Current Estimates of the Cure Fraction: A Feasibility Study of Statistical Cure for Breast and Colorectal Cancer

Margaret R. Stedman, Eric J. Feuer, Angela B. Mariotto

Correspondence to: Margaret R. Stedman, PhD, MPH, National Cancer Institute, Division of Cancer Control and Population Sciences, Surveillance Research Program, 9609 Medical Center Drive, Bethesda, MD 20892-9765 (e-mail: margaret.stedman@nih.gov).

Background

The probability of cure is a long-term prognostic measure of cancer survival. Estimates of the cure fraction, the proportion of patients “cured” of the disease, are based on extrapolating survival models beyond the range of data. The objective of this work is to evaluate the sensitivity of cure fraction estimates to model choice and study design.

Methods

Data were obtained from the Surveillance, Epidemiology, and End Results (SEER)-9 registries to construct a cohort of breast and colorectal cancer patients diagnosed from 1975 to 1985. In a sensitivity analysis, cure fraction estimates are compared from different study designs with short- and long-term follow-up. Methods tested include: cause-specific and relative survival, parametric mixture, and flexible models. In a separate analysis, estimates are projected for 2008 diagnoses using study designs including the full cohort (1975–2008 diagnoses) and restricted to recent diagnoses (1998–2008) with follow-up to 2009.

Results

We show that flexible models often provide higher estimates of the cure fraction compared to parametric mixture models. Log normal models generate lower estimates than Weibull parametric models. In general, 12 years is enough follow-up time to estimate the cure fraction for regional and distant stage colorectal cancer but not for breast cancer. 2008 colorectal cure projections show a 15% increase in the cure fraction since 1985.

Discussion

Estimates of the cure fraction are model and study design dependent. It is best to compare results from multiple models and examine model fit to determine the reliability of the estimate. Early-stage cancers are sensitive to survival type and follow-up time because of their longer survival. More flexible models are susceptible to slight fluctuations in the shape of the survival curve which can influence the stability of the estimate; however, stability may be improved by lengthening follow-up and restricting the cohort to reduce heterogeneity in the data.

J Natl Cancer Inst Monogr 2014;49:244–254

Cure is difficult to identify at the individual level because for certain cancers, there can be recurrences many years after periods of being symptom free. However, cure can be identified at the population level as the fraction of cancer patients (cure fraction), who have an observed mortality similar to the general population after a long follow-up period. This concept of cure, known as statistical cure, represents a long-term prognostic measure of the chances of being cured, and it is a desirable statistic for the patient and clinician.

Several models exist to estimate statistical cure [eg, mixture (1), nonmixture (2), and flexible (3) models] and all of these models rely on the length of follow-up to attain a stable estimate of statistical cure (4). Although cancer registries may have long histories of incidence and follow-up data, recent advances in treatment, screening, and changes in staging criteria impact survival and complicate the estimation of statistical cure. Furthermore, there is interest in predicting the cure fraction for patients recently diagnosed with minimal follow-up information.

There has been some research to determine the minimal length of follow-up needed to obtain a stable cure estimate. Yu and colleagues (5) found that length of follow-up time needs to be at least two thirds the median survival time in the uncured patients to attain a stable cure estimate. Tai et al. (6) empirically derived the minimum follow-up time for 49 different cancer sites using median survival time and found that the minimal length of follow-up is site specific. For example, colon cancer was estimated to require a minimum of 12.2 years and breast cancer, a minimum of 36.2 years (6). However, median survival depends on the shape of the survival curve and the type of outcome measured (relative or cause-specific survival) and none of these studies has considered the sensitivity of the cure fraction to follow-up time while varying model assumptions.

In this paper, we use the Surveillance Epidemiology and End Results (SEER) registry data to perform a sensitivity analysis of the cure fraction estimate to model choice, survival outcome, and study design. We compare the impact of follow-up time on parametric mixture and flexible parametric cure models and on relative and cause-specific measures of cancer survival. Using the most recent SEER data, we project the cure fraction by stage and site for breast and colorectal cancer patients diagnosed in 2008.
First, we describe the data and methods used to perform a sensitivity analysis of the cure fraction with respect to follow-up time, model, and survival type using SEER breast and colorectal cancer data. Next, we provide results from the sensitivity analysis and future projections of the cure fraction for patients diagnosed in 2008. Finally, we conclude with some guidance on projecting the cure fraction using these models.

**Data and Methods**

**Data and Cohort Definitions**

We used the SEER site recode variables to identify patients diagnosed at ages 45–74 with colorectal cancer (SEER Site Recode International Classification of Diseases for Oncology [ICD-O]-3 definition: C180-C189, C260, C199, and C209) and at ages 45–64 with female breast cancer (SEER Site Recode ICD-O-3 definition: C500-C509) within the SEER-9 Registries. Cancers of the colon and rectum were combined because they were assumed to have similar cure fractions (7). Colorectal cancer has less incident cases than breast cancer, so a broader age range was included to increase the size of the colorectal cohort. We selected malignant cases and excluded patients diagnosed at death and secondary tumors. All data were obtained from the November 2011 submission, using SEER*Stat software, version 8.02, available at our website: http://seer.cancer.gov/seerstat.

We constructed four cohorts of breast cancer and colorectal cancer patients. The first two cohorts included female breast and colorectal cancer patients diagnosed between 1975 and 1985 and two different follow-up periods: 12 years (follow-up through December 1986) and 35 years (follow-up through December 2009). Presuming that having a longer follow-up time would optimize accuracy of the cure estimate, data from the 12- and 35-year observation periods are referred to as the “test” and “optimal” cohorts, respectively. To obtain estimates of the cure fraction that reflect the most recently diagnosed patients (diagnosed in 2008), we constructed two other cohorts: patients diagnosed with female breast and colorectal cancer from 1975 to 2008 (“full cohort”) and patients diagnosed with breast and colorectal cancer from 1998 to 2008 (“recent cohort”). All patients were followed from diagnosis until December 2009. Since entrance into the registry and length of follow-up depend upon the date of diagnosis (longer follow-up for the earlier diagnoses), there is confounding between the length of follow-up and the diagnosis year. To reduce confounding and also project the cure fraction for patients diagnosed in the most recent years, we adjusted for diagnosis year in all analyses.

**Relative and Cause-Specific Survival**

Cancer survival may be estimated either as relative survival (the ratio of all-cause survival and expected survival) or cancer-specific survival [based on cause of death from the death certificate (8)]. When examining a plot of cancer survival, the cure fraction is the point in the survival curve where the curve begins to level off and the slope is predicted to be zero. Sometimes, this occurs well beyond the observed data (see Figure 1). From a relative survival perspective, cure occurs when the observed survival of the cancer population matches the expected survival of the general population. From a cause-specific perspective, cure occurs when the cancer population is no longer dying from cancer and the hazard of dying of cancer is estimated to be zero. This concept of cure measured from cause-specific survival is sometimes referred to as “clinical cure” (9).

We obtained relative survival [by the Ederer II (10) method] and cause-specific survival estimates stratified by diagnosis year and SEER historic stage (localized, regional, distant, and all stages combined). Cause of death is based on the SEER cause-specific death classification, which classifies deaths related to the cancer diagnosis using patient death certificate information, tumor location, and comorbidities (http://seer.cancer.gov/causespecific/) (8). Relative survival can give biased estimates for localized stage breast cancer because expected survival estimated from the US population life tables underestimate all-cause mortality in this subgroup (8).

**Statistical Cure: Parametric Mixture Cure Model**

Although there are several models available to estimate cure, this study focuses on the parametric mixture cure model and the flexible parametric model, both of which have been adapted to population-based cancer survival analysis. The parametric mixture cure model predicts cancer survival, \( S(t) \), as a mixture of the cure fraction, \( \epsilon \), and the uncured fraction, \( 1 - \epsilon \):

\[
S(t) = \epsilon + (1 - \epsilon) G(t; \xi, \theta)
\]

The uncured fraction is assumed to follow a survival function, \( G(t; \xi, \theta) \), where the distribution can be modeled by either semiparametric or fully parametric distributions with parameters \( \xi \) and \( \theta \), including the Weibull, log normal, log logistic, and Gompertz distributions. The median survival time for the uncured can be derived from \( \xi \) and \( \theta \) (see Table 1). For this analysis, we examine the more commonly used log normal and Weibull survival distributions. The interpretation of the parameters, \( \xi \) and \( \theta \), depend on the survival function, \( G(t) \), see Table 1. For the log normal distribution, the parameters \( \xi \) and \( \theta \) equal the mean and SD of the survival function, respectively. For the Weibull distribution, \( \xi \) and \( \theta \) are parameterizations of the shape, \( \lambda \), and scale, \( \rho \) (1).
The parameters, \( c, \xi, \) and \( 0, \) may depend on covariates, \( X, \) such as diagnosis year. Covariates are introduced into the models with the parameters: \( \beta_1, \beta_2, \) and \( \beta_3. \)

\[
\xi = [1 + \exp(-\beta_0 - \beta_1 X)]^{-1}
\]

\[
\theta = \exp(\beta_{00} + \beta_{01} X)
\]

Accuracy of the cure estimate is improved by adjusting for covariates but it is also influenced by choice of parametric distribution and overall fit to the data. As in Huang et al. (11), we allowed the cure fraction, \( c, \) and the median survival in the uncured, \( \xi, \) to change simultaneously as a function of diagnosis year, however additional covariates could be included in the model. We estimated the parametric mixture cure model by maximum likelihood estimation methods using CanSurv software (12) available at http://surveillance.cancer.gov/cansurv/. CanSurv software reparameterizes the survival function to simplify computation of the parameters (1) (see Table 1 for results for some common survival functions). From the mixture cure models, the cure fraction estimate for patients diagnosed in a given year, \( Y, \) can be estimated as

\[
c = [1 + \exp(-\beta_0 - \beta_1 Y)]^{-1}
\]

Relative and cause-specific survival estimates were imported to CanSurv software to estimate cure from the mixture cure model. The models were stratified by cancer site and historic stage and adjusted for diagnosis year (11), a continuous covariate, so that we could obtain site-, stage-, and year-specific estimates. We estimated median survival of the uncured assuming either a Weibull or log normal survival function. Parameter estimates were exported to SAS software to calculate projections of the cure fraction and respective confidence intervals (CIs) for patients diagnosed in a given year (see equation 1 for the formula for the cure fraction). The variance of the cure fraction was estimated by the Delta Method (see Appendix A).

### Statistical Cure: Flexible Parametric Cure Model

Flexible parametric cure models predict the cure fraction by modeling the log cumulative excess hazard. These models were recently developed by Andersson et al. for relative survival (3). Excess mortality is the difference between the hazard functions for all-cause mortality and the expected mortality of the reference population: \( b(t) - b^*(t) = \lambda(t). \) Integrating these hazards gives more stable estimates of the cumulative mortality, \( H(t), \) the cumulative expected mortality, \( H^*(t), \) and the cumulative excess hazard, \( \Lambda(t) \) at time \( t, \) such that \( \Lambda(t) = H(t) - H^*(t). \)

The flexible parametric survival model predicts the log cumulative excess hazard, \( \ln(\Lambda(t)), \) using restricted cubic splines, \( \{\ln(t); \gamma_i\}, \) of the log survival time, \( t, \) where \( \gamma_i \) contains the parameters of the spline function:

\[
\ln(\Lambda(t)) = \{\ln(t); \gamma_i\}
\]

This is adapted to predict relative survival, \( R(t), \)

\[
\ln[-\ln(R(t))] = \{\ln(t); \gamma_{i0}\}
\]

where \( \ln(\Lambda(t)) = \ln(-\ln(R(t))). \) Cure is determined when the cumulative excess hazard rate levels off, i.e., when there is no difference between the mortality rate of the cancer population and the general population and excess mortality is zero. In flexible parametric survival models, this can be accomplished by constraining the log cumulative excess hazard function to have a zero slope after a certain point in time.

This is implemented by specifying the \( K \) knots from the spline function in reverse order \( \{k_1, \ldots, k_K\} \) where the last spline parameter is restricted to be zero \( \{\gamma_{i0} = 0\} \) (3). Thus, the location of the last knot determines when cure occurs. As recommended by Andersson et al. (3), we set the last knot at the last observed time point. Relative survival is estimated as:

\[
R(t) = \exp[-\exp(\gamma_{00} + \gamma_{01} v_1(\ln(t)) + \gamma_{02} v_2(\ln(t)) + \gamma_{03} v_3(\ln(t)) + \ldots + \gamma_{0k} v_k(\ln(t)))]
\]

In the backwards spline function, \( \gamma_{01}, \ldots, \gamma_{0k-1} \) are the parameters and \( v_1(\ln(t)), \ldots, v_k(\ln(t)) \) are the basis functions defined as:

\[
v_j(\ln(t)) = \left(\frac{k_j - \ln(t)}{k_k - k_j}\right)^+, \quad \lambda_j = \frac{k_{j+1} - k_j}{k_k - k_j}, \quad \lambda_k = 1,
\]

where

\[
\epsilon = \exp[-\exp(\gamma_{00})]
\]
To adjust for a covariate such as diagnosis year (Y) in the model, we added the covariate, Y, as follows:

\[
R(t) = \exp[-\exp(\gamma_0 + \beta(1 + Y)) + \gamma_0 \psi_1(\ln(t)) + \gamma_0 \psi_2(\ln(t)) + \gamma_0 \psi_3(\ln(t)) + \ldots + \gamma_0 \psi_k(\ln(t))]
\]

so that cure may be predicted for any given diagnosis year, Y:

\[
\epsilon = \exp[-\exp(\gamma_0 + \beta(1 + Y))]
\]

The variance of the cure fraction is approximated by the Delta Method (see Appendix B). See Appendix C for a formula for the median survival in the uncured. Time varying effects may also be added to more complex models (3).

The flexible cure model can be estimated using the stpm2 package (13) available in STATA v12. The package is designed to model relative survival. Although it is possible to flexibly model cause-specific survival (14), software for this method has not been fully tested and will not be included in this analysis. Accuracy of the flexible model depends on the placement of the knots for the splines and length of follow-up of the available data (3).

STATA was used to estimate the median survival for the uncured, the cure fraction, and CIs for the flexible parametric models. As explained by Lambert and Royston (13), we applied the stpm2 package to each cohort (test, optimal, full, recent). Expected monthly survival probabilities from the SEER life tables were exported to STATA, converted to yearly mortality rates, and included as the baseline hazard, \(H(t) = -\ln[S(t)^{12}]\) in the flexible model.

### Sensitivity Analyses to Length of Follow-up, Model, and Survival Type

We investigated the sensitivity of cure fraction estimates to length of follow-up by comparing 1985 cure fraction estimates from the 12-year test cohort to those from the 35-year optimal cohort. We compared results from the cause-specific and relative survival outcomes for each of the models. Additionally, all results were evaluated for fit to their respective observed cancer survival curve. By definition, the cure fraction estimate should be at or below the tail of the observed survival curve (5). Models with 1985 cure estimates significantly above the 25-year observed survival estimates were considered poor fits to the data.

### Sensitivity of 2008 Projected Cure Fraction to Study Design

Since the cure fraction is a function of diagnosis year in the parametric mixture and flexible parametric models, the predicted cure fraction was extrapolated to 2008 by setting \(Y = 2008\) in equations 1 and 3. Two different study designs were tested. In one design, patients were selected from diagnosis years 1998–2008 (recent cohorts). However, in these cohorts, the lengths of follow-up may not be long enough to provide stable estimates of the cure fraction. So, in a second design, the number of diagnosis years was expanded to include 1975–2008 (full cohorts). While the full cohort provides sufficient follow-up, a linear trend in diagnosis year over this long interval may not be appropriate. 2008 projections were compared from both study designs for cause-specific and relative survival parametric mixture and flexible models.

### Results

#### Sensitivity to Length of Follow-up: Comparison of 1985 Estimates From the Test and Optimal Cohorts

The 1985 cure fraction for colorectal cancer ranged between 55% and 78% for localized stage, 41% and 48% for regional stage, 3% for distant stage, and 42% and 50% for all stages combined (Table 2, optimal cohort). With the exception of localized stage, cure estimates are fairly stable from the optimal cohort and similar to estimates obtained using only 12 years of data (test cohort). For localized stage, there is more variation across models and the estimates have wider CIs than other stages. For regional stage, there is less variation across models and tighter confidence intervals; however, the Weibull and flexible models from the test cohort tend to slightly overestimate the cure fraction compared to the 25-year observed cancer survival (Weibull = 48%, Flexible = 49%, 25-year relative survival = 44%). For distant stage disease, two of the models did not converge. The cure fraction was very small, slightly above zero.

The 1985 breast cancer cure fraction estimates showed more variability and ranged between 57% and 67% (cause-specific) for localized stage, 35% and 43% for regional stage, 3% and 6% for distant stage, and 43% and 57% for all stages combined (Table 2, optimal cohort). Test cohort estimates for breast were less reliable and less precise than the colorectal estimates. In general, test estimates exceeded the optimal estimates, indicating that 12 years of data is not enough to obtain an accurate prediction of the cure fraction. Weibull and flexible estimates tended to be higher than lognormal estimates and all cases of the flexible model for the test cohort exceeded the 25-year cancer survival estimate. Distant stage results were similarly influenced by small numbers of at risk patients at the end of the survival curve (fewer than 10 survivors 25 years postdiagnosis). Combining all stages decreased the stability of the estimate. Although most models converged, test estimates fluctuated widely between 23% and 64%. Relative survival estimates were higher than cause-specific estimates for the optimal cohort and these differences were most apparent for localized stage.

#### 2008 Projections of the Cure Fraction: Sensitivity to Diagnosis Year and Length of Follow-up

To illustrate modeling the cure fraction as a function of calendar time, we plotted the cure fraction and median survival time estimates by diagnosis year from each survival model using the test cohort of patients diagnosed with regional colorectal cancer in 1975–1985 and followed through 1986. For the parametric mixture models, there is compensation between the cure fraction estimates and median survival times, where models with higher cure fractions have lower median survival times followed through 1986. For the parametric mixture models, there is compensation between the cure fraction estimates and median survival times, where models with higher cure fractions have lower median survival times followed through 1986. For the parametric mixture models, there is compensation between the cure fraction estimates and median survival times, where models with higher cure fractions have lower median survival times followed through 1986. For the parametric mixture models, there is compensation between the cure fraction estimates and median survival times, where models with higher cure fractions have lower median survival times followed through 1986.
Table 2. 1985 cure fraction (95% confidence interval) projections by stage, years follow-up, cancer survival type, and survival function for colorectal and breast cancer patients, diagnosed between 1975 and 1985 and followed through 1986 (test cohort) and 2009 (optimal cohort)*

<table>
<thead>
<tr>
<th>Site</th>
<th>Stage</th>
<th>25-year cancer survival</th>
<th>Log normal</th>
<th>Weibull</th>
<th>Relative survival</th>
<th>Log normal</th>
<th>Weibull</th>
<th>Flexible</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Test cohort</td>
<td>Optimal cohort</td>
<td>Test cohort</td>
<td>Optimal cohort</td>
<td>Test cohort</td>
<td>Optimal cohort</td>
<td>Test cohort</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Colorectal cancer</td>
<td>L</td>
<td>72% (70%, 74%)</td>
<td>53%</td>
<td>55%</td>
<td>77%</td>
<td>65%</td>
<td>77%</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>43% (41%, 45%)</td>
<td>44%</td>
<td>41%</td>
<td>50%</td>
<td>42%</td>
<td>44%</td>
<td>41%</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>3% (2%, 5%)</td>
<td>3%</td>
<td>3%</td>
<td>3%</td>
<td>3%</td>
<td>2%</td>
<td>3%</td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>47% (45%, 48%)</td>
<td>44%</td>
<td>42%</td>
<td>49%</td>
<td>44%</td>
<td>48%</td>
<td>41%</td>
</tr>
<tr>
<td>Breast cancer</td>
<td>L</td>
<td>74% (72%, 76%)</td>
<td>66%</td>
<td>57%</td>
<td>—</td>
<td>67%</td>
<td>80%</td>
<td>72%</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>42% (39%, 44%)</td>
<td>40%</td>
<td>35%</td>
<td>62%</td>
<td>40%</td>
<td>41%</td>
<td>35%</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>7% (4%, 10%)</td>
<td>0%</td>
<td>4%</td>
<td>4%</td>
<td>5%</td>
<td>5%</td>
<td>0%</td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>56% (55%, 58%)</td>
<td>31%</td>
<td>43%</td>
<td>61%</td>
<td>51%</td>
<td>58%</td>
<td>23%</td>
</tr>
</tbody>
</table>

* Plotted values are in bold print. D = distant; L = local; R = regional.
† Cure model did not fully converge.
‡ Cure estimate is significantly above net survival (one-sided Z-test, \(P < .05\)).
§ 25-year net survival estimate is based on less than 25 years of data.
survival and a steep trajectory for the cure fraction. The optimal and recent cohorts are both subsets of the full cohort, however modeling the full cohort does not match results from the more restricted cohorts (test and recent). These plots demonstrate how cure, median survival, and diagnosis year are interconnected, however these results are limited to a few models of regional stage colorectal cancer.

We then investigated results from the full and recent cohorts by model and study design. The 2008 cure projections for colorectal cancer ranged 75%–92% for localized stage, 54%–72% for regional stage, 4%–15% for distant stage, and 53%–67% for all stages combined (Table 3, recent cohort). For breast cancer, the 2008 cure projections were 90%–96% (cause-specific) for localized stage, 65%–96% for regional stage, 1%–26% for distant stage, and 66%–88% for all stages combined based on results from the recent cohort. There is much more variability in the 2008 cure fraction estimates compared to the 1985 cure fraction estimates. Cure fraction estimates from the full cohort tended to be higher than estimates from the recent cohort, as displayed in Figure 3. Cohort differences were greatest for the lognormal models. There was more stability in the full cohort cure fraction estimates across models and more precise confidence intervals, but this is more likely due to the sample sizes than accuracy of the estimates. The flexible model did not converge for most cases of the full cohort. When all stages were combined, stability improved across models.

**Discussion**

Cure fractions are difficult to estimate because they imply extrapolations beyond the observed survival time. They vary depending on cancer site, stage, survival outcome type, length of follow-up, and
Table 3. 2008 cure (95% confidence interval) projections by stage, years follow-up, cancer survival type, and survival function for colorectal and breast cancer patients diagnosed between 1998 and 2008 (recent cohort) or 1975–2008 (full cohort) and followed up through 2009

<table>
<thead>
<tr>
<th>Site</th>
<th>Stage</th>
<th>Log normal Recent cohort</th>
<th>Log normal Full cohort</th>
<th>Weibull Recent cohort</th>
<th>Weibull Full cohort</th>
<th>Relative survival Recent cohort</th>
<th>Relative survival Full cohort</th>
<th>Flexible Recent cohort</th>
<th>Flexible Full cohort</th>
</tr>
</thead>
<tbody>
<tr>
<td>Colorectal cancer</td>
<td>L</td>
<td>75% (55%, 95%)</td>
<td>83% (83%, 84%)</td>
<td>80% (58%, 100%)</td>
<td>92% (83%, 100%)</td>
<td>92% (90%, 93%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>56% (49%, 62%)</td>
<td>65% (64%)</td>
<td>54% (44%, 63%)</td>
<td>63% (57%, 70%)</td>
<td>72% (70%, 74%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>4% (2%, 6%)</td>
<td>7% (10%)</td>
<td>4% (2%, 6%)</td>
<td>7% (5%, 9%)</td>
<td>15% (14%, 17%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>53% (50%, 56%)</td>
<td>60% (62%)</td>
<td>53% (49%, 57%)</td>
<td>60% (57%, 63%)</td>
<td>67% (66%, 68%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breast cancer</td>
<td>L</td>
<td>90% (84%, 96%)</td>
<td>96% (92%)</td>
<td>92% (92%, 92%)</td>
<td>96% (96%, 96%)</td>
<td>96% (96%, 96%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>65% (53%, 77%)</td>
<td>96% (75%)</td>
<td>67% (55%, 79%)</td>
<td>85% (80%, 89%)</td>
<td>84% (82%, 86%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>2% (0%, 6%)</td>
<td>17% (19%)</td>
<td>1% (1%)</td>
<td>13% (4%)</td>
<td>26% (23%, 30%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>66% (57%, 75%)</td>
<td>85% (84%)</td>
<td>70% (60%, 79%)</td>
<td>86% (83%, 89%)</td>
<td>88% (87%, 89%)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Plotted values are in bold print. D = distant; L = local; R = regional.
† Cure model did not fully converge.
for diagnosis year by adding a time varying covariate to the spline function (3), however interactions between two continuous predictors are complex to implement and interpret. Simplified models with a categorical time varying covariate for diagnosis year show only slight modifications to the cure estimate and small adjustments to the median survival time across diagnosis years. Another assumption of the flexible model is that cure is reached on or before the placement of the last knot. Andersson et al. (3) explored the sensitivity of the flexible cure model to knot placement and found that the cure estimate and median survival time were fairly robust to the number of knots and their placement; however, they recommend using these models where cure is observed within the given follow-up time. An advantage to the flexible model is that one does not have to define a specific distribution for the survival curve. An advantage to the parametric cure models is that one does not have to specify when cure occurs.

Both the parametric mixture and flexible models rely on the Newton-Raphson algorithm to solve for parameter estimates. This method occasionally fails to converge in cases where there is a flat likelihood or where the data are multimodal. We found our models did not converge for some cases of localized and distant stage disease and for the full cohort where the flexible model was applied. For the parametric mixture model, we eased the convergence criteria (12) and for the flexible model, starting values were estimated from a linear function of log time (13); however, neither approach improved convergence of the models.

Li et al. (4) and Yu et al. (5) separately evaluated the identifiability of the cure fraction and recommended that cure estimates should be based on large cohorts with long follow-up data. Sites prone to flat likelihood estimates (eg, models that fail to converge) may be stratified or adjusted to encourage greater homogeneity in the data. An advantage to the data in the SEER program is the availability of large cohorts of data with long-term follow-up. With the optimal cohort including 35 years of data, stratified or adjusted to encourage greater homogeneity in the data. An advantage to the data in the SEER program is the availability of large cohorts of data with long-term follow-up. With the optimal cohort including 35 years of data, stratified or adjusted to encourage greater homogeneity in the data. An advantage to the data in the SEER program is the availability of large cohorts of data with long-term follow-up. With the optimal cohort including 35 years of data, stratified or adjusted to encourage greater homogeneity in the data.

In these analyses, we compared study designs with different follow-up times to determine if cure fractions could be accurately predicted with minimal follow-up time so that we could provide more recent estimates. For colorectal cancer, 12 years was considered adequate to achieve a cure estimate, however we found a slight overestimate in cure for most cases with shorter follow-up time indicating that we may be missing a few events in the 12-year follow-up period. For breast cancer, shorter follow-up time exaggerated the overestimation of cure due increased numbers of unobserved cancer deaths.
Yu et al. (20) compared several cure models including the mixture, nonmixture, and flexible models for relative survival using grouped and individual data. Under ideal conditions, all models produced similar results. Less than ideal conditions where cure could not be identified or the model was misspecified led to inconsistent results and biased cure estimates. We agree with this assessment although it was limited to relative survival models with Weibull and flexible distribution functions. They recommend the flexible model in cases where the parametric model fails. In contrast, our findings indicated that the flexible model always produce higher estimates of the cure fraction and we do not recommend reporting cure where there is instability in the results from parametric models. Instead we prefer to improve stability of the parametric mixture model through stratification and adjustment for covariates, such as diagnosis year. The authors also noted that differences in scale of the time interval can influence estimates from grouped survival data. We did not test this in our models, as all survival time was specified in yearly intervals.

Based on our findings, we recommend: running multiple models of different parametric families to determine the stability of the cure fraction estimate, examining plots for fit of the mixture cure model and for a leveling off of the survival curve tail to support existence of cure, and stratifying and adjusting the models for important covariates to improve homogeneity in the data. Cohorts should span a limited time frame to avoid major changes in screening or treatment of the disease that could introduce heterogeneity in the data. Choice of cause-specific or relative survival outcomes should be determined by weighing prior knowledge of the matching life table to the accuracy of cause of death information to avoid introducing bias (8). Length of follow-up should be determined by the median survival time and optimized to capture the majority of cancer deaths in the observation period.

Cure models are potentially useful for projecting long-term survival estimates for cancer populations at the time of diagnosis. However, modelers should be cautious in interpreting these estimates under less ideal conditions, including shorter follow-up time, and small sample sizes. Special considerations should also be made about how to incorporate covariates into the model, model selection, and study design. For example, we see stable cases, such as regional stage colorectal cancer: between 40% and 50% in 1985 increasing to 55%–65% by 2008. These estimates are realistic and can be obtained with the data available. However, other cancer sites may require more follow-up time to attain a stable estimate.

**Appendix A: Variance and CI for Statistical Cure from Parametric Mixture Model**

Given the parameter estimates \( \hat{\beta}_o \) and \( \hat{\beta}_1 \), consider the transformation \( g(\hat{\beta}_0, \hat{\beta}_1) \). The approximate variance of \( g(\hat{\beta}_0, \hat{\beta}_1) \) can be found by the Delta Method:

\[
\text{var}(g(\hat{\beta}_0, \hat{\beta}_1)) = GVG,
\]

where \( G = [g_{11}, g_{12}] \) so that

\[
g_{11} = \frac{\partial g(\hat{\beta}_0, \hat{\beta}_1)}{\partial \hat{\beta}_0} \bigg|_{\hat{\beta}_0=\beta_0, \hat{\beta}_1=\beta_1} = \frac{-\exp(-\hat{\beta}_0 - \hat{\beta}_1 X)}{[1 + \exp(-\hat{\beta}_0 - \hat{\beta}_1 X)]^2} X
\]

\[
g_{12} = \frac{\partial g(\hat{\beta}_0, \hat{\beta}_1)}{\partial \hat{\beta}_1} \bigg|_{\hat{\beta}_0=\beta_0, \hat{\beta}_1=\beta_1} = \frac{-\exp(-\hat{\beta}_0 - \hat{\beta}_1 X)}{[1 + \exp(-\hat{\beta}_0 - \hat{\beta}_1 X)]^2} X
\]

and \( V \) is the variance-covariance matrix of \( (\hat{\beta}_0, \hat{\beta}_1) \). Then,

\[
\hat{\text{var}}[g(\hat{\beta}_0, \hat{\beta}_1)] = [g_{11}]^2 \hat{\text{var}}(\hat{\beta}_0) + [g_{12}]^2 \hat{\text{var}}(\hat{\beta}_1) + 2[g_{11}][g_{12}] \hat{\text{cov}}(\hat{\beta}_0, \hat{\beta}_1)
\]

where,

\[
A = \exp(-\hat{\beta}_0 - \hat{\beta}_1 X).
\]

The Wald test statistic for testing \( H_0: g(\hat{\beta}_0, \hat{\beta}_1) = 0 \) vs the two-sided alternative is

\[
Z = \frac{g(\hat{\beta}_0, \hat{\beta}_1)}{\sqrt{\text{var}[g(\hat{\beta}_0, \hat{\beta}_1)]}}
\]

where \( Z \) is approximately Standard Normal Distribution. A \((1 - \alpha) \times 100\% \) CI for \( g(\hat{\beta}_0, \hat{\beta}_1) = \epsilon \) is given by

\[
g(\hat{\beta}_0, \hat{\beta}_1) \pm Z_{\alpha/2} \sqrt{\text{var}[g(\hat{\beta}_0, \hat{\beta}_1)]},
\]

where \( Z_{\alpha/2} \) is the upper \( \alpha/2 \)-th percentile of the Standard Normal Distribution.

**Appendix B: Variance and CI for Statistical Cure From Flexible Parametric Model**

Given the parameter estimates \( \hat{\gamma}_0 \) and \( \hat{\beta} \), consider the transformation \( g(\hat{\gamma}_0, \hat{\beta}) \). The approximate variance of \( g(\hat{\gamma}_0, \hat{\beta}) \) can be found by the Delta Method:

\[
\hat{\text{var}}[g(\hat{\gamma}_0, \hat{\beta})] = GV G,
\]

where \( G = [g_{11}, g_{12}] \) so that

\[
g_{11} = \frac{\partial g(\hat{\gamma}_0, \hat{\beta})}{\partial \gamma_0} \bigg|_{\gamma_0=\gamma_0^o, \beta=\beta} = \exp[-\exp(\hat{\gamma}_0 + \hat{\beta} X)] [-\exp(\hat{\gamma}_0 + \hat{\beta} X)]
\]

\[
g_{12} = \frac{\partial g(\hat{\gamma}_0, \hat{\beta})}{\partial \beta} \bigg|_{\gamma_0=\gamma_0^o, \beta=\beta} = \exp[-\exp(\hat{\gamma}_0 + \hat{\beta} X)] [-\exp(\hat{\gamma}_0 + \hat{\beta} X)] X
\]

where \( X \) is a given covariate.
and $V$ is the variance-covariance matrix of $(\hat{\gamma}_{00}, \hat{\beta})$. Then,

$$\text{var}[g(\hat{\gamma}_{00}, \hat{\beta})] = [g_{11}]^2 \text{var}(\hat{\gamma}_{00}) + [g_{12}]^2 \text{var}(\hat{\beta})$$

$$+ 2[g_{11}][g_{12}] \text{cov}(\hat{\gamma}_{00}, \hat{\beta})$$

$$= [B^* \exp(2B)] \text{var}(\hat{\gamma}_{00}) + [(BX)^T \exp(2B)] \text{var}(\hat{\beta})$$

$$+ 2[B^* X \exp(2B)] \text{cov}(\hat{\gamma}_{00}, \hat{\beta}),$$

where

$$B = -\exp(\hat{\gamma}_{00} + \hat{\beta} X).$$

The Wald test statistic for testing $H_0: g(\hat{\gamma}_{00}, \hat{\beta}) = 0$ vs the two-sided alternative is

$$Z = \frac{g(\hat{\gamma}_{00}, \hat{\beta})}{\sqrt{\text{var}[g(\hat{\gamma}_{00}, \hat{\beta})]}},$$

where $Z$ is approximately Standard Normal Distribution. $A(1 - \alpha) \times 100\%$ CI for $g(\gamma_{00}, \beta) = c$ is given by

$$g(\hat{\gamma}_{00}, \hat{\beta}) \pm Z_{\frac{\alpha}{2}} \sqrt{\text{var}[g(\hat{\gamma}_{00}, \hat{\beta})]},$$

where $Z_{\frac{\alpha}{2}}$ is the upper $\frac{\alpha}{2}$th percentile of the Standard Normal Distribution.

### Appendix C: Median Survival in the Uncured Fraction for the Flexible Parametric Model without Time Dependent Covariates

The flexible parametric model adjusting for diagnosis year, $Y$, is as follows:

$$R(t) = \exp[-\exp(\gamma_0 + \beta(Y) + \gamma_{01} v_1(t) + \gamma_{02} v_2(t) + \gamma_{03} v_3(t) + \gamma_{04} v_4(t) + \gamma_{05} v_5(t) + \gamma_{06} v_6(t) + \gamma_{07} v_7(t) + \gamma_{08} v_8(t) + \gamma_{09} v_9(t) + \gamma_{10} v_{10}(t)))]$$

where $\beta$, $\gamma_0$, ..., $\gamma_{10}$, and $v_1(t)$, $v_2(t)$, ..., $v_{10}(t)$ are as defined in “Data and Methods.” We rewrite this as a mixture model (3):

$$R(t) = c + (1 - c) \frac{F_t(t) - c}{1 - c}$$

where,

$$c = \exp[-\exp(\gamma_0 + \beta(Y))].$$

and

$$F_t(t) = \exp[\gamma_0 v_1(t) + \gamma_0 v_2(t) + \gamma_0 v_3(t) + \gamma_0 v_4(t) + \gamma_0 v_5(t) + \gamma_0 v_7(t) + \gamma_0 v_8(t) + \gamma_0 v_9(t) + \gamma_0 v_{10}(t)]$$

so that the uncured fraction is assumed to follow a survival distribution function, $F_t(t) - c$. To estimate the median survival in the uncured fraction, we set $F_t(t) - c = 0.5$ and solve for $t$.

Note that in the STATA output spline parameters $\gamma_{02}$, ..., $\gamma_{07}$ are orthogonalized and the notation is reversed from the written formulas, so that $\gamma_0$, $\lambda$, and $\gamma$ are $\gamma_{0k}$, $\lambda_k$, and $c_k$ respectively.

### References

11. Huang L, Cronin KA, Johnson KA, Mariotto AB, Feuer EJ. Improved survival time: what can survival cure models tell us about population-based survival improvements in late-stage colorectal, ovarian, and testicular cancer? Cancer. 2008;112(10):2289–2300.


**Notes**

We thank Dr Therese Andersson for her helpful explanation of the flexible model and the STATA stpm2 software package.

**Affiliations of authors:** Surveillance Research Program, Division of Cancer Control and Population Sciences, National Cancer Institute, National Institutes of Health, Bethesda, MD (MRS, EJF, ABM).