

Calculation of Survival Time Fields

Study cut-off date is a pre-determined date for the submission (e.g. 12/31/2010 for data with “complete” follow-up through 2010). All deaths up to this date should be recorded in the data through death clearance linkages. For example, the SEER November 2012 submission databases will contain complete death data through 2010, therefore, 12/31/2010 will be the study cut-off date.

Survival fields based on date of last contact:

Date of last contact = min(date of last contact, study cutoff date). Any follow-up beyond the study cut-off will be ignored for these fields.

Survival Months: a 4-digit field containing complete months of survival.

$$\text{Survival Months} = \text{floor}((\text{date last contact} - \text{date dx}) / \text{days in a month})$$

Days in a month = 365.24/12. Both dates will be complete when this is calculated – see algorithm below for rules to impute missing date components.

Survival Months Flag: a 1-digit field, see below for Flag Values.

Survival fields using presumed alive assumption:

Date of last contact = min(date of last contact, study cutoff date) for patients with vital status of dead or unknown.

Date of last contact = study cutoff date for patients coded as alive.

Survival Months - Presumed Alive: a 4-digit field containing complete months of survival.

$$\text{Survival Months} = \text{floor}((\text{date last contact} - \text{date dx}) / \text{days in a month})$$

Days in a month = 365.24/12. Both dates will be complete – see algorithm below for rules to impute missing date components.

Survival Months Flags – Presumed Alive: a 1-digit field, see below for Flag Values.

Flag Values:

0=complete dates and 0 days of survival; (i.e. date last contact = date dx)

1=complete dates and more than 0 days of survival; (i.e. date last contact > date dx)

2=incomplete dates and could be zero days of survival (i.e. all components that are known in both dates are equal, e.g. 99/99/2010 and 12/31/2010)

3=incomplete dates and must be more than 0 days of survival (i.e. at least one difference in known date components, e.g. 11/99/2010 and 12/31/2010)

9=Unknown/could not be calculated

Blank=Not coded

Note: If any tumor has flag = 1 or 3 (survival > 0 days), then any prior tumor must also be coded as survival > 0 days.

Calculation of Survival Time Fields

Algorithm for assigning values for missing components of dates:

All records for a patient are considered when assigning values to missing date components. Each record is processed in chronological order based on date of diagnosis and sequence number.

Step 1: If any date has a missing day with known year and month, day is set to the midpoint of the earliest possible value and the latest possible value. The earliest possible value = max(1, day of diagnosis for a prior tumor that occurred in the same month and year). The latest possible value = min(number of days in that calendar month, day of diagnosis of a subsequent tumor or day of last contact in the same month and year) .

$$\text{Midpoint} = \text{floor}((\text{day1} + \text{day2}) / 2)$$

Step 2: At this point all days have been assigned for dates with known month and year. If any date has a missing month with known year, month and day are set to the midpoint of the earliest possible value and the latest possible value. The earliest possible value = max(Jan 1, month/day of diagnosis for a prior tumor that occurred in the same year). The latest possible value = min(Dec 31, month/day of diagnosis of a subsequent tumor or last contact in the same year) .

$$\text{Midpoint} = \text{floor}((\text{date1} + \text{date2}) / 2)$$

Months is set to 9999 and the flag is set to 9 if year of diagnosis or year of last contact is missing, or date of last contact is prior to date of diagnosis.