Interpretation: This measure represents the average survival for the years included, with more weight on the earliest years for estimating longer survival. It is most appropriate when *k* is relatively short, such as 5 or 10 years. As the survival length increases, survival estimates are increasingly affected by patients diagnosed further in the past.

Modified Period Survival

The period method as proposed by Brenner and Gefeller [1,2] has been modified slightly for application to SEER data. The differences between the method originally proposed and the one being implemented in this report is described in detail in [3]. The modifications to the original method lead to a small difference in the estimate of period survival.

Inclusion in the analysis: Cases diagnosed in the (*current data year* - *x*) that survive (*x*-1) years after diagnosis are included in the estimation of *x*-year interval survival. Conditional interval survival estimates (conditioned on surviving to the beginning of the interval) are then multiplied together to obtain cumulative period survival. For example, when the current data year is 2000, all cases diagnosed in 1999 are included in the estimation of 1-year survival, cases diagnosed in 1998 that survived at least until the beginning of their second year after diagnosis are included in the estimate of conditional 2-year survival, and so on.
Calculation of survival: The appropriate conditional overall survival and expected survival are calculated for each year of diagnosis. The cumulative estimate is comprised of interval estimates from different cohorts.
Interpretation: Gives more current estimates of long-term survival then the complete method by estimating each *x*-year conditional survival using only the most recent diagnosis year with follow-up information for x years.

Modeling of Survival

The modeling approach is based on the method presented in Haukulinen and Tenkanen [4]. Let S_{iy} be the conditional i^{th} year interval survival for cases diagnosed in calendar year y. The basic regression model is

$$\log(-\log[S_{iy}]) = \alpha_i + \beta y.$$

This formulation is a Cox proportional hazards model on the hazard of dying of cancer with $\exp(\alpha_i)$ representing the effect of surviving the i^{th} year after diagnosis conditioned on surviving (i+1) after diagnosis and $\exp(\beta y)$ represents the multiplicative effect of diagnosis year on the hazard of dying of cancer. Note that this model assumes that $\log(-\log[S_{iy}])$ changes linearly over time. One possible extension of the basic model is the addition of one or more join points in the time trend to allow for changes in the linear trend. For example one join point model at time Y is written as

$$\log(-\log[S_{iy}]) = \alpha_i + \beta_1 y + \beta_2 (y - Y)^+,$$

where $(y-Y)^+$ is 0 if y □Y and (y-Y) if y>Y. Generally, a Y values that provides the best fit to the data is selected.

Inclusion in the analysis: Cases diagnosed in a specified number of years before the current data year.
Calculation of survival: Model parameter estimates are used to project survival for an individual diagnosed in the current calendar year.
Interpretation: Modeling allows for the extrapolation of observed rates to predict what survival will look like in the future.

The model used to obtain a survival estimate can vary in a number of ways including the type of parameterization, the method of projection and the number of diagnosis years used in the fitting the model. The chapter reports results from several different model variations.

To estimate 5-year survival, 5 and 10 years before the current calendar year are used, i.e. for current data year 2000, cases diagnosed between 1995-1999 and between 1990-1999. To estimate 10-year survival, 10 years of data are used. Model results are projected two ways to estimate survival for cases diagnosed in the current calendar year: a flat projection that extrapolates the fitted trend to the current data year 2000, and a trend projection that extrapolated the fitted trend to the current calendar year 2003. When interpreting the survival estimate as the survival for cases diagnosed in the current calendar year, the flat projection method assumes that the trend in survival is flat from the current data year to the current calendar year (2000-2003). The flat projection provides a more conservative projection than one where we continue to model fitted trends through 2003. Tables 1 and 2 list the types of models used to calculate 5 and 10-year relative survival in this chapter.

COMPARISON OF METHODS

A number of validation studies were performed to compare each method's ability to predict survival for patients diagnosed in the current calendar year. This chapter reports the results for female breast cancer as an example. A cohort estimate of survival was estimated for cases diagnosed in each year between 1983 and 1995. This cohort estimate is what we call "observed" for each calendar year. The validation study uses only data that would be available in the corresponding calendar year to estimate survival using the cohort, complete, period and modeling methods. For example in the calendar year 1995 a survival analysis would include cases diagnosed in 1991 and earlier with follow-up information through 1992. Using only that data that would have been available in a particular calendar year, 5-year survival was estimated for all the methods presented in Table 1. Figure 1 shows how well each method would have predicted current calendar year survival for 1983 to 1995.

Modeling option 1, include cases diagnosed in the 5 years prior to the current data year with a flat projection from the current data year to the current calendar year, best predicts the survival experiences of cases diagnosed in the current calendar year without overestimating future improvements in survival. The cohort method, which estimates 5-year survival from patients diagnosed 5 years before the current data year and 8 years

before the current calendar year, does the worst at predicting future survival. The complete and period methods fall below the modeling, with the period method consistently giving a better prediction than the complete method.

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;50:211-16.

Cronin KA, Mariotto A, Scoppa S, Green D, Clegg L. Differences between Brenner et al. and NCI methods for calculating period survival. SRAB Tech Report # 2003-02. (http://srab.cancer.gov/reports/)

Haukulinen T, Tenkanen L. Regression analysis of relative survival rates. *Appl Statist* 1987;36:309-17.

TABLE 1. CURRENT ESTIMATE OF 5-YEAR SURVIVAL BY CANCER SITE AND STAGE OF DISEASE*

	Cohort	Complete	Period	Modeling**		
				Opt 1	Opt 2	Opt 3
All sites (male and female)	62.7%	63.5%	64.9%	65.5%	64.7%	66.6%
All sites (male)	61.8%	63.2%	64.8%	65.3%	64.4%	66.6%
All sites (female)	63.6%	63.9%	64.9%	65.0%	65.0%	66.3%
Colon and Rectum (male and female)						
All sites	61.0%	62.8%	64.7%	66.0%	63.4%	64.2%
Localized	88.7%	89.7%	90.1%	89.9%	89.0%	88.2%
Regional	65.2%	66.6%	67.2%	69.4%	66.7%	67.3%
Distant	8.7%	9.8%	11.6%	12.3%	9.7%	10.3%
Breast (female)						
All sites	86.6%	87.1%	88.1%	89.0%	88.3%	89.6%
Localized	97.2%	97.2%	97.3%	97.2%	97.3%	97.8%
Regional	78.2%	79.3%	80.5%	83.1%	81.7%	83.8%
Distant	22.2%	24.8%	26.9%	29.6%	26.2%	28.5%
Ovary (female)						
All sites	53.4%	53.4%	53.3%	55.1%	56.4%	59.1%
Localized	95.0%	95.1%	95.6%	96.3%	95.6%	95.8%
Regional	76.5%	74.0%	71.3%	81.4%	82.8%	84.4%
Distant	30.9%	31.6%	31.9%	34.1%	35.8%	39.9%
Prostate (male)						
All sites	97.8%	98.5%	99.6%	99.8%	99.7%	99.9%
Localized/Regional	100.%	100.%	100.%	100.%	100.%	100.%
Distant	34.3%	31.1%	32.5%	30.5%	27.8%	23.8%

***Data Source:** SEER 9, Nov. 2002 submission with diagnosis years through 1999 and follow-up through 2000.

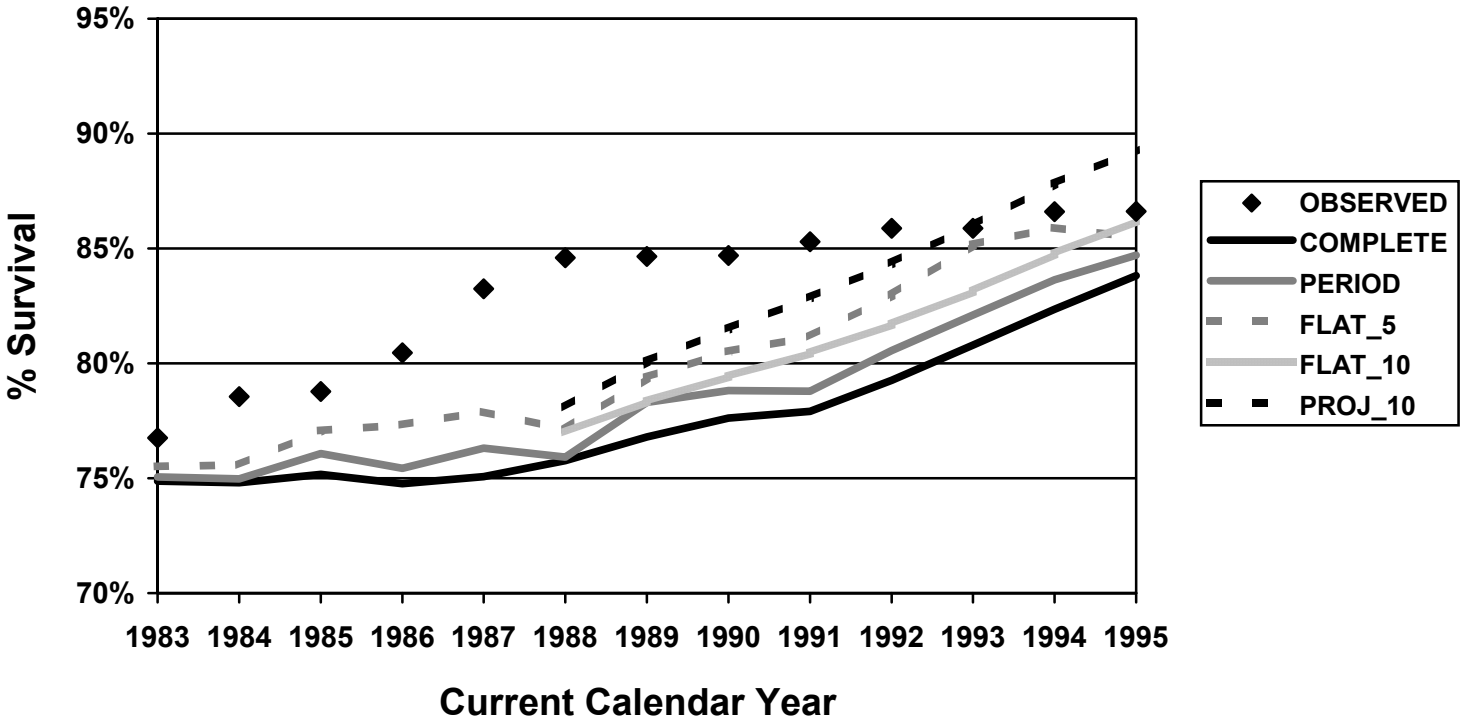
****Modeling Alternatives**

Opt 1. Include cases diagnosed in the 5 years prior to the current data year with a flat projection from 2000 to 2003.

Opt 2. Include cases diagnosed in the 10 years prior to the current data year with a flat projection from 2000 to 2003.

Opt 3 Include cases diagnosed in the 10 years prior to the current data year with the trend projected to 2003.

FIGURE 1. COMPARISON OF THE ABILITY OF SIX METHODS TO PREDICT SURVIVAL IN THE CURRENT CALENDAR YEAR: FEMALE BREAST CANCER*



***Data Source:** SEER 9, Nov. 2002 submission with diagnosis years through 1999 and follow-up through 2000.