Complications in Survival Analysis

Elizabeth Koehler
VICC Cancer Biostatistics Center

February 20, 2009
1. Intro

2. Describing Survival Data

3. Regression Models

4. Competing Risks

5. Summary
Upcoming Workshops

- March 20: Heidi Chen
  Analysis of MALDI-TOF Data

- April 17: Tatsuki Koyama
  Analysis for Biomarker Discovery

- May 15: Andrew Yi and Pengcheng Lu
  Data Mining in GEO
 Goals

Who is here?
Goals:

1. To demonstrate and describe complications in survival analysis.
Goals

Who is here?

Goals:

1. To demonstrate and describe complications in survival analysis
2. To briefly outline solutions for dealing with these complications
What is Survival Analysis?

Suppose you were investigating a drug for treating cancer.

- recruit a cohort of patients
What is Survival Analysis?

Suppose you were investigating a drug for treating cancer.

- recruit a cohort of patients
- place some on the standard treatment and some on the investigated treatment
What is Survival Analysis?

Suppose you were investigating a drug for treating cancer.

- recruit a cohort of patients
- place some on the standard treatment and some on the investigated treatment
- follow them for 6mos
- record the length of time until they die, and their status
What is Survival Analysis?

Suppose you were investigating a drug for treating cancer.

- recruit a cohort of patients
- place some on the standard treatment and some on the investigated treatment
- follow them for 6mos
- record the length of time until they die, and their status

What methods can be used to evaluate time-to-event data?
Complications in Survival Analysis

Time to Death when Death Observed

- Standard
- New
Complications in Survival Analysis

Time to Death when Death Observed

<table>
<thead>
<tr>
<th></th>
<th>Standard</th>
<th>New</th>
</tr>
</thead>
<tbody>
<tr>
<td>37.3%</td>
<td>48.6%</td>
<td></td>
</tr>
</tbody>
</table>
Survival analysis appropriately accounts for the outcomes we are unable to observe.
Among those still alive, what is the risk of death?

![Graph showing survival data over days]
Among those still alive, what is the risk of death?
Complications in Survival Analysis
More formally...

Survivorship:

\[ S(t) = Pr(T > t) \]  \hspace{1cm} (1)

This survivorship function is estimated using the Kaplan-Meier method.
# Kaplan-Meier Method

<table>
<thead>
<tr>
<th>Days</th>
<th>At Risk</th>
<th>Event</th>
<th>Survivorship Est</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.22</td>
<td>20.00</td>
<td>1.00</td>
<td>$1 \times \frac{19}{20} = 0.95$</td>
</tr>
<tr>
<td>2.04</td>
<td>19.00</td>
<td>0.00</td>
<td>0.95</td>
</tr>
<tr>
<td>2.37</td>
<td>18.00</td>
<td>1.00</td>
<td>$1 \times \frac{19}{20} \times \frac{17}{18} = 0.90$</td>
</tr>
<tr>
<td>2.79</td>
<td>17.00</td>
<td>1.00</td>
<td>$1 \times \frac{19}{20} \times \frac{17}{18} \times \frac{16}{17} = 0.84$</td>
</tr>
</tbody>
</table>
Kaplan-Meier Curves

Kaplan-Meier Curve

Kaplan-Meier Curve by Group

Complications in Survival Analysis
Kaplan-Meier Curves

Kaplan-Meier Curve

Days

Kaplan-Meier Curve by Group

log-rank p-value = 0.0289

Days

Elizabeth Koehler VICC Cancer Biostatistics Center
Complications in Survival Analysis
Testing KM Curves

Log-rank test is essentially a chi-squared test

\[
Q = \frac{\sum_{i=1}^{m} (d_{1i} - \hat{e}_{1i})}{\sum_{i=1}^{m} \hat{v}_{1i}}
\]

(2)

- the last time has the same weight as the first time
Testing KM Curves

Log-rank test is essentially a chi-squared test

\[ Q = \frac{\sum_{i=1}^{m} (d_{1i} - \hat{e}_{1i})}{\sum_{i=1}^{m} \hat{v}_{1i}} \]  

the last time has the same weight as the first time

- tends to emphasize differences at the end of the survival curve
Testing KM Curves

Log-rank test is essentially a chi-squared test

\[ Q = \frac{\sum_{i=1}^{m} (d_{1i} - \hat{e}_{1i})}{\sum_{i=1}^{m} \hat{v}_{1i}} \]  

- the last time has the same weight as the first time
- tends to emphasize differences at the end of the survival curve
- most powerful under proportional hazards
Complication 1: Crossing KM Curves
Complication 1: Crossing KM Curves

Problem:
Log-rank tests have no power to detect differences in curves when the KM curves cross

Solution:
Can’t fix this problem with weights, best addressed in regression model
What is a hazard?

- hazard: instantaneous death rate at time $t$ given alive up to time $t$, $h(t)$

- hazard ratio: ratio of death rates in one group of patients compared to another
Cox Proportional Hazards Model

\[ h(t, x, \beta) = h_0(t)r(x, \beta) \]  

\( h_0(t) \) = how the hazard changes over time, baseline hazard  
\( r(x, \beta) \) = how hazard changes as a function of subject covariates  

- Cox suggested \( r(x, \beta) = e^{\beta_1 x} \) (over 23,000 citations)
What are Proportional Hazards?

\[
\text{Hazard Ratio} = \frac{h(t, x_1, \beta)}{h(t, x_0, \beta)} = \frac{h_0(t)e^{\beta_1 x_1}}{h_0(t)e^{\beta_1 x_0}} \tag{4}
\]
What are Proportional Hazards?

\[
\text{Hazard Ratio} = \frac{h(t, x_1, \beta)}{h(t, x_0, \beta)} = \frac{h_0(t) e^{\beta_1 x_1}}{h_0(t) e^{\beta_1 x_0}}
\]  \hspace{1cm} (4)

\[
HR = e^{\beta_1 (x_1 - x_0)}
\]  \hspace{1cm} (5)
What are Proportional Hazards?

\[
\text{HazardRatio} = \frac{h(t, x_1, \beta)}{h(t, x_0, \beta)} = \frac{h_0(t) e^{\beta_1 x_1}}{h_0(t) e^{\beta_1 x_0}}
\]

\[
HR = e^{\beta_1 (x_1 - x_0)}
\]

\[
\ln(HR) = \beta_1 (x_1 - x_0)
\]
What are Proportional Hazards?

\[ \text{Hazard Ratio} = \frac{h(t, x_1, \beta)}{h(t, x_0, \beta)} = \frac{h_0(t) e^{\beta_1 x_1}}{h_0(t) e^{\beta_1 x_0}} \]  

\[ HR = e^{\beta_1 (x_1 - x_0)} \]  

\[ \ln(HR) = \beta_1 (x_1 - x_0) \]  

- ratio of hazard functions remains constant over time
Complications in Survival Analysis
Complication 2: Nonproportional Hazards

- if difference in the log hazards changes over time, we will only get an average
What to do?

- check proportional hazard assumption
- stratify on variables
- allow effects to vary with time
Complications in Survival Analysis
Time-Varying Covariates
Time-Varying Covariates
Complication 3: Competing Risks

Both KM and Cox PH methods operate under the assumption the censoring mechanism is not related to the endpoint.

What if endpoint is time to recurrence?
Complication 3: Competing Risks

Both KM and Cox PH methods operate under the assumption the censoring mechanism is not related to the endpoint.

What if endpoint is time to recurrence?
- Include deaths as events?
Complication 3: Competing Risks

Both KM and Cox PH methods operate under the assumption the censoring mechanism is not related to the endpoint.

What if endpoint is time to recurrence?
- Include deaths as events?
- Include deaths as censored?
Complications in Survival Analysis
- censoring CR does not allow direct interpretation of the survival probabilities
- including CR as event does not allow for distinguishing the contributions of each risk type

Fine and Gray have provided methods comparable to the log-rank test and Cox proportional hazards.
Competing Risks Regression

- Cox PH, censoring CR: once the CR is observed, the subject is removed from the risk set.
- CR regression: once CR is observed, subject is still contained in the risk set, but their contribution to it diminishes as time progresses. This weight is related to the survivorship function.
Questions to Consider

Do the KM curves cross?
Is it reasonable to assume treatment has the same effect in the beginning of follow-up that it has in the end?
Are there variables whose effect I am not interested in, but I believe may have different baseline hazards?

log hazard stratified by center

Elizabeth Koehler VICC Cancer Biostatistics Center
Complications in Survival Analysis
Are there variables that change over time?
Are subjects being censored for reasons other than administrative ones? (esp National Death Index data)

Are there multiple endpoints of interest?
■ Are subjects being censored for reasons other than administrative ones? (esp National Death Index data)

■ Are there multiple endpoints of interest?

If yes to any of these, you should consider seeking help from a biostatistician!
Thanks especially to Nipun Merchant whose data created the examples for this presentation.

Contact Info:
Elizabeth Koehler
elizabeth.koehler@vanderbilt.edu
322-7193