Session 18
Estimating net survival – past and present

Paul W Dickman\textsuperscript{1} and Paul C Lambert\textsuperscript{1,2}

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

Cancer survival: principles, methods and applications
LSHTM
July 2014
Choose a measure and then choose an estimator

- Crude survival (real world probabilities) [measure]
  - Choice of estimators in either cause-specific or relative survival setting. [See session 19]
- Net survival (hypothetical world probabilities) [measure]
  - Cause-specific setting
    - Censor the survival times of those who die of other causes and apply standard estimators (e.g., Kaplan-Meier).
  - Relative survival setting (Berkson 1942 [1])
    - Ederer I (1961 [2])
    - Ederer II (1959 [3])
    - Hakulinen (1982 [4])
    - Pohar Perme (2012 [5])

‘Net survival’ is not a new concept

- The distinction between net probabilities and crude probabilities has a long history in the theory of competing risks.
- Not new in the field of population-based cancer survival. Jacques Estève and colleagues (1990) [6] were very explicit that net survival is a theoretical measure that can be estimated by either cause-specific survival or relative survival.
- ‘Although this is hardly explicit in [Ederer et al. 1961 [2]], the intention of the originators of this concept was to estimate net survival’ [page 531]

A major breakthrough; Pohar Perme et al. 2012 [5]

- Although the concept of net survival is not new, we only recently fully understood what it is (and what it isn’t).
- Pohar Perme et al. 2012 [5] described net survival (the measure) and the various estimators in a formal framework.
- They showed that net survival is

$$\text{net survival} = \frac{1}{n} \sum_{i=1}^{n} \frac{S_i(t)}{S^*_i(t)}.$$ 

- Which is not the same as

$$\text{relative survival} = \frac{\frac{1}{n} \sum_{i=1}^{n} S_i(t)}{\frac{1}{n} \sum_{i=1}^{n} S^*_i(t)}.$$
Comments

- The Pohar Perme method does not estimate relative survival (the marginal observed divided by the marginal expected survival). It does, however, estimate net survival in a relative survival setting (rather than a cause-specific setting). Here, I believe, Maja and I agree.
- Among the estimators in a relative survival setting, Pohar Perme is the only one that is an unbiased estimator of net survival. This is indisputable.
- Where we disagree, is that I view all of the relative survival based estimators as estimators of net survival.
- Bias in Ederer II is usually small in practice. [Here we agree]

Why would one use a biased estimator (even small bias) when an unbiased alternative exists?

- I’m not crazy! I am a great fan of the Pohar Perme estimator.
- However, I don’t think the Ederer II and model-based estimators are as biased as some would have you believe.
- My main message has been ‘previous estimates made with Ederer II as not necessarily greatly biased’.
- My other message has been that, if moving to the new estimator is a non-trivial exercise (e.g., need to introduce new software) then one should be aware that the benefit (reduced bias) is quite small.
- If you want to do more than nonparametric estimation of net survival then modelling has advantages.

Four estimators in a relative survival setting

Is it a relative survival ratio?
Is it intended as an estimator of net survival?

<table>
<thead>
<tr>
<th>Estimator</th>
<th>Maja</th>
<th>Paul</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pohar Perme</td>
<td>X</td>
<td>✓</td>
</tr>
<tr>
<td>Ederer I</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Ederer II</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Hakulinen</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

* Used as an estimator of net survival, but is biased
In estimating net survival, cancer registries should abandon all classical methods and adopt the new Pohar-Perme estimator.

Unfortunately, due to inherent biases, most of the statistical methods used to estimate net survival are quite inaccurate.”

Great errors may occur …

We see no reason to favor any classically used method such as Ederer I, Hakulinen, Ederer II or the univariable excess-rate regression models because, unlike the PPE, they are all biased”

- Similar views in other papers [7, 8, 9, 10].

“… we believe Roche et al. misrepresent the bias in the preferred classical methods in a manner that may unduly cause alarm.”

“… it would be a great pity if the scientific community dismissed methodologically sound and important research because of Roche and coworker’s claims of “great errors.”

“… the comparisons by Roche et al. are not objective; they grossly overstate the magnitude of the bias in the Ederer II method in a manner that could mislead and alarm the research community.”

“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.”
Relative survival is an estimator of net survival

- We view relative survival as an approach to estimating net survival.
- Pohar Perme [5] showed that relative survival, as it is usually calculated, is a biased estimator of net survival. This is indisputable.
- This has lead to some people viewing relative survival as a separate and distinct quantity to net survival. We disagree.
- Relative survival is a method for estimating net survival. It is biased, but the bias in practice is so small (Ederer II or modelling approach) that it can be ignored (see later).

Relative survival was designed to estimate net survival

- The concept of relative survival, the ratio of observed to expected survival, was introduced by Berkson in 1942 [1], although he did not use the term ‘relative survival’.
- He proposed relative survival as an estimator for ‘the survival so far as cancer is concerned’ [1], the concept that is today known as net survival.
- Although the term ‘net survival’ was not used in the early literature, Berkson and Ederer viewed both cause-specific survival and relative survival as estimators of net survival, a view that we share.

Definition of relative survival

- In their seminal article from 1961 [2], Ederer and colleagues defined the ‘relative survival rate’ as ‘the ratio of the observed survival rate in a group of patients, during a specified interval, to the expected survival rate. The expected survival rate is that of a group similar to the patient group in such characteristics as age, sex, and race, but free of the specific disease under study’. [their emphasis]
- Relative survival is a ratio rather than a rate, and observed and expected survival are proportions rather than rates, but we otherwise use this same definition.
- We define relative survival as the the ratio of the all-cause survival of the patients to the (all-cause) survival that would be expected in the absence of the specific disease under study.
Choice of approach for calculating relative survival

- Each of the following approaches provide reasonable estimates of 5-year net survival for most applications:
  - Model-based
  - Ederer II (age standardised)
  - Pohar Perme
- We can show, in extreme scenarios, that the Pohar Perme method is highly variable but this is not a concern for most practical applications.
- Ederer II is theoretically biased, but the bias is so small that it doesn’t matter in practice (provided one age-standardises).
- We prefer model-based estimation, but any of the estimators can be used.

Choice of approach for estimating net survival

- We view the Pohar Perme estimator as an estimator of net survival in a relative survival framework; it doesn’t estimate relative survival (the marginal observed divided by the marginal expected) but it uses a relative survival, rather than a cause-specific, approach.
- The Pohar Perme method estimates the following quantity

\[ R_s(t) = \frac{1}{n} \sum_{i=1}^{n} \frac{S_i(t)}{S_i^*(t)}. \]

- It can be interpreted as the marginal net survival under the same two assumptions (presented in the next section) that are required to interpret relative survival as net survival.

Figure 1 from Hakulinen (1977) [12]
Choosing the relative survival method for cancer survival estimation

Timo Hakulinen
Finnish Cancer Registry, Liisankatu 21 B, SF-00170 Helsinki 17, Finland

ABSTRACT

The theoretical and empirical results show a good agreement between the method suggested in 1959 by Ederer and Heise (the so-called Ederer II method) and the gold standard. This result is in part due to the fact that as follow-up time increases the conditional survival becomes increasingly more independent of age. Moreover, (annual) survival ratios become increasingly more independent of age. Furthermore, the empirical results suggested that the use of the method by Ederer and Heise, multiplication of the annual relative survival ratios, instead of direct standardisation, should be considered in future applications. This would be particularly important for the long-term follow-up when age-specific relative survival is not available in the oldest age categories.

Conclusion: The use of the method by Ederer and Heise, multiplication of the annual relative survival ratios, instead of direct standardisation, should be considered in future applications. This would be particularly important for the long-term follow-up when age-specific relative survival is not available in the oldest age categories.

Methods: The gold standard for the cumulative relative survival ratio is the weighted average of the observed survival ratios. This is the so-called Ederer I method. The method by Ederer and Heise is the so-called Ederer II method. The two methods are essentially equivalent if survival probability is estimated for a limited time horizon.

RESULTS:

The observed survival ratios were estimated from the population-based Finnish Cancer Registry were studied for the different relative survival ratios, instead of direct standardisation, should be considered in future applications. This would be particularly important for the long-term follow-up when age-specific relative survival is not available in the oldest age categories.

CANCER Survival Corrected for Heterogeneity in Patient Withdrawal

Timo Hakulinen
Finnish Cancer Registry, Liisankatu 21 B, SF-00170 Helsinki 17, Finland

Informative right-censoring

- To make it possible for statistical analysis we make the crucial assumption that, conditional on the values of any explanatory variables, censoring is unrelated to future prognosis.
- The statistical methods used for survival analysis assume that the prognosis for an individual censored at time t will be no different from those individuals who were alive at time t and were under follow-up past time t.
- One way to think of this is that, conditional on the values of any explanatory variables, the individuals censored at time t should be a random sample of the individuals at risk at time t.
- This is known as noninformative censoring. Under this assumption, there is no need to distinguish between the different reasons for right-censoring.
Informative right-censoring 2

- When withdrawal from follow-up is associated with prognosis, this is known as informative censoring and standard methods of analysis will result in biased estimates.
- Common methods for controlling for informative censoring are to stratify or condition on those explanatory factors on which censoring depends.
- Determining whether or not censoring is informative is not a statistical issue — it must be made based on subject matter knowledge.

Life table estimates of patient survival

Women diagnosed 1994 - 2001 with follow-up to the end of 2002

<table>
<thead>
<tr>
<th>Interval</th>
<th>Number at risk</th>
<th>Deaths</th>
<th>Censored</th>
<th>Effective number at risk</th>
<th>Interval-specific observed survival</th>
<th>Cumulative observed survival</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>N</td>
<td>D</td>
<td>W</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1559</td>
<td>209</td>
<td>0</td>
<td>1559.0</td>
<td>0.86594</td>
<td>0.86594</td>
</tr>
<tr>
<td>2</td>
<td>1350</td>
<td>125</td>
<td>177</td>
<td>1261.5</td>
<td>0.90091</td>
<td>0.78014</td>
</tr>
<tr>
<td>3</td>
<td>1048</td>
<td>58</td>
<td>172</td>
<td>962.0</td>
<td>0.93971</td>
<td>0.73310</td>
</tr>
<tr>
<td>4</td>
<td>818</td>
<td>32</td>
<td>155</td>
<td>740.5</td>
<td>0.95679</td>
<td>0.70142</td>
</tr>
<tr>
<td>5</td>
<td>631</td>
<td>23</td>
<td>148</td>
<td>557.0</td>
<td>0.95871</td>
<td>0.67246</td>
</tr>
<tr>
<td>6</td>
<td>460</td>
<td>10</td>
<td>130</td>
<td>395.0</td>
<td>0.97468</td>
<td>0.65543</td>
</tr>
<tr>
<td>7</td>
<td>320</td>
<td>5</td>
<td>129</td>
<td>255.5</td>
<td>0.98043</td>
<td>0.64261</td>
</tr>
<tr>
<td>8</td>
<td>186</td>
<td>3</td>
<td>134</td>
<td>119.0</td>
<td>0.97479</td>
<td>0.62641</td>
</tr>
<tr>
<td>9</td>
<td>49</td>
<td>1</td>
<td>48</td>
<td>25.0</td>
<td>0.96000</td>
<td>0.60135</td>
</tr>
</tbody>
</table>

Life table estimates of patient survival

Women diagnosed 1994 - 2001 with follow-up to the end of 2002

<table>
<thead>
<tr>
<th>Interval</th>
<th>Number at risk</th>
<th>Deaths</th>
<th>Censored</th>
<th>Effective number at risk</th>
<th>Interval-specific observed survival</th>
<th>Cumulative observed survival</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>N</td>
<td>D</td>
<td>W</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1559</td>
<td>209</td>
<td>0</td>
<td>1559.0</td>
<td>0.86594</td>
<td>0.87472</td>
</tr>
<tr>
<td>2</td>
<td>1350</td>
<td>125</td>
<td>177</td>
<td>1261.5</td>
<td>0.90091</td>
<td>0.90829</td>
</tr>
<tr>
<td>3</td>
<td>1048</td>
<td>58</td>
<td>172</td>
<td>962.0</td>
<td>0.93971</td>
<td>0.94772</td>
</tr>
<tr>
<td>4</td>
<td>818</td>
<td>32</td>
<td>155</td>
<td>740.5</td>
<td>0.95679</td>
<td>0.96459</td>
</tr>
<tr>
<td>5</td>
<td>631</td>
<td>23</td>
<td>148</td>
<td>557.0</td>
<td>0.95871</td>
<td>0.98284</td>
</tr>
<tr>
<td>6</td>
<td>460</td>
<td>10</td>
<td>130</td>
<td>395.0</td>
<td>0.97468</td>
<td>0.98848</td>
</tr>
<tr>
<td>7</td>
<td>320</td>
<td>5</td>
<td>129</td>
<td>255.5</td>
<td>0.98043</td>
<td>0.99772</td>
</tr>
<tr>
<td>8</td>
<td>186</td>
<td>3</td>
<td>134</td>
<td>119.0</td>
<td>0.97479</td>
<td>1.00000</td>
</tr>
<tr>
<td>9</td>
<td>49</td>
<td>1</td>
<td>48</td>
<td>25.0</td>
<td>0.96000</td>
<td>1.00000</td>
</tr>
</tbody>
</table>
Necessary assumptions for relative survival to estimate net survival

- For RSR to estimate net survival, we need
  - appropriate population life tables; and
  - independence (conditional on covariates) of cancer and non-cancer mortality.
- What do we mean by ‘independence’ in this context? We mean that there are no factors associated with both cancer and non-cancer mortality other than those factors that are adjusted for in both the analysis and in the population mortality file.
- The Pohar Perme estimator addresses point 2 in a non-parametric setting, but assumption 1 must still be assessed (for all RSR-based estimators).

Necessary assumptions for cause-specific survival to estimate net survival

- For cause-specific survival to estimate net survival, we need
  - accurate cause of death information; and
  - independence (conditional on covariates) of cancer and non-cancer mortality.
- What do we mean by ‘independence’ in this context? We mean that there are no factors associated with both cancer and non-cancer mortality other than those factors for which we have adjusted in the model.

Assumptions required for estimating crude probabilities of death

- To estimate crude probabilities of death in a cause-specific setting, we do not require the independence assumption (we only require correct classification of cause of death).
- Similarly, the independence assumption is not required to estimate crude probabilities of death in a relative survival framework (we only require appropriate population life tables).
Summary: Ederer I, Ederer II, and Hakulinen

- Expected survival can be thought of as being calculated for a cohort of patients from the general population matched by age, sex, and period. The methods differ regarding how long each individual is considered to be ‘at risk’ for the purpose of estimating expected survival.

- Ederer I: the matched individuals are considered to be at risk indefinitely (even beyond the closing date of the study). The time at which a cancer patient dies or is censored has no effect on the expected survival.

- Ederer II: the matched individuals are considered to be at risk until the corresponding cancer patient dies or is censored.

- Hakulinen: if the survival time of a cancer patient is censored then so is the survival time of the matched individual. However, if a cancer patient dies the matched individual is assumed to be ‘at risk’ until the closing date of the study.

Illustration of expected survival estimation

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
<th>J</th>
<th>K</th>
<th>L</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
<td>16</td>
</tr>
<tr>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>17</td>
</tr>
<tr>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>17</td>
<td>18</td>
</tr>
<tr>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>17</td>
<td>18</td>
<td>19</td>
</tr>
<tr>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>17</td>
<td>18</td>
<td>19</td>
<td>20</td>
</tr>
<tr>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>17</td>
<td>18</td>
<td>19</td>
<td>20</td>
<td>21</td>
</tr>
<tr>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>17</td>
<td>18</td>
<td>19</td>
<td>20</td>
<td>21</td>
<td>22</td>
</tr>
<tr>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>17</td>
<td>18</td>
<td>19</td>
<td>20</td>
<td>21</td>
<td>22</td>
<td>23</td>
</tr>
</tbody>
</table>

The Pohar Perme estimator


- Pohar Perme et al. illustrate that existing relative survival-based estimators do not exactly estimate marginal net survival, and presented a new unbiased estimator.

- The new approach estimates net survival as internally age-standardized relative survival, but without the need to perform separate calculations for each age strata.

- The method addresses the specific scenario where we wish to estimate net survival, non-parametrically, for a single cohort (i.e., all ages). Pohar Perme et al. remark in their paper that a model-based approach is also possible.
Net survival 1

- Overall mortality made up of two components.
  \[ h_i(t) = h_i^*(t) + \lambda_i(t) \]
- Net survival for an individual \(i\)
  \[ S_{Ni}(t) = \exp \left( - \int_0^t \lambda_i(u) du \right) \]
- The overall (marginal) net survival for a cohort is
  \[ S_N(t) = \frac{1}{N} \sum_{i=1}^{N} S_{Ni}(t) \]
- That is, overall (marginal) net survival is the average of the net survival over all individuals.

Net survival 2

- This implies that the marginal net hazard is
  \[ \lambda_N(t) = \frac{\sum_{i=1}^{N} S_{Ni}(t) \lambda_i(t)}{\sum_{i=1}^{N} S_{Ni}(t)} \]
- This is a weighted average of the individual excess mortality rates with weights equal to the individual net survival.
- Note that in a multivariable model for relative survival, we can estimate the marginal net survival
  \[ \hat{S}_N(t) = \frac{1}{N} \sum_{i=1}^{N} \hat{R}_i(t) \]
  where \(\hat{R}_i(t)\) is the model-based estimate of relative survival for individual \(i\).
- However, the aim of the Pohar Perme approach is to estimate the marginal net survival without modelling.

Ederer II estimate (ignoring age group)

- The hazard rate for the Ederer II estimator can be written as
  \[ \lambda_{E2}(t) = \frac{\sum_{i=1}^{N} S_{Oi}(t) \lambda_i(t)}{\sum_{i=1}^{N} S_{Oi}(t)} \]
- Recall that the marginal net hazard is actually
  \[ \lambda_N(t) = \frac{\sum_{i=1}^{N} S_{Ni}(t) \lambda_i(t)}{\sum_{i=1}^{N} S_{Ni}(t)} \]
  \(\lambda_{E2}\) is called observable net survival. The weights are \(S_{Oi}(t)\) (all-cause survival) rather than \(S_{Ni}(t)\) (net survival).
- \(\lambda_{E2}\) is the same as what would be estimated in a cause-specific setting treating deaths due to other causes as censored.
Ederer I/Hakulinen estimates (ignoring age group)

- Ederer II estimator can also be written,
  \[ \lambda_{E2}(t) = \lambda_O(t) - \frac{\sum_{i=1}^{N} S_{Oi}(t) h_i^*(t)}{\sum_{i=1}^{N} S_{Oi}(t)} \]
- The hazard rate for the Ederer I estimate can be written as
  \[ \lambda_{E1}(t) = \lambda_O(t) - \frac{\sum_{i=1}^{N} S_i^*(t) h_i^*(t)}{\sum_{i=1}^{N} S_i^*(t)} \]
- Hakulinen is the same under non-informative censoring.
- Thus the current estimators (Ederer I and II and Hakulinen) do not exactly estimate net survival.

Summary of current estimators

- When net survival is equal over ages (and sex etc.) then the different estimators give identical estimates.
- When expected survival is the same for all subjects, the different estimators give identical estimates.
- This is why we see less variation between methods when estimating within age groups.
- This is why there is less variation between the different estimators of age-standardized estimates.

The New Estimator (adapted to discrete time)

- The estimate of net survival is in a hypothetical world.
- To be at risk at time \( t \), an individual has
  - not died of their cancer.
  - not died of other causes.
- Compared to the hypothetical world,
  - The number at risk is too small (because people die due to causes other than cancer).
  - The number of events (deaths due to cancer) is too small.
- Solution: weight by inverse of expected survival.
- Pohar Perme estimates in continuous time. Needs to be adapted to discrete time.
The New Estimator

- **k**: Interval length
- **wij**: Weight for \(i^{th}\) subject in \(j^{th}\) interval.
- **dij**: Event indicator for \(i^{th}\) subject in \(j^{th}\) interval.
- **dij**: Expected deaths for \(i^{th}\) subject in \(j^{th}\) interval (\(dij = hijyij\)).
- **yij**: Time at risk for \(i^{th}\) subject in \(j^{th}\) interval.

Estimate cumulative weighted excess hazard

\[
H_{wj} = \sum_j k \left( \sum_i wijdij - \sum_i wijdij^* \right) / \sum_i wijyij
\]

- Weights are the inverse of expected survival and vary as a function of follow-up time.
- The weights have the effect of increasing the sample still at risk to account for the expected proportion of patients lost due to mortality due to other causes.

Some comments on Pohar Perme

- Provides an internally age standardized estimate of relative survival without the need to estimate relative survival separately by age group.
- Strictly, the estimate is also standardized over other variables in the population mortality file (usually calendar year and sex).
- If comparing net survival between groups, the Pohar Perme estimate will not account for different age distributions in the different groups. Still need to externally age standardize.
- The method is designed for continuous time and is unbiased for continuous time. Properties unclear when we have discrete time.

Comparison of approaches (ignoring age group)
**Comparison of approaches (ignoring age group)**

The Pohar Perme estimate is more variable than the standard estimators but only an issue in practice for old patients and/or long follow-up.

This is related to the use of weights. Older people carry high weight, particularly as follow-up increases. Their exit can lead to large changes in the estimates.

**A simulation study (submitted 2014)**

- We performed a simulation study to assess bias and coverage of the different methods of estimating age-standardised relative/net survival.
- We chose to simulate scenarios where there is substantial variation by age. In practice most cancers have far less variation.
- Compared age-standardized relative survival using Ederer II and Pohar Perme at 5, 10 and 15 years.

**Scenario 1** The excess mortality rate varies by age in the first year, but after one year there is no variation in excess mortality by age. Excess mortality rate ratio in first year is 1.051 per yearly increase in age.

**Scenario 2** The excess mortality rate varies by age throughout the follow-up time with an excess mortality rate ratio of 1.03 per yearly increase in age.
Scenario 1: External Age Standardisation: 5 years

<table>
<thead>
<tr>
<th>Method</th>
<th>Bias</th>
<th>MSE</th>
<th>Coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pohar Perme</td>
<td>-0.09</td>
<td>0.315</td>
<td>95.5</td>
</tr>
<tr>
<td>Ederer 2 (Brenner)</td>
<td>0.16</td>
<td>0.272</td>
<td>94.8</td>
</tr>
<tr>
<td>Ederer 2 (Standardized)</td>
<td>-0.09</td>
<td>0.265</td>
<td>95.1</td>
</tr>
<tr>
<td>Ederer 2 (Standardized)</td>
<td>0.26</td>
<td>0.321</td>
<td>92.3</td>
</tr>
<tr>
<td>stpm2 (grouped)</td>
<td>-0.10</td>
<td>0.256</td>
<td>95.3</td>
</tr>
<tr>
<td>Model based (continuous)</td>
<td>-0.10</td>
<td>0.265</td>
<td>95.3</td>
</tr>
</tbody>
</table>
Scenario 1: External Age Standardisation: 10 years

- **Pohar Perme**
  - Bias: -0.13
  - MSE: 1.408
  - Coverage: 93.0%

- **Ederer 2 (Brenner)**
  - Bias: 0.14
  - MSE: 0.378
  - Coverage: 93.7%

- **Ederer 2 (Standardized)**
  - Bias: -0.19
  - MSE: 0.493
  - Coverage: 93.9%

- **Ederer 2 (Standardized)**
  - Bias: 0.37
  - MSE: 0.576
  - Coverage: 90.6%

- **stpm2 (grouped)**
  - Bias: 0.07
  - MSE: 0.695
  - Coverage: 93.3%

- **Model based (continuous)**

---

Scenario 1: External Age Standardisation: 15 years

- **Pohar Perme**
  - Bias: -0.30
  - MSE: 8.745
  - Coverage: 89.6%

- **Ederer 2 (Brenner)**
  - Bias: 0.13
  - MSE: 0.470
  - Coverage: 93.0%

- **Ederer 2 (Standardized)**
  - Bias: -0.28
  - MSE: 1.074
  - Coverage: 93.0%

- **Ederer 2 (Standardized)**
  - Bias: 0.22
  - MSE: 0.842
  - Coverage: 93.0%

- **stpm2 (grouped)**
  - Bias: 0.03
  - MSE: 1.528
  - Coverage: 92.6%

- **Model based (continuous)**

---

Scenario 2: External Age Standardisation: 5 years

- **Pohar Perme**
  - Bias: -0.10
  - MSE: 0.330
  - Coverage: 94.8%

- **Ederer 2 (Brenner)**
  - Bias: 0.93
  - MSE: 1.108
  - Coverage: 53.1%

- **Ederer 2 (Standardized)**
  - Bias: 0.10
  - MSE: 0.286
  - Coverage: 96.0%

- **Ederer 2 (Standardized)**
  - Bias: 0.64
  - MSE: 0.661
  - Coverage: 75.6%

- **stpm2 (grouped)**
  - Bias: 0.0081
  - MSE: 0.281334
  - Coverage: 0.1%

- **Model based (continuous)**

---

Dickman and Lambert | Population-Based Cancer Survival | LSHTM, June 2014

---

Dickman and Lambert | Population-Based Cancer Survival | LSHTM, June 2014

---

Dickman and Lambert | Population-Based Cancer Survival | LSHTM, June 2014
Scenario 2: External Age Standardisation: 10 years

<table>
<thead>
<tr>
<th>Method</th>
<th>Bias</th>
<th>MSE</th>
<th>Coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pohar Perme</td>
<td>-0.16</td>
<td>1.156</td>
<td>93.8</td>
</tr>
<tr>
<td>Ederer 2 (Brenner)</td>
<td>1.91</td>
<td>3.962</td>
<td>10.7</td>
</tr>
<tr>
<td>Ederer 2 (Standardized)</td>
<td>0.21</td>
<td>0.526</td>
<td>95.2</td>
</tr>
<tr>
<td>Ederer 2 (Standardized)</td>
<td>0.94</td>
<td>1.272</td>
<td>68.0</td>
</tr>
<tr>
<td>stpm2 (grouped)</td>
<td>0.0290</td>
<td>0.558352</td>
<td>0.1</td>
</tr>
<tr>
<td>Model based (continuous)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Scenario 2: External Age Standardisation: 15 years

<table>
<thead>
<tr>
<th>Method</th>
<th>Bias</th>
<th>MSE</th>
<th>Coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pohar Perme</td>
<td>-0.23</td>
<td>5.404</td>
<td>91.1</td>
</tr>
<tr>
<td>Ederer 2 (Brenner)</td>
<td>2.80</td>
<td>8.269</td>
<td>1.2</td>
</tr>
<tr>
<td>Ederer 2 (Standardized)</td>
<td>0.29</td>
<td>1.053</td>
<td>93.7</td>
</tr>
<tr>
<td>Ederer 2 (Standardized)</td>
<td>1.10</td>
<td>1.887</td>
<td>73.8</td>
</tr>
<tr>
<td>stpm2 (grouped)</td>
<td>0.0074</td>
<td>1.049332</td>
<td>0.1</td>
</tr>
<tr>
<td>Model based (continuous)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Summary of simulation

- Under these extreme scenarios, bias in Ederer II is negligible when age standardising.
- A theoretical bias exists, but in practice this can be ignored.
- There is more variation in the Pohar Perme estimate (particularly at 10 and 15 years).
- For the Ederer II method to be unbiased there needs to be either,
  - no variation in expected survival within age groups (certainly not true)
  - no variation in relative survival within age groups (unlikely to be true)
- However, the variation in both of these will be reduced by estimating separately within age groups.
Summary of simulation 2

- The ‘cost’ of the bias when using Ederer II can effectively be ignored, but there is a ‘benefit’ in using Ederer II in terms of precision.
- The difference in the methods is small at 5 years.
- Modelling (continuous age) also performs well.

The oldest age group

- The methods differ most for the oldest age group (often 75+).
- One can question the utility of estimating long-term net survival of elderly patients (the hypothetical world is very different from the real world).
- Large variation in expected survival in this age group and possibly also large variation in net survival.
- The numbers at risk as follow-up time increases will reduce proportionately more than other age groups due to higher mortality due to other cases and to cancer. In the Pohar Perme method, these individuals have a lot of weight.

Scenario 1: Age 75+

<table>
<thead>
<tr>
<th></th>
<th>5 Years</th>
<th>10 Years</th>
<th>15 years</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pohar Perme</strong></td>
<td>-0.1590</td>
<td>-0.1613</td>
<td>-0.6053</td>
</tr>
<tr>
<td></td>
<td>2.1021</td>
<td>14.3727</td>
<td>101.8314</td>
</tr>
<tr>
<td></td>
<td>96.2</td>
<td>92.5</td>
<td>91.4</td>
</tr>
<tr>
<td><strong>Ederer II</strong></td>
<td>-0.1137</td>
<td>-0.3198</td>
<td>-0.5318</td>
</tr>
<tr>
<td></td>
<td>1.5983</td>
<td>3.7299</td>
<td>10.0422</td>
</tr>
<tr>
<td></td>
<td>96.1</td>
<td>95.0</td>
<td>95.1</td>
</tr>
<tr>
<td><strong>Model Based</strong></td>
<td>0.8137</td>
<td>0.9559</td>
<td>0.5888</td>
</tr>
<tr>
<td>(Grouped Age)</td>
<td>2.0951</td>
<td>3.8231</td>
<td>6.8827</td>
</tr>
<tr>
<td></td>
<td>90.0</td>
<td>92.2</td>
<td>94.8</td>
</tr>
<tr>
<td><strong>Model Based</strong></td>
<td>-0.5241</td>
<td>-0.0224</td>
<td>0.0451</td>
</tr>
<tr>
<td>(Continuous Age)</td>
<td>1.1701</td>
<td>5.7305</td>
<td>14.7700</td>
</tr>
<tr>
<td></td>
<td>94.9</td>
<td>93.3</td>
<td>92.7</td>
</tr>
</tbody>
</table>

Bias, MSE and Coverage
Scenario 2: Age 75+

<table>
<thead>
<tr>
<th></th>
<th>5 Years</th>
<th>10 Years</th>
<th>15 years</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pohar Perme</strong></td>
<td>-0.2124</td>
<td>-0.2210</td>
<td>-0.3109</td>
</tr>
<tr>
<td></td>
<td>2.4165</td>
<td>11.5840</td>
<td>61.9570</td>
</tr>
<tr>
<td></td>
<td>95.0</td>
<td>94.3</td>
<td>92.8</td>
</tr>
<tr>
<td><strong>Ederer II</strong></td>
<td>0.5060</td>
<td>0.9602</td>
<td>1.2465</td>
</tr>
<tr>
<td></td>
<td>1.9939</td>
<td>4.3176</td>
<td>9.9081</td>
</tr>
<tr>
<td></td>
<td>94.4</td>
<td>92.7</td>
<td>92.7</td>
</tr>
<tr>
<td><strong>Model Based</strong></td>
<td>1.9387</td>
<td>2.5769</td>
<td>2.8401</td>
</tr>
<tr>
<td>(Grouped Age)</td>
<td>5.1811</td>
<td>9.1152</td>
<td>13.3959</td>
</tr>
<tr>
<td></td>
<td>66.8</td>
<td>67.2</td>
<td>80.1</td>
</tr>
<tr>
<td><strong>Model Based</strong></td>
<td>0.1237</td>
<td>0.0649</td>
<td>-0.0117</td>
</tr>
<tr>
<td>(Continuous Age)</td>
<td>1.6218</td>
<td>4.6727</td>
<td>10.2586</td>
</tr>
<tr>
<td></td>
<td>94.0</td>
<td>95.2</td>
<td>94.6</td>
</tr>
</tbody>
</table>

Bias, MSE and Coverage

What approach is preferable for the following analyses?

- 5-year net survival, age-specific estimates for a recent period.
- 5-year net survival for all ages for a recent period.
- Temporal trends in 5-year net survival (age-standardised) within registry.
- Comparison of 5-year net survival (age-standardised) with other populations.
- Study prognostic factors for cancer survival.

5-year net survival, age-specific estimates for a recent period

- Doesn’t really matter; can use Ederer II, Pohar Perme, or model-based.
- Pohar Perme is theoretically unbiased but can use the other approaches in practice (bias is negligible).
- The small bias in Ederer II will depend on the magnitude of the variation (by age) in survival within each age category.
- That is, usually only an issue in the oldest age group and for longer follow-up.
5-year net survival for all ages for a recent period

- Unlike the previous scenario (age-specific), we must now account for age in the estimation procedure. That is, Ederer II applied to all patients may result in a non-negligible bias.
- The Pohar Perme estimator was designed specifically for this scenario. It accounts for age by weighting.
- Can also account for age by stratification. That is, Ederer II internally age standardised.
- Can also account for age by modelling; model-based estimation.
- We prefer modelling since it has the potential to do more than just estimate the 5-year net survival.

Temporal trends in 5-year net survival (age-standardised) within registry

- Use either Ederer II or Pohar Perme if sufficient data, or Brenner alternative (with Ederer II) if data are sparse.
- Use age distribution in last period as standard so the age-standardised estimates for that period represent the actual survival for patients diagnosed in that period.
- Use world standard cancer population if you wish to compare to published data that also use that standard (or if you want others to be able to make comparisons with your data).
- We like modelling!

Comparison of 5-year net survival (age-standardised) with other populations

- Use age distribution in your registry as standard if you wish to maximise the relevance of the estimates for your specific population.
- Use world standard cancer population if you wish to compare to published data that also use that standard (or if you want others to be able to make comparisons with your data).
- Be aware of differences in registration and inclusion criteria (e.g., definition of multiple primaries). These will have a greater impact on the results than the choice between, for example, Ederer II and Pohar Perme.
Study prognostic factors for cancer survival

- Model! But which model?
- Cause-specific or excess?
  - Cox (only for cause-specific)
  - Poisson
  - Flexible parametric models
  - Other models

References


References 2


