CRC cause-specific survival

Many of CRC-SPIN v1.0's published natural history modeling outputs were generated prior to a 2013 CRC survival update.[1] CRC-SPIN v1.0's initial CRC survival model was based on relative survival using SEER survival data from 1975 to 1979 (prior to the diffusion of colorectal cancer screening).[2] This data was imported into CAN*SURV to generate proportional hazards models stratified by location (colon and rectum) and American Joint Committee on Cancer (AJCC) CRC stage (stages I through IV), with sex and age as covariates.[3]

In 2013, CISNET updated the CRC survival methodology,[1,3] and since that time, all three models (CRC-SPIN, SimCRC, and MISCAN) use the same survival method.[4] The updated survival methodology is additive-hazards-based with time-varying covariables and takes into account improved survival for more recently diagnosed CRC relative to those diagnosed in 1975 due to improvements in therapy and surveillance methods. This method is used to determine CRC survival based on sex, stage at diagnosis, age at diagnosis, and year of diagnosis, with the optional inclusion of race for those models that can incorporate race as a risk factor (ie, SimCRC). This consistency reduced the between-model variability related to CRC survival, allowing investigators to directly compare differences in model results primarily due to natural history structure and assumptions, although it introduces the possibility of systemic bias if inherent issues exist within the survival approach. Ultimately, although the CISNET teams provide a general theoretical discussion into their methodology, an enumeration of the parameters themselves and a detailed description of how they are implemented in their models is not published.[1]

CRC-AIM implements cause-specific survival as a set of parametric regression equations that model survival probabilities, stratified by location and AJCC CRC stage, as a function of sex and age at diagnosis. To compare the survival outcomes of CRC-AIM to those of the CISNET models, we generated survival curves that mimicked the timeframes of SEER data used by CISNET both before and after their survival update in 2013. We generated parametric cause-specific survival curves using SEER data from 1975 to 1979 and compared natural history outcomes to publicly available values across multiple CISNET publications before the survival update. Additionally, we generated a similar set of curves using 2000-2003 SEER data and compared natural history outcomes to results described in the 2016 USPSTF modeling report, after the survival update. Both sets of comparisons are described in greater detail in the main manuscript.

Here, we include detailed descriptions of the SEER queries that generated data for the parametric linear regression models and comprehensive information about those models that were used to create the cause-specific survival curves. Specifically, for each condition—using the SEER data timeframes before and after the survival update, subdivided by AJCC CRC stage and location—we include the survival curves themselves along with model selection details and fitted model diagnostics and parameter estimates. Here is a complete list of the figures and tables that correspond to each condition:

- 1975-1979 SEER data:
    - Stage I: colon (**Figure 1**, **Table 1**); rectum (**Figure 2**, **Table 2**)
    - Stage II: colon (**Figure 3**, **Table 3**); rectum (**Figure 4**, **Table 4**)

- Stage III: colon (**Figure 5**, **Tables 5-6**); rectum (**Figure 6**, **Table 7**)
- Stage IV: colon (**Figure 7**, **Table 8**); rectum (**Figure 8**, **Tables 9-10**)

- 2000-2003 SEER data:
  - Stage I: colon (**Figure 9**, **Table 11**); rectum (**Figure 10**, **Tables 12-13**)
  - Stage II: colon (**Figure 11**, **Table 14**); rectum (**Figure 12**, **Tables 15-16**)
  - Stage III: colon (**Figure 13**, **Tables 17-18**); rectum (**Figure 14**, **Tables 19-20**)
  - Stage IV: colon (**Figure 15**, **Tables 21-22**); rectum (**Figure 16**, **Tables 23-24**)

SEER queries

We queried the SEER database to extract data from the 1975-1979 and 2000-2003 timeframes using the criteria described below. The case listing file was saved as a comma-separated values (CSV) file and imported into JMP v13.0 (SAS Institute) for statistical analysis. Furthermore, we describe how the SEER data was modified in JMP prior to fitting the regression models (**Table 25**).

SEER query for 1975-1979 survival

Software

Surveillance Research Program, National Cancer Institute SEER*Stat software (www.seer.cancer.gov/seerstat) version 8.3.5. 10/05/2018

Data

Surveillance, Epidemiology, and End Results (SEER) Program (www.seer.cancer.gov) SEER*Stat Database: Incidence – SEER 9 Regs Research Data, Nov 2017 Sub (1973-2015) <Katrina/Rita Population Adjustment> - Linked To County Attributes – Total U.S., 1969-2016 Counties, National Cancer Institute, DCCPS, Surveillance Research Program, released April 2018, based on the November 2017 submission.

Statistic

Cause-Specific Survival

Definition of Cause of Death: Dead due to cancer using SEER cause-specific death classification
Missing/Unknown COD: Exclude From Analysis
Selection
Select Only: Malignant Behavior, Known Age
Exclude: All Death Certificate Only and Autopsy Only, Alive with No Survival Time
Exclusions to Match the Expected Survival Table: Age Values Not Found in Table, Invalid Year, Values Not Found for Other Variables in Table
Case Selection
{Site and Morphology.Site recode ICD-O-3/WHO 2008} = ' Colon and Rectum'
AND {Site and Morphology.Histologic Type ICD-O-3} = 8000-8001,8010,8020,8140,8210-8211,8220-8221,8260-8263,8480-8482,8490
AND {Race, Sex, Year Dx, Registry, County.Year of diagnosis} = '1975','1976','1977','1978','1979'
Multiple Primary Selection: First Primary Only (Sequence Number 0 or 1)
Parameters
Pre-calculated Duration: Survival Months (from complete dates)
Study Cutoff: Dec 2015
Censor When Attained Age Exceeds Expected Table Max
Display: Case Listing
Intervals: Number: 999, Months Per: 1
Table
Case Listing:

- SEER historic stage A
- 2-Digit NS EOD part 1 (1973-1982)
- AJCC 5th Ed Schrag Code 1975-1979
- Age recode with <1 year olds
- Survival months
- Survival months flag
- Vital status recode (study cutoff used)
- Sex
- COD to site recode
- COD to site rec KM
- SEER cause-specific death classification
- SEER summary stage 1977 (1995-2000)
- SEER other cause of death classification
- Site Colon vs Rectum

SEER query for 2000-2003 survival
Software
Surveillance Research Program, National Cancer Institute SEER*Stat software (www.seer.cancer.gov/seerstat) version 8.3.5. 10/05/2018
Data
Surveillance, Epidemiology, and End Results (SEER) Program (www.seer.cancer.gov)
SEER*Stat Database: Incidence – SEER 9 Regs Research Data, Nov 2017 Sub

(1973-2015) <Katrina/Rita Population Adjustment> - Linked To County Attributes – Total U.S., 1969-2016 Counties, National Cancer Institute, DCCPS, Surveillance Research Program, released April 2018, based on the November 2017 submission.

Statistic

Cause-Specific Survival

Definition of Cause of Death: Dead due to cancer using SEER cause-specific death classification

Missing/Unknown COD: Exclude From Analysis

Selection

Select Only: Malignant Behavior, Known Age

Exclude: All Death Certificate Only and Autopsy Only, Alive with No Survival Time

Exclusions to Match the Expected Survival Table: Age Values Not Found in Table, Invalid Year, Values Not Found for Other Variables in Table

Case Selection

{Site and Morphology.Site recode ICD-O-3/WHO 2008} = ' Colon and Rectum'

AND {Site and Morphology.Histologic Type ICD-O-3} = 8000-8001,8010,8020,8140,8210-8211,8220-8221,8260-8263,8480-8482,8490

AND {Race, Sex, Year Dx, Registry, County.Year of diagnosis} = '2000','2001','2002','2003'

Multiple Primary Selection: First Primary Only (Sequence Number 0 or 1)

Parameters

Pre-calculated Duration: Survival Months (from complete dates)

Study Cutoff: Dec 2015

Censor When Attained Age Exceeds Expected Table Max

Display: Case Listing

Intervals: Number: 999, Months Per: 1

Table

Case Listing:

- AJCC 5th Ed Schrag Code 1988-2003
- Age recode with <1 year olds
- Survival months
- Survival months flag
- Vital status recode (study cutoff used)
- Sex
- COD to site recode
- COD to site rec KM
- SEER cause-specific death classification
- SEER summary stage 1977 (1995-2000)
- SEER summary stage 2000 (2001-2003)
- SEER other cause of death classification
- Site Colon vs Rectum

AJCC staging for 1975-1979 SEER data

SEER registry data prior to 1988 uses historic staging criteria that categorize cancer as local, regional, and distant. However, the AJCC staging system uses tumor, lymph node, and metastasis information to stage cancer. Since the CISNET CRC models use AJCC stages I through IV, we needed to convert pre-1988 SEER registry data. We used code developed by Deborah Schrag, coauthor on the CISNET 2013 survival update,[1] which essentially is a complicated user-defined variable that uses 12 standard SEER variables to recode stage according to the AJCC 5th Edition AJCC Cancer Staging Manual.[5] Although the AJCC staging manual staging is currently in its eighth edition,[6] the changes since the fifth edition primarily involve subgrouping stages II through IV and providing prognosis-related details.[7]

Functional implementation of survival models in CRC-AIM

There are some technical issues related to survival assumptions that must be considered when implementing a screening overlay. To illustrate, consider an example of an individual with two synchronous preclinical CRCs, one stage I and the other stage III. According to the documentation of CRC-SPIN's natural history model, the first cancer to become clinically detectable—whichever cancer's sojourn time expires first—determines CRC-related survival.[3] If the stage I cancer's sojourn time expires before the stage III cancer, the stage I cancer would determine survival. However, if a cancer is detected through screening, CRC-SPIN assumes the detection of all synchronous cancers and that all existing cancer is treated.[3] In this example, both the stage I and stage III cancers would be observed by an endoscopist. The CRC-SPIN authors do not explicitly state which cancer dictates the choice of survival function for this scenario, or

when the survival function would be applied (ie, after the ST expiration for the stage I cancer or for the stage III cancer). If the stage III cancer dictates survival, the screened individual may likely have a shorter lifespan compared to natural history, which was dictated by the stage I survival function.

This issue also extends to the removal of adenomas through screening. Consider a scenario in which an individual has a large, fast-growing adenoma A1 and a preclinical cancer pCRC2, which arose from another adenoma A2. In the natural history model, assume that pCRC2 transitions to clinical detection CRC2 through expiration of sojourn time into a stage I CRC, after which the survival method would be applied. In a hypothetical screening scenario, the adenoma that led to the CRC (A2) may have been removed but A1 could have been missed. This adenoma could hypothetically transition to preclinical CRC (pCRC1) and then a stage IV clinically detected CRC (CRC1) within a year after the appearance of the non-existent stage I CRC (CRC2) in the “parallel universe” natural history model. This individual would likely die sooner in the screening scenario from the stage IV CRC than the parallel universe individual that had stage I CRC in the natural history arm.

Although these corner-case scenarios can be addressed in numerous ways, the specific approaches undertaken by CRC-SPIN is not publicly described.

In CRC-AIM, if a cancer is detected by screening, the survival method is determined by the CRC stage diagnosed upon detection. This approach potentially confers a survival benefit if an earlier-stage CRC is detected relative to the CRC stage present when the sojourn time expires in a parallel universe. To prevent bias resulting from automatically assigning a worse survival outcome for screen-detected CRC

compared to natural history, the survival function will be applied only when the sojourn time would have otherwise expired for the screen-detected cancer. In other words, the CRC survival functions in CRC-AIM will not be implemented during the lead-time of the cancer. For CRC-AIM, upon clinical detection, the cancer stage and location are used to define the survival model. Based on the survival model, age at diagnosis and sex may be used to determine the cumulative probability of survival for each yearly interval after diagnosis. We restricted the number of intervals to 100 (the maximum age of an individual minus the minimum age one can be diagnosed with CRC, or 120 minus 20). A random uniform (0,1) distribution is used as an inverse CDF lookup to determine CRC-based survival years. If the random uniform distribution exceeds the maximum probability, then the maximum survival years are applied.

Linear regression modeling

We fit five separate parametric linear regression models for each SEER dataset (1975-1979 and 2000-2003), AJCC stage (stages I to IV), and location (colon versus rectum). These models were based on five different distributions to describe survival time—Weibull, lognormal, exponential, Fréchet, and loglogistic. Model effects were sex and age at CRC diagnosis and statistical significance of an effect was based on the Wald test. Age at diagnosis was subdivided into 20-49, 50-59, 60-69, 70-79, and ≥80 years, consistent with the categories used by CISNET.[1] Right-censoring was performed as appropriate. Any subject with a time of death that was reliably recorded as 0 months was recoded to 0.5 months to prevent model-fitting issues. Model selection was based

on smallest Akaike information criterion (AICc) value for the fitted distribution across the five separate models for a given combination of AJCC stage and location.

CRC-AIM uses the CDF to describe the cumulative transition probability of CRC-specific death as a function of sex and age at CRC diagnosis. Regression-based coefficients are multiplied by an indicator function if sex or age at diagnosis criteria are met (sex indicator is 1 if met, -1 if not met; age at diagnosis indicator is 1 if met, 0 if not met) and are linearly combined with an intercept, represented by λ in Weibull models and μ in loglogistic, lognormal, and Fréchet models..

For 1975-1979 and 2000-2003 Stage I colon cancers, the Weibull model was selected. The probability of transition to CRC-specific death at or before time *t* is given by:

$$F(t;\, k, \lambda) = 1 - e^{-(\frac{t}{\lambda})^k} \text{ for } t \geq 0.$$

with *k* as the shape parameter, and λ as the scale parameter. Since $k < 1$ in each of the two Weibull models, the failure rate decreases over time, representing the curative effect one would expect in a proportion of diagnosed and treated early-stage cancer.

For 1975-1979 Stage IV rectal-based cancers and 2000-2003 Stage I rectal-based cancers, the loglogistic model was selected:

$$F(t;\, \mu, \sigma) = \Phi_{logis}\left[\frac{\log_{(t)} - \mu}{\sigma}\right] \text{ for } t > 0$$

where

$$\Phi_{logis}(z) = \frac{1}{1 + e^{-z}}$$

For 1975-1979 Stage I, II, and III rectal-based cancers; Stage II colon-based cancers; and 2000-2003 Stage II, III, and IV colon- and rectal-based cancers, the lognormal model was selected:

$$F(t; \mu, \sigma) = \Phi_{nor}\left[\frac{\log_{(t)} - \mu}{\sigma}\right] \text{ for } t > 0$$

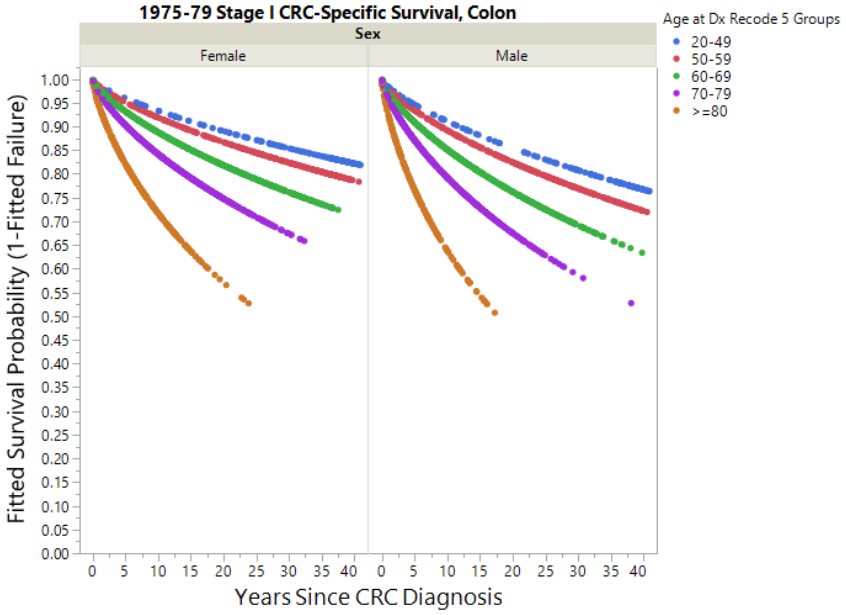
where

$$\Phi_{nor}(z) = \int_{-\infty}^{z} \Phi_{nor}(w)dw$$

For 1975-1979 Stage III and IV colon cancer, the Fréchet model was selected:

$$F(t; \mu, \sigma) = exp\left[-exp\left(-\frac{log_{(t)} - \mu}{\sigma}\right)\right] \text{ for } t > 0$$

Figure 1. Modeled cause-specific survival for 1975-79 stage I colorectal cancer (CRC) in the colon, stratified by age at diagnosis and sex. CRC-specific survival is represented as percent survival by years since diagnosis. Note: The abnormal residual behavior at low probability is due to recoding survival month from 0 months to 0.5 months—the abnormal plot observations have survival month values of 0.5—and has little practical impact on survival, which is based on yearly increments.



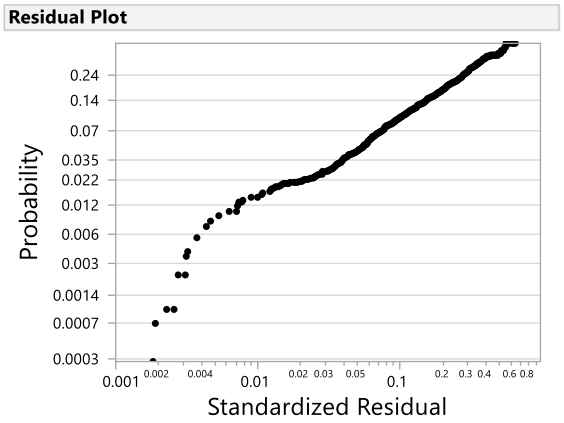
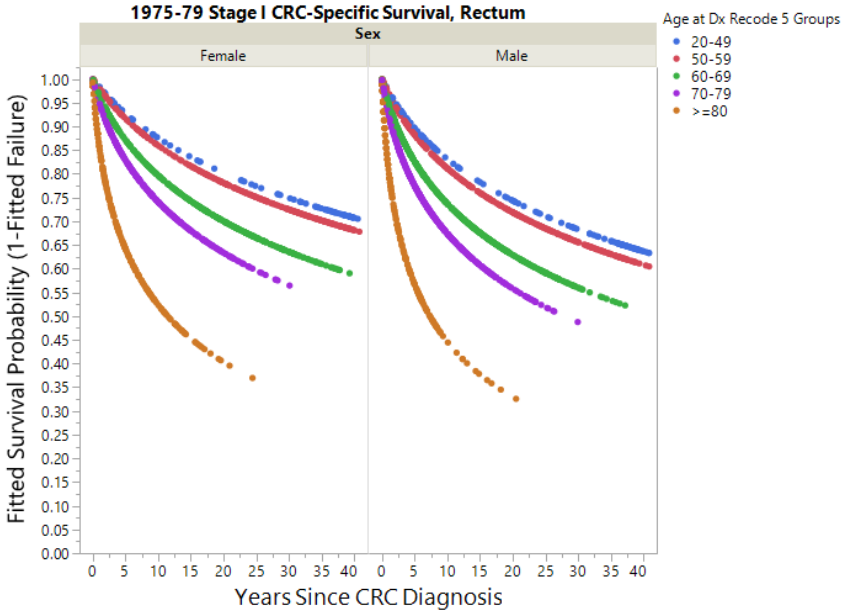


Figure 2. Modeled cause-specific survival for 1975-79 stage I colorectal cancer (CRC) in the rectum, stratified by age at diagnosis and sex. CRC-specific survival is represented as percent survival by years since diagnosis. Note: The abnormal residual behavior at low probability is due to recoding survival month from 0 months to 0.5 months—the abnormal plot observations have survival month values of 0.5—and has little practical impact on survival, which is based on yearly increments.



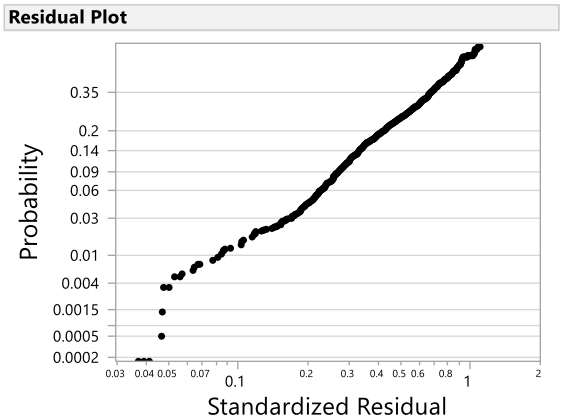
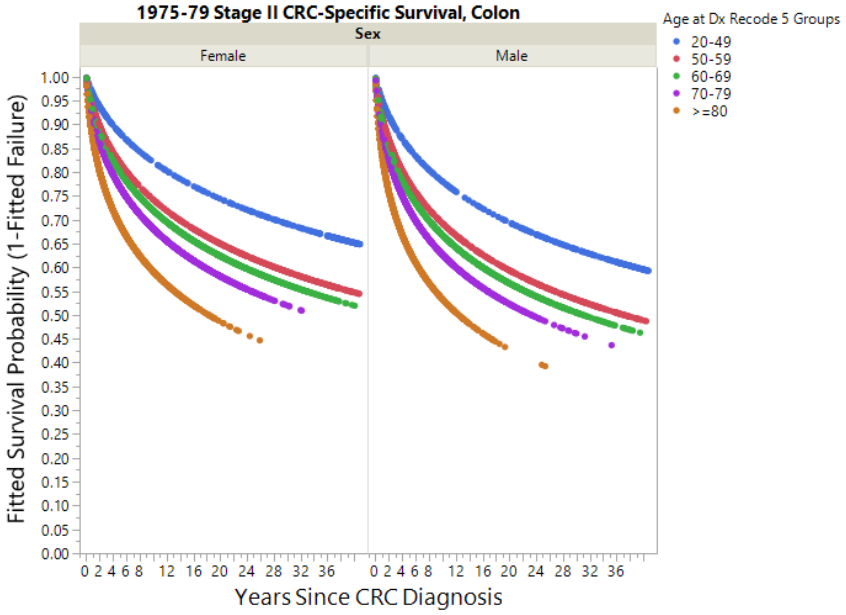


Figure 3. Modeled cause-specific survival for 1975-79 stage II colorectal cancer (CRC) in the colon, stratified by age at diagnosis and sex. CRC-specific survival is represented as percent survival by years since diagnosis. Note: The abnormal residual behavior at low probability is due to recoding survival month from 0 months to 0.5 months—the abnormal plot observations have survival month values of 0.5—and has little practical impact on survival, which is based on yearly increments.



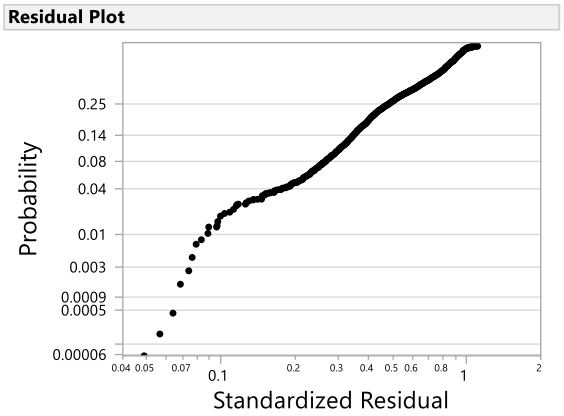
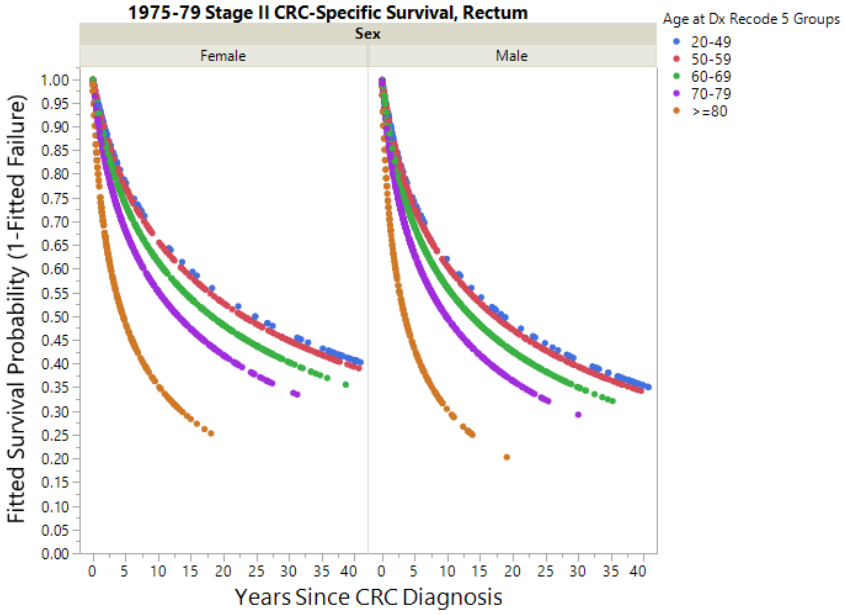


Figure 4. Modeled cause-specific survival for 1975-79 stage II colorectal cancer (CRC) in the rectum, stratified by age at diagnosis and sex. CRC-specific survival is represented as percent survival by years since diagnosis. Note: The abnormal residual behavior at low probability is due to recoding survival month from 0 months to 0.5 months—the abnormal plot observations have survival month values of 0.5—and has little practical impact on survival, which is based on yearly increments.



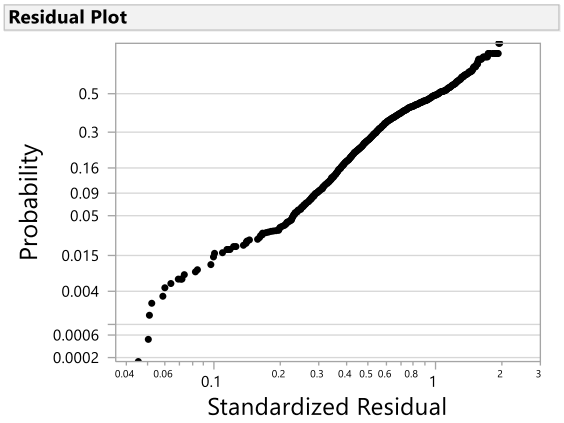
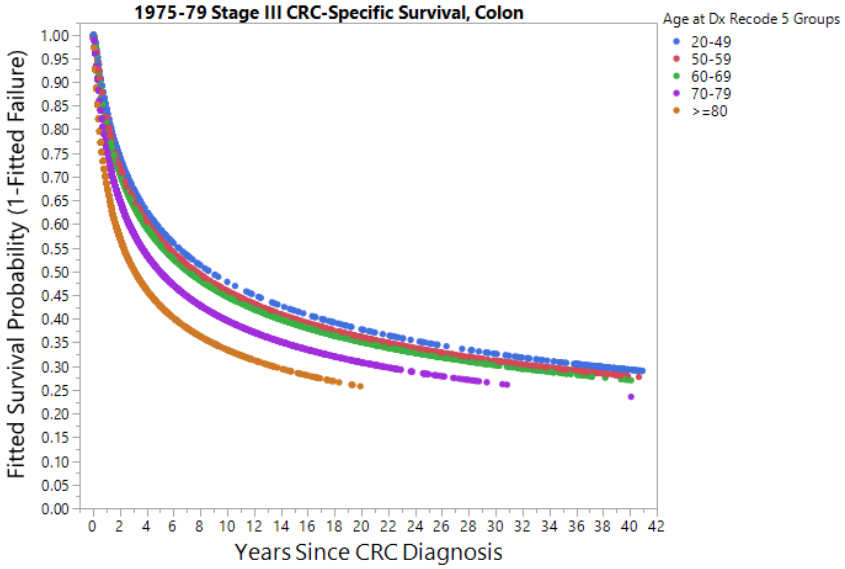


Figure 5. Modeled cause-specific survival for 1975-79 stage III colorectal cancer (CRC) in the colon, stratified by age at diagnosis. CRC-specific survival is represented as percent survival by years since diagnosis. Note: The abnormal residual behavior at low probability is due to recoding survival month from 0 months to 0.5 months—the abnormal plot observations have survival month values of 0.5—and has little practical impact on survival, which is based on yearly increments.



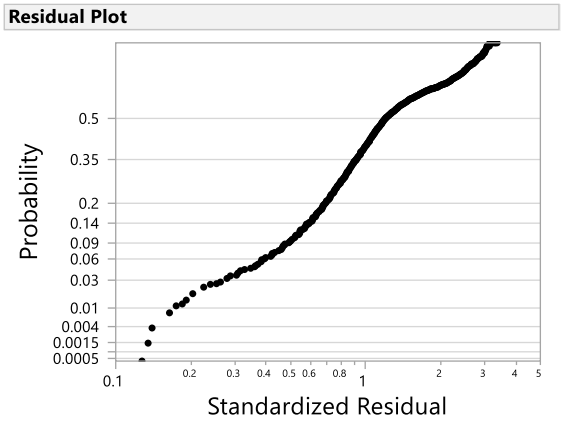
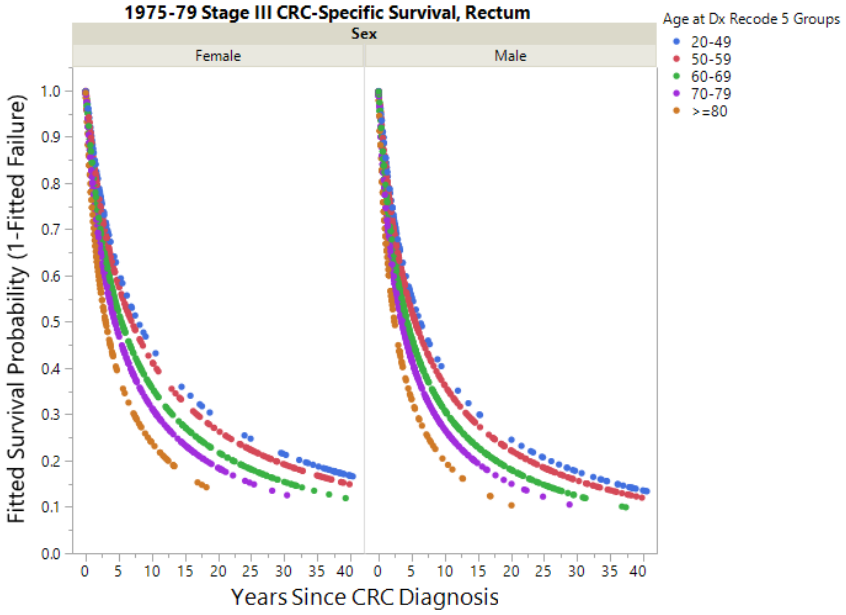


Figure 6. Modeled cause-specific survival for 1975-79 stage III colorectal cancer (CRC) in the rectum, stratified by age at diagnosis and sex. CRC-specific survival is represented as percent survival by years since diagnosis. Note: The abnormal residual behavior at low probability is due to recoding survival month from 0 months to 0.5 months—the abnormal plot observations have survival month values of 0.5—and has little practical impact on survival, which is based on yearly increments.



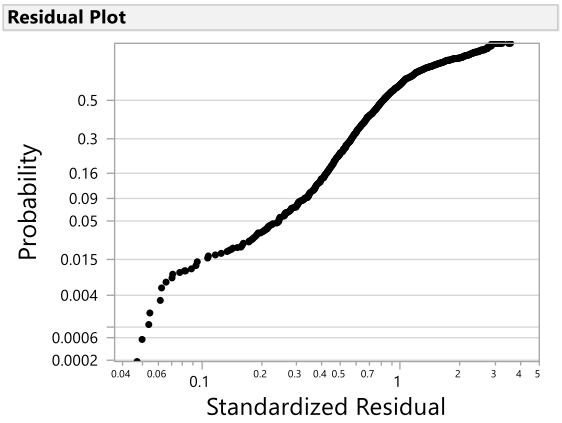
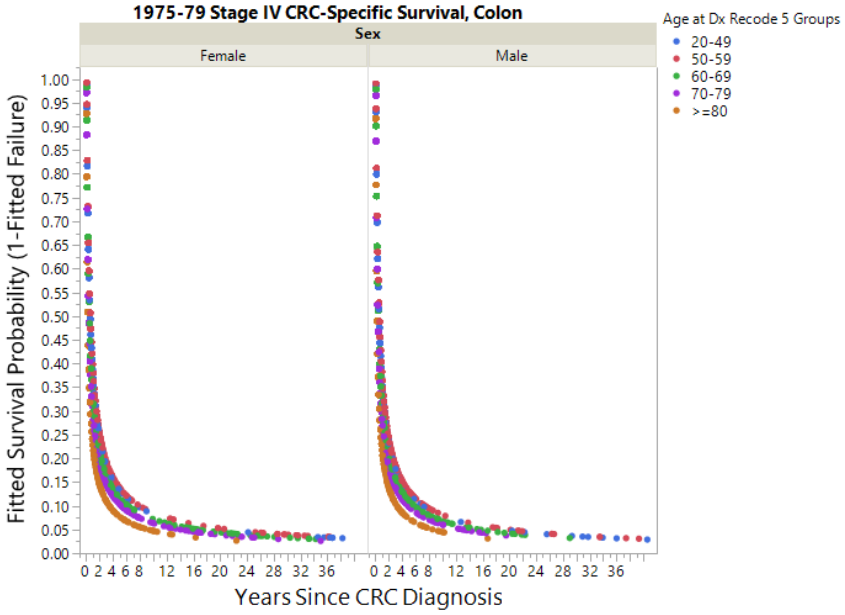


Figure 7. Modeled cause-specific survival for 1975-79 stage IV colorectal cancer (CRC) in the colon, stratified by age at diagnosis and sex. CRC-specific survival is represented as percent survival by years since diagnosis. Note: The abnormal residual behavior at low probability is due to recoding survival month from 0 months to 0.5 months—the abnormal plot observations have survival month values of 0.5—and has little practical impact on survival, which is based on yearly increments.



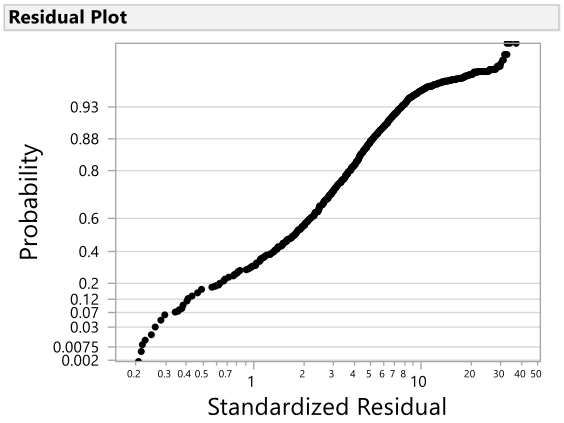
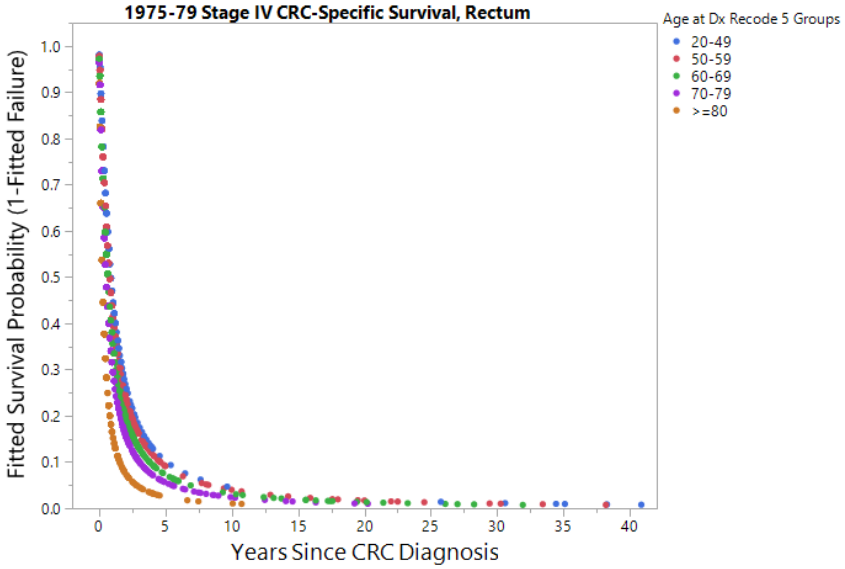


Figure 8. Modeled cause-specific survival for 1975-79 stage IV colorectal cancer (CRC) in the rectum, stratified by age at diagnosis. CRC-specific survival is represented as percent survival by years since diagnosis. Note: The abnormal residual behavior at low probability is due to recoding survival month from 0 months to 0.5 months—the abnormal plot observations have survival month values of 0.5—and has little practical impact on survival, which is based on yearly increments.



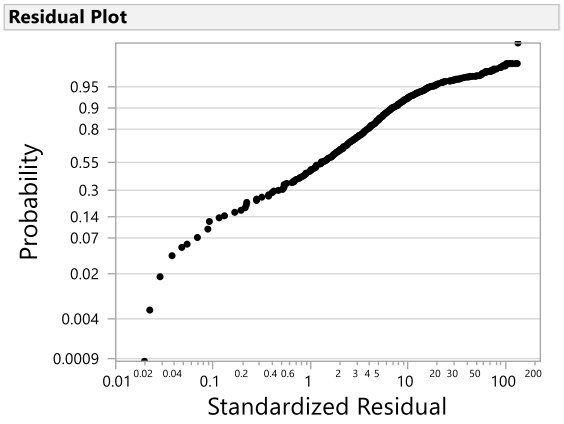
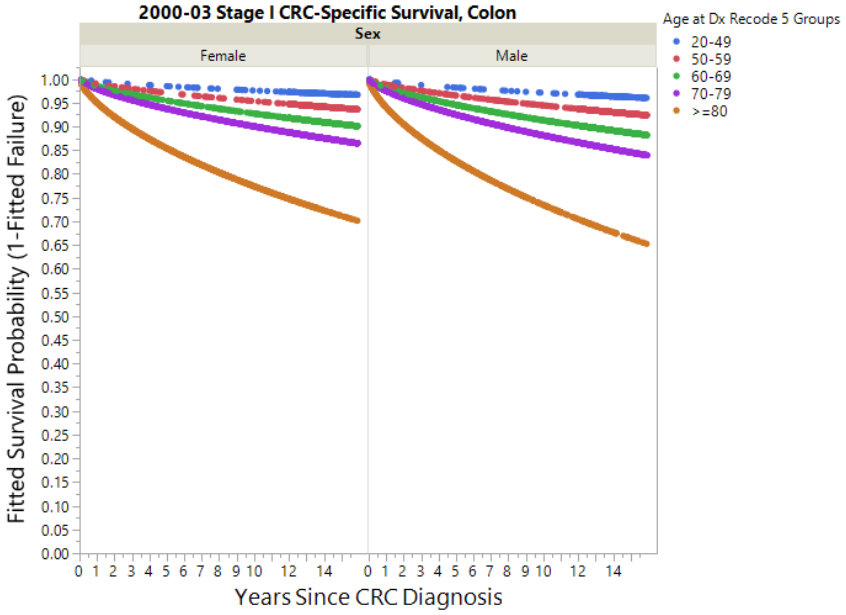


Figure 9. Modeled cause-specific survival for 2000-03 stage I colorectal cancer (CRC) in the colon, stratified by age at diagnosis and sex. CRC-specific survival is represented as percent survival by years since diagnosis. Note: The abnormal residual behavior at low probability is due to recoding survival month from 0 months to 0.5 months—the abnormal plot observations have survival month values of 0.5—and has little practical impact on survival, which is based on yearly increments.



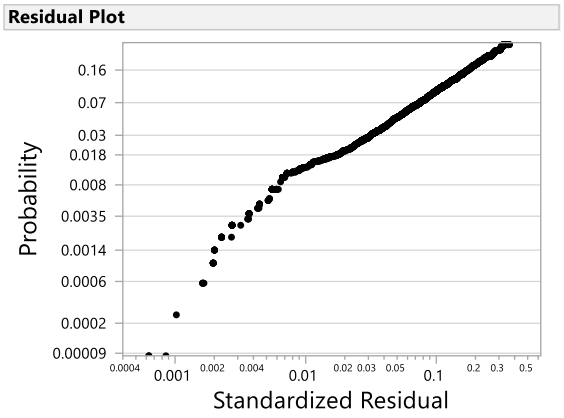
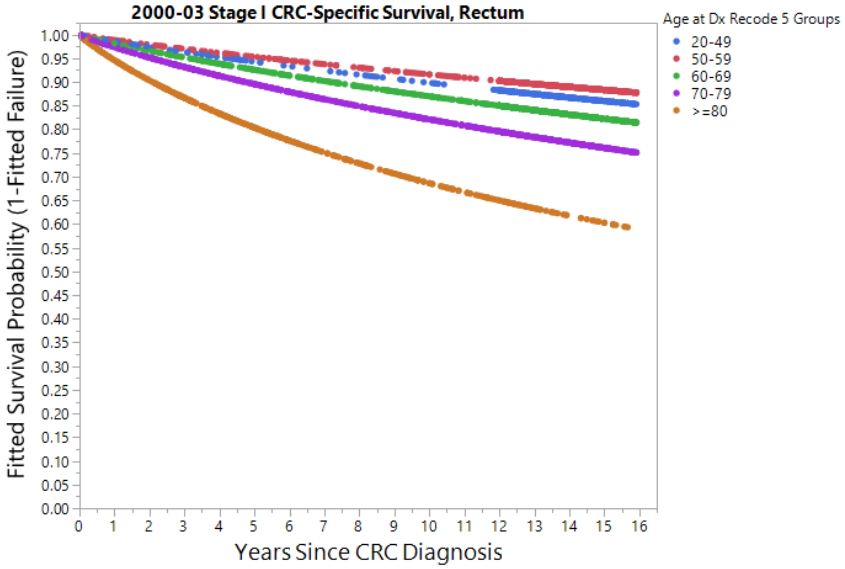


Figure 10. Modeled cause-specific survival for 2000-03 stage I colorectal cancer (CRC) in the rectum, stratified by age at diagnosis. CRC-specific survival is represented as percent survival by years since diagnosis. Note: The abnormal residual behavior at low probability is due to recoding survival month from 0 months to 0.5 months—the abnormal plot observations have survival month values of 0.5—and has little practical impact on survival, which is based on yearly increments.



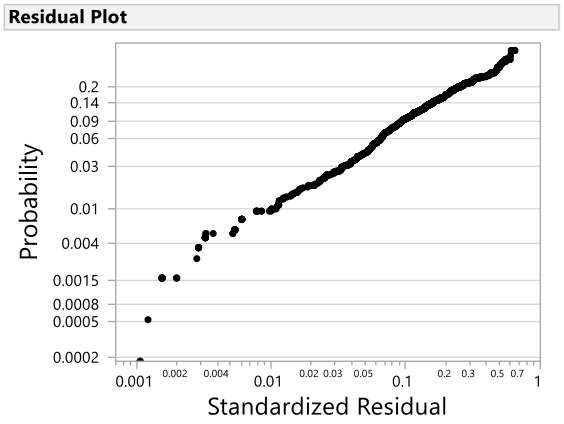
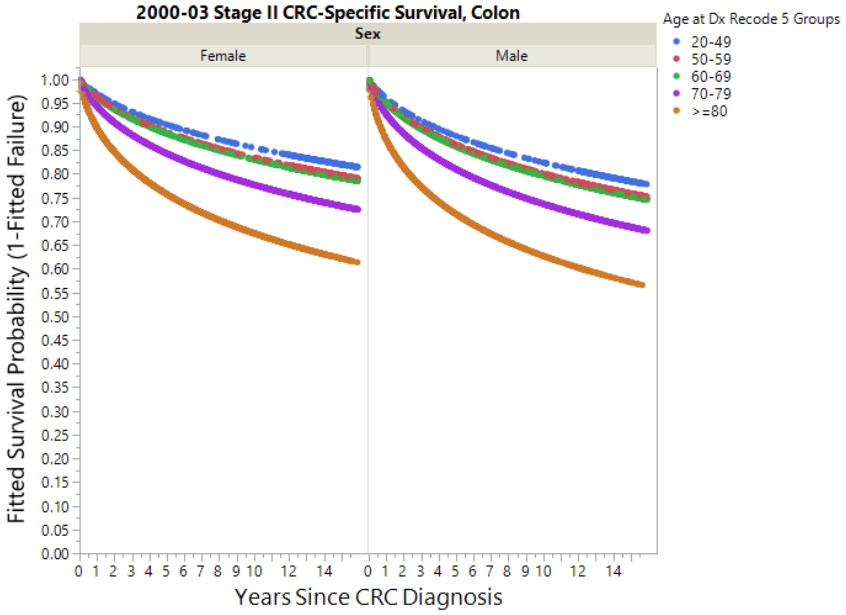


Figure 11. Modeled cause-specific survival for 2000-03 stage II colorectal cancer (CRC) in the colon, stratified by age at diagnosis and sex. CRC-specific survival is represented as percent survival by years since diagnosis. Note: The abnormal residual behavior at low probability is due to recoding survival month from 0 months to 0.5 months—the abnormal plot observations have survival month values of 0.5—and has little practical impact on survival, which is based on yearly increments.



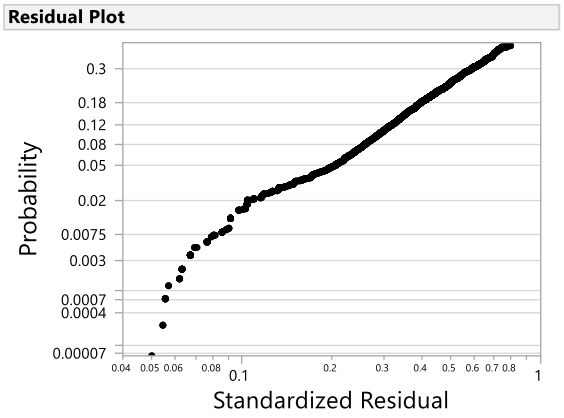
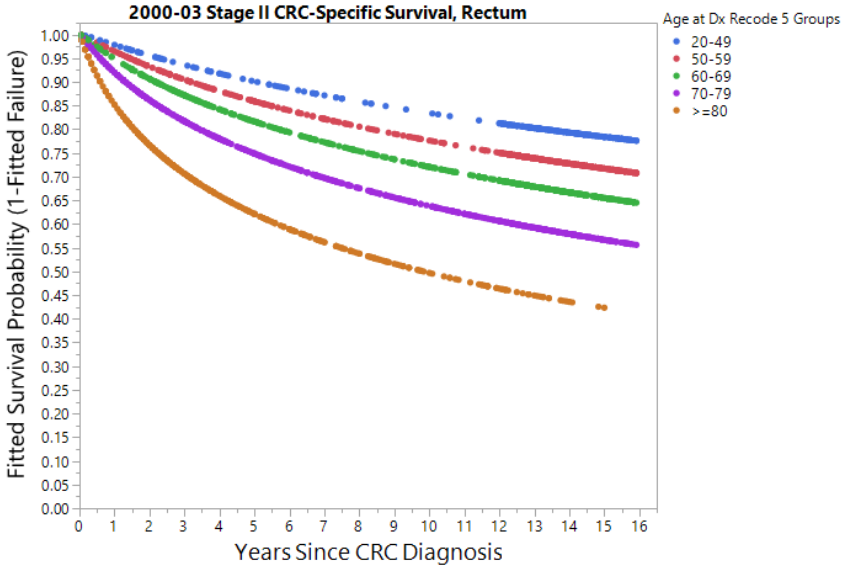


Figure 12. Modeled cause-specific survival for 2000-03 stage II colorectal cancer (CRC) in the rectum, stratified by age at diagnosis. CRC-specific survival is represented as percent survival by years since diagnosis. Note: The abnormal residual behavior at low probability is due to recoding survival month from 0 months to 0.5 months—the abnormal plot observations have survival month values of 0.5—and has little practical impact on survival, which is based on yearly increments.



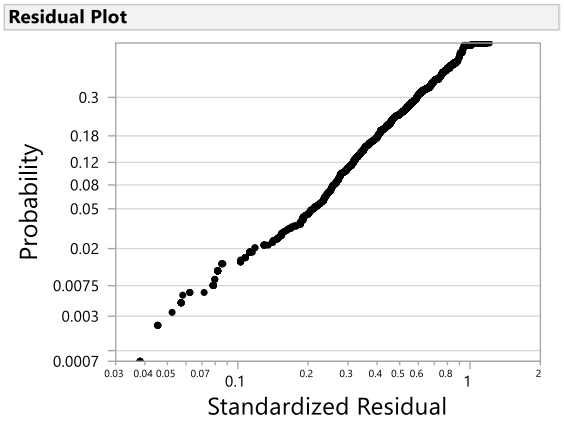
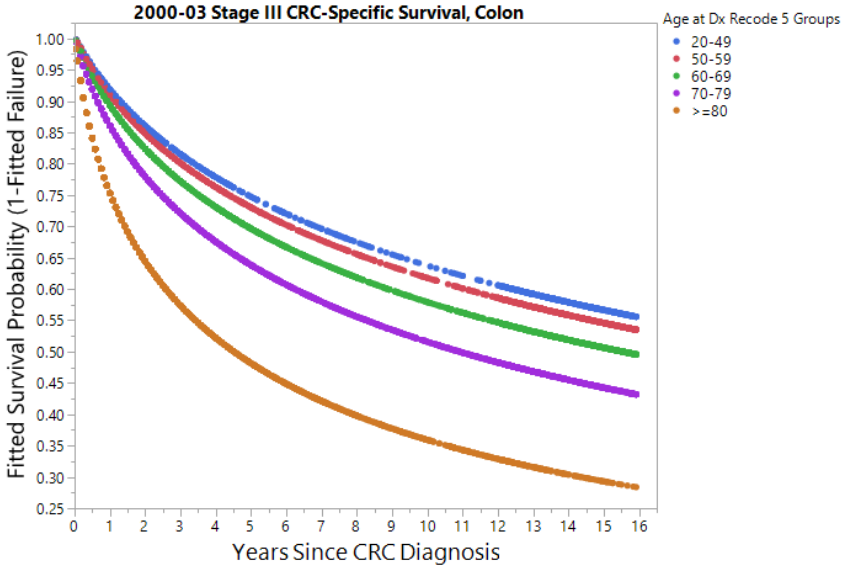


Figure 13. Modeled cause-specific survival for 2000-03 stage III colorectal cancer (CRC) in the colon, stratified by age at diagnosis and sex. CRC-specific survival is represented as percent survival by years since diagnosis. Note: The abnormal residual behavior at low probability is due to recoding survival month from 0 months to 0.5 months—the abnormal plot observations have survival month values of 0.5—and has little practical impact on survival, which is based on yearly increments.



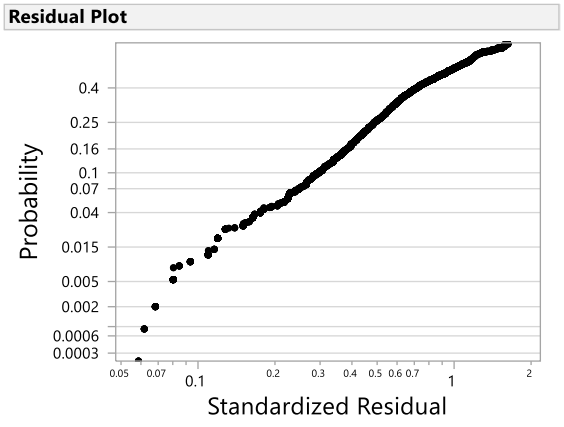
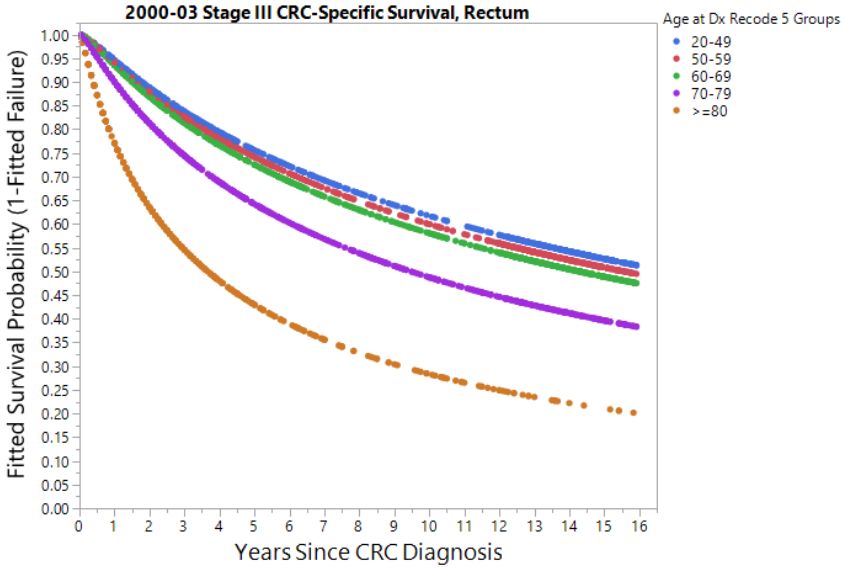


Figure 14. Modeled cause-specific survival for 2000-03 stage III colorectal cancer (CRC) in the rectum, stratified by age at diagnosis. CRC-specific survival is represented as percent survival by years since diagnosis. Note: The abnormal residual behavior at low probability is due to recoding survival month from 0 months to 0.5 months—the abnormal plot observations have survival month values of 0.5—and has little practical impact on survival, which is based on yearly increments.



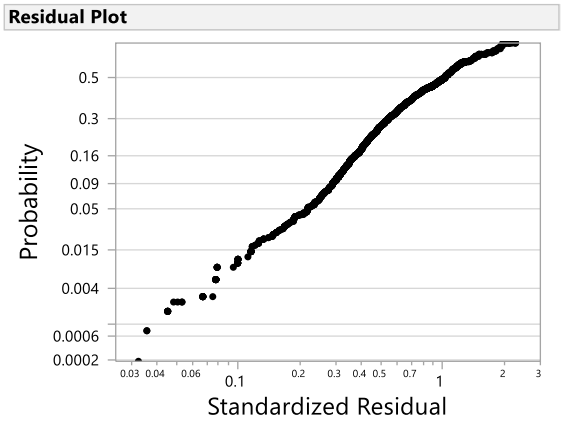
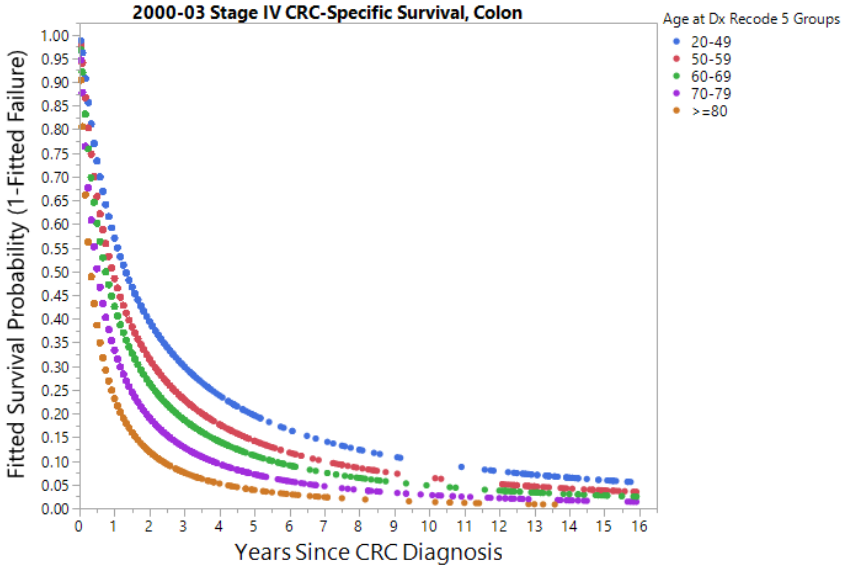


Figure 15. Modeled cause-specific survival for 2000-03 stage IV colorectal cancer (CRC) in the colon, stratified by age at diagnosis. CRC-specific survival is represented as percent survival by years since diagnosis. Note: The abnormal residual behavior at low probability is due to recoding survival month from 0 months to 0.5 months—the abnormal plot observations have survival month values of 0.5—and has little practical impact on survival, which is based on yearly increments.



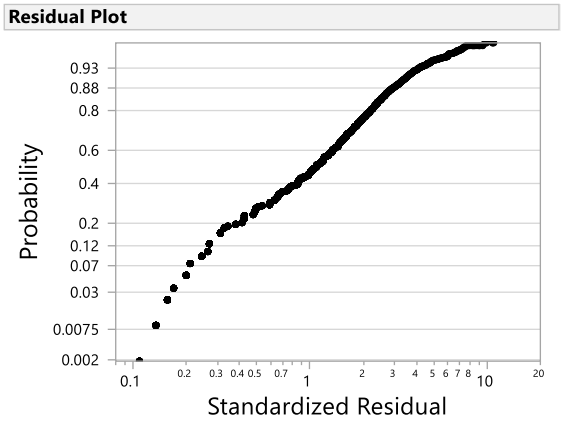
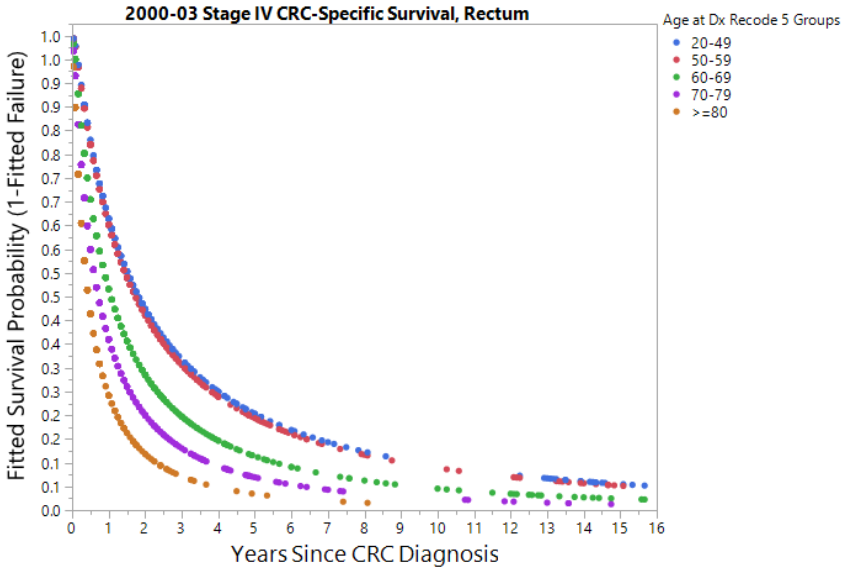


Figure 16. Modeled cause-specific survival for 2000-03 stage IV colorectal cancer (CRC) in the rectum, stratified by age at diagnosis. CRC-specific survival is represented as percent survival by years since diagnosis.



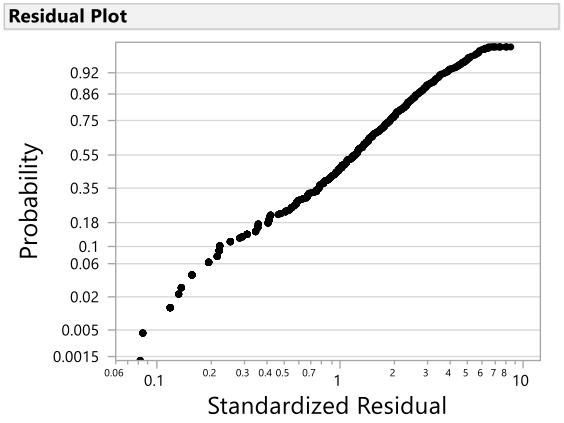


Table 1. Model selection details and fitted model diagnostics and parameter estimates for cause-specific survival for 1975-79 stage I colorectal cancer (CRC) in the colon. The highlighted distribution was selected because it resulted in the smallest Akaike information criterion (AICc) and the parameter estimates correspond to that distribution. Parameter δ = 1/κ. The Wald confidence interval was used.

1975-79 Stage I CRC, Colon: Model Comparison				
Distribution	**AICc**			
Weibull	6825.0			
Lognormal	6864.6			
Exponential	6895.8			
Frechet	6936.9			
Loglogistic	6829.9			
Observation Used	3555			
Uncensored Values	667			
Right Censored Values	2888			
Whole Model Test				
ChiSquare	**DF**	**Prob>Chisq**		
134.9306	5	<.0001		
Parameter Estimates				
Term	**Estimate**	**Std Error**	**Lower 95%**	**Upper 95%**
Intercept	4.787517	0.092399	4.606419	4.968616
Sex[Female]	0.858742	0.169591	0.526351	1.191134
Age at Dx Recode 5 Groups[20-49]	0.586693	0.117597	0.356207	0.817178
Age at Dx Recode 5 Groups[50-59]	0.135153	0.095296	-0.051624	0.321931
Age at Dx Recode 5 Groups[60-69]	-0.354845	0.093406	-0.537918	-0.171773
Age at Dx Recode 5 Groups[70-79]	0.200183	0.051827	0.098603	0.301763
δ	1.324690	0.046260	1.234021	1.415358
Wald Tests				
Source	**Nparm**	**DF**	**Wald ChiSquare**	**Prob>ChiSq**
Sex	1	1	14.9188	0.0001
Age at Dx Recode 5 Groups	4	4	145.6113	<.0001

Table 2. Model selection details and fitted model diagnostics and parameter estimates for cause-specific survival for 1975-79 stage I colorectal cancer (CRC) in the rectum. The highlighted distribution was selected because it resulted in the smallest Akaike information criterion (AICc) and the parameter estimates correspond to that distribution. The Wald confidence interval was used.

1975-79 Stage I CRC, Rectum: Model Comparison				
Distribution	**AICc**			
Weibull	7812.9			
Lognormal	7787.6			
Exponential	7933.7			
Frechet	7849.7			
Loglogistic	7792.9			
Observation Used	2951			
Uncensored Values	870			
Right Censored Values	2081			
Whole Model Test				
ChiSquare	**DF**	**Prob>Chisq**		
196.0059	5	<.0001		
Parameter Estimates				
Term	**Estimate**	**Std Error**	**Lower 95%**	**Upper 95%**
Intercept	3.795728	0.078583	3.641709	3.949748
Sex[Female]	0.222715	0.053987	0.116902	0.328528
Age at Dx Recode 5 Groups[20-49]	0.911051	0.171131	0.575640	1.246462
Age at Dx Recode 5 Groups[50-59]	0.741344	0.111520	0.522769	0.959920
Age at Dx Recode 5 Groups[60-69]	0.171881	0.095995	-0.016266	0.360028
Age at Dx Recode 5 Groups[70-79]	-0.244835	0.101088	-0.442965	-0.046706
δ	2.273770	0.060525	2.155143	2.392396
Wald Tests				
Source	**Nparm**	**DF**	**Wald ChiSquare**	**Prob>ChiSq**
Sex	1	1	17.0182	<.0001
Age at Dx Recode 5 Groups	4	4	206.1739	<.0001

Table 3. Model selection details and fitted model diagnostics and parameter estimates for cause-specific survival for 1975-79 stage II colorectal cancer (CRC) in the colon. The highlighted distribution was selected because it resulted in the smallest Akaike information criterion (AICc) and the parameter estimates correspond to that distribution. The Wald confidence interval was used.

1975-79 Stage II CRC, Colon: Model Comparison				
Distribution	**AICc**			
Weibull	26345.42			
Lognormal	26170.66			
Exponential	27542.08			
Frechet	26271.67			
Loglogistic	26247.34			
Observation Used	9215			
Uncensored Values	3172			
Right Censored Values	6043			
Whole Model Test				
ChiSquare	**DF**	**Prob>Chisq**		
210.33	5	<.0001		
Parameter Estimates				
Term	**Estimate**	**Std Error**	**Lower 95%**	**Upper 95%**
Intercept	3.600079	0.046776	3.508399	3.691758
Sex[Female]	0.192775	0.033208	0.127689	0.257860
Age at Dx Recode 5 Groups[20-49]	0.913622	0.108666	0.700641	1.126604
Age at Dx Recode 5 Groups[50-59]	0.207249	0.078138	0.054102	0.360396
Age at Dx Recode 5 Groups[60-69]	0.026601	0.062580	-0.096054	0.149255
Age at Dx Recode 5 Groups[70-79]	-0.260595	0.059840	-0.377880	-0.143311
δ	2.610969	0.036533	2.539365	2.682572
Wald Tests				
Source	**Nparm**	**DF**	**Wald ChiSquare**	**Prob>ChiSq**
Sex	1	1	33.6997	<.0001
Age at Dx Recode 5 Groups	4	4	194.7110	<.0001

Table 4. Model selection details and fitted model diagnostics and parameter estimates for cause-specific survival for 1975-79 stage II colorectal cancer (CRC) in the rectum. The highlighted distribution was selected because it resulted in the smallest Akaike information criterion (AICc) and the parameter estimates correspond to that distribution. The Wald confidence interval was used.

1975-79 Stage II CRC, Rectum: Model Comparison				
Distribution	**AICc**			
Weibull	10796.0836			
Lognormal	10639.6158			
Exponential	11120.6735			
Frechet	10682.5047			
Loglogistic	10686.3145			
Observation Used	3016			
Uncensored Values	1432			
Right Censored Values	1584			
Whole Model Test				
ChiSquare	**DF**	**Prob>Chisq**		
141.3549	5	<.0001		
Parameter Estimates				
Term	**Estimate**	**Std Error**	**Lower 95%**	**Upper 95%**
Intercept	2.527056	0.054762	2.419724	2.634389
Sex[Female]	0.143037	0.042876	0.059002	0.227073
Age at Dx Recode 5 Groups[20-49]	0.538863	0.144432	0.255782	0.821944
Age at Dx Recode 5 Groups[50-59]	0.465769	0.089596	0.290165	0.641373
Age at Dx Recode 5 Groups[60-69]	0.229271	0.076962	0.078427	0.380115
Age at Dx Recode 5 Groups[70-79]	-0.099496	0.079690	-0.255685	0.056693
δ	2.042596	0.041566	1.961128	2.124063
Wald Tests				
Source	**Nparm**	**DF**	**Wald ChiSquare**	**Prob>ChiSq**
Sex	1	1	11.1294	0.0008
Age at Dx Recode 5 Groups	4	4	145.4905	<.0001

Table 5. Model selection details and fitted model diagnostics and parameter estimates for cause-specific survival for 1975-79 stage III colorectal cancer (CRC) in the colon. The highlighted distribution was selected because it resulted in the smallest Akaike information criterion (AICc) and the parameter estimates correspond to that distribution. The Wald confidence interval was used.

1975-79 Stage III CRC, Colon: Model Comparison				
Distribution	**AICc**			
Weibull	23746.7629			
Lognormal	23031.5681			
Exponential	25518.8877			
Frechet	22867.6911			
Loglogistic	23169.7505			
Observation Used	6205			
Uncensored Values	3667			
Right Censored Values	2538			
Whole Model Test				
ChiSquare	**DF**	**Prob>Chisq**		
139.7007	4	<.0001		
Parameter Estimates				
Term	**Estimate**	**Std Error**	**Lower 95%**	**Upper 95%**
Intercept	0.972896	0.034369	0.905534	1.040258
Age at Dx Recode 5 Groups[20-49]	0.382867	0.083996	0.218238	0.547496
Age at Dx Recode 5 Groups[50-59]	0.259110	0.060790	0.139964	0.378256
Age at Dx Recode 5 Groups[60-69]	0.178915	0.052158	0.076687	0.281143
Age at Dx Recode 5 Groups[70-79]	-0.174654	0.051553	-0.275695	-0.073612
σ	2.200354	0.026582	2.148254	2.252454
Wald Tests				
Source	**Nparm**	**DF**	**Wald ChiSquare**	**Prob>ChiSq**
Age at Dx Recode 5 Groups	4	4	139.0287	<.0001

Table 6. Preliminary model parameter estimates for the cause-specific survival for 1975-79 stage III colorectal cancer (CRC) in the colon. Parameter estimates correspond to the highlighted distribution in **Table S14**, demonstrating that sex is not a significant covariate. The Wald confidence interval was used.

1975-79 Stage III CRC, Colon: Preliminary Model Comparison				
Parameter Estimates				
Term	**Estimate**	**Std Error**	**Lower 95%**	**Upper 95%**
Intercept	0.973078	0.034460	0.905538	1.040618
Sex[Female]	-0.002081	0.028599	-0.058135	0.053972
Age at Dx Recode 5 Groups[20-49]	0.382792	0.084002	0.218150	0.547433
Age at Dx Recode 5 Groups[50-59]	0.258967	0.060821	0.139760	0.378175
Age at Dx Recode 5 Groups[60-69]	0.178682	0.052256	0.076263	0.281102
Age at Dx Recode 5 Groups[70-79]	-0.174549	0.051573	-0.275630	-0.073468
σ	2.200347	0.026582	2.148247	2.252448
Wald Tests				
Source	**Nparm**	**DF**	**Wald ChiSquare**	**Prob>ChiSq**
Sex	1	1	0.0053	0.9420
Age at Dx Recode 5 Groups	4	4	137.6711	<.0001

Table 7. Model selection details and fitted model diagnostics and parameter estimates for cause-specific survival for 1975-79 stage III colorectal cancer (CRC) in the rectum. The highlighted distribution was selected because it resulted in the smallest Akaike information criterion (AICc) and the parameter estimates correspond to that distribution. The Wald confidence interval was used.

1975-79 Stage III CRC, Rectum: Model Comparison				
Distribution	**AICc**			
Weibull	11003.1542			
Lognormal	10616.6864			
Exponential	11431.3049			
Frechet	10616.7219			
Loglogistic	10621.7841			
Observation Used	2640			
Uncensored Values	1764			
Right Censored Values	876			
Whole Model Test				
ChiSquare	**DF**	**Prob>Chisq**		
64.0900	5	<.0001		
Parameter Estimates				
Term	**Estimate**	**Std Error**	**Lower 95%**	**Upper 95%**
Intercept	1.549157	0.041109	1.468586	1.629729
Sex[Female]	0.113776	0.035208	0.044770	0.182783
Age at Dx Recode 5 Groups[20-49]	0.409712	0.101801	0.210185	0.609238
Age at Dx Recode 5 Groups[50-59]	0.280006	0.068825	0.145111	0.414901
Age at Dx Recode 5 Groups[60-69]	0.031083	0.061877	-0.090193	0.152359
Age at Dx Recode 5 Groups[70-79]	-0.171752	0.066091	-0.301289	-0.042216
σ	1.675403	0.029953	1.616696	1.734110
Wald Tests				
Source	**Nparm**	**DF**	**Wald ChiSquare**	**Prob>ChiSq**
Sex	1	1	10.4428	0.0012
Age at Dx Recode 5 Groups	4	4	57.8493	<.0001

Table 8. Model selection details and fitted model diagnostics and parameter estimates for cause-specific survival for 1975-79 stage IV colorectal cancer (CRC) in the colon. The highlighted distribution was selected because it resulted in the smallest Akaike information criterion (AICc) and the parameter estimates correspond to that distribution. DNC = did not converge. The Wald confidence interval was used.

1975-79 Stage IV CRC, Colon: Model Comparison				
Distribution	**AICc**			
Weibull	14411.1196			
Lognormal	12728.1446			
Exponential	DNC			
Frechet	12705.7556			
Loglogistic	12712.6132			
Observation Used	6853			
Uncensored Values	6209			
Right Censored Values	644			
Whole Model Test				
ChiSquare	**DF**	**Prob>Chisq**		
268.2056	5	<.0001		
Parameter Estimates				
Term	**Estimate**	**Std Error**	**Lower 95%**	**Upper 95%**
Intercept	-1.361909	0.019978	-1.401066	-1.322753
Sex[Female]	0.036882	0.016846	0.003865	0.069900
Age at Dx Recode 5 Groups[20-49]	0.259513	0.052021	0.157554	0.361473
Age at Dx Recode 5 Groups[50-59]	0.312021	0.036765	0.239963	0.384078
Age at Dx Recode 5 Groups[60-69]	0.068881	0.030781	0.008550	0.129211
Age at Dx Recode 5 Groups[70-79]	-0.111060	0.030466	-0.170773	-0.051347
σ	1.372644	0.013103	1.346962	1.398326
Wald Tests				
Source	**Nparm**	**DF**	**Wald ChiSquare**	**Prob>ChiSq**
Sex	1	1	4.7934	0.0286
Age at Dx Recode 5 Groups	4	4	264.7558	<.0001

Table 9. Model selection details and fitted model diagnostics and parameter estimates for cause-specific survival for 1975-79 stage IV colorectal cancer (CRC) in the rectum. The highlighted distribution was selected because it resulted in the smallest Akaike information criterion (AICc) and the parameter estimates correspond to that distribution. DNC = did not converge. The Wald confidence interval was used.

1975-79 Stage IV CRC, Rectum: Model Comparison				
Distribution	**AICc**			
Weibull	5536.4062			
Lognormal	5070.5617			
Exponential	DNC			
Frechet	5254.9737			
Loglogistic	5065.0192			
Observation Used	2408			
Uncensored Values	2212			
Right Censored Values	196			
Whole Model Test				
ChiSquare	**DF**	**Prob>Chisq**		
143.4434	4	<.0001		
Parameter Estimates				
Term	**Estimate**	**Std Error**	**Lower 95%**	**Upper 95%**
Intercept	-0.509357	0.031974	-0.572025	-0.446688
Age at Dx Recode 5 Groups[20-49]	0.415409	0.084990	0.248831	0.581987
Age at Dx Recode 5 Groups[50-59]	0.316781	0.056604	0.205838	0.427723
Age at Dx Recode 5 Groups[60-69]	0.125683	0.050379	0.026942	0.224424
Age at Dx Recode 5 Groups[70-79]	-0.096889	0.053525	-0.201797	0.008018
σ	0.785458	0.013837	0.758339	0.812577
Wald Tests				
Source	**Nparm**	**DF**	**Wald ChiSquare**	**Prob>ChiSq**
Age at Dx Recode 5 Groups	4	4	149.8985	<.0001

Table 10. Preliminary model parameter estimates for the cause-specific survival for 1975-79 stage IV colorectal cancer (CRC) in the rectum. Parameter estimates correspond to the highlighted distribution in **Table S18**, demonstrating that sex is not a significant covariate. The Wald confidence interval was used.

1975-79 Stage IV CRC, Rectum: Preliminary Model Comparison				
Parameter Estimates				
Term	**Estimate**	**Std Error**	**Lower 95%**	**Upper 95%**
Intercept	-0.511735	0.032056	-0.574563	-0.448906
Sex[Female]	-0.027858	0.028592	-0.083898	0.028181
Age at Dx Recode 5 Groups[20-49]	0.413433	0.084969	0.246896	0.579970
Age at Dx Recode 5 Groups[50-59]	0.312648	0.056746	0.201428	0.423869
Age at Dx Recode 5 Groups[60-69]	0.122855	0.050451	0.023973	0.221737
Age at Dx Recode 5 Groups[70-79]	-0.095562	0.053518	-0.200456	0.009332
σ	0.785255	0.013834	0.758140	0.812370
Wald Tests				
Source	**Nparm**	**DF**	**Wald ChiSquare**	**Prob>ChiSq**
Sex	1	1	0.9493	0.3299
Age at Dx Recode 5 Groups	4	4	144.2795	<.0001

Table 11. Model selection details and fitted model diagnostics and parameter estimates for cause-specific survival for 2000-03 stage I colorectal cancer (CRC) in the colon. The highlighted distribution was selected because it resulted in the smallest Akaike information criterion (AICc) and the parameter estimates correspond to that distribution. Parameter δ = 1/κ. The Wald confidence interval was used.

2000-03 Stage I CRC, Colon: Model Comparison				
Distribution	**AICc**			
Weibull	6712.5586			
Lognormal	6732.2328			
Exponential	6820.2702			
Frechet	6778.3830			
Loglogistic	6714.0577			
Observation Used	6054			
Uncensored Values	645			
Right Censored Values	5409			
Whole Model Test				
ChiSquare	**DF**	**Prob>Chisq**		
241.2994	5	<.0001		
Parameter Estimates				
Term	**Estimate**	**Std Error**	**Lower 95%**	**Upper 95%**
Intercept	5.879682	0.163628	5.558976	6.200387
Sex[Female]	0.132082	0.057253	0.019869	0.244295
Age at Dx Recode 5 Groups[20-49]	1.609704	0.352986	0.917864	2.301544
Age at Dx Recode 5 Groups[50-59]	0.649916	0.180608	0.295931	1.003902
Age at Dx Recode 5 Groups[60-69]	-0.011486	0.141261	-0.288353	0.265381
Age at Dx Recode 5 Groups[70-79]	-0.486255	0.128291	-0.737700	-0.234809
δ	1.430413	0.052380	1.327749	1.533077
Wald Tests				
Source	**Nparm**	**DF**	**Wald ChiSquare**	**Prob>ChiSq**
Sex	1	1	5.3222	0.0211
Age at Dx Recode 5 Groups	4	4	204.3100	<.0001

Table 12. Model selection details and fitted model diagnostics and parameter estimates for cause-specific survival for 2000-03 stage I colorectal cancer (CRC) in the rectum. The highlighted distribution was selected because it resulted in the smallest Akaike information criterion (AICc) and the parameter estimates correspond to that distribution. The Wald confidence interval was used.

2000-03 Stage I CRC, Rectum: Model Comparison				
Distribution	**AICc**			
Weibull	4427.7791			
Lognormal	4431.6023			
Exponential	4439.5690			
Frechet	4462.4581			
Loglogistic	4425.4890			
Observation Used	2827			
Uncensored Values	439			
Right Censored Values	2388			
Whole Model Test				
ChiSquare	**DF**	**Prob>Chisq**		
241.2994	5	<.0001		
Parameter Estimates				
Term	**Estimate**	**Std Error**	**Lower 95%**	**Upper 95%**
Intercept	4.247524	0.106285	4.039210	4.455839
Age at Dx Recode 5 Groups[20-49]	0.470342	0.163954	0.148998	0.791685
Age at Dx Recode 5 Groups[50-59]	0.703278	0.131565	0.445416	0.961140
Age at Dx Recode 5 Groups[60-69]	0.159296	0.110116	-0.056527	0.375118
Age at Dx Recode 5 Groups[70-79]	-0.256080	0.104809	-0.461501	-0.050659
δ	1.107932	0.048396	1.013078	1.202787
Wald Tests				
Source	**Nparm**	**DF**	**Wald ChiSquare**	**Prob>ChiSq**
Age at Dx Recode 5 Groups	4	4	96.8821	<.0001

Table 13. Preliminary model parameter estimates for the cause-specific survival for 2000-03 stage I colorectal cancer (CRC) in the rectum. Parameter estimates correspond to the highlighted distribution in **Table S20**, demonstrating that sex is not a significant covariate. The Wald confidence interval was used.

2000-03 Stage I CRC, Rectum: Preliminary Model Comparison				
Parameter Estimates				
Term	**Estimate**	**Std Error**	**Lower 95%**	**Upper 95%**
Intercept	4.258682	0.107017	4.048934	4.468431
Sex[Female]	0.087242	0.059925	-0.030209	0.204693
Age at Dx Recode 5 Groups[20-49]	0.476808	0.163836	0.155696	0.797920
Age at Dx Recode 5 Groups[50-59]	0.713406	0.131665	0.455347	0.971464
Age at Dx Recode 5 Groups[60-69]	0.171302	0.110311	-0.044904	0.387507
Age at Dx Recode 5 Groups[70-79]	-0.260059	0.104704	-0.465276	-0.054843
σ	1.106348	0.048312	1.011658	1.201038
Wald Tests				
Source	**Nparm**	**DF**	**Wald ChiSquare**	**Prob>ChiSq**
Sex	1	1	2.1195	0.1454
Age at Dx Recode 5 Groups	4	4	98.7242	<.0001

Table 14. Model selection details and fitted model diagnostics and parameter estimates for cause-specific survival for 2000-03 stage II colorectal cancer (CRC) in the colon. The highlighted distribution was selected because it resulted in the smallest Akaike information criterion (AICc) and the parameter estimates correspond to that distribution. The Wald confidence interval was used.

2000-03 Stage II CRC, Colon: Model Comparison				
Distribution	**AICc**			
Weibull	16392.3346			
Lognormal	16356.1742			
Exponential	16909.8138			
Frechet	16423.1679			
Loglogistic	16372.8197			
Observation Used	8258			
Uncensored Values	1866			
Right Censored Values	6392			
Whole Model Test				
ChiSquare	**DF**	**Prob>Chisq**		
204.1411	5	<.0001		
Parameter Estimates				
Term	**Estimate**	**Std Error**	**Lower 95%**	**Upper 95%**
Intercept	4.501293	0.072495	4.359205	4.643381
Sex[Female]	0.182476	0.042847	0.098496	0.266455
Age at Dx Recode 5 Groups[20-49]	0.619266	0.136899	0.350949	0.887583
Age at Dx Recode 5 Groups[50-59]	0.377763	0.106751	0.168534	0.586992
Age at Dx Recode 5 Groups[60-69]	0.324346	0.089388	0.149149	0.499542
Age at Dx Recode 5 Groups[70-79]	-0.222634	0.076338	-0.372255	-0.073013
δ	2.831562	0.052886	2.727907	2.935217
Wald Tests				
Source	**Nparm**	**DF**	**Wald ChiSquare**	**Prob>ChiSq**
Sex	1	1	18.1369	<.0001
Age at Dx Recode 5 Groups	4	4	206.2314	<.0001

Table 15. Model selection details and fitted model diagnostics and parameter estimates for cause-specific survival for 2000-03 stage II colorectal cancer (CRC) in the rectum. The highlighted distribution was selected because it resulted in the smallest Akaike information criterion (AICc) and the parameter estimates correspond to that distribution. The Wald confidence interval was used.

2000-03 Stage II CRC, Rectum: Model Comparison				
Distribution	**AICc**			
Weibull	5542.9467			
Lognormal	5520.6179			
Exponential	5607.6837			
Frechet	5566.6658			
Loglogistic	5524.4337			
Observation Used	2183			
Uncensored Values	663			
Right Censored Values	1520			
Whole Model Test				
ChiSquare	**DF**	**Prob>Chisq**		
117.9298	4	<.0001		
Parameter Estimates				
Term	**Estimate**	**Std Error**	**Lower 95%**	**Upper 95%**
Intercept	3.467962	0.084494	3.302357	3.633568
Age at Dx Recode 5 Groups[20-49]	0.960835	0.164340	0.638735	1.282935
Age at Dx Recode 5 Groups[50-59]	0.497604	0.122020	0.258449	0.736758
Age at Dx Recode 5 Groups[60-69]	0.114634	0.111704	-0.104301	0.333570
Age at Dx Recode 5 Groups[70-79]	-0.391419	0.105896	-0.598971	-0.183867
δ	2.187381	0.067674	2.054742	2.320019
Wald Tests				
Source	**Nparm**	**DF**	**Wald ChiSquare**	**Prob>ChiSq**
Age at Dx Recode 5 Groups	4	4	119.1489	<.0001

Table 16. Preliminary model parameter estimates for the cause-specific survival for 2000-03 stage II colorectal cancer (CRC) in the rectum. Parameter estimates correspond to the highlighted distribution in **Table S24**, demonstrating that sex is not a significant covariate. The Wald confidence interval was used.

2000-03 Stage II CRC, Rectum: Preliminary Model Comparison				
Parameter Estimates				
Term	**Estimate**	**Std Error**	**Lower 95%**	**Upper 95%**
Intercept	3.467825	0.084915	3.301394	3.634256
Sex[Female]	-0.000969	0.059721	-0.118020	0.116081
Age at Dx Recode 5 Groups[20-49]	0.960789	0.164363	0.638643	1.282934
Age at Dx Recode 5 Groups[50-59]	0.497535	0.122094	0.258236	0.736834
Age at Dx Recode 5 Groups[60-69]	0.114536	0.111868	-0.104721	0.333793
Age at Dx Recode 5 Groups[70-79]	-0.391428	0.105897	-0.598983	-0.183874
σ	2.187376	0.067674	2.054737	2.320016
Wald Tests				
Source	**Nparm**	**DF**	**Wald ChiSquare**	**Prob>ChiSq**
Sex	1	1	0.0003	0.9871
Age at Dx Recode 5 Groups	4	4	118.0013	<.0001

Table 17. Model selection details and fitted model diagnostics and parameter estimates for cause-specific survival for 2000-03 stage III colorectal cancer (CRC) in the colon. The highlighted distribution was selected because it resulted in the smallest Akaike information criterion (AICc) and the parameter estimates correspond to that distribution. The Wald confidence interval was used.

2000-03 Stage III CRC, Colon: Model Comparison				
Distribution	**AICc**			
Weibull	23078.6953			
Lognormal	22728.0115			
Exponential	23978.1744			
Frechet	22747.3516			
Loglogistic	22845.1864			
Observation Used	7326			
Uncensored Values	3284			
Right Censored Values	4042			
Whole Model Test				
ChiSquare	**DF**	**Prob>Chisq**		
344.5836	4	<.0001		
Parameter Estimates				
Term	**Estimate**	**Std Error**	**Lower 95%**	**Upper 95%**
Intercept	2.536067	0.036069	2.465373	2.606762
Age at Dx Recode 5 Groups[20-49]	0.541576	0.081362	0.382109	0.701043
Age at Dx Recode 5 Groups[50-59]	0.426254	0.066332	0.296246	0.556262
Age at Dx Recode 5 Groups[60-69]	0.206712	0.058163	0.092714	0.320710
Age at Dx Recode 5 Groups[70-79]	-0.146281	0.053108	-0.250372	-0.042191
σ	2.204045	0.030060	2.145128	2.262963
Wald Tests				
Source	**Nparm**	**DF**	**Wald ChiSquare**	**Prob>ChiSq**
Age at Dx Recode 5 Groups	4	4	359.3299	<.0001

Table 18. Preliminary model parameter estimates for the cause-specific survival for 2000-03 stage III colorectal cancer (CRC) in the colon. Parameter estimates correspond to the highlighted distribution in **Table S26**, demonstrating that sex is not a significant covariate. The Wald confidence interval was used.

2000-03 Stage III CRC, Colon: Preliminary Model Comparison				
Parameter Estimates				
Term	**Estimate**	**Std Error**	**Lower 95%**	**Upper 95%**
Intercept	2.534884	0.036061	2.464206	2.605561
Sex[Female]	0.057506	0.029917	-0.001131	0.116142
Age at Dx Recode 5 Groups[20-49]	0.547783	0.081409	0.388224	0.707343
Age at Dx Recode 5 Groups[50-59]	0.434101	0.066450	0.303862	0.564341
Age at Dx Recode 5 Groups[60-69]	0.213685	0.058264	0.099489	0.327881
Age at Dx Recode 5 Groups[70-79]	-0.149754	0.053122	-0.253871	-0.045638
σ	2.203402	0.030051	2.144504	2.262300
Wald Tests				
Source	**Nparm**	**DF**	**Wald ChiSquare**	**Prob>ChiSq**
Sex	1	1	3.6947	0.0546
Age at Dx Recode 5 Groups	4	4	361.1771	<.0001

Table 19. Model selection details and fitted model diagnostics and parameter estimates for cause-specific survival for 2000-03 stage III colorectal cancer (CRC) in the rectum. The highlighted distribution was selected because it resulted in the smallest Akaike information criterion (AICc) and the parameter estimates correspond to that distribution. The Wald confidence interval was used.

2000-03 Stage III CRC, Rectum: Model Comparison				
Distribution	**AICc**			
Weibull	9203.2670			
Lognormal	9061.5251			
Exponential	9279.4864			
Frechet	9086.8547			
Loglogistic	9103.2215			
Observation Used	2685			
Uncensored Values	1248			
Right Censored Values	1437			
Whole Model Test				
ChiSquare	**DF**	**Prob>Chisq**		
164.9559	4	<.0001		
Parameter Estimates				
Term	**Estimate**	**Std Error**	**Lower 95%**	**Upper 95%**
Intercept	2.355839	0.043877	2.269841	2.441836
Age at Dx Recode 5 Groups[20-49]	0.470034	0.083975	0.305445	0.634622
Age at Dx Recode 5 Groups[50-59]	0.389915	0.073731	0.245406	0.534424
Age at Dx Recode 5 Groups[60-69]	0.303914	0.073070	0.160700	0.447128
Age at Dx Recode 5 Groups[70-79]	-0.107169	0.072685	-0.249628	0.035290
σ	1.751901	0.038754	1.675944	1.827857
Wald Tests				
Source	**Nparm**	**DF**	**Wald ChiSquare**	**Prob>ChiSq**
Age at Dx Recode 5 Groups	4	4	174.3351	<.0001

Table 20. Preliminary model parameter estimates for the cause-specific survival for 2000-03 stage III colorectal cancer (CRC) in the rectum. Parameter estimates correspond to the highlighted distribution in **Table S28**, demonstrating that sex is not a significant covariate. The Wald confidence interval was used.

2000-03 Stage III CRC, Rectum: Preliminary Model Comparison				
Parameter Estimates				
Term	**Estimate**	**Std Error**	**Lower 95%**	**Upper 95%**
Intercept	2.360388	0.044059	2.274034	2.446742
Sex[Female]	0.048828	0.038920	-0.027454	0.125110
Age at Dx Recode 5 Groups[20-49]	0.472463	0.083950	0.307924	0.637002
Age at Dx Recode 5 Groups[50-59]	0.395694	0.073836	0.250978	0.540410
Age at Dx Recode 5 Groups[60-69]	0.309335	0.073151	0.165962	0.452708
Age at Dx Recode 5 Groups[70-79]	-0.108186	0.072639	-0.250555	0.034183
σ	1.750808	0.038728	1.674903	1.826713
Wald Tests				
Source	**Nparm**	**DF**	**Wald ChiSquare**	**Prob>ChiSq**
Sex	1	1	1.5740	0.2096
Age at Dx Recode 5 Groups	4	4	175.9034	<.0001

Table 21. Model selection details and fitted model diagnostics and parameter estimates for cause-specific survival for 2000-03 stage IV colorectal cancer (CRC) in the colon. The highlighted distribution was selected because it resulted in the smallest Akaike information criterion (AICc) and the parameter estimates correspond to that distribution. The Wald confidence interval was used.

2000-03 Stage IV CRC, Colon: Model Comparison				
Distribution	**AICc**			
Weibull	12638.3198			
Lognormal	11934.9884			
Exponential	14039.0074			
Frechet	12219.7368			
Loglogistic	12048.4684			
Observation Used	5175			
Uncensored Values	4668			
Right Censored Values	507			
Whole Model Test				
ChiSquare	**DF**	**Prob>Chisq**		
403.4377	4	<.0001		
Parameter Estimates				
Term	**Estimate**	**Std Error**	**Lower 95%**	**Upper 95%**
Intercept	-0.376556	0.022984	-0.421604	-0.331508
Age at Dx Recode 5 Groups[20-49]	0.653115	0.055073	0.545174	0.761055
Age at Dx Recode 5 Groups[50-59]	0.319011	0.046280	0.228303	0.409719
Age at Dx Recode 5 Groups[60-69]	0.086081	0.042726	0.002340	0.169823
Age at Dx Recode 5 Groups[70-79]	-0.292951	0.039731	-0.370821	-0.215080
σ	1.558274	0.016318	1.526292	1.590256
Wald Tests				
Source	**Nparm**	**DF**	**Wald ChiSquare**	**Prob>ChiSq**
Age at Dx Recode 5 Groups	4	4	422.7788	<.0001

Table 22. Preliminary model parameter estimates for the cause-specific survival for 2000-03 stage IV colorectal cancer (CRC) in the colon. Parameter estimates correspond to the highlighted distribution in **Table S30**, demonstrating that sex is not a significant covariate. The Wald confidence interval was used.

2000-03 Stage IV CRC, Colon: Preliminary Model Comparison				
Parameter Estimates				
Term	**Estimate**	**Std Error**	**Lower 95%**	**Upper 95%**
Intercept	-0.376659	0.022984	-0.421708	-0.331610
Sex[Female]	0.010891	0.022191	-0.032603	0.054385
Age at Dx Recode 5 Groups[20-49]	0.653505	0.055077	0.545556	0.761454
Age at Dx Recode 5 Groups[50-59]	0.320078	0.046330	0.229272	0.410883
Age at Dx Recode 5 Groups[60-69]	0.087155	0.042781	0.003305	0.171004
Age at Dx Recode 5 Groups[70-79]	-0.292947	0.039730	-0.370815	-0.215079
σ	1.558236	0.016317	1.526255	1.590218
Wald Tests				
Source	**Nparm**	**DF**	**Wald ChiSquare**	**Prob>ChiSq**
Sex	1	1	0.2409	0.6236
Age at Dx Recode 5 Groups	4	4	420.4701	<.0001

Table 23. Model selection details and fitted model diagnostics and parameter estimates for cause-specific survival for 2000-03 stage IV colorectal cancer (CRC) in the rectum. The highlighted distribution was selected because it resulted in the smallest Akaike information criterion (AICc) and the parameter estimates correspond to that distribution. The Wald confidence interval was used.

2000-03 Stage IV CRC, Rectum: Model Comparison				
Distribution	**AICc**			
Weibull	4760.9838			
Lognormal	4604.2003			
Exponential	4958.6266			
Frechet	4801.3275			
Loglogistic	4614.0333			
Observation Used	1680			
Uncensored Values	1519			
Right Censored Values	161			
Whole Model Test				
ChiSquare	**DF**	**Prob>Chisq**		
188.1231	4	<.0001		
Parameter Estimates				
Term	**Estimate**	**Std Error**	**Lower 95%**	**Upper 95%**
Intercept	-0.170694	0.036309	-0.241859	-0.099529
Age at Dx Recode 5 Groups[20-49]	0.588259	0.079296	0.432842	0.743676
Age at Dx Recode 5 Groups[50-59]	0.540838	0.068976	0.405646	0.676029
Age at Dx Recode 5 Groups[60-69]	0.048875	0.067135	-0.082708	0.180458
Age at Dx Recode 5 Groups[70-79]	-0.343684	0.067486	-0.475954	-0.211414
σ	1.434866	0.026359	1.383203	1.486530
Wald Tests				
Source	**Nparm**	**DF**	**Wald ChiSquare**	**Prob>ChiSq**
Age at Dx Recode 5 Groups	4	4	200.4892	<.0001

Table 24. Preliminary model parameter estimates for the cause-specific survival for 2000-03 stage IV colorectal cancer (CRC) in the rectum. Parameter estimates correspond to the highlighted distribution in **Table S32**, demonstrating that sex is not a significant covariate. The Wald confidence interval was used.

2000-03 Stage IV CRC, Rectum: Preliminary Model Comparison				
Parameter Estimates				
Term	**Estimate**	**Std Error**	**Lower 95%**	**Upper 95%**
Intercept	-0.163711	0.036939	-0.236111	-0.091311
Sex[Female]	0.037664	0.036799	-0.034460	0.109789
Age at Dx Recode 5 Groups[20-49]	0.590337	0.079294	0.434923	0.745750
Age at Dx Recode 5 Groups[50-59]	0.545970	0.069134	0.410469	0.681470
Age at Dx Recode 5 Groups[60-69]	0.053490	0.067261	-0.078340	0.185320
Age at Dx Recode 5 Groups[70-79]	-0.343140	0.067464	-0.475366	-0.210914
σ	1.434349	0.026350	1.382704	1.485994
Wald Tests				
Source	**Nparm**	**DF**	**Wald ChiSquare**	**Prob>ChiSq**
Sex	1	1	1.0476	0.3061
Age at Dx Recode 5 Groups	4	4	200.9993	<.0001

Table 25. Coding details for derived variables used for the SEER survival analysis. Survival analysis was performed using survival years (where a month value of 0 was recoded as 0.5) versus age and sex factors, stratified by site and American Joint Committee on Cancer (AJCC) stage. Survival models were selected based on smallest Akaike information criterion (AICc). Analysis performed using JMP v13.0 (SAS Institute).

New Column Name	Description	Code
"Age at Dx Recode 5 Groups"	Re-classifies SEER*STAT "Age recode with <1 year olds" (age at Diagnosis, 17 groups) to 5 groups	Match(:Name("Age recode with <1 year olds"), "01-04 years", "<20", "10-14 years", "<20", "15-19 years", "<20", "20-24 years", "20-49", "25-29 years", "20-49", "30-34 years", "20-49", "35-39 years", "20-49", "40-44 years", "20-49", "45-49 years", "20-49", "50-54 years", "50-59", "55-59 years", "50-59", "60-64 years", "60-69", "65-69 years", "60-69", "70-74 years", "70-79", "75-79 years", "70-79", "80-84 years", ">=80", "85+ years", ">=80", Empty())
"Site: C vs R"	Recodes SEER*STAT "Site Colon vs Rectum" to either Colon or Rectum	Match(:Name("Site recode ICD-O-3/WHO 2008"), "Appendix", "Colon", "Ascending Colon","Colon", "Cecum", "Colon", "Descending Colon", "Colon", "Hepatic Flexure", "Colon", "Large Intestine, NOS", "Colon", "Rectosigmoid Junction","Rectum", "Rectum", "Rectum", "Sigmoid Colon", "Colon", "Splenic Flexure", "Colon", "Transverse Colon", "Colon", Empty())

New Column Name	Description	Code
“Exclude Based on AJCC”	Filter column to exclude subjects where AJCC is Unstaged	Match(:Name("AJCC 5th Ed Schrag Code"), "AJCC 5th Ed Stage I", "NO", "AJCC 5th Ed Stage II", "NO", "AJCC 5th Ed Stage III", "NO", "AJCC 5th Ed Stage IV", "NO", "AJCC 5th Ed Unstaged", "YES", Empty())
“Exclude Age too Young”	Filter column to exclude subjects who are younger than the youngest CISNET Age to be at risk of colorectal cancer (<Age 20)	Match(:Age at Dx Recode 5 Groups, "<20", "YES", ">=80", "NO", "20-49", "NO", "50-59", "NO", "60-69", "NO", "70-79", "NO", Empty())
“ExcludeIs1.Survival.Months”	Filter column to exclude from analysis where survival values could not reliably be calculated and there could be 0 days of follow-up. Uses SEER*STAT survival months flag column. If survival month dates were incomplete and there could be 0 days of follow-up, then data was excluded. We included data where complete dates were available.	Match(:Survival months flag, "Complete dates are available and there are 0 days of survival", "0", "Complete dates are available and there are more than 0 days of survival", "0", "Incomplete dates are available and there cannot be zero days of follow-up", "0", "Incomplete dates are available and there could be zero days of follow-up", "1", Empty())
“Surv.Mo.0.as.05”	When sufficient documentation existed for a recorded 0-month survival, the 0-month survival was recoded as 0.5 month survival (so it is not censored in survival analysis). Note: this can alter the type of parametric regression that is ultimately selected.	If(:ExcludeIs1.Survival.Months == "0" & :Survival months == 0, 0.5, If(:ExcludeIs1.Survival.Months == "0" & :Survival months != 0, :Survival months, Empty()))
“Surv.Year.MonthRecoded”	Calculates survival years using the “Surv.Mo.0.as.05” column	:Surv.Mo.0.as0.5 / 12

New Column Name	Description	Code
“CensorEquals1”	If the individual is recorded as still be alive at the end of the follow-up or has died of other causes, then the individual is considered as a Censor candidate (pending not being excluded due to other factors listed above).	Match(:Name("SEER cause-specific death classification"), "Alive or dead of other cause", 1, "Dead (attributable to this cancer dx)", 0, Empty())

537

References

1. Rutter CM, Johnson EA, Feuer EJ, Knudsen AB, Kuntz KM, Schrag D. Secular trends in colon and rectal cancer relative survival. *J Natl Cancer Inst.* 2013;105(23):1806-1813.
2. Rutter CM, Savarino JE. An evidence-based microsimulation model for colorectal cancer: validation and application. *Cancer Epidemiol Biomarkers Prev.* 2010;19(8):1992-2002.
3. CISNET Colorectal Cancer Collaborators. RAND Corporation (CRC-SPIN), 2015. HI.001.03112015.70373:https://cisnet.cancer.gov/colorectal/profiles.html.
4. Knudsen AB, Zauber AG, Rutter CM, et al. Estimation of Benefits, Burden, and Harms of Colorectal Cancer Screening Strategies: Modeling Study for the US Preventive Services Task Force. *JAMA.* 2016;315(23):2595-2609.
5. Schrag D. AJCC Staging: Staging Colon Cancer Patients using the TNM system from 1975 to the present. 2007; https://www.mskcc.org/departments/epidemiology-biostatistics/epidemiology/ajcc-staging.
6. Amin MB, Greene FL, Edge SB, et al. The Eighth Edition AJCC Cancer Staging Manual: Continuing to build a bridge from a population-based to a more "personalized" approach to cancer staging. *CA Cancer J Clin.* 2017;67(2):93-99.
7. Tong GJ, Zhang GY, Liu J, et al. Comparison of the eighth version of the American Joint Committee on Cancer manual to the seventh version for colorectal cancer: A retrospective review of our data. *World J Clin Oncol.* 2018;9(7):148-161.