

Improved Estimates of Cancer-Specific Survival Rates From Population-Based Data

Nadia Howlader, Lynn A. G. Ries, Angela B. Mariotto, Marsha E. Reichman, Jennifer Ruhl, Kathleen A. Cronin

Manuscript received January 8, 2010; revised August 20, 2010; accepted August 25, 2010.

Correspondence to: Nadia Howlader, MS, Data Analysis and Interpretation Branch, Surveillance Research Program, Division of Cancer Control and Population Sciences, National Cancer Institute, 6116 Executive Blvd, Ste 504, Bethesda, MD 20892-8315 (e-mail: howlader@mail.nih.gov).

Background Accurate estimates of cancer survival are important for assessing optimal patient care and prognosis. Evaluation of these estimates via relative survival (a ratio of observed and expected survival rates) requires a population life table that is matched to the cancer population by age, sex, race and/or ethnicity, socioeconomic status, and ideally risk factors for the cancer under examination. Because life tables for all subgroups in a study may be unavailable, we investigated whether cause-specific survival could be used as an alternative for relative survival.

Methods We used data from the Surveillance, Epidemiology, and End Results Program for 2 330 905 cancer patients from January 1, 1992, through December 31, 2004. We defined cancer-specific deaths according to the following variables: cause of death, only one tumor or the first of multiple tumors, site of the original cancer diagnosis, and comorbidities. Estimates of relative survival and cause-specific survival that were derived by use of an actuarial method were compared.

Results Among breast cancer patients who were white, black, or of Asian or Pacific Islander descent and who were older than 65 years, estimates of 5-year relative survival (107.5%, 106.6%, and 103.0%, respectively) were higher than estimates of 5-year cause-specific survival (98.6%, 95% confidence interval [CI] = 98.4% to 98.8%; 97.4%, 95% CI = 96.2% to 98.2%; and 99.2%, 95% CI = 98.4%, 99.6%, respectively). Relative survival methods likely underestimated rates for cancers of the oral cavity and pharynx (eg, for white cancer patients aged ≥ 65 years, relative survival = 54.2%, 95% CI = 53.1% to 55.3%, and cause-specific survival = 60.1%, 95% CI = 59.1% to 60.9%) and the lung and bronchus (eg, for black cancer patients aged ≥ 65 years, relative survival = 10.5%, 95% CI = 9.9% to 11.2%, and cause-specific survival = 11.9%, 95% CI = 11.2% to 12.6%), largely because of mismatches between the population with these diseases and the population used to derive the life table. Socioeconomic differences between groups with low and high status in relative survival estimates appeared to be inflated (eg, corpus and uterus socioeconomic status gradient was 13.3% by relative survival methods and 8.8% by cause-specific survival methods).

Conclusion Although accuracy of the cause of death on a death certificate can be problematic for cause-specific survival estimates, cause-specific survival methods may be an alternative to relative survival methods when suitable life tables are not available.

J Natl Cancer Inst 2010;102:1584–1598

Cancer survival estimation is essential for assessing prognosis and improvements in cancer care (eg, performance of a new cancer drug and management of patients). Cancer survival can be measured in several ways depending on the question being examined. Most often, researchers or patients are interested in net survival, which filters out the effect of mortality from causes other than the disease in question (by treating deaths from other causes as censored observations) and estimates the probability of surviving cancer in the absence of other causes of death. Net survival provides a useful measure for tracking survival across time and for comparing populations with different life expectancies. Two methods are available to calculate net survival:

cause-specific survival and relative survival (ie, the ratio of observed survival to the expected survival from a comparable group in the general population). If reliable information on the cause of death is available, an analysis of cause-specific survival can be performed by use of standard survival methods in which deaths attributed to the disease of interest are treated as events, and deaths from other causes are treated as censored observations. Analyses of cause-specific survival are used in clinical trial settings in which reliable information on the cause of death is available. However, in population-based cancer registries, cause of death is obtained from death certificates, which are often inaccurately reported (1). For example, cancer at the

site of recurrence or metastasis may be reported as the cause of death instead of cancer at the primary site. Therefore, analysis of cancer registry data usually relies on relative survival methods that do not require cause of death information. Relative survival is calculated by use of life tables and is defined as the ratio of the observed to expected survival rates (2,3). Observed survival is estimated by use of the actuarial or Kaplan–Meier method and uses all causes of death as an event. The expected survival is obtained from national life table and is compared with the observed survival after matching for age, sex, year of diagnosis, and race. For example, if the 5-year observed all-cause survival rate was 80.3% for women diagnosed with malignant breast cancer, and the expected survival rate was 90.3% for a comparable population with similar characteristics (eg, sex, age, and race), then the 5-year relative survival rate for breast cancer patients would be 89.0% (ie, the observed all-cause survival rate divided by the expected survival rate). The relative survival rate in the example represents an adjusted survival rate that does not take into account causes of death other than breast cancer. Additional information about net, cause-specific, and relative survival can be found at <http://srab.cancer.gov/survival/measures.html>.

The National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) Program currently collects cancer incidence and survival information from 17 geographic areas that encompass nearly 26% of the total US population (4); the SEER program is considered a benchmark for cancer survival surveillance in the United States. SEER-reported survival estimates generally use the relative survival method (2). Expected rates for relative survival calculation are derived from a life table of the general population (including white and black individuals and those of other races). The reliability of relative survival is diminished when life tables do not represent mortality of all groups in the population under study. This lack of representativeness occurs when other-cause mortality in the cancer patient cohort under study differs from that of the reference population as a result of various factors that are associated with other-cause mortality either positively (eg, use of screening or higher socioeconomic status) or negatively (eg, lower socioeconomic status or smoking behavior). The Centers for Disease Control and Prevention's National Center for Health Statistics publishes life tables for white persons and black persons but not for those from other racial groups such as American Indian or Asian groups. Moreover, a National Center for Health Statistics study (5) that compared self-reported race or ethnicity from a survey with race reported on death certificates found a high degree of race misclassification for minorities. As a consequence, data in the numerator (total death) and denominator (total population) offset each other, and so mortality rates were underestimated for American Indians by 21% and Asian or Pacific Islanders by 11% (5). This problem with racial misclassification poses additional challenge for the estimation of life tables for these special populations. As a result, a relative survival approach carries a strong potential for biased estimates for minority subgroups. Cause-specific survival is an alternative approach that has not been used systematically in population-based registries because of concerns about misclassification of cause of deaths on death certificates (6) and variations in how specific cause of deaths are associated with cancer diagnoses under different classification schemes.

CONTEXT AND CAVEATS

Prior knowledge

Accurate estimates of cancer survival are important for various analyses of patient treatment and prognosis. Relative survival (a ratio of observed and expected survival rates) can provide such estimates but requires that the population used to generate the expected survival rates match that of the specific cancer being studied. Life tables for all subgroups may not be available.

Study design

Estimates from cause-specific survival methods were compared with those of relative survival methods to determine whether cause-specific survival could be used when life tables for a specific subgroup are not available. Data were from the Surveillance, Epidemiology, and End Results Program.

Contribution

For breast cancer, relative survival estimates were higher than cause-specific survival estimates. Relative survival methods likely underestimated rates for few other cancers, largely because of mismatches between the population with that cancer and the population used to derive the life table. However, for most cancers, the relative survival approach produced accurate estimates that were similar to the estimates produced by cause-specific survival approach.

Implications

Although cause-specific survival estimates use cause of death from death certificates, which can be problematic, such methods may be an alternative to relative survival methods when suitable life tables are not available.

Limitations

Cause of death information and race or ethnicity from death certificates may not be accurate.

From the Editors

For these reasons, we have developed a new classification variable, the SEER cause-specific death classification variable, to obtain estimates for cause-specific survival. This variable includes mainly cancer and some noncancer causes of deaths (eg, AIDS and/or site-related diseases).

Study Subjects and Methods

Study Populations

We used data from 13 geographic SEER areas that include the states of Alaska (Alaska Natives), Connecticut, Hawaii, Iowa, New Mexico, and Utah; the metropolitan areas of Atlanta, Detroit, Seattle–Puget Sound, San Francisco–Oakland, Los Angeles, San Jose–Monterey, and Rural Georgia. Mortality data were obtained from the National Center for Health Statistics, which collects data from all state vital health reporting systems and records detailed cause of deaths from death certificates (<http://www.cdc.gov/nchs/>). Data were included for diagnoses from January 1, 1992, through December 31, 2004, and for follow-up of patients from January 1, 1992, through December 31, 2005. Total number of study subjects was 2 330 905. We excluded cancer patients whose initial diagnosis was found on the death certificate or at autopsy, patients who were not under active follow-up or alive with no survival time, and

patients who had no or unknown microscopic confirmation of their cancer. Patients with unknown or missing cause of death also were excluded (representing <1% of overall cancer patients) because it would not be possible to classify such deaths as a death caused by the cancer (total number of patients who were excluded were 17 396, and the total number of patients included in this study were 2 330 905). Cancer site and morphology were coded according to the *International Classification of Diseases for Oncology*, Second Edition (*ICD-O-2*) or Third Edition (*ICD-O-3*), depending on the year of diagnosis. Age in years when the patient was first diagnosed with cancer was obtained from the SEER database (7).

Race information on newly diagnosed cancer patients was derived for white persons, black persons, American Indians or Alaska Natives, and Asians or Pacific Islanders. Survival rates for the seven main Asian or Pacific Islander groups are presented in this report, including Asian Indian or Pakistani, Chinese, Filipino, Korean, Japanese, Vietnamese, and Hawaiian persons (8). In addition, the other Asian group included individuals of Laotian, Hmong, Kampuchean, Thai, and Asian, not otherwise specified, descent. Similarly, the other Pacific Islander group included individuals of Micronesian, Chamorro, Guamanian, Tahitian, Samoan, Tongan, Melanesian, Fiji Islander, New Guinean, and Pacific Islander, not otherwise specified, descent.

The SEER historic staging scheme provided information for in situ and invasive cancers, with the invasive cancers being divided into the following four stage categories: localized to the primary tumor site, tumor with regional spread or metastases to regional lymph node, tumor with distant metastases, or unknown stage.

Poverty data (obtained from 2000 census data) that were collected at the census-tract level was used as a surrogate for socioeconomic disparity. Cut points that were based on empiric research and policy relevance (9,10) were used to create the tertile distribution for poverty level (ie, <10% for high socioeconomic status, 10%–19.99% for medium socioeconomic status, and ≥20% for low socioeconomic status). Areas with a poverty level of 20% or higher were considered to be severely disadvantaged. Tract-level information from 2000 census data was not used before 1996 for this study, and missing data at the tract level ranged from 2.0% to 0.6% during the study period. Estimates of relative survival and cause-specific survival were computed (as described below) for all three socioeconomic status categories.

SEER Cause-Specific Death Classification Variable

The SEER cause-specific death classification variable is defined by taking into account cause of death in conjunction with sequence of tumor occurrence (ie, only one tumor or the first of multiple tumors), site of the original cancer diagnosis, and comorbidities (eg, AIDS and/or site-related diseases), with the aim of capturing deaths that were related to the specific cancer but were not coded as such to provide guidance as to which deaths should be “attributable” to a specific cancer diagnosis. To determine the optimal classification, cause-specific survival estimates for different potential classification schemes were compared with relative survival estimates in situations with accurate life tables for cancer cohorts in this study.

The *ICD*, Tenth Edition (*ICD-10*) was used to code cause of deaths beginning on January 1, 1999; the *ICD*, Ninth Edition (*ICD-9*)

was used from January 1, 1979, through December 31, 1998; and *ICD*, Eighth Edition (*ICD-8*) was used from January 1, 1975, through December 31, 1978. Previous studies (6,11) have shown that the accuracy of reported underlying cause of death varies substantially, depending on the site of initial diagnosis. Also, comorbidities may add ambiguity to the cause of death. For example, certain cancer types—such as Kaposi sarcoma and non-Hodgkin lymphoma, and less frequently cancers of the lung, mouth, cervix, and digestive system—are more likely to occur in people who are infected with HIV. Thus, AIDS (codes B210–B219) and HIV deaths (codes B20 and B22–B24), as well as deaths related to diseases of the specific site, are considered in the SEER cause-specific death classification variable. Because there is more ambiguity in the cause of death for patients diagnosed with more than one cancer, the SEER cause-specific death classification variable uses two different schemas as described below, depending on whether the index tumor is an individual’s first and only tumor or the first of multiple tumors. Note that SEER multiple primary tumors are defined as the diagnosis of two or more independent reportable neoplasms in an individual (as opposed to multiple concurrent tumors or a sequence of recurrence events).

Cancer Patients Diagnosed With One and Only One Cancer.

For this group of patients, a death was classified as being related to the specific cancer if it was a cancer death attributed to the same cancer site, a cancer death from within the general organ system as specified by *ICD-O-3* (eg, oral cavity and pharynx compared with lip), a cancer death from all other malignant cancers, or a death from AIDS with cancer. The rationale for including all malignant cancers was that among patients diagnosed with only one cancer, cause of death that was coded to another cancer site likely was a misclassification (eg, death from metastatic disease). Also, for some cancer sites, deaths coded as HIV related and deaths attributed to noncancer diseases that were related to the site of first cancer diagnosis (eg, death from liver cirrhosis for patient diagnosed with primary liver cancer) were contributing causes to the SEER cause-specific death classification variable.

First Cancer for Patients With More Than One Cancer Diagnosis.

For individuals who had more than one cancer, cause-specific survival was calculated only for the first cancer diagnosed. The contributing causes of deaths to the SEER cause-specific death classification variable were a cancer death that was attributed to the particular cancer (same site as the first diagnosis), a cancer death that was attributed to the general organ system of the study site, or a cancer death that was attributed to multiple cancers with unknown primary. Deaths from cancer at a selected site that were coded as being attributed to AIDS and cancer, to HIV-related causes, or to noncancer diseases that were related to the site of first cancer diagnosis were also classified as deaths related to the first cancer. Deaths from all other malignant cancers were not classified as cancer deaths and, therefore, were considered to be censored events in cause-specific survival calculations. The rationale for this exclusion was that if a patient was diagnosed with colorectal cancer with a subsequent diagnosis of breast cancer and then died of the latter diagnosis, it was likely that the death was attributed correctly to breast cancer. A detailed description of the *ICD* codes used to

define the SEER cause-specific death classification variable can be found on the SEER Web site (<http://seer.cancer.gov/causespecific/index.html>). Detailed *ICD-10* codes for selected cancers in this study are in Appendix Tables 1 and 2. The SEER cause-specific death classification variable is now available in SEER*Stat, version 6.4.4, software (12) and can be used to generate cause-specific survival estimates.

Statistical Analysis

Relative survival was calculated by actuarial method as the ratio of observed all-cause survival to expected survival (2). Expected survival rates were calculated by use of 1990 and 2000 US decennial life tables that were matched on age, sex, year of diagnosis, and race (white, black, and other). Because life tables are not available for more specific races (eg, American Indians or Alaska Natives and Asian American persons), the life table for other races was used to generate relative survival estimates for these race groups. This practice generally is not advised but was used in this study for comparison.

Cause-specific survival was calculated by use of the actuarial method and used the SEER cause-specific death classification variable as the endpoint. Deaths that were attributed to the cancer according to the defined SEER cause-specific death classification variable were treated as deaths (Appendix Tables 1 and 2), and other deaths were considered censoring events. Survival times were measured in months and were censored at the date of a patient being lost to follow-up, the date of death from causes not considered as deaths due to the cancer, or on December 31, 2005, whichever occurred first. Standard errors were generated by use of Greenwood's formula (13). The 95% confidence intervals (CIs) for the survivorship function were produced with a formula that was based on log-log transformation (14).

It should be noted that the cause-specific survival estimate that treated deaths from other causes as a censoring event was a valid estimate. Because the analyses did not provide evidence to the contrary, we propose that this assumption was reasonable but acknowledge that it remains unverified.

Results

A total of 2 330 905 individuals were diagnosed with an in situ or malignant cancer in the 13 SEER areas from January 1, 1992, through December 31, 2004. We determined the contributions of different causes of deaths by use of the SEER cause-specific death classification variable for selected cancer sites among people with only one cancer diagnosis (Table 1) or more than one cancer diagnosis (Table 2). It should be noted that more than 90% of cancer patients in the SEER database had only one tumor. The denominator of the SEER cause-specific death classification variable is the number of deaths from all causes (including cancer and noncancer deaths), and the numerator is the underlying number of "attributable" causes of deaths. For example, for patients diagnosed with one and only one cancer at a major site (including breast, prostate, and colorectal cancers), underlying causes of death were usually attributed to the same site (56.38% breast, 34.66% prostate, and 64.66% colorectal cancer). However, for smaller subsites such as lip, hypopharynx, and rectum, numerous deaths would have been

misattributed if the only cause of death contributing to the SEER cause-specific death classification variable was death from that particular primary cancer. For these sites, a large proportion of deaths were assigned to cancers of the general organ system. For example, for people diagnosed with only rectum cancer, 38.6% of deaths were attributed to sites within the colorectal system. The causes of death from AIDS and cancer or from HIV alone contributed fewer than 1% of total deaths for most sites, except for Kaposi sarcoma, for which approximately 74% of the deaths were attributed to these causes. In general, site-specific diseases contributed a negligible amount to the cancer-attributable deaths (0.00%–1.22%). A smaller percentage of the deaths among those with more than one primary cancer than among those with one primary tumor contributed to the SEER cause-specific death classification variable, as expected because multiple primary cancer patients have more competing "other-cause" death.

We plotted 5-year relative survival vs cause-specific survival estimates for 100 cancer sites in Figure 1. Overall, there was very good agreement between the relative survival and cause-specific survival estimates, with a larger variability for people diagnosed at age 65 years or older (for the full dataset, see <http://seer.cancer.gov/causespecific/index.html>). Next, we compared survival estimates from relative and cause-specific approaches for several cancer sites (Table 3). Among people diagnosed with any cancer at age 65 years or older, the relative survival approach yielded higher estimates for cancers at all sites combined (relative survival = 61.4% and cause-specific survival = 59.7%), for melanoma of the skin (relative survival = 96.2% and cause-specific survival = 89.4%), for breast cancer (relative survival = 91.7% and cause-specific survival = 86.9%), for cancers of the corpus uterus and not otherwise specified (relative survival = 79.0% and cause-specific survival = 75.8%), and for prostate cancer (relative survival = 97.9% and cause-specific survival = 89.4%). The relative survival approach compared with the cause-specific survival approach yielded lower estimates for cancer of the oral cavity and pharynx (relative survival = 52.8% and cause-specific survival = 58.4%), for cancer of the cervix uteri (relative survival = 52.6% and cause-specific survival = 56.5%), and for myeloma (relative survival = 26.1% and cause-specific survival = 29.9%). Similar patterns were found when 10-year relative and cause-specific survival rates were compared (data not shown).

In some instances, relative survival that was based on US life tables may not accurately estimate the probability of cancer-specific survival (Tables 4 and 5). Among white patients who were aged 65 years or older at diagnosis of in situ breast cancer, the 5-year relative survival was 107.5% (Table 4), a value that was greater than 100% survival, whereas the 5-year cause-specific survival for this group was 98.6% (95% CI = 98.4% to 98.8%). It should be noted that when relative survival is 100% or higher, a confidence interval cannot be calculated. Similar overestimates with the relative survival approach compared with the cause-specific survival approach were observed among black persons (106.6% vs 97.4%, 95% CI = 96.2% to 98.2%, respectively) and among persons of Asian or Pacific Islander descent (103% vs 99%, 95% CI = 98.4% to 99.6%, respectively). We attributed these overestimates to the problem that mortality from other causes among patients with in situ breast cancer was not represented well in the life table used. Patients diagnosed with early-stage breast

Table 1. Causes of death contributing to the Surveillance, Epidemiology, and End Results cause-specific death classification variable for people diagnosed with one and only one tumor, 1992–2004*

Primary cancer diagnosis	Causes contributing to cause-specific death variable, % of total causes of deaths (95% CI)						Causes not contributing†, % of individuals (95% CI)	
	AIDS and cancer‡	HIV alone§	Cancer of same site	Cancer of same body system	Cancer of any other site	Site-specific disease	Other causes	Other causes
Breast	0.03 (0.03 to 0.03)	N/A	56.38 (56.29 to 56.47)	0.00 (0.00 to 0.00)	3.53 (3.50 to 3.56)	0.02 (0.02 to 0.02)	40.04 (39.95 to 40.13)	
Prostate	0.00 (0.00 to 0.00)	N/A	34.66 (34.57 to 34.75)	0.02 (0.02 to 0.02)	3.60 (3.57 to 3.63)	0.10 (0.09 to 0.11)	61.62 (61.53 to 61.71)	
Colon and rectum	0.01 (0.01 to 0.01)	N/A	64.66 (64.55 to 64.77)	0.60 (0.58 to 0.62)	4.83 (4.78 to 4.88)	1.18 (1.16 to 1.20)	28.72 (28.62 to 28.82)	
Colon	0.02 (0.02 to 0.02)	N/A	61.35 (61.22 to 61.48)	2.53 (2.49 to 2.57)	5.20 (5.14 to 5.26)	1.22 (1.19 to 1.25)	29.68 (29.56 to 29.80)	
Rectum	0.00 (0.00 to 0.00)	N/A	32.38 (32.14 to 32.62)	36.58 (36.34 to 36.82)	3.83 (3.73 to 3.93)	1.06 (1.01 to 1.11)	26.14 (25.92 to 26.36)	
Lung and bronchus	0.05 (0.05 to 0.05)	N/A	85.62 (85.54 to 85.70)	0.04 (0.04 to 0.04)	4.19 (4.15 to 4.23)	N/A	10.10 (10.04 to 10.16)	
Oral cavity and pharynx	0.13 (0.11 to 0.15)	0.30 (0.27 to 0.33)	56.83 (56.59 to 57.07)	4.10 (4.00 to 4.20)	12.91 (12.75 to 13.07)	N/A	25.73 (25.52 to 25.94)	
Tongue	0.06 (0.04 to 0.08)	0.25 (0.20 to 0.30)	41.58 (41.08 to 42.08)	26.08 (25.64 to 26.52)	11.84 (11.52 to 12.16)	N/A	20.19 (19.79 to 20.59)	
Lip	0.00 (0.00 to 0.00)	1.55 (1.35 to 1.75)	0.00 (0.00 to 0.00)	16.41 (15.82 to 17.00)	8.36 (7.92 to 8.80)	N/A	73.68 (72.98 to 74.38)	
Hypopharynx	0.00 (0.00 to 0.00)	0.15 (0.08 to 0.22)	8.62 (8.11 to 9.13)	61.27 (60.38 to 62.16)	12.71 (12.10 to 13.32)	N/A	17.25 (16.56 to 17.94)	
Esophagus	0.02 (0.01 to 0.03)	N/A	81.82 (81.52 to 82.12)	3.41 (3.27 to 3.55)	3.60 (3.45 to 3.75)	0.19 (0.16 to 0.22)	10.96 (10.72 to 11.20)	
Stomach	0.03 (0.02 to 0.04)	N/A	67.95 (67.69 to 68.21)	13.41 (13.22 to 13.60)	4.94 (4.82 to 5.06)	0.45 (0.41 to 0.49)	13.22 (13.03 to 13.41)	
Non-Hodgkin lymphoma	2.32 (2.26 to 2.38)	2.77 (2.71 to 2.83)	68.30 (68.13 to 68.47)	0.12 (0.11 to 0.13)	3.87 (3.80 to 3.94)	N/A	22.62 (22.47 to 22.77)	
Kaposi sarcoma	14.74 (14.38 to 15.10)	58.96 (58.46 to 59.46)	3.19 (3.01 to 3.37)	N/A	2.59 (2.43 to 2.75)	N/A	20.52 (20.11 to 20.93)	

* Thirteen registries from the Surveillance, Epidemiology, and End Results (SEER) Program include Atlanta, Connecticut, Detroit, Hawaii, Iowa, Los Angeles, New Mexico, San Francisco–Oakland, San Jose–Monterey, Rural Georgia, Seattle, Utah, Alaska Native. CI = confidence interval; N/A = not available.

† Causes not contributing to death represent censoring events.

‡ For people with only one cancer, AIDS and cancer is a contributing cause of death for all sites.

§ HIV death is a contributing cause only for selected sites.

Table 2. Causes of death contributing to the Surveillance, Epidemiology, and End Results cause-specific death classification variable for people diagnosed with more than one cancer (9.0% of overall cancer patients), 1992–2004*

Primary cancer diagnosis	Causes contributing to cause-specific death variable, % of total causes of deaths (95% CI)						Causes not contributing, † % of individuals (95% CI)
	AIDS and cancer, ‡	HIV alone, §	Cancer of same site of first diagnosis	Cancer of same body system	Cancer of unknown site,	Site-specific disease,	
Breast	N/A	N/A	32.96 (32.88 to 33.04)	0.00 (0.00 to 0.00)	5.10 (5.06 to 5.14)	0.00 (0.00 to 0.00)	61.94 (61.85 to 62.03)
Prostate	N/A	N/A	10.62 (10.56 to 10.68)	0.02 (0.02 to 0.02)	6.40 (6.36 to 6.44)	0.02 (0.02 to 0.02)	82.94 (82.87 to 83.01)
Colon and rectum	N/A	N/A	39.73 (39.62 to 39.84)	0.91 (0.89 to 0.93)	4.48 (4.43 to 4.53)	0.80 (0.78 to 0.82)	54.07 (53.96 to 54.18)
Colon	N/A	N/A	36.12 (35.99 to 36.25)	3.57 (3.52 to 3.62)	4.37 (4.32 to 4.42)	0.92 (0.89 to 0.95)	55.03 (54.90 to 55.16)
Rectum	N/A	N/A	16.53 (16.34 to 16.72)	26.90 (26.68 to 27.12)	4.83 (4.72 to 4.94)	0.45 (0.42 to 0.48)	51.30 (51.05 to 51.55)
Lung and bronchus	N/A	N/A	61.73 (61.63 to 61.83)	0.75 (0.73 to 0.77)	3.86 (3.82 to 3.90)	N/A	33.66 (33.56 to 33.76)
Oral cavity and pharynx	0.29 (0.26 to 0.32)	0.10 (0.08 to 0.12)	30.30 (30.07 to 30.53)	2.13 (2.06 to 2.20)	9.97 (9.82 to 10.12)	N/A	57.21 (56.97 to 57.45)
Tongue	0.41 (0.35 to 0.47)	0.41 (0.35 to 0.47)	16.53 (16.16 to 16.90)	23.97 (23.54 to 24.40)	9.09 (8.80 to 9.38)	N/A	49.59 (49.09 to 50.09)
Lip	0.00 (0.00 to 0.00)	0.00 (0.00 to 0.00)	0.00 (0.00 to 0.00)	7.23 (6.82 to 7.64)	6.02 (5.64 to 6.40)	N/A	86.75 (86.21 to 87.29)
Hypopharynx	0.00 (0.00 to 0.00)	0.00 (0.00 to 0.00)	4.67 (4.28 to 5.06)	37.38 (36.50 to 38.26)	7.48 (7.00 to 7.96)	N/A	50.47 (49.56 to 51.38)
Esophagus	N/A	N/A	46.72 (46.33 to 47.11)	4.92 (4.75 to 5.09)	4.51 (4.35 to 4.67)	1.23 (1.14 to 1.32)	42.62 (42.23 to 43.01)
Stomach	N/A	N/A	36.10 (35.84 to 36.36)	13.37 (13.18 to 13.56)	5.35 (5.23 to 5.47)	0.27 (0.24 to 0.30)	44.92 (44.65 to 45.19)
Non-Hodgkin lymphoma	0.86 (0.83 to 0.89)	0.86 (0.83 to 0.89)	38.47 (38.29 to 38.65)	0.19 (0.17 to 0.21)	5.17 (5.09 to 5.25)	N/A	54.45 (54.27 to 54.63)
Kaposi sarcoma	27.59 (27.13 to 28.05)	32.76 (32.28 to 33.24)	0.00 (0.00 to 0.00)	N/A	0.00 (0.00 to 0.00)	N/A	39.66 (39.16 to 40.16)

* Thirteen registries from the Surveillance, Epidemiology, and End Results (SEER) Program include Atlanta, Connecticut, Detroit, Hawaii, Iowa, Los Angeles, New Mexico, San Francisco–Oakland, San Jose–Monterey, Rural Georgia, Seattle, Utah, Alaska Native. CI = confidence interval; N/A = not available.

† Causes not contributing to death represent censoring events.

‡ For people with more than one cancer, AIDS and cancer is a contributing cause for selected sites.

§ HIV death is a contributing cause only for selected sites.

|| Multiple primary tumors of unknown sites (eg, codes C798, C80, C97, and D489).

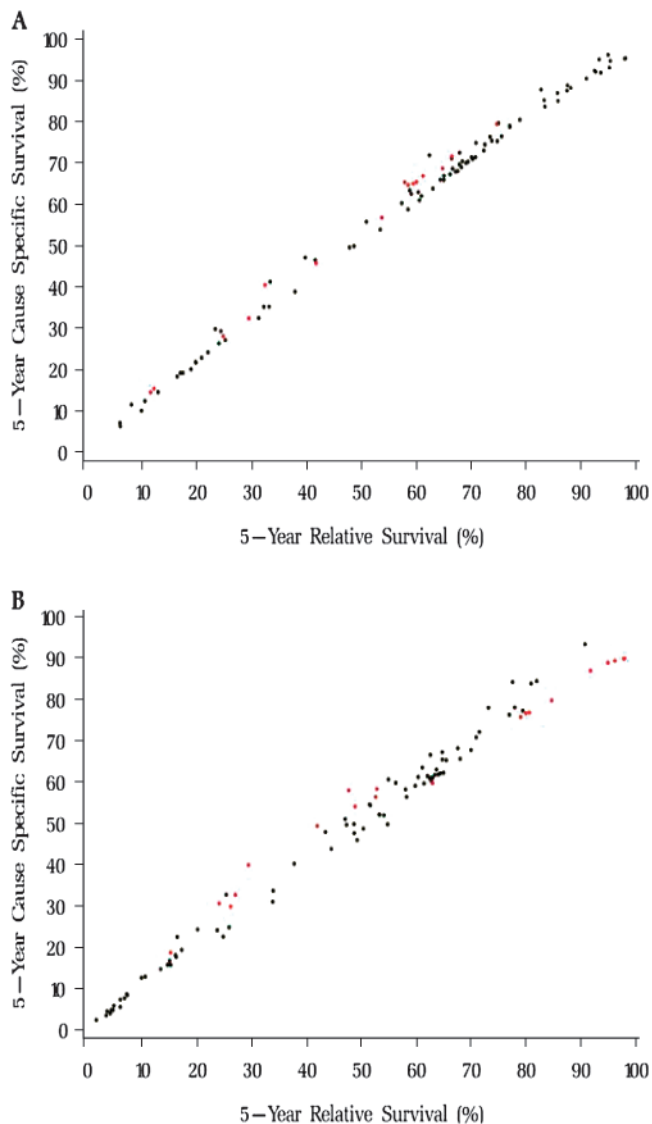


Figure 1. Five-year cause-specific survival vs relative survival for 100 cancer sites by age at diagnosis. **A)** Age at diagnosis of 50–64 years. **B)** Age at diagnosis of 65 years or older. The data source was 13 registries from the Surveillance, Epidemiology, and End Results Program. Patient diagnosed between 1992 and 2004 and followed through 2005. **Red dots** = cancer sites for which the difference between cause-specific survival and relative survival is greater than 3 percentage points and the 95% confidence intervals do not overlap; **black dots** = cancer sites for which the difference between cause-specific survival and relative survival is less than 3 percentage points and the 95% confidence intervals overlap.

cancer through a screening examination tend to be healthier and have a longer life expectancy than the general population because of the “healthy screener effect” (15,16). Therefore, relative survival estimates that are calculated by use of a population life table that underestimates the cohort’s expected survival are biased and overestimated survival. A similar argument could be applied to relative survival estimates for patients diagnosed with early-stage prostate cancer, a diagnosis often made by prostate-specific antigen screening (17,18).

Expected survival rates for cancers of the lung and bronchus and cancers of the oral cavity and pharynx that are based on life

tables derived from general population data are too high because of the underlying risk factors of smoking and/or alcohol use. Smoking is a major risk factor for lung and oral cancers, as well as other diseases (eg, heart disease), and so a cohort of lung cancer patients will contain a much higher proportion of smokers than the general population. Life expectancy for this cohort would be lower than for the general population. Relative survival estimates for lung and oral cancers that were calculated from US life tables were clearly underestimated, as shown in Table 4 for patients of all races (eg, for white cancer patients aged ≥ 65 years, relative survival = 54.2%, 95% CI = 53.1% to 55.3%, and cause-specific survival = 60.1%, 95% CI = 59.1% to 60.9%) and the lung and bronchus (eg, for black cancer patients aged ≥ 65 years, relative survival = 10.5%, 95% CI = 9.9% to 11.2%, and cause-specific survival = 11.9%, 95% CI = 11.2% to 12.6%). Relative survival estimates for American Indians/Alaska Natives (Table 4) were lower than cause-specific survival estimates for all sites, with the exception estimates for breast and prostate cancers. The expected rates that were used to calculate relative survival rates for American Indians/Alaska Natives in Table 4 were estimated with US life tables for races other than white and black (ie, other race). American Indians have a much lower life expectancy than most race groups (19). Therefore, expected survival rates that are generated with the other race category may be too optimistic, resulting in an underestimate of survival. The biggest discrepancy between relative survival and cause-specific survival estimates for American Indians/Alaska Natives was observed among older patients with localized colon and rectum cancer (relative survival = 71.7%, 95% CI = 61.5% to 79.6%; and cause-specific survival = 79.7%, 95% CI = 72.1% to 85.3%). The survival percentages for distant stage of disease across these races for lung cancer, for example, showed little deviation, presumably because survival was so poor for distant disease that other causes did not play a major role in mortality.

Socioeconomic variation in survival for selected cancer cohorts by the relative survival and cause-specific survival methods is shown in Table 5. Because socioeconomic status-based life tables are not available, life tables that are based on the general US population will give an expected survival rate that is likely too optimistic for low socioeconomic status areas and too pessimistic for high socioeconomic status areas, causing bias in opposite directions. For example, the 5-year survival rate for all sites combined was 72.0% (95% CI = 71.88% to 72.15%) by the relative survival approach and 69.9% (95% CI = 69.86% to 70.08%) by the cause-specific survival approach in high socioeconomic status areas (ie, areas with <10% poverty) because the expected survival rate in areas of high socioeconomic status was underestimated, resulting in an overestimation of the relative survival rate. Similarly, the expected survival rate of patients with low socioeconomic status may be overestimated by use of a general population life table, resulting in an underestimation of the relative survival rate in areas of low socioeconomic status (ie, for all sites combined, relative survival = 57.3% (95% CI = 56.96% to 57.55%); and cause-specific survival = 59.1%, 95% CI = 58.84% to 59.35%). This systematic underestimation of cancer survival for low socioeconomic status groups and overestimation of cancer survival for high socioeconomic status groups exaggerates the socioeconomic status gradient; for example, corpus and uterus cancer disparity measure was 13.3% by relative survival and 8.8% by

Table 3. Five-year relative survival (RS) and 5-year cause-specific survival (CSS) for selected cancer sites and age groups, 1992–2004*

Cancer site	50–64 Y			≥65 Y		
	5-y RS, % (95% CI)	5-y CSS, % (95% CI)	Dif.† %	5-y RS, % (95% CI)	5-y CSS, % (95% CI)	Dif.† %
All sites combined	69.4 (69.3 to 69.5)	70.3 (70.2 to 70.5)	−0.9	61.4 (61.2 to 61.5)	59.7 (59.6 to 59.8)	1.7
Oral cavity and pharynx	59.4 (58.5 to 60.3)	65.0 (64.2 to 65.9)	−5.6†	52.8 (51.7 to 53.8)	58.4 (57.5 to 59.2)	−5.6†
Lip	94.9 (92.3 to 96.7)	96.2 (94.7 to 97.3)	−1.3	90.7 (87.3 to 93.3)	93.4 (92.1 to 94.5)	−2.7
Tongue	60.0 (58.2 to 61.8)	65.4 (63.7 to 67.1)	−5.4†	48.8 (46.6 to 51.0)	54.1 (52.3 to 55.9)	−5.3†
Esophagus	16.9 (15.8 to 18.1)	19.2 (18.0 to 20.4)	−2.3	14.6 (13.7 to 15.5)	15.8 (15.0 to 16.7)	−1.2
Stomach	25.1 (24.0 to 26.2)	27.2 (26.1 to 28.3)	−2.1	23.6 (22.9 to 24.4)	24.2 (23.5 to 24.8)	−0.6
Colon and rectum	68.2 (67.7 to 68.7)	69.0 (68.6 to 69.5)	−0.8	64.0 (63.6 to 64.4)	61.9 (61.6 to 62.2)	2.1
Colon excluding rectum	67.1 (66.5 to 67.7)	68.0 (67.4 to 68.5)	−0.9	64.4 (63.9 to 64.8)	62.2 (61.8 to 62.5)	2.2
Rectum and rectosigmoid junction	70.2 (69.4 to 71.0)	71.0 (70.3 to 71.7)	−0.8	62.9 (62.2 to 63.7)	61.1 (60.5 to 61.7)	1.8
Liver and intrahepatic bile duct	11.5 (10.6 to 12.4)	14.5 (13.5 to 15.6)	−3.0†	6.7 (6.1 to 7.4)	7.7 (7.1 to 8.4)	−1.0
Liver	12.1 (11.2 to 13.1)	15.4 (14.3 to 16.6)	−3.3†	7.2 (6.6 to 8.0)	8.5 (7.8 to 9.2)	−1.3
Intrahepatic bile duct	5.9 (4.1 to 8.2)	7.1 (5.1 to 9.6)	−1.2	4.1 (3.1 to 5.4)	4.1 (3.1 to 5.3)	0.0
Pancreas	6.0 (5.4 to 6.5)	6.3 (5.8 to 6.9)	−0.3	3.4 (3.1 to 3.7)	3.5 (3.3 to 3.8)	−0.1
Larynx	66.5 (65.1 to 67.8)	71.6 (70.3 to 72.8)	−5.1†	64.7 (63.1 to 66.2)	67.3 (66.1 to 68.4)	−2.6
Lung and bronchus	17.4 (17.0 to 17.7)	19.2 (18.9 to 19.6)	−1.8	13.3 (13.1 to 13.5)	14.8 (14.6 to 15.0)	−1.5
Melanoma of the skin	95.2 (94.7 to 95.5)	93.1 (92.8 to 93.4)	2.1	96.2 (95.5 to 96.8)	89.4 (89.0 to 89.8)	6.8†
Breast	91.0 (90.8 to 91.2)	90.4 (90.2 to 90.6)	0.6	91.7 (91.4 to 92.0)	87.0 (86.8 to 87.2)	4.7†
Cervix uteri	64.8 (63.3 to 66.2)	68.7 (67.3 to 70.1)	−3.9†	52.5 (50.6 to 54.5)	56.4 (54.7 to 58.1)	−3.9†
Corpus and uterus, NOS	87.4 (86.9 to 88.0)	87.5 (87.0 to 88.0)	−0.1	79.0 (78.2 to 79.8)	75.8 (75.2 to 76.4)	3.2†
Corpus uteri	88.1 (87.6 to 88.7)	88.2 (87.7 to 88.7)	−0.1	79.9 (79.1 to 80.7)	76.7 (76.1 to 77.3)	3.2†
Uterus, NOS	33.1 (26.9 to 39.4)	35.2 (28.9 to 41.6)	−2.1	14.9 (11.1 to 19.4)	16.8 (12.8 to 21.3)	−1.9
Ovary	53.4 (52.3 to 54.5)	53.9 (52.8 to 55.0)	−0.5	33.7 (32.7 to 34.7)	31.1 (30.2 to 31.9)	2.6
Prostate	98.1 (97.9 to 98.3)	95.4 (95.2 to 95.5)	2.7†	97.9 (97.6 to 98.2)	89.9 (89.8 to 90.1)	8.0†
Urinary bladder	85.7 (85.0 to 86.4)	87.0 (86.4 to 87.5)	−1.3	76.9 (76.2 to 77.6)	76.3 (75.9 to 76.8)	0.6
Kidney and renal pelvis	66.6 (65.6 to 67.5)	68.7 (67.8 to 69.5)	−2.1	58.0 (57.0 to 59.0)	58.3 (57.5 to 59.0)	−0.3
Brain and ONS	16.3 (15.3 to 17.4)	18.4 (17.3 to 19.5)	−2.1	6.0 (5.4 to 6.6)	7.4 (6.7 to 8.1)	−1.4
Brain	12.8 (11.9 to 13.8)	14.6 (13.5 to 15.6)	−1.8	3.5 (3.1 to 4.1)	4.6 (4.0 to 5.2)	−1.1
Cranial nerves ONS	73.4 (67.5 to 78.4)	76.4 (71.1 to 80.8)	−3.0	54.9 (47.7 to 61.5)	60.7 (54.7 to 66.1)	−5.8
Thyroid	95.3 (94.5 to 96.0)	94.7 (94.1 to 95.3)	0.6	84.7 (82.8 to 86.4)	79.8 (78.5 to 81.0)	4.9†
Hodgkin lymphoma	74.6 (72.0 to 77.1)	79.5 (77.2 to 81.6)	−4.9†	51.6 (48.2 to 54.9)	54.4 (51.6 to 57.2)	−2.8
Non-Hodgkin lymphoma	67.8 (67.1 to 68.6)	69.7 (69.0 to 70.4)	−1.9	53.3 (52.5 to 54.0)	52.1 (51.5 to 52.7)	1.2
Myeloma	41.6 (40.1 to 43.1)	45.8 (44.3 to 47.3)	−4.2†	26.1 (25.1 to 27.1)	29.9 (28.9 to 30.9)	−3.8†
Leukemia	53.7 (52.6 to 54.8)	56.8 (55.7 to 57.9)	−3.1†	37.7 (36.8 to 38.5)	40.3 (39.6 to 41.0)	−2.6
Acute lymphocytic leukemia	20.8 (17.0 to 24.8)	22.9 (18.9 to 27.1)	−2.1	7.1 (4.8 to 10.0)	8.7 (6.3 to 11.6)	−1.6
Chronic lymphocytic leukemia	83.3 (81.7 to 84.7)	85.2 (83.9 to 86.4)	−1.9	71.5 (69.9 to 73.0)	72.2 (71.1 to 73.2)	−0.7
Acute myeloid leukemia	19.7 (18.1 to 21.4)	21.7 (20.0 to 23.5)	−2.0	4.3 (3.7 to 5.0)	4.8 (4.2 to 5.5)	−0.5

* Atlanta, Connecticut, Detroit, Hawaii, Iowa, Los Angeles, New Mexico, San Francisco–Oakland, San Jose–Monterey, Rural Georgia, Seattle, Utah, Alaska Native. CI = confidence interval; Dif. = difference; NOS = not otherwise specified; ONS = other nervous system.

† Survival differences of greater than 3% and no overlap between confidence intervals. Data source: Surveillance, Epidemiology, and End Results–13*, 1992–2004. Dif. = 5-year CSS minus RS.

Table 4. Five-year relative survival (RS) and cause-specific survival (CSS) rates by races, selected cancer sites and stages. Surveillance, Epidemiology, and End Results 13, 1992-2004*

Selected cancer cohort	White			Black			American Indian/Alaska Native			Asian or Pacific Islander		
	RS, % (95% CI)	CSS, % (95% CI)	Dif., %	RS, % (95% CI)	CSS, % (95% CI)	Dif., %	RS, % (95% CI)	CSS, % (95% CI)	Dif., %	RS, % (95% CI)	CSS, % (95% CI)	Dif., %
Breast												
In situ and <65 y	100.9†	99.7 (99.6 to 99.8)	1.2	99.9 (0.0 to 100.0)	98.8 (98.3 to 99.2)	1.1	98.4 (91.3 to 99.7)	98.1 (94.1 to 99.4)	0.3	100.0 (1.1 to 100.0)	99.7 (99.4 to 99.9)	0.2
In situ and ≥65 y	107.5†	98.6 (98.4 to 98.8)	8.9	106.6†	97.4 (96.2 to 98.2)	9.2	93.8 (51.2 to 99.4)	100.0+	-6.2	103.0+	99.2 (98.4 to 99.6)	3.8
Prostate												
Localized/regional and <65 y	101.3†	98.3 (98.2 to 98.4)	3.0	101.5†	97.4 (97.1 to 97.7)	4.1	93.4 (87.4 to 96.6)	97.4 (93.9 to 98.9)	-4.1	99.2 (97.4 to 99.8)	98.4 (97.8 to 98.8)	0.8
Localized/regional and ≥65 y	104.5†	94.8 (94.6 to 94.9)	9.8	101.4†	93.3 (92.8 to 93.8)	8.1	87.9 (80.2 to 92.8)	92.3 (88.3 to 94.9)	-4.4	100.0†	96.4 (95.9 to 96.8)	3.6
Colon and rectum												
Localized and <65 y	93.1 (92.6 to 93.6)	93.5 (93.0 to 93.8)	-0.3	88.3 (86.5 to 89.8)	89.2 (87.9 to 90.4)	-1.0	86.0 (79.5 to 90.6)	88.1 (82.4 to 92.0)	-2.1	93.1 (91.7 to 94.3)	94.7 (93.6 to 95.6)	-1.6
Localized and ≥65 y	89.6 (88.9 to 90.3)	85.5 (85.1 to 85.9)	4.1	81.8 (79.2 to 84.1)	80.3 (78.8 to 81.8)	1.5	71.7 (61.5 to 79.6)	79.7 (72.1 to 85.3)	-8.0	86.5 (84.4 to 88.2)	89.1 (87.9 to 90.1)	-2.6
Oral cavity and pharynx												
All stages and <65 y	67.2 (66.4 to 68.0)	71.6 (70.8 to 72.3)	-4.4	42.8 (40.8 to 44.8)	48.9 (47.0 to 50.9)	-6.1	52.2 (44.6 to 59.3)	59.5 (51.7 to 66.5)	-7.3	68.1 (66.0 to 70.1)	71.7 (69.6 to 73.6)	-3.6
All stages and ≥65 y	54.2 (53.1 to 55.3)	60.1 (59.1 to 60.9)	-5.8	35.0 (31.4 to 38.7)	40.0 (36.7 to 43.4)	-5.0	40.9 (28.6 to 52.7)	38.7 (27.9 to 49.3)	2.2	50.8 (47.2 to 54.2)	54.4 (51.1 to 57.5)	-3.6
Lung and bronchus												
All stages and <65 y	18.7 (18.3 to 19.0)	20.5 (20.1 to 20.8)	-1.8	13.9 (13.2 to 14.6)	15.5 (14.7 to 16.2)	-1.6	16.8 (13.6 to 20.3)	19.5 (16.0 to 23.3)	-2.7	18.4 (17.2 to 19.6)	20.7 (19.4 to 22.0)	-2.3
All stages and ≥65 y	13.5 (13.3 to 13.8)	15.0 (14.7 to 15.2)	-1.4	10.5 (9.9 to 11.2)	11.9 (11.2 to 12.6)	-1.4	7.3 (5.1 to 9.9)	10.0 (7.3 to 13.1)	-2.7	14.4 (13.6 to 15.3)	16.3 (15.4 to 17.1)	-1.9
Melanoma of the skin												
In situ and <65 y	2.6 (2.5 to 2.7)	2.9 (2.8 to 3.0)	-0.3	2.7 (2.4 to 3.1)	3.1 (2.7 to 3.4)	-0.4	1.3 (0.5 to 3.0)	1.6 (0.6 to 3.6)	-0.3	4.1 (3.5 to 4.6)	4.7 (4.2 to 5.4)	-0.7
In situ and ≥65 y	101.6†	99.7 (99.6 to 99.8)	1.9									
	109.1†	98.9 (98.7 to 99.1)	10.2									

* Example illustrating when expected survival is not representative of cancer cohorts. AD = absolute difference (= RS minus CSS). Blanks indicate that the numbers of case patients that were too small to show number for races other than white. The 5-year RS and CSS rates that were calculated by use of diagnosis years 1992-2004 with follow-up through 2005. Dif. = difference.

† When relative survival was 100% or higher, a confidence interval cannot be constructed.

Table 5. Five-year relative survival (RS) and cause-specific survival (CSS) rates by socioeconomic status (SES) status for selected cancer cohorts. Surveillance, Epidemiology, and End Results, 1996–2004*

Poverty level and cancer analyzed	RS, % (95% CI)	CSS, % (95% CI)
All sites combined		
High SES (poverty <10%)	72.02 (71.88 to 72.15)	69.97 (69.86 to 70.08)
Medium SES (poverty 10%–19.99%)	64.05 (63.82 to 64.28)	64.00 (63.81 to 64.19)
Low SES (poverty ≥20%)	57.26 (56.96 to 57.55)	59.10 (58.84 to 59.35)
Difference between high and low SES status†	14.76	10.87
Digestive system		
High SES (poverty <10%)	50.11 (49.76 to 50.46)	49.22 (48.92,49.51)
Medium SES (poverty 10%–19.99%)	44.45 (43.92 to 44.98)	45.13 (44.67 to 45.58)
Low SES (poverty ≥20%)	37.68 (37.05 to 38.32)	40.27 (39.69 to 40.85)
Difference between high and low SES status†	12.43	8.95
Female breast		
High SES (poverty <10%)	93.5 (93.27 to 93.72)	90.98 (90.81 to 91.14)
Medium SES (poverty 10%–19.99%)	89.48 (89.03 to 89.91)	88.00 (87.67 to 88.33)
Low SES (poverty ≥20%)	83.1 (82.42 to 83.75)	83.38 (82.85 to 83.89)
Difference between high and low SES status†	10.4	7.6
Corpus and uterus, not otherwise specified		
High SES (poverty <10%)	86.54 (85.66 to 87.19)	84.14 (83.62 to 84.65)
Medium SES (10%–19.99%)	81.61 (80.38 to 82.78)	81.29 (80.33 to 82.20)
Low SES (poverty ≥20%)	73.22 (71.46 to 74.88)	75.25 (73.81 to 76.62)
Difference between high and low SES status†	13.32	8.89
Lymphoma		
High SES (poverty <10%)	70.66 (70.02 to 71.28)	69.00 (68.48 to 69.52)
Medium SES (poverty 10%–19.99%)	64.00 (62.94 to 65.04)	64.64 (63.75 to 65.51)
Low SES (poverty ≥20%)	57.71 (56.32 to 59.07)	60.07 (58.83 to 61.28)
Difference between high and low SES status†	12.95	8.93

* The 5-year RS and CSS rates were calculated by use of diagnosis years 1996–2004 with follow-up through 2005. Poverty was used as a surrogate for SES for selected cancer cohorts.

† Confidence interval was not constructed for survival difference.

cause-specific survival. Cause-specific survival provided more accurate estimates across the socioeconomic status gradient by addressing differing rates of mortality as a result of other causes of death.

Cause-specific survival, which allows flexibility and the use of different endpoints, can provide insight into unique mortality patterns experienced by a cancer cohort. For example, when 5-year survival rates for patients diagnosed with non-Hodgkin lymphoma were examined by age at diagnosis, estimates of relative survival and of cause-specific survival differed substantially (Figure 2). When different endpoint definitions were used for the cause-specific survival measures including all malignant cancer deaths, non-Hodgkin lymphoma deaths, and cancer and AIDS deaths, the effect was particularly evident for patients aged 20–49 years. Because AIDS-associated non-Hodgkin lymphoma accounted for more than 50% of non-Hodgkin lymphoma-associated deaths among patients aged 30–40 years (data not shown), accounting for AIDS yielded dramatically different results for young adult patients. The cause-specific survival estimate was 78% (95% CI = 75.9% to 79.9%) for the endpoint non-Hodgkin lymphoma alone, compared with 57% (95% CI = 55.1% to 59.2%) for the endpoint all cancers and AIDS with cancer. These effects also were evident though less pronounced for other young and middle-aged adults. Relative survival and cause-specific survival estimates were similar when AIDS with cancer deaths were included as endpoints.

The cause-specific survival approach uses the same definition across all different race groups, which allows the generation of survival estimates for racial subgroups and assessment of survivorship

experiences among these populations (20,21). We present (Figure 3) the most recent 5-year cause-specific survival rates for four major cancer sites (ie, breast, prostate, colon and rectum, and lung) among the following population groups: white, black, Native American,

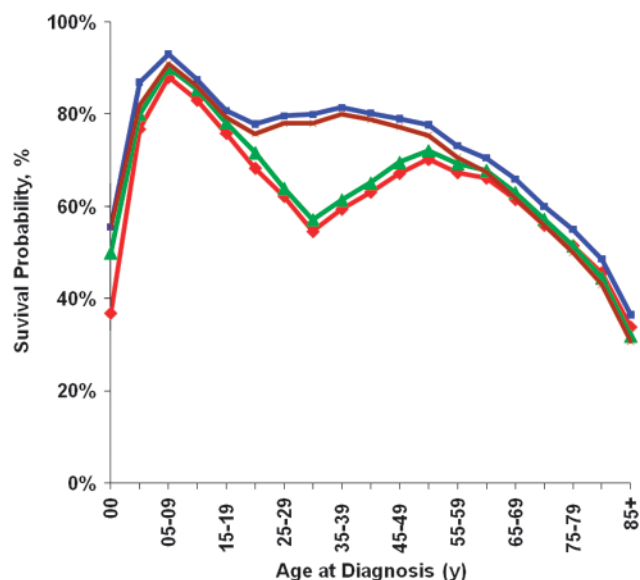


Figure 2. Five-year non-Hodgkin lymphoma cancer survival by age at diagnosis, 1992–2004. Red line = relative survival; blue line = cause-specific survival with non-Hodgkin lymphoma deaths; green line = cause-specific survival with cancer and AIDS death; brown line = cause-specific survival with all malignant cancer deaths.

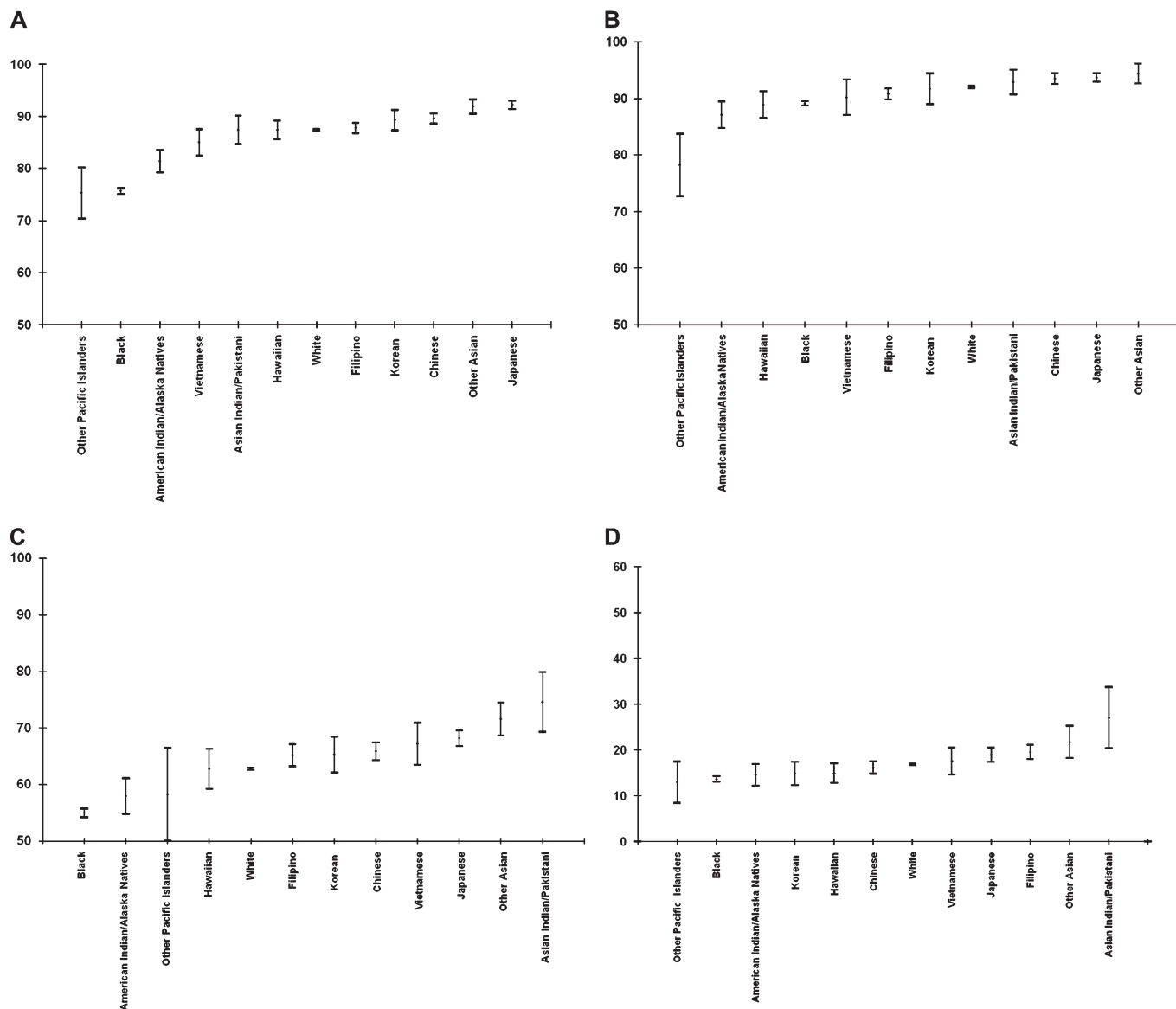


Figure 3. Five-year cause-specific survival probabilities for selected cancer sites by race and ethnicity. Data source is Surveillance, Epidemiology, and End Results (SEER-13, 1992–2004). **A)** Female breast cancer. **B)** Prostate cancer. **C)** Colon and rectum cancer (men and women). **D)** Lung and bronchus (men and women). Note that the y-axis in panel D is different from those in panels A, B, or C. Means (dots) and 95% confidence intervals (error bars) of the estimates are shown.

detailed Asian subcategories (eg, Japanese, Korean, Filipino, and Chinese) and major groups of Pacific Islanders such as Native Hawaiian. Cause-specific survival rates varied the most among patients with colon and rectum cancer (from 55%, 95% CI = 55.2% to 55.8%, among black persons to 75%, 95% CI = 69.3% to 79.9%, among Asian Indians or Pakistanis) and the least among prostate cancer patients (from 87%, 95% CI = 84.8% to 89.5% among Native Americans to 94%, 95% CI = 92.9% to 94.5%, among Japanese), for which survival rates generally were good. Japanese women had superior survival from breast cancer (92%, 95% CI = 91.4% to 92.9%), followed by Chinese women (90%, 95% CI = 88.6% to 90.6%) and Korean women (89%, 95% CI = 87.3% to 91.3%). We observed a survival pattern among racial groups in which Asian subgroups experienced the best survival rates, followed by white patients and black patients; Native American patients had the poorest survival rates. The numbers of patients in subgroups

such as other Asian or other Pacific Islanders were too small to provide reliable survival rates.

Discussion

We have developed a classification variable for cause of death associations with specific cancer diagnoses that appears to take into account likely misclassification of cause of death while not overly expanding the causes of death that are associated with each cancer diagnosis. For most of cancer sites, estimates obtained from relative and cause-specific approaches were similar (Table 3) because life tables were fairly representative of other-cause mortality for most cohorts in this analysis. However, in several situations, one approach provided more reliable results than the other.

We found that the cause-specific survival approach for reporting survival was sometimes advantageous. For heavily screened

populations, different socioeconomic strata, populations with strong risk factors for cancer and other diseases, and minority racial subgroups, cause-specific survival rates are likely to provide more accurate survival estimates than relative survival rates because currently available life tables do not take these factors into account. Relative survival methods lack the flexibility to address potential sources of bias if the associated factors are not accounted for in life tables. Thus, cause-specific survival estimates may be considered as an option to relative survival estimates in such circumstances.

Relative survival methods have many strengths, including independence from potential miscoding of the underlying cause of death. Coding practices vary substantially among countries (22,23), making the cause-specific survival approach inappropriate for international comparisons. In addition, for cause-specific survival methods to provide reliable survival estimates across racial or ethnic groups, cause of death assignments must be uniform across the groups studied. Bias may occur when comparing cause-specific survival rates across diverse groups if various racial or ethnic groups have different rates of follow-up (24,25). Further studies are needed to explore this issue in depth.

This study had several limitations. There may be possible over- or underattribution of cancer as a cause of death when relying on death certificates and the omission of deaths from noncancer causes that are a consequence of treatment, either acutely or from late effects (26,27). Although these deaths are not cancer deaths in a biological sense, they nonetheless reflect the consequences of cancer. A future validation study that tracks down and verifies all medical information for a small subset of deaths could shed some light on this issue. For these reasons, relative survival is the measure of choice for reporting survival rates when international comparisons are made (28).

Problems with the use of relative survival methods, however, are particularly evident when generating survival statistics for minority groups. For example, among Native Americans, relative survival prognosis estimates that are based on life tables for “other race” are likely to be misleadingly grim. Attempts to establish life tables for Native Americans have met with limited success because race information on the death certificates has a high rate of inaccuracy (5) and because of the small population size and disperse location of this minority group.

In summary, we have developed a classification scheme that associates cause of death information with cancer diagnoses that appears to be consistent with relative survival statistics in most situations and provides particularly useful estimates of survival when life tables do not reflect mortality patterns accurately in the cancer population. Relative survival methods have the advantage of being independent of the accuracy of the reported cause of death, although they are limited by the availability of appropriate life tables. Conversely, cause-specific survival methods have the advantage of being independent of the existence of appropriate life tables but are limited by cause of death accuracy. Use of these two approaches, each in appropriate circumstances, should enable production of more informative survival statistics.

Appendix Table 1. Codes for Surveillance, Epidemiology, and End Results cause-specific death variable for selected cohorts of patients with one and only one tumor (sequence 00)*

Primary diagnosis	Cause of death					
	AIDS and cancer	HIV alone	Cancer of same diagnosis site	Cancer of same body system	Any other cancer†	Site-specific disease
Breast	B210–B219	N/A	C50, D05, D24, D486	C445, D225, D485	C000–C444, C446–C499, C510–D049, D060–D224, D226–D239, D250–D484, D487–D489	N610–N649
Prostate	B210–B219	N/A	C61, D075, D291, D40	C60, C62, C63	C000–C599, C640–D074, D076–D290, D292–D399, D410–D489	N400–N509
Colon and rectum	B210–B219	N/A	C18, C19, C20, C785, D010–D012, D12, D374–D375	C17, C21, C26, D371–D373, D376–D379	C000–C169, C220–C259, C270–C784, C786–D009, D013–D119, D130–D370, D380–D489	K200–K319, K350–K389, K510–K579, K620–K639, K650–K669, K920–K929
Colon	B210–B219	N/A	C18, D010, D120–D126, D374	C17, C19–C21, C26, C785, D011–D012, D127–D129, D371–373, D375–D379	C000–C169, C220–C259, C270–C784, C786–D009, D013–D119, D130–D370, D380–D489	K200–K319, K350–K389, K510–K579, K620–K639, K650–K669, K920–K929
Rectum	B210–B219	N/A	C19, C20, C785, D011–D012, D127–D128, D375	C17–C18, C21, C26, D010, D120–D126, D129, D371–D374, D376–D379	C000–C169, C220–C259, C270–C784, C786–D009, D013–D119, D130–D370, D380–D489	K200–K319, K350–K389, K510–K579, K620–K639, K650–K669, K920–K929
Lung and bronchus	B210–B219	N/A	C34, C780, D022, D143, D381	C32, C33, C39, D144, D15, D380, D382–D389	C000–C319, C350–C389, C400–C779, C781–D021, D023–D142, D145–D149, D160–D379, D390–D489	N/A

(Appendix Table continues)

Appendix Table 1 Continued.

Primary diagnosis	Cause of death					
	AIDS and cancer	HIV alone	Cancer of same diagnosis site	Cancer of same body system	Any other cancer†	Site-specific disease
Oral cavity and pharynx	B210–B219	B20, B22–B24	C00–C15, C31–C32, C760, D000, D10–D11, D370	C410–C411, C440, C443–C444, C449, C490, C499, D030, D033, D034, D040, D043, D044, D210, D220, D223, D224, D230, D233, D234	C160–C309, C330–C409, C412–C439, C441–C442, C445–C448, C450–C489, C491–C498, C500–C759, C761–C999, D001–D029, D031–D032, D035–D039, D041–D042, D045–D099, D120–D209, D211–D219, D221–D222, D225–D229, D231–D232, D235–D369, D371–D489	N/A
Tongue	B210–B219	B20, B22–B24	C02, D101	C00–C01, C03–C15, C31–C32, C410–C411, C440, C443–C444, C449, C490, C499, C760, D000, D030, D033, D034, D040, D043, D044, D100, D102–D111, D210, D220, D223, D224, D230, D233, D234, D370	C160–C309, C330–C409, C412–C439, C441–C442, C445–C448, C450–C489, C491–C498, C500–C759, C761–C999, D001–D029, D031–D032, D035–D039, D041–D042, D045–D099, D120–D209, D211–D219, D221–D222, D225–D229, D231–D232, D235–D369, D371–D489	N/A
Lip	B210–B219	B20, B22–B24	C01, D000, D100, D370	C00, C02–C15, C31–C32, C410–C411, C440, C443–C444, C449, C490, C499, C760, D030, D033, D034, D040, D043, D044, D101–D111, D210, D220, D223, D224, D230, D233, D234	C160–C309, C330–C409, C412–C439, C441–C442, C445–C448, C450–C489, C491–C498, C500–C759, C761–C999, D001–D029, D031–D032, D035–D039, D041–D042, D045–D099, D120–D209, D211–D219, D221–D222, D225–D229, D231–D232, D235–D369, D371–D489	N/A
Hypopharynx	B210–B219	B20, B22–B24	C13, D107	C00–C12, C14–C15, C31–C32, C410–C411, C440, C443–C444, C449, C490, C499, C760, D000, D030, D033, D034, D040, D043, D044, D100–D109, D111, D210, D220, D223, D224, D230, D233, D234, D370	C160–C309, C330–C409, C412–C439, C441–C442, C445–C448, C450–C489, C491–C498, C500–C759, C761–C999, D001–D029, D031–D032, D035–D039, D041–D042, D045–D099, D120–D209, D211–D219, D221–D222, D225–D229, D231–D232, D235–D369, D371–D489	N/A
Esophagus	B210–B219	N/A	C15, D001, D130, D377	C16, C26, D371–D376, D378–D379	C000–C149, C170–C259, C270–D000, D002–D129, D131–D370, D380–D489	K200–K319, K510–K579, K920–K929
Stomach	B210–B219	N/A	C16, D002, D131, D371	C14, C15, C26, D131, D372–D379	C000–C139, C170–C259, C270–D001, D003–D130, D132–D370, D380–D489	K200–K319, K510–K579, K920–K929
Non-Hodgkin lymphoma	B210–B219	B20, B22–B24	C77, C81–C96, D360, D470 D361–D369, D471–D479		C000–C769, C780–C809, C970–D359, D370–D469, D480–D489	N/A
Kaposi sarcoma	B210–B219	B20, B22–B24	C46		C000–C459, C470–D489	N/A

* From the *International Classification of Diseases, Version 10*; N/A = not available.

† This category includes any malignant cancer deaths other than primary or related site of diagnosis.

Appendix Table 2. Codes for Surveillance, Epidemiology, and End Results cause-specific death variable for selected cohorts of patients with more than one tumor (sequence 01)*

Primary diagnosis	Cause of death				Cancer of unknown site†	Site-specific disease
	AIDS and cancer	HIV alone	Cancer of same first diagnosis site	Cancer of same body system		
Breast	N/A	N/A	C50, D05, D24, D486	C445, D225, D485	C798, C800–C809, C970–C979, D489	N610–N649
Prostate	N/A	N/A	C61, D075, D291, D40	C60, C62, C63	C798, C800–C809, C970–C979, D489	N400–N509
Colon and rectum	N/A	N/A	C18, C19, C20, C785, D010–D012, D12, D374–D375	C17, C21, C26, D371–D373, D376–D379	C798, C800–C809, C970–C979, D489	K200–K319, K350–K389, K510–K579, K620–K639, K650–K669, K920–K929
Colon	N/A	N/A	C18, D010, D120–D126, D374	C17, C19–C21, C26, C785, D011–D012, D127–D129, D371–D373, D375–D379	C798, C800–C809, C970–C979, D489	K200–K319, K350–K389, K510–K579, K620–K639, K650–K669, K920–K929
Rectum	N/A	N/A	C19, C20, C785, D011–D012, D127–D128, D375	C17–C18, C21, C26, D010, D120–D126, D129, D371–D374, D376–D379	C798, C800–C809, C970–C979, D489	K200–K319, K350–K389, K510–K579, K620–K639, K650–K669, K920–K929
Lung and bronchus	N/A	N/A	C34, C780, D022, D143, D381	C32, C33, C39, D144, D15, D380, D382–D389	C798, C800–C809, C970–C979, D489	N/A
Oral cavity and pharynx	B210–219	B20, B22–B24	C00–C15, C31–C32, C760, D000, D10–D11, D370	C410–C411, C440, C443–C444, C449, C490, C499, D030, D033, D034, D040, D043, D044, D210, D220, D223, D224, D230, D233, D234	C798, C800–C809, C970–C979, D489	N/A
Tongue	B210–219	B20, B22–B24	C02, D101	C00–C01, C03–C15, C31–C32, C410–C411, C440, C443–C444, C449, C490, C499, C760, D000, D030, D033, D034, D040, D043, D044, D100, D102–D111, D210, D220, D223, D224, D230, D233, D234, D370	C798, C800–C809, C970–C979, D489	N/A
Lip	B210–219	B20, B22–B24	C01, D000, D100, D370	C00, C02–C15, C31–C32, C410–C411, C440, C443–C444, C449, C490, C499, C760, D030, D033, D034, D040, D043, D044, D101–D111, D210, D220, D223, D224, D230, D233, D234	C798, C800–C809, C970–C979, D489	N/A
Hypopharynx	B210–219	B20, B22–B24	C13, D107	C00–C12, C14–C15, C31–C32, C410–C411, C440, C443–C444, C449, C490, C499, C760, D000, D030, D033, D034, D040, D043, D044, D100–D106, D108–D109, D11, D210, D220, D223, D224, D230, D233, D234, D370	C798, C800–C809, C970–C979, D489	N/A
Esophagus	N/A	N/A	C15, D001, D130, D377	C16, C26, D371–D376, D378–D379	C798, C800–C809, C970–C979, D489	K200–K319, K510–K579, K920–K929
Stomach	N/A	N/A	C16, D002, D131, D371	C14, C15, C26, D131, D372–D379	C798, C800–C809, C970–C979, D489	K200–K319, K510–K579, K920–K929
Non-Hodgkin lymphoma	B210–219	B20, B22–B24	C77, C81–C96, D360, D470	D361–D369, D471–D479	C798, C800–C809, C970–C979, D489	N/A
Kaposi sarcoma	B210–219	B20, B22–B24	C46		C798, C800–C809, C970–C979, D489	N/A

* International Classification of Diseases, Version 10; N/A = not available.

† Second malignant neoplasm of other specified or malignant neoplasm without specification of site.

References

1. Begg CB, Schrag D. Attribution of deaths following cancer treatment. *J Natl Cancer Inst.* 2002;94(14):1044–1045.
2. Ederer F, Axtell LM, Cutler SJ. The relative survival rate: a statistical methodology. *Natl Cancer Inst Monogr.* 1961;6:101–121.
3. Esteve J, Benhamou E, Croasdale M, Raymond L. Relative survival and the estimation of net survival: elements for further discussion. *Stat Med.* 1990;9(5):529–538.
4. National Cancer Institute. *Division of Cancer Control and Population Sciences, Surveillance Research Program.* <http://seer.cancer.gov>. Accessed August 11, 2008.
5. Rosenberg HM, Maurer JD, Sorlie PD, et al. Quality of death rates by race and Hispanic origin: a summary of current research, 1999. *Vital Health Stat 2.* 1999;71(128):1–13.
6. Percy C, Stanek E III, Gloeckler L. Accuracy of cancer death certificates and its effect on cancer mortality statistics. *Am J Public Health.* 1981;71(3):242–250.
7. Surveillance, Epidemiology, and End Results (SEER) Program. SEER*Stat Database: Incidence—SEER 17 Regs, November 2007 submission (1973–2005 varying)—Linked to County Attributes—Total U.S., 1969–2005 Counties. National Cancer Institute, DCCPS, Surveillance Research Program, Cancer Statistics Branch, released March 2008, based on the November 2007 submission. 2008. <http://www.seer.cancer.gov>. Accessed August 11, 2008.
8. Miller BA, Chu KC, Hankey BF, Ries LA. Cancer incidence and mortality patterns among specific Asian and Pacific Islander populations in the U.S. *Cancer Causes Control.* 2008;19(3):227–256.
9. Singh GK, Miller BA, Hankey BF, Edwards BK. *Area Socioeconomic Variations in U.S. Cancer Incidence, Mortality, Stage, Treatment, and Survival, 1975–1999.* NCI Cancer Surveillance Monograph Series, Number 4. Bethesda, MD: National Cancer Institute; 2003. NIH Publication No. 03-5417.
10. Krieger N, Chen JT, Waterman PD, Soobader MJ, Subramanian SV, Carson R. Geocoding and monitoring of U.S. socioeconomic inequalities in mortality and cancer incidence: does the choice of area-based measure and geographic level matter?: the Public Health Disparities Geocoding Project. *Am J Epidemiol.* 2002;156(5):471–482.
11. Boer R, Ries L, van Ballegooijen M, Feuer E, Legler J, Habbema D. *Ambiguities in Calculating Cancer Patient Survival: The SEER Experience for Colorectal and Prostate Cancer.* Bethesda, MD: Statistical Research and Applications Branch, National Cancer Institute; 2003. Technical Report No. 2003-05. <http://srab.cancer.gov/reports>. Accessed August 11, 2008.
12. SEER*Stat Software. *National Cancer Institute, DCCPS, Surveillance Research Program.* <http://www.seer.cancer.gov/seerstat>. Version 6.4.4. Accessed August 11, 2008.
13. Greenwood M. The natural duration of cancer. In: *Reports on Public Health and Medical Subjects.* Vol 33. London, UK: HM Stationery Office; 1926:1–26.
14. Hall WJ, Wellner JA. Confidence bands for a survival curve from censored data. *Biometrika.* 1980;67(1):133–143.
15. Dignam JJ, Huang L, Ries L, Reichman M, Mariotto A, Feuer E. Estimating breast cancer-specific and other-cause mortality in clinical trial and population-based cancer registry cohorts. *Cancer.* 2009;115(22):5272–5283.
16. Berry DA, Baines CJ, Baum M, et al. Flawed inferences about screening mammography's benefit based on observational data. *J Clin Oncol.* 2009;27(4):639–640.
17. Brenner H, Arndt V. Long-term survival rates of patients with prostate cancer in the prostate-specific antigen screening era: population-based estimates for the year 2000 by period analysis. *J Clin Oncol.* 2005;23(3):441–447.
18. Welch HG, Schwartz LM, Woloshin S. Are increasing 5-year survival rates evidence of success against cancer? *JAMA.* 2000;283(22):2975–2978.
19. DHHS Indian Health Services Facts on Indian Health Disparities. *Factsheet.* Rockville, MD: Department of Health and Human Services; 2006.
20. Clegg LX, Li FP, Hankey BF, Chu K, Edwards BK. Cancer survival among U.S. whites and minorities: a SEER (Surveillance, Epidemiology, and End Results) Program population-based study. *Arch Intern Med.* 2002;162(17):1985–1993.
21. Ries LAG, Young JL, Keel GE, Eisner MP, Lin YD, Horner M-J, eds. *SEER Survival Monograph: Cancer Survival Among Adults: U.S. SEER Program, 1988–2001, Patient and Tumor Characteristics.* Bethesda, MD: National Cancer Institute, SEER Program; 2007. NIH Publication No. 07–6215.
22. Percy C, Dolman A. Comparison of the coding of death certificates related to cancer in seven countries. *Public Health Rep.* 1978;93(4):335–350.
23. Percy C, Muir C. The international comparability of cancer mortality data. Results of an international death certificate study. *Am J Epidemiol.* 1989;129(5):934–946.
24. Pineda MD, White E, Kristal AR, Taylor V. Asian breast cancer survival in the U.S.: a comparison between Asian immigrants, U.S.-born Asian Americans and Caucasians. *Int J Epidemiol.* 2001;30(5):976–982.
25. Parkin DM, Khlat M. Studies of cancer in migrants: rationale and methodology. *Eur J Cancer.* 1996;32A(5):761–771.
26. Sarfati D, Blakely T, Pearce N. Measuring cancer survival in populations: relative survival vs cancer-specific survival. *Int J Epidemiol.* 2010;39(17):598–610.
27. Welch HG, Black WC. Are deaths within 1 month of cancer-directed surgery attributed to cancer? *J Natl Cancer Inst.* 2002;94(14):1066–1070.
28. Coleman MP, Quaresma M, Berrino F, et al. Cancer survival in five continents: a worldwide population-based study (CONCORD). *Lancet Oncol.* 2008;9(8):730–756.

Funding

Division of Cancer Control and Population Sciences, Surveillance Research Program, National Cancer Institute, National Institutes of Health.

Notes

The authors had full responsibility for the design of the study, the collection of the data, the analysis and interpretation of the data, the decision to submit the manuscript for publication, and the writing of the manuscript.

The authors thank Gretchen Keel and Rachidi Aminou of Information Management Services, Inc (IMS) for programming assistance. The authors would also like to thank Dr Brenda Edwards of the Surveillance, Epidemiology, and End Results Program for providing a very helpful review of the manuscript.

Affiliation of authors: Division of Cancer Control and Population Sciences, Surveillance Research Program, National Cancer Institute, National Institutes of Health, Bethesda, MD.