Extending the Cox Model - Multiple Endpoints, Time-Varying Analyses

Stephen McDonald
Limitations of Cox model

- Assumes
  - Proportional hazard constant
    - not constant hazard
  - Single endpoint
    - Which everyone reaches (eventually)
  - Continuous followup
- Strength: does not require definition of underlying mortality pattern
Agenda

- Gaps
- Time varying factors
- Multiple end points
- Going parametric
- Restricted mean survival time
- Stratification
- Shared frailty
Standard survival data

Event ★
Censored ○

Group 1
Group 2
Reality - single terminal endpoint
Reality - non-terminal endpoints

Group 1
Group 2
Group 3
Group 4

Time
Setting up survival analysis is key

- Simple approach to setting up assumes single line data
- Specify
  - Entry date, exit date, failure variable,
    - `-stset txperiod, failure(txstatus) scale(365.25)`
  - `-stset transplantperiod, failure(transplantstatus) scale(365.25)`
- Better: Entry date, scale, id
Basics

Kaplan-Meier survival estimates

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.50</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.75</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Cox regression -- Breslow method for ties

| t   | Haz. Ratio | Std. Err. | z    | P>|z| | [95% Conf. Interval] |
|-----|------------|-----------|------|-----|---------------------|
| LR era |
| 1963- | 2.200397   | 0.0746714 | 23.24 | 0.000 | 2.059795 - 2.351717 |
| 1973- | 2.189555   | 0.0524473 | 32.72 | 0.000 | 2.089136 - 2.298801 |
| 1983- | 1.484001   | 0.0331121 | 17.69 | 0.000 | 1.420501 - 1.550339 |
| 1993- | 1 (base)   |           |      |       |                     |
| 2003- | 0.611893   | 0.0181903 | -15.33 | 0.000 | 0.517293 - 0.688657 |
| 2013- | 0.5091195  | 0.0246626 | -14.72 | 0.000 | 0.4348713 - 0.5814418 |
But ....things change over time

- Post transplant graft failure rates vary
- Relationships may also vary over time
Setting up survival analysis is key!
Splitting followup time

```
. split yearsfollowup, at(1 2 5 10)
(67,322 observations (episodes) created)

. tab years

<table>
<thead>
<tr>
<th>observation</th>
<th>Freq.</th>
<th>Percent</th>
<th>Cum.</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>26,094</td>
<td>29.37</td>
<td>29.37</td>
</tr>
<tr>
<td>1</td>
<td>22,486</td>
<td>25.47</td>
<td>54.84</td>
</tr>
<tr>
<td>2</td>
<td>20,521</td>
<td>23.16</td>
<td>78.00</td>
</tr>
<tr>
<td>5</td>
<td>15,575</td>
<td>17.59</td>
<td>95.59</td>
</tr>
<tr>
<td>10</td>
<td>8,992</td>
<td>9.91</td>
<td>100.00</td>
</tr>
</tbody>
</table>

Total | 55,606 | 100.00 |

. stata

-> set endtransplantdate, id(transplantid) failure(transplantstatus) origin(time transplantdate) scale(365.25)

id: transplantid
failure event: transplantstatus != 0 & transplantstatus < .
obsv. time interval: [endtransplantdate[ n-1], endtransplantdate]
exit on or before: failure
for analysis: (time-origin)/365.25
origin: time transplantdate

95606 total observations
0 exclusions

95606 observations remaining, representing
29884 subjects
19794 failures in single-failure-per-subject data
223800.022 total analysis time at risk and under observation

at risk from t = 0
earliest observed entry t = 0
last observed exit t = 47.63313
```
Rates vary

```
.strate years

failure _d: transplantstatus
analysis time _t: (endtransplantdate-origin)/365.25
origin: time transplantdate
id: transplantid
```

Estimated rates and lower/upper bounds of 95% confidence intervals
(95606 records included in the analysis)

<table>
<thead>
<tr>
<th>years~p</th>
<th>D</th>
<th>Y</th>
<th>Rate</th>
<th>Lower</th>
<th>Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>4507</td>
<td>2.4e+04</td>
<td>0.189176</td>
<td>0.183733</td>
<td>0.194780</td>
</tr>
<tr>
<td>1</td>
<td>888</td>
<td>2.1e+04</td>
<td>0.041359</td>
<td>0.038726</td>
<td>0.044171</td>
</tr>
<tr>
<td>2</td>
<td>2291</td>
<td>5.4e+04</td>
<td>0.042536</td>
<td>0.040829</td>
<td>0.044313</td>
</tr>
<tr>
<td>5</td>
<td>3236</td>
<td>6.0e+04</td>
<td>0.053752</td>
<td>0.051932</td>
<td>0.055637</td>
</tr>
<tr>
<td>10</td>
<td>4442</td>
<td>6.4e+04</td>
<td>0.068930</td>
<td>0.066932</td>
<td>0.070987</td>
</tr>
</tbody>
</table>
Comparison over time

- We know that outcomes have improved over time
- The K-M graph suggests that the improvement may not be constant across all categories

```stata
data.list 1/5
foreach year of num 1 2 3 4 5
    stcox ib(3).tx era if yearsfol==`year'
est store group`year'
end
```

```stata
est table *, cform b(%.3g)
```

<table>
<thead>
<tr>
<th>Variable</th>
<th>group1</th>
<th>group2</th>
<th>group5</th>
<th>group10</th>
</tr>
</thead>
<tbody>
<tr>
<td>tx_era</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1963-</td>
<td>4.45</td>
<td>1.99</td>
<td>1.48</td>
<td>1.1</td>
</tr>
<tr>
<td>1973-</td>
<td>4.24</td>
<td>2.02</td>
<td>1.19</td>
<td>.993</td>
</tr>
<tr>
<td>1983-</td>
<td>2.07</td>
<td>1.42</td>
<td>1.29</td>
<td>1.15</td>
</tr>
<tr>
<td>tx_era</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2003-</td>
<td>.927</td>
<td>.833</td>
<td>.866</td>
<td>.993</td>
</tr>
<tr>
<td>2013-</td>
<td>.700</td>
<td>.967</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

end of do-file
Other splitting

- The “origin” is not the same as entry

```
. stset endtransplantdate, failure(transplantstatus) origin(dateofbirth) entry(transplantdate) scale(365.25) id(transplantid)
```

- id: transplantid
- failure event: transplantstatus != 0 & transplantstatus < .
- obs. time interval: (endtransplantdate[_n-1], endtransplantdate]
- enter on or after: time transplantdate
- exit on or before: failure
- t for analysis: (time-origin)/365.25
- origin: time dateofbirth

95606 total observations
0 exclusions

95606 observations remaining, representing
26094 subjects
15364 failures in single-failure-per-subject data
223800.022 total analysis time at risk and under observation

at risk from t = 0
earliest observed entry t = .9965777
last observed exit t = 94.05613
Splitting illustrated
Other approaches to TVC

- These examples use a “piecewise” approach to followup
  - i.e. categorical

- Stata allows you to define a continuous TVC
  - `-stcox x x , tvc(var) texp(exp)–`
  - Can also be used as a test of PH.....
What can we do now?

- Look at effects of variables at different age
  - For example,
    - does the marked improvement in the first year extend across all age groups equally?
    - Is it the age at transplantation, or the prevalent age that determines risk of rejection

- Split on other factors
  - Change in therapy
Multiple failures

- Previous analyses have concentrated on “terminal” outcomes
- Cox model can be extended to multiple failures per individual
  - There are alternatives (e.g. multilevel Poisson)
- Multiple situations - here looking at multiple failures of a single type with ordinal nature
  - Other possibilities include multiple possible failures (cf competing risk)
Example

```
<table>
<thead>
<tr>
<th>id</th>
<th>sequenceno</th>
<th>pdsequence</th>
<th>pddate</th>
<th>endpddate</th>
<th>pdperiod</th>
<th>pdstatus</th>
</tr>
</thead>
<tbody>
<tr>
<td>P9001321</td>
<td>1</td>
<td>1</td>
<td>26feb1977</td>
<td>28Feb1978</td>
<td>367</td>
<td>1</td>
</tr>
<tr>
<td>P9001277</td>
<td>1</td>
<td>1</td>
<td>07aug1977</td>
<td>12sep1977</td>
<td>36</td>
<td>1</td>
</tr>
<tr>
<td>P9001266</td>
<td>1</td>
<td>1</td>
<td>20apr1977</td>
<td>16may1977</td>
<td>26</td>
<td>1</td>
</tr>
<tr>
<td>P9001255</td>
<td>1</td>
<td>1</td>
<td>11may1972</td>
<td>07oct1972</td>
<td>149</td>
<td>1</td>
</tr>
<tr>
<td>P9001244</td>
<td>1</td>
<td>1</td>
<td>10dec1980</td>
<td>24mar1981</td>
<td>104</td>
<td>1</td>
</tr>
<tr>
<td>P9002231</td>
<td>1</td>
<td>1</td>
<td>02sep1977</td>
<td>30sep1977</td>
<td>28</td>
<td>1</td>
</tr>
<tr>
<td>P9002220</td>
<td>1</td>
<td>1</td>
<td>17jan1984</td>
<td>18sep1984</td>
<td>245</td>
<td>1</td>
</tr>
<tr>
<td>P9002209</td>
<td>1</td>
<td>1</td>
<td>18jun1976</td>
<td>07sep1976</td>
<td>81</td>
<td>1</td>
</tr>
<tr>
<td>P9002198</td>
<td>1</td>
<td>1</td>
<td>13jan1971</td>
<td>09mar1971</td>
<td>55</td>
<td>1</td>
</tr>
<tr>
<td>P9002198</td>
<td>15</td>
<td>2</td>
<td>10feb1986</td>
<td>10jun1986</td>
<td>618</td>
<td>1</td>
</tr>
<tr>
<td>P9002198</td>
<td>18</td>
<td>3</td>
<td>15nov1986</td>
<td>25jul1988</td>
<td>155</td>
<td>1</td>
</tr>
<tr>
<td>P9002187</td>
<td>3</td>
<td>1</td>
<td>27aug1975</td>
<td>29jan1976</td>
<td>81</td>
<td>1</td>
</tr>
<tr>
<td>P9002176</td>
<td>1</td>
<td>1</td>
<td>14jan1977</td>
<td>05apr1977</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```

Gap
In practice

```
. stset endpdate, time0( pddate) scale(365.25) id( id) fail( pdstatus) exit(time .) origin(pddate)

    id:  id
  failure event:  pdstatus != 0 & pdstatus < .
obs. time interval:  (pddate, endpdate]
exit on or before:  time .
t for analysis:  (time-origin)/365.25
    origin:  time pddate

39510  total observations
     0  exclusions
```

```
39510  observations remaining, representing
33803  subjects
28963  failures in multiple-failure-per-subject data
70551.919  total analysis time at risk and under observation

  at risk from t = 0
earliest observed entry t = 0
last observed exit t = 42.38741
```
Outcomes

- In many cases the risk of (recurrent) failure is NOT independent of the first occurrence
- Stratify on failure order - assume people who have had a “failure” are at higher risk of subsequent failure
Technique failure analysis
What’s next

- Beat CKD statistics workshop
  - Sydney, June 2017
- Beat CKD workshop
  - Adelaide, August 2017