

Differences Between Brenner et al. and NCI Methods for Calculating Period Survival

Kathleen Cronin, Angela Mariotto, Steve Scoppa, Don Green, Lynn Clegg

The application of period survival to the SEER data led to several modifications to the methodology proposed for calculating period survival by Brenner et al.(1,2) These modifications were introduced to address data issues specific to SEER and what we believe to be improvement on the original algorithm. The differences between the Brenner et al. and NCI methods have a minor effect on the resulting period survival estimates. This report summarizes the difference between the two methods and presents a few examples demonstrating the impact on the estimates for period survival.

We use the following terminology in this report.

Current data year – the most recent year for which SEER data is available. This lags behind the current calendar year by 3 years. For example in calendar year 2003, the last year of data included in SEER reporting is 2000.

Time scales:

- Survival time – time since diagnosis for each patient
- Calendar time – calendar year

Data

NCI

Survival estimates produced by the SEER program do not include patients diagnosed in the current data year. For example, in calendar year 2003 and current data year 2000 no cases diagnosed in 2000 would be included in the survival estimate. However follow-up information on cases diagnosed before the year 2000 would be included in the survival estimate. To calculate 10-year period survival, data would be used for cases diagnosed in 1990 to 1999 with follow-up information through 2000. Exclusion of cases diagnosed in the current data year is based on a long-standing belief that there is incomplete mortality information for cases diagnosed in that last year of follow-up.

Brenner et al.

Cases diagnosed in the current data year (i.e. 2000) are included in the estimation of one-year survival.

Inclusion Criterion

NCI

To estimate the conditional probability of surviving x-years conditioned on surviving to the beginning of year x, we consider all cases diagnosed in

$$\text{calendar year} = \text{current data year} - x$$

and include only those cases that survive (x-1) years from diagnoses.

Note that although we consider all cases diagnosed within a particular calendar year, the inclusion criteria is based on survival time from diagnosis not calendar year. Figure 1 shows an example for the current data year of 2000, cases diagnosed in 1999 would be included in the estimate for 1-year survival. Similarly cases diagnosed in 1998 that survived until the beginning of their second year past diagnosis would be included in the estimate of 2-year survival.

Brenner et al.

This approach includes all cases that survive until the current data year. Therefore the inclusion criteria is based on calendar time rather than survival time. Each individual that dies during the current data year counts as a death in the interval

$x = (\text{current data year} - \text{diagnosis year})$ if the month of death is before the month of diagnosis and in the interval

$x = (\text{current data year} - \text{diagnosis year} + 1)$ if the month of death is after the month of diagnosis.

Figure 2 demonstrates cases diagnosed in 1999 that would be included/excluded in current data year 2000 period survival calculation. Note that in this case some deaths would count as a death in the 1-year survival interval while others would count in the 2-year survival interval.

Comparison of Approaches

Figures 1 and 2 demonstrate the differences between the two methods by classifying the same 6 patients under both methods. Patients 1 and 4 are included in the NCI estimate but excluded in the Brenner et al. estimate because they did not survive until the current data year (i.e. 2000). The inclusion method used in the Brenner et al. method could lead to differential exclusion of cases with shorter survival time, since cases diagnosed in 1999 with very short survival are less likely to survive until year 2000. Thus introducing a bias toward longer survival estimates. The NCI method includes all cases diagnosed in 1999 in calculating 1-year survival.

With the NCI method cases diagnosed in 1999 would only contribute to 1-year survival estimates in current data year 2000, partial year follow-up information is ignored. The Brenner et al. method utilized partial year follow-up information. For example patients 5 and 6 in Figure 2 survive into their second year after diagnosis and therefore contribute to the number at risk for 2-year survival. The Brenner et al. method takes advantage of this partial year follow-up to get estimates of survival that are slightly more up to date.

Finally, Brenner et al. uses patients diagnosed in 1999 and followed up through 2000 in both the 1-year and 2-year estimates. Patient 3 in Figure 2 is an event in calculating 1-year survival since

they survived less than 1 year. Patient 2 is counted as at risk in year 1 and an event in year 2 since they survived the first year and died in the second year. The NCI method would also count patient 3 as an event in year 1. However patient 2 would only count at risk for the first year and not as an event.

Calculation of Expected Survival

Let E_{ix} be the cumulative expected survival from diagnosis through the end of survival year x for individual i .

NCI

The calculation of expected survival for interval x is based on all individuals that had the potential to contribute to x -year survival. As described above, all individuals diagnosed in (current calendar year $- x$) can potentially contribute, although only individuals that survive to the beginning of the x^{th} -year after diagnosis actually contribute to the numerator of the relative survival calculation.

$$\text{Expected survival for interval } x = \sum_{i=1}^n E_{ix} / \sum_{i=1}^n E_{i(x-1)}$$

Where n represents all cases diagnosed in (current data year $-x$) whether or not they survived until the beginning of the x^{th} -year after diagnosis. This definition of interval expected survival has the property that for a given cohort of patients the product of interval expected survival equals the cumulative expected survival estimate for any year. The method is referred to as *backed out expected survival* because the interval survival is backed out from the cumulative expected survival estimates.

Brenner et al.

When calculating expected survival for year x , only individuals that survive to the current data year are included in the calculation.

$$\text{Expected survival for interval } x = \sum_{i=1}^n (E_{ix} / E_{i(x-1)})$$

Where there are n individuals included in the calculation of x -year interval relative survival.

Conclusion

The difference described above has little effect on the estimate of period survival. To demonstrate the magnitude of difference between the NCI and Brenner et al. approaches we

present a table of the two approaches applied to breast, colon and rectum, and prostate cancer. Table 1 shows 10 year period survival estimates using the NCI and Brenner et al. methods by historical stage for the cancer sites considered.

References

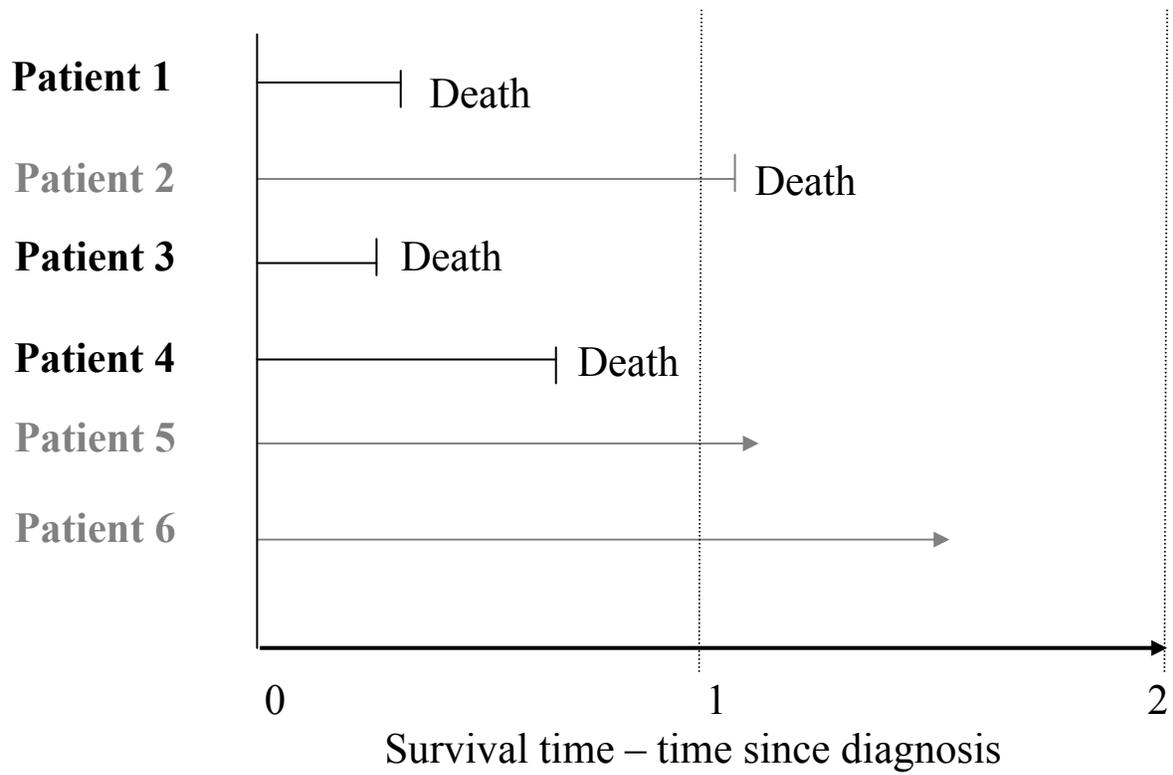
Brenner H, Gefeller O. An alternative approach to monitoring cancer patient survival. *Cancer* 1996;78:2004-10

Brenner H, Gefeller O. Deriving more up-to-date estimates of long-term patient survival. *J Clin Epidemiol* 1997;50:211-216

Table 1 . Comparison of NCI and Brenner et al. methods for calculating 10 year period survival using SEER 2003 submission for the SEER 9

Cancer site/stage	NCI Estimate	Brenner et al. Estimate
Female breast / all stages	0.805	0.806
Female breast / localized	0.936	0.934
Female breast / regional	0.672	0.668
Female breast / distant	0.124	0.122
Prostate / all stages	0.978	0.963
Prostate / local & regional	1.000	1.000
Prostate / distant	0.197	0.193
Colon and rectum / all stages	0.562	0.583
Colon and rectum / localized, male and female	0.810	0.831
Colon and rectum / regional, male and female	0.585	0.593
Colon and rectum / distant, male and female	0.070	0.083

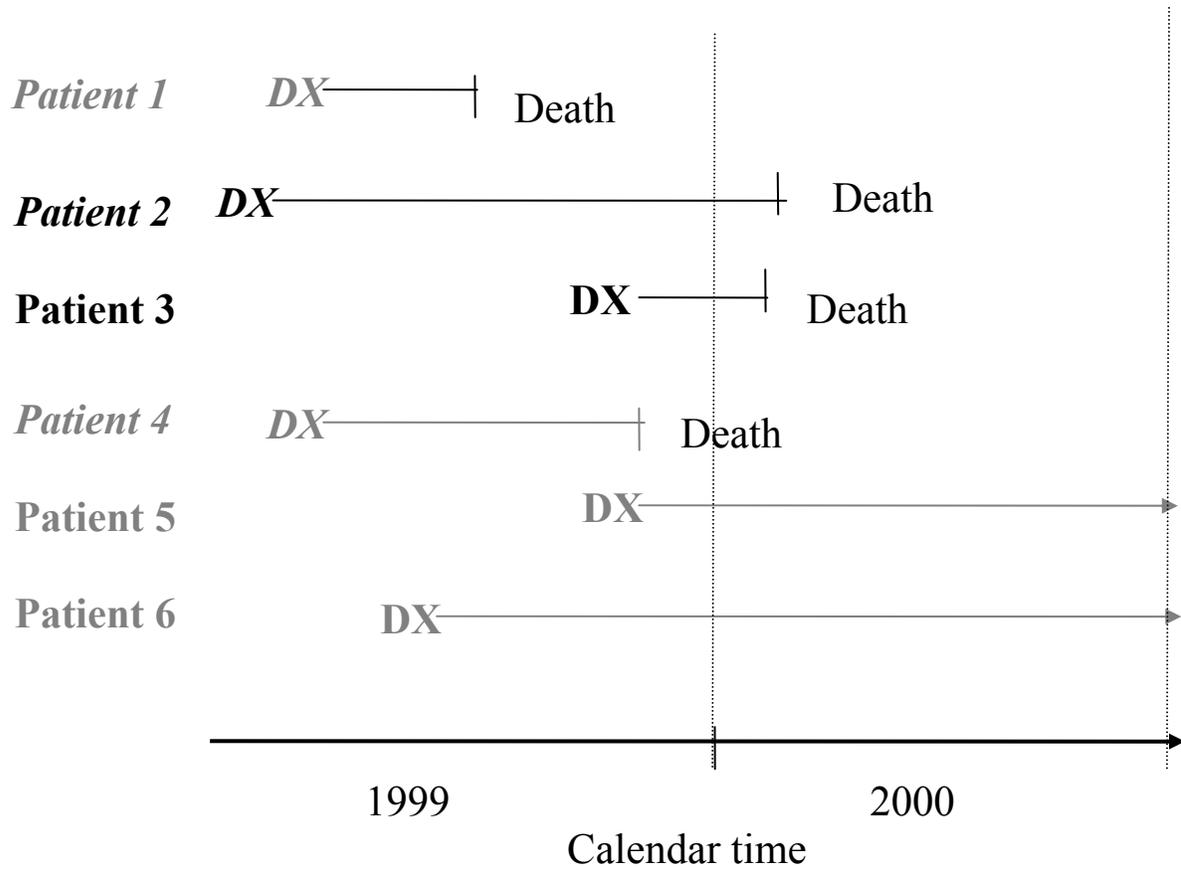
Figure 1. NCI Period Survival For Diagnosis Year 1999 and Data Year 2000



Included as event – 1 year survival

Included at risk

Figure 2. Brenner et al. Period Survival For Diagnosis Year 1999 and Data Year 2000



Excluded

Included as event – 1 year survival

Included as event – 2 year survival

Included at risk