A marginal model for relative survival

Paul C Lambert\textsuperscript{1,2}, Elisavet Syriopoulou\textsuperscript{1}, Mark J Rutherford\textsuperscript{1}

\textsuperscript{1}Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden
\textsuperscript{2}Department of Health Sciences, University of Leicester, UK

International Biometric Society 2020
Excess Mortality and Relative Survival

Excess Mortality

\[ h(t|X_i) = h^*(t|X_i) + \lambda(t|X_i) \]

- All cause mortality rate partitioned into two components
  - Expected mortality rate
  - Excess mortality rate

Relative Survival

\[ R(t|X_i) = \exp\left( -\int_0^t \lambda(u|X_i) \, du \right) \]

Interpreted as net survival under assumptions.
Excess Mortality and Relative Survival

**Excess Mortality**

\[ h(t|X_i) = h^*(t|X_i) + \lambda(t|X_i) \]

- All cause mortality rate partitioned into two components
  - Expected mortality rate
  - Excess mortality rate

Interpreted as net survival under assumptions.
Excess Mortality

\[ h(t|X_i) = h^*(t|X_i) + \lambda(t|X_i) \]

- All cause mortality rate partitioned into two components
  - Expected mortality rate
  - Excess mortality rate

Relative Survival

\[ R(t|X_i) = \exp \left( -\int_0^t \lambda(u|X_i) \, du \right) \]

Interpreted as net survival under assumptions.
Excess Mortality

\[ h(t|X_i) = h^*(t|X_i) + \lambda(t|X_i) \]

- All cause mortality rate partitioned into two components
  - Expected mortality rate
  - Excess mortality rate
Excess Mortality and Relative Survival

Excess Mortality

\[ h(t|X_i) = h^*(t|X_i) + \lambda(t|X_i) \]

- All cause mortality rate partitioned into two components
  - Expected mortality rate
  - Excess mortality rate

Relative Survival

\[ R(t|X_i) = \exp \left( - \int_0^t \lambda(u|X_i) du \right) \]

- Interpreted as net survival under assumptions.
Marginal Relative Survival

- Often interested in a summary measure.

\[
R_m(t) = E_X [R(t|X_i)]
\]

Estimation of \( R_m(t) \) is:

\[
\hat{R}_m(t) = \frac{1}{N} \sum_{i=1}^{N} \hat{R}(t|X_i)
\]

Effect of age usually non-linear. Assumption of proportional excess hazards usually not appropriate.
Often interested in a summary measure.

### Marginal Relative Survival

\[ R_m(t) = E_x [R(t|X_i)] \]

Estimated in modelling framework by,

### Estimation of \( R_m(t) \)

\[ \hat{R}_m(t) = \frac{1}{N} \sum_{i=1}^{N} \hat{R}(t|X_i) \]
Marginal Relative Survival

- Often interested in a summary measure.

\[
R_m(t) = E_X [R(t|X_i)]
\]

- Estimated in modelling framework by,

\[
\hat{R}_m(t) = \frac{1}{N} \sum_{i=1}^{N} \hat{R}(t|X_i)
\]

- Effect of age usually non-linear. Assumption of proportional excess hazards usually not appropriate.
A model without covariates

- A reasonable all-cause parametric model with no covariates should give similar estimates to the non-parametric Kaplan-Meier estimate
A reasonable all-cause parametric model with no covariates should give similar estimates to the non-parametric Kaplan-Meier estimate.

A relative survival parametric model with no covariates is not similar to the non-parametric (Pohar Perme) estimate.
A model without covariates

All-cause Survival

Relative Survival

Years from diagnosis

All-cause Survival

Years from diagnosis
A model without covariates

All-cause Survival

Relative Survival

Kaplan-Meier

RP model with no covariates

Weibull model (no covariates)

Pohar Perme

RP model (no covariates)
A model without covariates

All-cause Survival

Relative Survival

- Kaplan-Meier
- RP model with no covariates
- Weibull model (no covariates)

- Pohar Perme
- RP model (no covariates)
A model without covariates

All-cause Survival

- Kaplan-Meier
- RP model with no covariates
- Weibull model (no covariates)

Relative Survival

- Pohar Perme
- RP model (no covariates)
Why the discrepancy?

- Consider a model with no covariates for the excess mortality,

\[ h(t|X_i) = h^*(t|X_i) + \lambda(t) \]
Why the discrepancy?

- Consider a model with no covariates for the excess mortality,

$$h(t|X_i) = h^*(t|X_i) + \lambda(t)$$

- All cause mortality rate varies between individuals due to variation in expected mortality rate
Why the discrepancy?

- Consider a model with no covariates for the excess mortality,

\[ h(t | X_i) = h^*(t | X_i) + \lambda(t) \]

- All cause mortality rate varies between individuals due to variation in expected mortality rate
- The model assumes excess mortality rate is constant between individuals
Why the discrepancy?

- Consider a model with no covariates for the excess mortality,

\[ h(t|X_i) = h^*(t|X_i) + \lambda(t) \]

- All cause mortality rate varies between individuals due to variation in expected mortality rate
- The model assumes excess mortality rate is constant between individuals
- Which is not what we want to estimate.....

**Marginal hazard function**

\[ \lambda_m(t) = \frac{E_x[R(t|X)\lambda(t|X)]}{E_x[R(t|X)]} \]
A marginal model

Marginal Model

\[ h_m(t) = h_m^*(t) + \lambda_m(t) \]
A marginal model

\[ h_m(t) = h_m^*(t) + \lambda_m(t) \]

- How to define \( h_m^*(t) \)?
- \( \lambda_m(t) \) is the hazard in the hypothetical situation where it is not possible to die from other causes.
A marginal model

**Marginal Model**

\[ h_m(t) = h^*_m(t) + \lambda_m(t) \]

- How to define \( h^*_m(t) \)?
- \( \lambda_m(t) \) is the hazard in the hypothetical situation where it is not possible to die from other causes.
- As time increases those more likely to die from other causes increasingly underrepresented.
- Estimation needs to account for this through incorporation of weights.
Weights

- Weights the same as in non-parametric Pohar Perme estimator.

**Expected survival at time** $t$

$$S^*(t|X_i) = \exp \left(- \int_0^t h^*(u|X_i)du \right)$$

**Weights**

$$w_i^*(t) = \frac{1}{S^*(t|X_i)}$$
Weights

- Weights the same as in non-parametric Pohar Perme estimator.

**Expected survival at time** $t$

$$S^*(t|X_i) = \exp \left( - \int_0^t h^*(u|X_i)du \right)$$

**Weights**

$$w_i^*(t) = \frac{1}{S^*(t|X_i)}$$

- Weights are....
  - Incorporated into likelihood
  - Used to calculate weighted marginal expected mortality rate.
\[ \ln L_i = d_i \ w_i^*(t_i) \ \ln [h_m^*(t_i) + \lambda_m(t_i | \gamma)] - \int_0^{t_i} w_i^*(u) \ \lambda_m(u | \gamma) \, du \]
\[
\ln L_i = d_i w_i^*(t_i) \ln [h^*_m(t_i) + \lambda_m(t_i|\gamma)] - \int_0^{t_i} w_i^*(u) \lambda_m(u|\gamma) \, du
\]
Likelihood

\[ \ln L_i = d_i \ w_i^* (t_i) \ \ln (h_m^*(t_i)) + \lambda_m(t_i|\gamma) - \int_0^{t_i} w_i^*(u) \ \lambda_m(u|\gamma) \, du \]

Marginal expected hazard

\[ h_m^*(t_i) = \frac{\sum_{j \in R(t_i)} w_j^*(t_i) \ h^*(t_i|X_j)}{\sum_{j \in R(t_i)} w_j^*(t_i)} \]
Likelihood

\[ \ln L_i = d_i \ w_i^*(t_i) \ \ln [h_m^*(t_i) + \lambda_m(t_i|\gamma)] - \int_0^{t_i} w_i^*(u) \ \lambda_m(u|\gamma) \, du \]

Marginal expected hazard

\[ h_m^*(t_i) = \frac{\sum_{j \in \mathcal{R}(t_i)} \left[ w_j^*(t_i) \right] h^*(t_i|X_j)}{\sum_{j \in \mathcal{R}(t_i)} w_j^*(t_i)} \]
Approximation

- We approximate the integral in the likelihood by splitting the time-scale into a number of intervals and assume that the weight is constant within each interval.
We approximate the integral in the likelihood by splitting the time-scale into a number of intervals and assume that the weight is constant within each interval. This means that standard software to fit relative survival models can be used.
Approximation

- We approximate the integral in the likelihood by splitting the time-scale into a number of intervals and assume that the weight is constant within each interval.
- This means that standard software to fit relative survival models can be used.
- The software needs to be able to,
  - incorporate weights
  - incorporate delayed entry
We approximate the integral in the likelihood by splitting the time-scale into a number of intervals and assume that the weight is constant within each interval.

This means that standard software to fit relative survival models can be used.

The software needs to be able to,

- incorporate weights
- incorporate delayed entry

When choosing number of split points there is a balance between accuracy versus computational efficiency.

Stata software mrsprep calculates weights and restructures the data.
<table>
<thead>
<tr>
<th></th>
<th>Time</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>5</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Pohar Perme</td>
<td>-0.0012</td>
<td>0.0006</td>
<td>0.0033</td>
<td></td>
</tr>
<tr>
<td></td>
<td>95.1</td>
<td>96.1</td>
<td>95.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>215.703</td>
<td>323.095</td>
<td>505.115</td>
<td></td>
</tr>
<tr>
<td>Conditional Model</td>
<td>0.0292</td>
<td>0.0590</td>
<td>0.0584</td>
<td></td>
</tr>
<tr>
<td></td>
<td>45.9</td>
<td>10.3</td>
<td>15.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1058.174</td>
<td>3806.964</td>
<td>3802.487</td>
<td></td>
</tr>
<tr>
<td>Regression standardization*</td>
<td>0.0007</td>
<td>0.0014</td>
<td>0.0030</td>
<td></td>
</tr>
<tr>
<td></td>
<td>94.6</td>
<td>97.6</td>
<td>96.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>194.753</td>
<td>309.939</td>
<td>449.500</td>
<td></td>
</tr>
<tr>
<td>Marginal model</td>
<td>0.0007</td>
<td>0.0009</td>
<td>0.0035</td>
<td></td>
</tr>
<tr>
<td></td>
<td>95.7</td>
<td>96.9</td>
<td>95.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>194.764</td>
<td>307.189</td>
<td>498.378</td>
<td></td>
</tr>
</tbody>
</table>

**bias, Coverage, MSE**

* 0.5% of models did not converge
## Simulation

<table>
<thead>
<tr>
<th></th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Pohar Perme</td>
<td>-0.0012</td>
</tr>
<tr>
<td></td>
<td>95.1</td>
</tr>
<tr>
<td></td>
<td>215.703</td>
</tr>
<tr>
<td>Conditional Model</td>
<td>0.0292</td>
</tr>
<tr>
<td></td>
<td>45.9</td>
</tr>
<tr>
<td></td>
<td>1058.174</td>
</tr>
<tr>
<td>Regression standardization*</td>
<td>0.0007</td>
</tr>
<tr>
<td></td>
<td>94.6</td>
</tr>
<tr>
<td></td>
<td>194.753</td>
</tr>
<tr>
<td>Marginal model</td>
<td>0.0007</td>
</tr>
<tr>
<td></td>
<td>95.7</td>
</tr>
<tr>
<td></td>
<td>194.764</td>
</tr>
</tbody>
</table>

**bias, Coverage, MSE**

* 0.5% of models did not converge
## Simulation

<table>
<thead>
<tr>
<th></th>
<th>Time 1</th>
<th>Time 5</th>
<th>Time 10</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pohar Perme</strong></td>
<td>-0.0012</td>
<td>0.0006</td>
<td>0.0033</td>
</tr>
<tr>
<td></td>
<td>95.1</td>
<td>96.1</td>
<td>95.2</td>
</tr>
<tr>
<td></td>
<td>215.703</td>
<td>323.095</td>
<td>505.115</td>
</tr>
<tr>
<td><strong>Conditional Model</strong></td>
<td>0.0292</td>
<td>0.0590</td>
<td>0.0584</td>
</tr>
<tr>
<td></td>
<td>45.9</td>
<td>10.3</td>
<td>15.0</td>
</tr>
<tr>
<td></td>
<td>1058.174</td>
<td>3806.964</td>
<td>3802.487</td>
</tr>
<tr>
<td><strong>Regression standardization</strong></td>
<td>0.0007</td>
<td>0.0014</td>
<td>0.0030</td>
</tr>
<tr>
<td></td>
<td>94.6</td>
<td>97.6</td>
<td>96.9</td>
</tr>
<tr>
<td></td>
<td>194.753</td>
<td>309.939</td>
<td>449.500</td>
</tr>
<tr>
<td><strong>Marginal model</strong></td>
<td>0.0007</td>
<td>0.0009</td>
<td>0.0035</td>
</tr>
<tr>
<td></td>
<td>95.7</td>
<td>96.9</td>
<td>95.2</td>
</tr>
<tr>
<td></td>
<td>194.764</td>
<td>307.189</td>
<td>498.378</td>
</tr>
</tbody>
</table>

**bias, Coverage, MSE**

* 0.5% of models did not converge
External (Age) standardization and contrasts

- When comparing different populations there is a need to standardize to the same covariate distribution.
- Common to standardize to external age distribution.
  - E.g. The International Cancer Survival Standard (ICSS)
- Let $p_i^a$ be the proportion in the age group to which the $i^{th}$ individual belongs
- Let $p_i^R$ be the corresponding proportion in the reference population.
Weights can be defined to upweight or downweight individual relative to the reference population.

**Incorporating weights**

\[
\begin{align*}
    w^a_i &= \frac{p^R_i}{p^a_i} \\
    w_i(t) &= w^a_i w^*_i(t)
\end{align*}
\]

- Enables externally age-standardized estimates to be obtained without the need to model, or stratify by, age.
- When modelling covariates (e.g. different regions/countries, socio-economic groups, time-periods or sexes) weights should be calculated separately within subgroups.
Data distributed with strs Stata package.
I will compare relative survival of males and females.
Need to age standardize to same age distribution (ICSS).
Use Flexible Parametric (Royston-Parmar) models.
Proportional Excess Hazards

Relative Survival vs Years from diagnosis for Males and Females.
Melanoma Example

Marginal excess hazard ratio = 0.71 (0.59, 0.85)

Proportional Excess Hazards

Marginal excess hazard ratio = 0.71 (0.59, 0.85)

Paul C Lambert
Marginal Relative Survival
August 2020
External Age Standardization

**Difference in Marginal Relative Survival**

- Y-axis: Relative Survival
- X-axis: Years from diagnosis
- The graph shows the difference in marginal relative survival over time.

**Marginal excess hazard ratio**

- Y-axis: Marginal excess hazard ratio
- X-axis: Years from diagnosis
- The graph illustrates the marginal excess hazard ratio over time.
Conclusion

- Enables estimation of (externally) standardized marginal relative survival without the need to model or stratify by age (or other covariates affecting expected mortality rates).
- Approach enables further adjustment of relative survival models using IPW (See Betty Syrioupoulou’s talk).
- Useful way to obtain summary measure, but conditional model is useful for more detailed comparisons of population groups.