Population-based cancer survival analysis

Paul W. Dickman
Department of Medical Epidemiology and Biostatistics
Karolinska Institutet
Stockholm, Sweden
paul.dickman@mep.ki.se

22 June 2004
Biostatistical Centre, Institute of Biomedical Informatics,
Faculty of Medicine, Ljubljana, Slovenia

Slides available at http://www.pauldickman.com/teaching/

Overview

- Estimating cancer patient survival using data collected by population-based cancer registries
  - Net survival
  - Cause-specific survival
  - Relative survival
- Estimating survival using a period (as opposed to cohort) approach

Key concepts in population-based cancer survival analysis

- Patient survival can be measures in terms of mortality rates or survival proportions; the two are mathematically related.
- For cancer patients, we can study all-cause mortality (or observed survival) but mortality due to cancer is usually of greater interest.
- Net survival — the proportion of patients who would have survived t years or more following diagnosis in the hypothetical situation where the disease of interest were the only possible cause of death.
  - Net survival is a hypothetical quantity which can be estimated using, e.g., the cause-specific survival proportion or the relative survival ratio.
- Cause-specific survival is the analog of cause-specific mortality — only those deaths which can be attributed to the cancer in question are considered to be events, while all other deaths are considered censorings.

Disadvantages of cause-specific survival

- Using cause-specific survival to estimate net survival requires that reliably coded information on cause of death is available.
- Even when cause of death information is available to the cancer registry via death certificates, it is often vague and difficult to determine whether or not cancer is the primary cause of death.
- How do we classify, for example, deaths due to treatment complications or suicide?
- Consider a man diagnosed with prostate cancer and treated with estrogen who dies following a myocardial infarction. Do we classify this death as 'due entirely to prostate cancer' or 'due entirely to other causes'?
- Welch [1] studied deaths among surgically treated cancer patients that occurred within one month of diagnosis. They found that 41% of deaths were not attributed to the coded cancer.

Relative survival

- Can instead estimate excess mortality: the difference between observed (all-cause) and expected mortality.

$$\begin{array}{ccccc} \text{excess} & = & \text{observed} & - & \text{expected} \\ \text{mortality} & & \text{mortality} & & \text{mortality} \end{array}$$

- 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.
- It is usual to estimate the expected survival proportion from nationwide (or statewide) population life tables stratified by age, sex, calendar time, and, where applicable, race [2].
- Although these tables include the effect of deaths due to the cancer being studied, Ederer et al. [3] showed that this does not, in practice, affect the estimated survival proportions.

- A major advantage of relative survival (excess mortality) is that information on cause of death is not required, thereby circumventing problems with the inaccuracy [4] or nonavailability of death certificates.
- We obtain a measure of the excess mortality experienced by patients diagnosed with cancer, irrespective of whether the excess mortality is directly or indirectly attributable to the cancer.
- Deaths due to treatment complications or suicide are examples of deaths which may be considered indirectly attributable to cancer.

Cervical cancer diagnosed in New Zealand 1994 – 2001 Life table estimates of survival

Women diagnosed Jan 1994 - June 2001 with follow-up to June 2002

I	N	D	W	Effective number at risk	Interval-specific observed survival	Cumulative observed survival	Cumulative expected survival	Interval-specific relative survival	Cumulative relative survival
1	1559	209	0	1559.0	0.86594	0.86594	0.98996	0.87472	0.87472
2	1350	125	177	1261.5	0.90091	0.78014	0.98192	0.90829	0.79450
3	1048	58	172	962.0	0.93971	0.73310	0.97362	0.94772	0.75296
4	818	32	155	740.5	0.95679	0.70142	0.96574	0.96459	0.72630
5	631	23	148	557.0	0.95871	0.67246	0.95766	0.96679	0.70218
6	460	10	130	395.0	0.97468	0.65543	0.94972	0.98284	0.69013
7	320	5	129	255.5	0.98043	0.64261	0.94198	0.98848	0.68219
8	186	3	134	119.0	0.97479	0.62641	0.93312	0.98405	0.67130
9	49	1	48	25.0	0.96000	0.60135	0.91869	0.97508	0.65457

Cervical cancer diagnosed in New Zealand 1994 – 2001 Survival estimates stratified by age at diagnosis

		1-year			5-year		
Age	N	Obs.	Exp.	Relative	Obs.	Exp.	Relative
0–44	708	93.6	99.9	93.7	84.0	99.5	84.4
45–54	315	87.6	99.7	87.9	66.6	98.2	67.8
55–64	205	84.4	99.3	85.0	57.9	95.6	60.5
65–74	197	80.7	98.2	82.2	45.6	89.8	50.7
75+	134	59.0	93.3	63.2	28.4	68.9	41.2
All	1559	86.6	99.0	87.5	67.2	95.8	70.2

Issues with relative survival

- The central issue in estimating relative survival is defining a 'comparable group from the general population' and estimating expected survival.
- If not all of the excess mortality is due to the cancer then the relative survival ratio will underestimate net survival (overestimate excess mortality).
- For example, patients diagnosed with smoking-related cancers will experience excess mortality, compared to the general population, due to both the cancer and other smoking related conditions.
- Should the patients be a selected group from the general population, for example, with respect to social class, the national population might not be an appropriate comparison group.

Relative excess risk (RER)

- The concept of excess mortality and the use of relative excess risk as a measure of association are familiar to many epidemiologists.
- Consider a clinical trial where patients are randomised to receive either drug A or drug B, and followed up for one month to determine the proportion of each group who experienced a specific adverse event.
- If 8% of the patients on drug A experienced the event and 12% of the patients on drug B experienced the event, we would say that the relative risk (RR) of experiencing the event for drug B compared to drug A is $12/8 = 1.5$.
- We would conclude that the patients on drug B experienced the event 50% more often than the patients who received drug A.
- But is the relative risk the best measure of the relative merits of the two drugs?

- Consider the additional information that the adverse event rate in a control group, who received a placebo, was 6%.
- The excess risk is therefore 8-6=2% among those who received drug A and 12-6=6% among those who received drug B.
- The excess risk among the patients who received drug B was 3 times higher than the excess risk among the patients who received drug A.
- We say that the relative excess risk (RER) is 3.
- The analogous measure to cause-specific mortality would be if we were able to classify each of the adverse events as being 'due to the treatment' or 'not due to the treatment' and then compared the proportion of adverse events due to the treatment between the groups.

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.

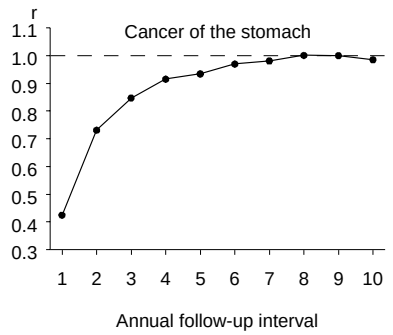


Figure 1: Plots of the annual (interval-specific) relative survival ratios (r) for males and females diagnosed with cancer of the stomach in Finland 1985–1994 and followed up to the end of 1995.

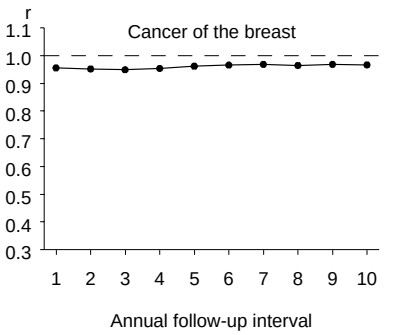


Figure 2: Plots of the annual (interval-specific) relative survival ratios (r) for females diagnosed with cancer of the breast in Finland 1985–1994 and followed up to the end of 1995.

- 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.
- I believe that plots of interval-specific survival are more informative than plots of cumulative survival.

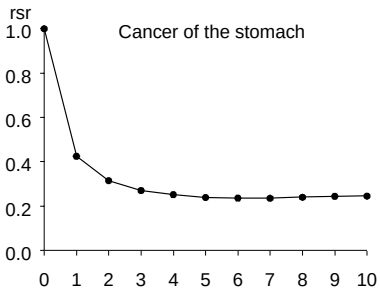


Figure 3: Plots of estimated cumulative relative survival ratios (rsr) for males and females diagnosed with cancer of the stomach in Finland 1985–1994 and followed up to the end of 1995.

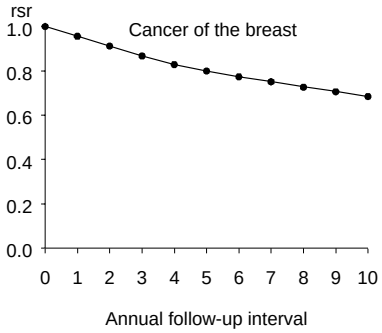


Figure 4: Plots of estimated cumulative relative survival ratios (rsr) for females diagnosed with cancer of the breast Finland 1985–1994 and followed up to the end of 1995.

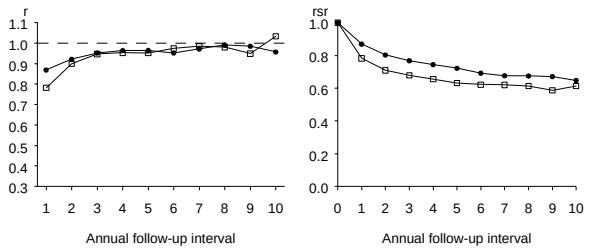


Figure 5: Plots of the interval-specific (r) and cumulative (rsr) relative survival ratios for cancer of the urinary bladder for males (•) and females (□) diagnosed in Finland during 1985–1994 and followed up to the end of 1995.

Estimation using a period approach

- The life tables typically used in cancer survival analysis are what demographers refer to a cohort life tables.
- In demography, a cohort life table is constructed by following all individuals born during one time period until all have died and keeping track of how many have died at different ages.
- A period life table uses information on individuals alive in each age class at one cross-section of time.
- For example, expectation of life at birth is calculated as if one person lived through all ages with age-specific survival probabilities of the current year.

- The period approach was first suggested for estimating cancer patient survival by Hermann Brenner in 1996 [5].
- For period estimation of cancer patient survival, the estimates of interval-specific survival are based only on those patients at risk during the period of interest.
- As such, a patient can contribute to the 5th life table interval without contributing to the first 4.

Cervical cancer diagnosed in New Zealand 1994 – 2001
Period estimates of survival for Jan. 2000 – Dec. 2001

Interval	N	D	W	Interval-specific relative survival	Cumulative relative survival	Patients Diagnosed
0.0 - 1.0	510	41	91	0.90347	0.90347	1999-2001
1.0 - 2.0	542	25	180	0.94052	0.84973	1998-2000
2.0 - 3.0	490	16	180	0.95929	0.81513	1997-1999
3.0 - 4.0	442	12	139	0.96896	0.78983	1996-1998
4.0 - 5.0	425	11	142	0.96980	0.76598	1995-1997
5.0 - 6.0	398	7	138	0.98151	0.75182	1994-1996
6.0 - 7.0	253	4	126	0.98665	0.74178	1993-1995
7.0 - 8.0	123	1	122	0.99388	0.73724	1992-1994

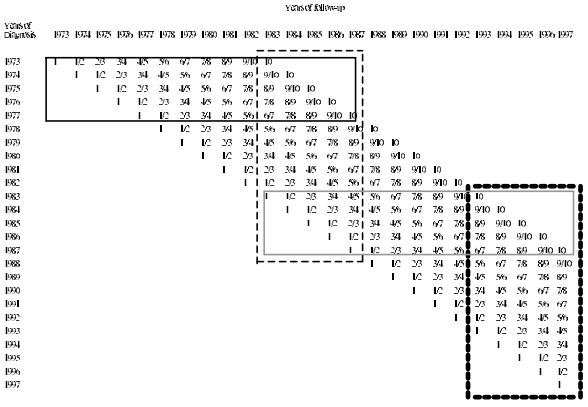
The approach was heavily criticised when first suggested

- A potential problem with the period method is that estimates may be too optimistic if, for example, survival improves during the first year following diagnosis but this is offset by a decrease in survival during subsequent years.

Table 1: Hypothetical data illustrating a potential problem with the period method. The table shows interval-specific survival estimates.

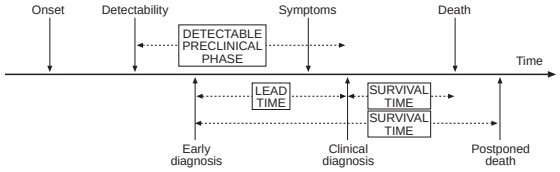
Interval	cohort 1991–1995 (fu to 2000)	cohort 1996–2000 (fu to 2005)	period 1996–2000 (fu to 2000)
1	0.54	0.73	0.73
2	0.82	0.77	0.78
3	0.88	0.85	0.87
4	0.97	0.89	0.95
5	0.98	0.95	0.98
Π	0.37	0.40	0.46

- Studies have shown that this is not a concern in practice [6, 7, 8, 9].
- Some people find it difficult to characterize the set of patients to whom the rates refer — demographic (index) thinking is required.
- The period approach detects trends in patient survival earlier than the cohort approach.
- Brenner and Hakulinen [8] demonstrated that, at least for Finnish data, period analysis is a useful tool for predicting long-term survival of cancer patients using the most up-to-date data available.
- Relative survival estimates made using the period approach can be modelled in the usual manner.



Population-based cancer patient survival

- In population-based survival analysis, survival time is measured from the date of diagnosis to the date of death.



- Diagnosis does not occur at the same point in the natural history for every patient, so care is required when interpreting survival estimates.

References

[1] Welch HG, Black WC. Are deaths within 1 month of cancer-directed surgery attributed to cancer? *J Natl Cancer Inst* 2002;**94**:1066–70.

[2] Berkson J, Gage RP. Calculation of survival rates for cancer. *Proceedings of Staff Meetings of the Mayo Clinic* 1950;**25**:270–286.

[3] Ederer F, Axtell LM, Cutler SJ. The relative survival rate: A statistical methodology. *National Cancer Institute Monograph* 1961;**6**:101–121.

[4] Percy CL, Stanek E, Gloeckler L. Accuracy of cancer death certificates and its effect on cancer mortality statistics. *American Journal of Public Health* 1981;**71**:242–250.

[5] Brenner H, Gefeller O. An alternative approach to monitoring cancer patient survival. *Cancer* 1996;**78**:2004–2010.

[6] Brenner H, Gefeller O, Stegmaier C, Ziegler H. More up-to-date monitoring of long-term survival rates by cancer registries: an empirical example. *Methods Inf Med* 2001;**40**:248–52.

[7] Brenner H, Hakulinen T. Long-term cancer patient survival achieved by the end of the 20th century: most up-to-date estimates from the nationwide Finnish cancer registry. *British Journal of Cancer* 2001;**85**:367–371.

[8] Brenner H, Hakulinen T. Up-to-date long-term survival curves of patients with cancer by period analysis. *Journal of Clinical Oncology* 2002;**20**:826–832.

[9] Brenner H, Gefeller O, Hakulinen T. Period analysis for 'up-to-date' cancer survival data: theory, empirical evaluation, computational realisation and applications. *European Journal of Cancer* 2004;**40**:326–35.