A TUTORIAL ON ACCOUNTING FOR COMPETING RISKS IN SURVIVAL ANALYSIS

Rebecca Scherzer, PhD
Principal Research Statistician
Kidney Health Research Collaborative
San Francisco VA Medical Center
University of California at San Francisco

Clinical Research Statistical Methods Seminar
May 8th, 2017
OUTLINE

• Background
  • When does the problem occur, when does it matter?

• Methods and illustrations
  • Survival curves and other graphical methods
  • Regression models
  • Number-needed-to-treat (NNT)

• Interpretation
  • Cause-specific hazard versus sub-distribution hazard:
    • which to use and when?

• Discussion
  • Best practices and caveats
  • Limitations and research gaps
  • Further reading and resources
BACKGROUND

- Clinical research studies often record the time to more than one outcome:
  - Examples: death, cardiovascular disease (CVD), end stage renal disease (ESRD)

- A competing event is one that precludes the occurrence of the event of interest:
  - Example: after transplant or death, patient is no longer at risk for primary outcome of interest (ESRD or CVD).
If a patient experiences a competing event, standard survival analysis methods treat that patient as *censored* for the outcome of interest (e.g., ESRD or CVD).

**Why is this a problem?**

- Kaplan-Meier curves overestimate the incidence of the outcome over time
- Cox models inflate the relative differences between groups, resulting in biased hazard ratios
ALTERNATIVES TO STANDARD METHODS:

- **Survival curves:** Cumulative Incidence Function (CIF)
  - Non-parametric CIF
  - Fine-Gray (1999) CIF
  - Inverse probability weighting (IPW) corrected Kaplan-Meier

- **Options for regression models:**
  - Sub-distribution hazard ratio (SHR)
    - Fine-Gray (1999)
    - Klein-Andersen (2005)
  - Cause-specific hazard ratio (CHR)

- **Number-needed-to-treat (NNT):**
  - Gouskova et al (2014)
FINE-GRAY (FG) MODEL
METHODS:
PLOTTING THE CUMULATIVE INCIDENCE

• In each case, we code the event categories as follows:
  • event=0: censored, event=1: outcome of interest, event=2: competing event.

<table>
<thead>
<tr>
<th>Non-parametric:</th>
<th>Fine-Gray:</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SAS</strong></td>
<td></td>
</tr>
<tr>
<td>proc lifetest; time year*event(0) /</td>
<td>proc phreg; model year*event(0)=x /</td>
</tr>
<tr>
<td>eventcode=1; run;</td>
<td>eventcode=1; run;</td>
</tr>
<tr>
<td><strong>Stata</strong></td>
<td></td>
</tr>
<tr>
<td>stset year, failure(event==1)</td>
<td>stset year, failure(event==1)</td>
</tr>
<tr>
<td>stcrreg, compete(event==2)</td>
<td>stcrreg x, compete(event==2)</td>
</tr>
<tr>
<td>stcurve, cif</td>
<td>stcurve, cif</td>
</tr>
<tr>
<td><strong>R</strong></td>
<td></td>
</tr>
<tr>
<td>library(cmprsk) plot (cuminc(year,</td>
<td>library(cmprsk)</td>
</tr>
<tr>
<td>event, cencode=0))</td>
<td>result&lt;- crr(year, event, x,</td>
</tr>
<tr>
<td></td>
<td>failcode=1, cencode=0)</td>
</tr>
<tr>
<td></td>
<td>plot(predict(result, x))</td>
</tr>
</tbody>
</table>
ILLUSTRATION:
NON PARAMETRIC ESTIMATION GIVES VISUAL COMPARISON
OF CUMULATIVE RISK OF CVD AND DEATH:
ILLUSTRATION:
COMPARISON OF CUMULATIVE INCIDENCE ESTIMATES BY WALKING SPEED, CVD VS. DEATH:
METHODS: CALCULATION OF SUB-DISTRIBUTION HAZARD RATIO (SHR):

- **Stata:**
  - `stset year, id(idno) failure (event==1)`
  - `stcrreg x, compete(event==2)`
- **SAS:**
  - `proc phreg;`
  - `model year*event(0)=x / eventcode=1;`
  - `run;`
- **R:**
  - `library(cmprsk)`
  - `crr(year, event, x, failcode=1,censcode=0)`
METHODS: 
CALCULATION OF CAUSE-SPECIFIC HAZARD RATIO (CHR)

- Stata:
  - `stset year, id(idno) failure (event==1)`
  - `stcox x`

- SAS:
  - `proc phreg;`
  - `model year*event(0,2)=x / eventcode=1;`
  - `run;`

- R:
  - `coxph(formula=Surv (year, event=="1") ~x)"`
COMPARISON OF MODELS SHOWS INFLATED HAZARD RATIOS FOR COX CHR VERSUS FG SHR

• Example 1: slower walking speed and risk of CVD

<table>
<thead>
<tr>
<th>Method</th>
<th>Hazard Ratio</th>
<th>95% Hazard Ratio Confidence Limits</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fine-Gray SHR</td>
<td>1.69</td>
<td>1.29-2.21</td>
<td>0.0001</td>
</tr>
<tr>
<td>Cox CSH</td>
<td>2.82</td>
<td>2.12-3.76</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

• Example 2: elevated biomarker and risk of ESRD

<table>
<thead>
<tr>
<th>Method</th>
<th>Hazard Ratio</th>
<th>95% Hazard Ratio Confidence Limits</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fine-Gray SHR</td>
<td>1.15</td>
<td>1.09-1.22</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Cox CSH</td>
<td>1.18</td>
<td>1.11-1.25</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>
ILLUSTRATION:
COMPARISON OF CUMULATIVE CVD INCIDENCE ESTIMATES BY WALKING SPEED, COX VERSUS FINE-GRAY MODEL:
METHODS:
NUMBER-NEEDED-TO TREAT (NNT)

• NNT is the reciprocal of the absolute risk difference:
  • Example: AR=5% => NNT=20, means that treating 20 patients would prevent one case of disease

• In the presence of competing risks, Gouskova et al (2014) define the NNT at time $t$ using the CIF from the Fine-Gray model:

$$NNT(t) = \frac{1}{CIF^{Ctl}(t) - CIF^{Trt}(t)}$$
METHODS:
ESTIMATE NNT USING CIF FROM FINE-GRAY MODEL:

• Example 1: Suppose a drug is available that can increase walking speed. How many patients must we treat to prevent CVD, in the presence of competing risk of death?
  
  • CIF for slow walkers at year 10 = 0.38
  • CIF for fast walkers at year 10 = 0.25
  • AR = 0.38 – 0.25 = 0.13 => NNT at 10 years = 8

• Example 2: Suppose a drug is available that can reduce biomarker levels. How many patients must we treat to prevent ESRD, in the presence of competing risk of death?
  
  • CIF for elevated biomarker at year 5 = 0.117
  • CIF for normal biomarker at year 5 = 0.102
  • AR = 0.015 => NNT at 5 years = 67
ILLUSTRATION:
ESTIMATION OF NNT OVER TIME:

Example 1: walk speed and CVD

Example 2: biomarker and ESRD

NNT=8

NNT=67
WHEN DO COX AND FG RESULTS DIFFER?

- If competing event is frequent
- If competing event occurs early
- Effect of censoring proportion …
- Effect of event time correlation …
Table 4

Comparison of competing risks regression models examining treatment and two covariates for competing outcomes in prostate cancer (RTOG 8610)

<table>
<thead>
<tr>
<th>Event type (death)</th>
<th>Model Effect Estimates</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cox CSH</td>
<td>Fine-Gray SDH</td>
<td>Klein-Andersen</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CHR</td>
<td>SHR</td>
<td>SHR</td>
<td>SHR</td>
<td>SHR</td>
</tr>
<tr>
<td></td>
<td>95% CI</td>
<td>95% CI</td>
<td>95% CI</td>
<td>95% CI</td>
<td>95% CI</td>
</tr>
<tr>
<td>A. Prostate Cancer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADT (vs RT only)</td>
<td>0.67</td>
<td>0.66</td>
<td>0.67</td>
<td>0.49–0.93</td>
<td></td>
</tr>
<tr>
<td>Age*</td>
<td>0.89</td>
<td>0.75</td>
<td>0.79</td>
<td>0.63–1.00</td>
<td></td>
</tr>
<tr>
<td>Grade 2 vs 1</td>
<td>1.84</td>
<td>1.83</td>
<td>1.87</td>
<td>1.04–3.23</td>
<td></td>
</tr>
<tr>
<td>Grade 3 vs 1</td>
<td>2.87</td>
<td>2.83</td>
<td>2.94</td>
<td>1.66–4.98</td>
<td></td>
</tr>
<tr>
<td>B. Other causes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADT (vs RT only)</td>
<td>1.13</td>
<td>1.26</td>
<td>1.20</td>
<td>0.85–1.51</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>2.02</td>
<td>1.93</td>
<td>1.88</td>
<td>1.60–2.57</td>
<td></td>
</tr>
<tr>
<td>Grade 2 vs 1</td>
<td>0.87</td>
<td>0.75</td>
<td>0.82</td>
<td>0.59–1.28</td>
<td></td>
</tr>
<tr>
<td>Grade 3 vs 1</td>
<td>0.91</td>
<td>0.60</td>
<td>0.61</td>
<td>0.62–1.35</td>
<td></td>
</tr>
<tr>
<td>All deaths</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADT (vs RT only)</td>
<td>0.88</td>
<td></td>
<td></td>
<td>0.71–1.09</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>1.36</td>
<td></td>
<td></td>
<td>1.15–1.61</td>
<td></td>
</tr>
<tr>
<td>Grade 2 vs 1</td>
<td>1.13</td>
<td></td>
<td></td>
<td>0.83–1.55</td>
<td></td>
</tr>
<tr>
<td>Grade 3 vs 1</td>
<td>1.44</td>
<td></td>
<td></td>
<td>1.06–1.97</td>
<td></td>
</tr>
</tbody>
</table>

* per 10 year increment in age

Adapted from T4, Dignam et al. 2012, Clin Cancer Res
EFFECT OF CENSORING ON HR:

Scenario: 2x CVD rate in Group B vs. Group A, same death rate in both groups

Adapted from T3, Dignam et al. 2012, Clin Cancer Res
INDEPENDENT EVENT TIMES:
SCENARIO 1: 33% CENSORING, CVD & DEATH EVENT TIMES UNCORRELATED

Hazard Ratio: Group B vs. A

- Cox CVD
  - FG CVD
  - Cox Death
  - FG Death

Adapted from T1, Dignam et al. 2012, Clin Cancer Res
CORRELATED EVENT TIMES
SCENARIO 2: 33% CENSORING, CVD & DEATH EVENT TIMES CORRELATED ($r=0.6$)

Adapted from T2, Dignam et al. 2012, Clin Cancer Res
### Recommendations for Analyzing Competing Risk Survival Data

- Cumulative incidence functions (CIFs) should be used to estimate the incidence of each of the different types of competing risks. Do not use the Kaplan-Meier estimate of the survival function for this purpose.

- Researchers need to decide whether the research objective is on addressing etiologic questions or on estimating incidence or predicting prognosis.

- Use the Fine-Gray subdistribution hazard model when the focus is on estimating incidence or predicting prognosis in the presence of competing risks.

- Use the cause-specific hazard model when the focus is on addressing etiologic questions.

- In some settings, both types of regression models should be estimated for each of the competing risks to permit a full understanding of the effect of covariates on the incidence and the rate of occurrence of each outcome.

---

*Austin et al, 2016*
DISCUSSION

• Caveats:
  • Interpretation can be difficult: effect of covariate on CSH may be different (even opposite!) effect on incidence.
  • Still need to check proportional hazard assumption, just as with ordinary Cox models
  • Non-informative censoring assumption:
    • probability of event should be unrelated to mechanism of censoring
    • length of follow-up should not depend on a patient’s medical condition
  • Best practices:
    • Do the usual regression checks: check for outliers and influential data points, assess linearity, collinearity, etc.
    • Use CIF plots and other visualization to examine covariate effects for each event type
DISCUSSION

• Limitations:
  • When running competing risk models, standard software has fewer options for stratification, shared frailty, tests of model fit, and variable selection methods.

• Research and software gaps:
  • Optimal method for reweighting
  • Left or interval censoring and truncation
  • Censoring assumptions: effect of competing risk on subsequent events (preclude versus change probability)
FURTHER READING AND RESOURCES

• Software:
  • https://cran.r-project.org/web/packages/cmprsk/cmprsk.pdf
  • www.stata.com/manuals13/ststcrreg.pdf
  • https://cran.r-project.org/web/packages/mstate/vignettes/Tutorial.pdf

• References:
  • Peter C. Austin, Douglas S. Lee and Jason P. Fine. Introduction to the Analysis of Survival Data in the Presence of Competing Risks Circulation. 2016;133:601-609, originally published February 8, 2016