

[2] Coviello V, Boggess M. Cumulative incidence estimation in the presence of competing risks. *The Stata Journal* 2004;**4**:103–112.

[3] Putter H, Fiocco M, Geskus RB. Tutorial in biostatistics: competing risks and multi-state models. *Statistics in Medicine* 2007;**26**:2389–2430.

[4] Hinchliffe SR, Lambert P. Flexible parametric modelling of cause-specific hazards to estimate cumulative incidence functions. *BMC Medical Research Methodology* 2013;**13**:13.

[5] Hinchliffe SR, Lambert P. Extending the flexible parametric survival model for competing risks. *The Stata Journal* 2013;**13**:344–355.

[6] de Wreede L, Fiocco M, Putter H. mstate: An r package for the analysis of competing risks and multi-state models. *Journal of Statistical Software* 2011;**38**.

[7] Gray R. A class of k-sample tests for comparing the cumulative incidence of a competing risk. *The Annals of Statistics* 1988;**16**:1141–1154.

[8] Geskus RB. Cause-specific cumulative incidence estimation and the fine and gray model under both left truncation and right censoring. *Biometrics* 2011;**67**:39–49.

[9] Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. *Journal of the American Statistical Association* 1999;**446**:496–509.

[10] Zhou B, Fine J, Laird G. Goodness-of-fit test for proportional subdistribution hazards model. *Stat Med* 2013;.

[11] Zhou B, Latouche A, Rocha V, Fine J. Competing risks regression for stratified data. *Biometrics* 2011;**67**:661–670.

[12] Jeong JH, Fine JP. Parametric regression on cumulative incidence function. *Biostatistics* 2007;**8**:184–196.

[13] Jeong JH, Fine JP. Direct parametric inference for the cumulative incidence function. *Applied Statistics* 2006;**55**:187–200.

[14] Lambert PC, Royston P. Further development of flexible parametric models for survival analysis. *The Stata Journal* 2009;**9**:265–290.

[15] Royston P, Lambert PC. *Flexible parametric survival analysis in Stata: Beyond the Cox model*. Stata Press, 2011.

[16] Gouskova NA, Kundu S, Imreyb PB, Fine JP. Number needed to treat for time-to-event data with competing risks. *Statistics in Medicine* 2013;**(in press)**.