Welcome to Regstat!

- Welcome to what we hope will be the first in a series of annual workshops organised by MEB in statistical methods for registry-based epidemiology.
- This year’s theme is statistical methods for population-based cancer patient survival.
- Today’s workshop has been organised in collaboration with FMS (Föreningen för Medicinsk Statistik). The organisers gratefully acknowledge financial support from Cancerfonden (The Swedish Cancer Society).
- In my presentation I will attempt to simultaneously introduce the field as well as the program for today’s workshop.
- Reprints available at http://pauldickman.com/regstat/

How might we measure the prognosis of cancer patients?

- Total mortality (among the patients).
- Our interest is typically in net mortality (mortality associated with a diagnosis of cancer).
- Cause-specific mortality provides an estimate of net mortality.
- Excess mortality provides an estimate of net mortality.

\[
\text{excess} = \frac{\text{total}}{\text{expected}} - \frac{\text{mortality}}{\text{mortality}}
\]

- Excess mortality is generally preferred when using data collected by population-based cancer registries [1].
Relative survival is the survival analog of excess mortality — the relative survival ratio is defined as the observed survival in the patient group divided by the expected survival of a comparable group from the general population.

\[
\text{relative survival ratio} = \frac{\text{observed survival proportion}}{\text{expected survival proportion}}
\]

Relative survival example

Table 1: Number of cases (N) and 5-year observed (\(p\)), expected (\(p^*\)), and relative (\(r\)) survival for males diagnosed with localised skin melanoma in Finland during 1985–1994.

<table>
<thead>
<tr>
<th>Age</th>
<th>N</th>
<th>(p)</th>
<th>(p^*)</th>
<th>(r)</th>
</tr>
</thead>
<tbody>
<tr>
<td>15–29</td>
<td>67</td>
<td>0.947</td>
<td>0.993</td>
<td>0.954</td>
</tr>
<tr>
<td>30–44</td>
<td>273</td>
<td>0.856</td>
<td>0.982</td>
<td>0.872</td>
</tr>
<tr>
<td>45–59</td>
<td>503</td>
<td>0.824</td>
<td>0.943</td>
<td>0.874</td>
</tr>
<tr>
<td>60–74</td>
<td>449</td>
<td>0.679</td>
<td>0.815</td>
<td>0.833</td>
</tr>
<tr>
<td>75+</td>
<td>200</td>
<td>0.396</td>
<td>0.505</td>
<td>0.784</td>
</tr>
</tbody>
</table>

- Note that relative survival controls for the fact that expected mortality depends on demographic characteristics (age, sex, etc.).
- In addition, relative survival may, and usually does, depend on such factors.

Applying relative survival to diseases other than cancer

- In order to interpret excess mortality as 'mortality due to the disease of interest' we need to accurately estimate expected mortality (the mortality that would have been observed in the absence of the disease).
- General population mortality rates may not satisfy this criteria.
- Excess mortality (compared to the general population) may nevertheless still be of interest.


More on the choice of measure for cancer patient survival

- The fact that excess mortality is generally preferred when using data collected by population-based cancer registries does not necessarily mean it is the uniformly best measure for all such studies (Kathy and others will talk more about this).

- Cause-specific mortality is the measure of choice in clinical trials assessing cancer patient survival. Should excess mortality be given greater consideration for trials?

- Additional (unpublished) analyses of data from a randomised trial of the effect of screening for prostate cancer using PSA and DRA showed a greater benefits of screening on excess mortality than cause-specific mortality.

- It is plausible that contact with health care professionals through screening benefits health (lowers mortality) via mechanisms other than those directly related to prostate cancer. Such benefits are captured with excess mortality but not cause-specific mortality.

Statistical cure

- The life table is a useful tool for describing the survival experience of the patients over a long follow-up period.

- In particular, an interval-specific relative survival ratio equal to one indicates that, during the specified interval, mortality in the patient group was equivalent to that of the general population.

- The attainment and maintenance of an interval-specific RSR of one indicates that there is no excess mortality due to cancer and the patients are assumed to be ‘statistically cured’.

- An individual is considered to be medically cured if he or she no longer displays symptoms of the disease.

- Statistical cure applies at a group, rather than individual, level.

Figure 1: Plots of the annual (interval-specific) relative survival ratios \((r)\) for males and females diagnosed with cancer of the stomach in Finland 1985–1994 and followed up to the end of 1995.
• Plots of the interval-specific RSR are also useful for assessing the quality of follow-up.

• If the interval-specific RSR levels out at a value greater than 1, this generally indicates that some deaths have been missed in the follow-up process.

• An interval-specific relative survival ratio of unity is generally not achieved for smoking-related cancers, such as cancer of the lung and kidney.

• Compared to the general population, these patients are subject to excess mortality due to the cancer in addition to excess mortality due to other conditions caused by smoking, such as cardiovascular disease.

• We’ll return to these concepts later when we discuss cure models.

---

**Interpreting relative survival estimates**

• The cumulative relative survival ratio can be interpreted as the proportion of patients alive after $i$ years of follow-up in the hypothetical situation where the cancer in question is the only possible cause of death.

• $1 - RSR$ can be interpreted as the proportion of patients who will die of cancer within $i$ years of follow-up in the hypothetical situation where the cancer in question is the only possible cause of death.

• We do not live in this hypothetical world. Estimates of the proportion of patients who will die of cancer in the presence of competing risks can also be made.

• Cronin and Feuer (2000) [5] showed how crude and net mortality could be estimated based on relative survival for grouped data (implemented in the Stata command `strs`) and Lambert et al [6] showed how individual-level estimates could be obtained.

---

**Estimating relative survival using a period approach**

• In 1996 Hermann Brenner suggested estimating cancer patient survival using a period, rather than cohort, approach [7].

• Time at risk is left truncated at the start of the period window and right censored at the end.

• This suggestion was initially met with scepticism although studies based on historical data [8] have shown that
  - period analysis provides very good predictions of the prognosis of newly diagnosed patients; and
  - highlights temporal trends in patient survival sooner than cohort methods.
Age-standardisation of relative survival

- The problem is more complex than age-standardisation of, for example, incidence rates since the age-distribution of the patients changes during follow-up.
- Which weights do we use and how does one interpret the resulting estimates?
- See the papers by Pokhrel et al and Brenner et al. [9, 10, 11, 12, 13].

Modelling excess mortality (relative survival)

- The hazard at time since diagnosis \( t \) for persons diagnosed with cancer is modelled as the sum of the known baseline hazard, \( \lambda^*(t) \), and the excess hazard due to a diagnosis of cancer, \( \nu(t) \) [14, 15, 16, 17, 18].

\[
\lambda(t) = \lambda^*(t) + \nu(t)
\]

- It is common to assume that the excess hazards are piecewise constant and proportional (although there are better approaches). Such models can be estimated in the framework of generalised linear models using standard statistical software (e.g., SAS, Stata, R) [14].
- Non-proportional excess hazards are common but can be incorporated by introducing follow-up time by covariate interaction terms.

Overview of approaches to modelling excess mortality

- Poisson regression (piecewise exponential) [14].
- Poisson regression with fine splitting and modelling the baseline excess hazard using splines or fractional polynomials [19, 20, 21, 22].
- Flexible parametric models on the log cumulative hazard scale [23].
- Analogue to the Cox model where no assumptions are made about the baseline excess hazard [24, 25].
- Cure models [26, 27, 28, 29, 30].
Modelling excess mortality using Poisson regression

- The model can be written as

\[
\ln(\mu_j - d_j^*) = \ln(y_j) + x \beta,
\]

where \(\mu_j = E(d_j)\), \(d_j^*\) the expected number of deaths, and \(y_j\) person-time.

- This implies a generalised linear model with outcome \(d_j\), Poisson error structure, link \(\ln(\mu_j - d_j^*)\), and offset \(\ln(y_j)\).

- Such models have previously been described by Breslow and Day (1987) [31, pp. 173–176] and Berry (1983) [17].

- The usual regression diagnostics (residuals, influence statistics) and method for assessing model fit for generalised linear models can be utilised.

- Hakulinen and Tenkanen [32] and Estève et al. [15] describe alternative approaches to fitting similar/identical models.

References


