Alternatives to Net and Relative Survival for Comparison of Survival between Populations

Peter Sasieni & Adam Brentnall
Wolfson Institute of Preventive Medicine
Queen Mary University of London
Problem

• To describe the survival in patients diagnosed with cancer reflecting only the mortality in excess of what they would have experienced in any case

• Two approaches
  • Cause-specific survival (death from cancer)
    • Problem 1: non-independence of causes of death
    • Problem 2: difficulty determining cause of death
  • Adjust for expected mortality
Why not simply use overall survival?

• Survival in patients diagnosed aged 75 plus will be much worse that in patients aged 55-74
  • Is that because elderly patients:
    • Aren’t treated properly?
    • Have co-morbidities and are more frail?
    • Simply die more often from completely unrelated diseases?

• Survival of cancer patients diagnosed in 2000-2004 much better than in those diagnosed in 1970-1974
  • Better treatment, earlier diagnosis
  • Fewer dying from cardiovascular disease, infections, ...
How should one adjust survival using expected rates?

• Method 1: Relative survival

\[ S_r(t) = \frac{S_O(t)}{S_E(t)} \]

\( r \)=relative

\( O \)=Observed

\( E \)=Expected
Example: Relative and cause-specific fatality

Cumulative Probability of Death in Men and Women Age 70+ Diagnosed with Localized Colorectal Cancer, 1985-2001, SEER 9 Registries

- Net probability of death (100 - relative survival)
- Crude probability of death, other
- Crude probability of death, cancer

Total Cumulative Mortality = (100 - Observed Survival)
What is net survival?

1. **Cause-specific survival**

   The survival that would be observed if the patients were only subject to the mortality from the disease of interest.

   If $T$ & $U$ competing survival times:
   - net-hazard
     \[ \lambda(t) = \lim P\{t \leq T < t + \Delta \mid T \geq t\} / \Delta \]
   - crude-hazard
     \[ \lambda^\#(t) = \lim P\{t \leq T < t + \Delta \mid T \geq t, U \geq t\} / \Delta \]

   Would like net-hazard but can only estimate crude-hazard.
What is net survival?

1. Cause-specific survival

2. Relative survival

The survival that would be observed if the patients were not subject to the mortality in the background population
Excess hazards

- Excess hazard is the difference between the observed and the expected hazard
  \[ \lambda_e(t) = \lambda_O(t) - \lambda_E(t) \]
- Note that the excess hazard is the logarithmic derivative of the relative survival:
  \[ S_r(t) = \frac{S_O(t)}{S_E(t)} \]
  where
  \[ \lambda_O(t) = -\frac{d \ln \{S_O(t)\}}{dt}; \quad \lambda_E(t) = -\frac{d \ln \{S_E(t)\}}{dt}; \]
  and hence \[ \lambda_e(t) = -\frac{d \ln \{S_O(t)/S_E(t)\}}{dt} \]
So....

• The excess hazard corresponds to the relative survival
Classical solution

• **Ederer-II** *(Ederer 1959, 1961)*
  • Estimate the relative survival:
    \[
    S_r(t) = \frac{S_O(t)}{S_E(t)}
    \]
    where \(S_O(t)\) is the observed (Kaplan-Meier) survival function, and
    \[
    \frac{d \ln\{S_E(t)\}}{dt} = -\sum_{i=1}^{n} Y_i(t) \lambda_{Ei}(t) / \sum_{i=1}^{n} Y_i(t)
    \]
    where \(Y_i(t)\) indicates whether or not the \(i^{th}\) individual is at risk at time \(t\)
  • It uses the expected hazard for the \(i^{th}\) individual only whilst that individual is at risk
For homogeneous data...

Net survival and relative survival are the same
But for heterogeneous data ...

- The mean relative survival:
  \[ \frac{1}{n} \sum_{i=1}^{n} S_{ri}(t) \]
  is **not** the same as

- The relative (mean) survival:
  \[ \frac{\sum_{i=1}^{n} S_{Oi}(t)}{\sum_{i=1}^{n} S_{Ei}(t)} \]

- **It is the mean relative survival that corresponds to the marginal net survival**
What happens with heterogeneous data?

• For the i’th individual we have $S_{ri}(t)$ corresponding to $\lambda_{ei}(t)$

• But how should we combine these to obtain an overall measure?
  • If the i’th individual dies at $T_i$ should we still try to estimate $S_{ri}(t)$ or $\lambda_{ei}(t)$ beyond $T_i$?
  • If we don’t then the overall estimate will depend on the expected mortality
Traditional approach to heterogeneity

- **Stratify**
  - Assume homogeneous within strata
  - Take a weighted average of the estimates within each strata
  - Note: traditionally the stratum-specific weights were fixed, but Brenner & Hakulinen (2003) allowed time-dependent weights
Problems with stratification

• If strata too broad then not homogeneous
• If strata too narrow then unable to estimate (for large t) because no one still at risk in stratum
Recent quotations

• “In estimating net survival, cancer registries should abandon all classical methods”
• “Due to inherent biases, most of the statistical methods used to estimate net survival are quite inaccurate.”
But ...

• If the excess hazard is homogenous within strata then the stratified Ederer-II estimator is consistent
  • The classical approaches are not so bad so long as one stratifies
How can we estimate the marginal net survival?

• Horvitz–Thompson / inverse probability weighting
  • Divide the indicator of “at-risk”, \( Y_i(t) \), by \( EY_i(t) = S_{Ei}(t) \)

• Pohar-Perme
  • Must use the same weights to estimate the “observed” hazard as well as the expected hazard
  • Yields a consistent estimators of the excess hazard and hence of the (marginal) net survival
Roche (2012) on Pohar-Perme

• “In estimating net survival, cancer registries should abandon all classical methods and adopt the new Pohar-Perme estimator.”

• “Due to inherent biases, most of the statistical methods used to estimate net survival are quite inaccurate.”

• “We see no reason to favour any classically used method ... because, unlike the PPE, they are all biased”
Dickman (2013) on Roche

• “The approach used by Roche et al. to calculate the `bias with the classical methods’ is fundamentally flawed.”
• “Researchers should also be aware that the lack of bias in the PP estimator comes at a price of higher variance.”
Also note:

• If the stratification is so fine that within strata the expected survival is homogenous, \( S_E(. | Z) = S_E(.) \), then
  • The stratified Pohar-Perme estimator is identical to the stratified Ederer-II estimator
Take a step back: What are we trying to do?
What are we trying to do?

• Compare the survival corresponding to the excess hazard in cancer patients in different populations
• Estimate the relative survival when it is the same as the net survival
Measures of net survival

• $S_E(. \mid z)$ is the expected survival conditional on covariates

• $H$ is the distribution of $Z$

• Define $S_r(. \mid z)=S_O(. \mid z)/S_E(. \mid z)$

• Functionals of $S_r$, $S_E$, $H$
  • $R(S_r, S_E, H)(t)$
Requirements of $R(S_r,S_E,H)$

1. It **estimates** the net survival when the net survival is homogeneous. If $S_r(\cdot \mid z) = S_r(\cdot)$ then
   
   $$R(S_r,S_E,H)(t) = S_r(t)$$

2. It is **invariant** under changes of the expected survival and the covariate distribution

   $$R(S_r,S_E,H) = R(S_r,S_{E*},H*)$$

3. **Ordering:** If $S_r(\cdot \mid z) < S_{r*}(\cdot \mid z)$ for all $z$, then

   $$R(S_r,S_E,H)(\cdot) < R(S_{r*},S_E,H)(\cdot)$$
Desirable properties

• Robustness
• **Precision** (efficient estimators will have small variance)
Families of measures

• Ratio of weighted average observed to weighted average expected survival

\[ E_H \{ w(t,Z)S_O(t|Z) \} / E_H \{ w(t,Z)S_E(t|Z) \} \]

• In order for the measure to depend on \( S_O \) only through \( S_r \) the weights must be proportional to \( 1/S_E \):

\[ E_H \{ v(t,Z)S_O(t|Z)/S_E(t|Z) \} / E_H \{ v(t,Z) \} = E_H \{ v(t,Z)S_r(t|Z) \} / E_H \{ v(t,Z) \} \]

• Weighted mean of the relative survival

• Note that \( v(t,z) \) must be proportional to \( h_0(z)/h(z) \) in order for the measure not to depend on \( H \).

\[ E_{H_0} \{ v^*(t,Z)S_r(t|Z) \} \mathrm{~where~} E_{H_0} \{ v^*(t,Z) \} = 1 \]
Families of measures

• Weighted mean of the relative survival

\[ R_{w}^{1} = \frac{E_{H}\{w(t,Z)S_{r}(t|Z) \, h_{0}(Z)/h(Z)\}}{E_{H}\{w(t,Z) \, h_{0}(Z)/h(Z)\}} \]
Families of measures (2\textsuperscript{nd} family)

• Weighted excess hazard

\[
R_w^2 = \exp \left\{ - \int_0^t \frac{E_H \{ w(u,Z) h_0(Z)/h(Z) \} \, d\Lambda_e (u|Z) \} \right\}
\]

Or

\[
R_v^2 = \exp \left\{ - \int_0^t \frac{E_H \{ v(u,Z) S_r(u|Z) \} \, d\Lambda_e (u|Z) \} \right\}
\]
Two families of measures

• Weighted mean of the relative survival

\[ R^1_w(t) = \frac{E_{H_0}\{w(t,Z)S_r(t|Z)\}}{E_{H_0}\{w(t,Z)\}} \]

• Weighted excess hazard

\[ R^2_w(t) = \exp \left\{ \int_0^t \frac{E_{H_0}\{w(u,Z)S_r(u|Z)d\Lambda_e(u|Z)\}}{E_{H_0}\{w(u,Z)S_r(u|Z)\}} \right\} \]

• NB The weights are not a function of \( S_r, S_p \) or \( H \)
Estimators

\[ Q^2_v = \exp \left\{ - \int_0^t \frac{\sum v_i(u)h_0(z_i)/h_n(z_i)Y_i(u)/S_{Ei}(u)\{dNi(u) - d\Lambda_{Ei}(u)\}}{\sum v_i(u)h_0(z_i)/h_n(z_i)Y_i(u)/S_{Ei}(u)} \right\} \]

Here: \( h_n \) is the “empirical density”
\( N_i(t) \) is the counting process (of death)

With: \( v=1 \) and \( h_0/h_n=1 \) we have the Pohar-Perme estimator

Note: \( h_0/h_n \) standardises inside the exponential
Estimators

\[ Q_v^2 = \exp \left\{ - \int_0^t \frac{\sum v_i(u)h_0(z_i)/h_n(z_i)Y_i(u)/S_{Ei}(u)\{dN_i(u) - d\Lambda_{Ei}(u)\}}{\sum v_i(u)h_0(z_i)/hn(z_i)Y_i(u)/SE_i(u)} \right\} \]

With: \( v_i(u) = S_{pi}(u) \) and \( h_0/h_n = 1 \) we have the Ederer-II estimator
Estimators

\[ Q_v^2 = \exp \left\{ - \int_0^t \frac{\sum v_i(u)h_0(z_i)/h_n(z_i)Y_i(u)/S_{E_i}(u)\{dN_i(u)-d\Lambda_{E_i}(u)\}}{\sum v_i(u)h_0(z_i)/h_n(z_i)Y_i(u)/S_{E_i}(u)} \right\} \]

When \( S_E = 1 \) (no competing risk):

• Both Ederer-II and Pohar-Perme give the Kaplan-Meier estimator, while \( Q_v^2 \) is a stratum weighted Kaplan-Meier estimator
Variance of ln(Q)

\[
\int_0^t J(u) \sum_{i=1}^n \left\{ v_i(u) \left( \frac{h_0}{hn}(z_i)/SE_i(u) \right)^2 \right\} dN_i(u)
\]

Where \( J(u) \) is an indicator of at least one individual at risk at \( u \).

In order to control the variance we want to counter balance the \( 1/SE_i \) term which could cause the variance to “blow up” when \( SE_i \) is very small (for some \( i \))

Set \( v_i(u) = S_{0i}(u) \) using a “standard” survival function
\( v_{i}(u) = S_{0i}(u) \): Choice of \( S_0 \)

- S-zero (not S-Oh)
- If \( S_0 = S_E \) (and \( h_0 = h \)) then have Ederer-II
- Want \( S_0 \) to be the minimum of \( S_E \) (or even \( S_O \)) in each of the populations being compared

- Also for robustness want \( S_0(t|z) \) to be zero for values of \( t \) for which \( S_E(t|z) \) can be very small for some \( z \) in one of the populations of interest
- But for precision do not want \( S_0(t|z) \) to be zero unnecessarily
What does $Q_{S_0}^2$ estimate?

- The ratio of observed to expected survival that would be observed in a standard population in which the covariate distribution at diagnosis matched the standard covariate distribution and the expected mortality matched the standard mortality
An estimator for $R^1$

$$\frac{\sum_{i=1}^{n} \{(S_{0i}/SE_i)(t))(h_0/hn)(z_i) \}Y_i(t)}{\hat{S}(t) \sum_{i=1}^{n} \{S_{0i}(t))(h_0/hn)(z_i) \}}$$

Here $\hat{S}(t)$ is the Kaplan-Meier estimator of the censoring distribution

Note: $EY=S_O$ so $Y_i/S_E$ is an “estimate” of the i’th relative survival. Hence this estimator can be viewed as a (very finely) stratified estimator (with stratification weights that are time-dependent)
EXAMPLE: SURVIVAL FROM BREAST CANCER WITH DIST. METS, USA 1973-2010 (n = 16,597)
EXAMPLE: SURVIVAL FROM BREAST CANCER WITH DIST. METS, USA 1973-2010 ($n = 16,597$)
SIMULATION: SETUP ($n = 2,000$)

<table>
<thead>
<tr>
<th>Population</th>
<th>Mortality rate aged 65-69</th>
<th>Mortality rate aged 70-85</th>
<th>Mortality rate aged 86+</th>
<th>Percent in Group 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: women USA 1980</td>
<td>Ref</td>
<td>Ref</td>
<td>Ref</td>
<td>60%</td>
</tr>
<tr>
<td>2: higher mortality</td>
<td>x1.2</td>
<td>x2.0</td>
<td>x4.0</td>
<td>70%</td>
</tr>
<tr>
<td>Standard population</td>
<td>x2.0</td>
<td>x4.0</td>
<td>x100.0</td>
<td>50%</td>
</tr>
</tbody>
</table>

- Two groups in population
  1. Aged 65 at diagnosis
  2. Aged 75 at diagnosis

- Same excess hazard all populations (3% per year)

- No censoring considered here (does not change findings)
SIMULATION RESULTS

(a) 5yr

(b) 10yr

(c) 15yr

(d) 20yr
## Results

<table>
<thead>
<tr>
<th>Measure</th>
<th>Pop</th>
<th>Mean bias (%)</th>
<th>SD (× 100)</th>
<th>Var((\hat{R}))/(R^2) (× 10000)</th>
<th>Mean bias (%)</th>
<th>SD (× 100)</th>
<th>Var((\hat{R}))/(R^2) (× 10000)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) 5-yr</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>net-strata</td>
<td>1</td>
<td>-0.1</td>
<td>1.2</td>
<td>1.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>net-brenner</td>
<td>1</td>
<td>-0.1</td>
<td>1.2</td>
<td>1.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>std-r2</td>
<td>1</td>
<td>0.0</td>
<td>1.0</td>
<td>1.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>std-r3</td>
<td>1</td>
<td>0.0</td>
<td>0.9</td>
<td>1.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>net-strata</td>
<td>2</td>
<td>-0.3</td>
<td>1.6</td>
<td>3.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>net-brenner</td>
<td>2</td>
<td>-0.2</td>
<td>1.6</td>
<td>3.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>std-r2</td>
<td>2</td>
<td>-0.1</td>
<td>1.2</td>
<td>1.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>std-r3</td>
<td>2</td>
<td>-0.1</td>
<td>1.0</td>
<td>1.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(b) 10-yr</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>net-strata</td>
<td>1</td>
<td>-0.2</td>
<td>1.6</td>
<td>5.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>net-brenner</td>
<td>1</td>
<td>-0.1</td>
<td>1.6</td>
<td>5.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>std-r2</td>
<td>1</td>
<td>-0.1</td>
<td>1.5</td>
<td>3.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>std-r3</td>
<td>1</td>
<td>0.0</td>
<td>1.7</td>
<td>4.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>net-strata</td>
<td>2</td>
<td>-0.8</td>
<td>2.8</td>
<td>15.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>net-brenner</td>
<td>2</td>
<td>-0.5</td>
<td>2.8</td>
<td>15.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>std-r2</td>
<td>2</td>
<td>-0.3</td>
<td>1.7</td>
<td>4.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>std-r3</td>
<td>2</td>
<td>-0.2</td>
<td>1.7</td>
<td>4.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(c) 15-yr</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>net-strata</td>
<td>1</td>
<td>0.0</td>
<td>1.9</td>
<td>13.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>net-brenner</td>
<td>1</td>
<td>0.1</td>
<td>1.9</td>
<td>13.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>std-r2</td>
<td>1</td>
<td>0.0</td>
<td>1.9</td>
<td>8.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>std-r3</td>
<td>1</td>
<td>0.1</td>
<td>2.1</td>
<td>9.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>net-strata</td>
<td>2</td>
<td>****</td>
<td>****</td>
<td>****</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>net-brenner</td>
<td>2</td>
<td>-0.8</td>
<td>25.4</td>
<td>2549.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>std-r2</td>
<td>2</td>
<td>-0.7</td>
<td>2.7</td>
<td>17.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>std-r3</td>
<td>2</td>
<td>-0.3</td>
<td>2.9</td>
<td>18.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(d) 20-yr</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>net-strata</td>
<td>1</td>
<td>0.6</td>
<td>2.2</td>
<td>51.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>net-brenner</td>
<td>1</td>
<td>1.0</td>
<td>2.2</td>
<td>52.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>std-r2</td>
<td>1</td>
<td>-0.2</td>
<td>2.4</td>
<td>28.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>std-r3</td>
<td>1</td>
<td>0.1</td>
<td>2.6</td>
<td>28.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>net-strata</td>
<td>2</td>
<td>****</td>
<td>****</td>
<td>****</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>net-brenner</td>
<td>2</td>
<td>7.3</td>
<td>130.1</td>
<td>188389.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>std-r2</td>
<td>2</td>
<td>-1.4</td>
<td>3.9</td>
<td>71.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>std-r3</td>
<td>2</td>
<td>-0.6</td>
<td>4.1</td>
<td>70.9</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>