

Disparities in All-cancer and Lung Cancer Survival by Social, Behavioral, and Health Status Characteristics in the United States: A Longitudinal Follow-up of the 1997-2015 National Health Interview Survey-National Death Index Record Linkage Study

Hyunjung Lee¹, Gopal K. Singh²

¹Department of Public Policy and Public Affairs, John McCormack Graduate School of Policy and Global Studies, University of Massachusetts Boston, Boston, MA, ²The Center for Global Health and Health Policy, Global Health and Education Projects, Inc., Riverdale, MD, USA

Most research on cancer patient survival uses registry-based (e.g., SEER) incidence and survival data that have limited socioeconomic status and health-risk information. In this study, we used the 1997-2015 National Health Interview Survey-National Death Index prospectively-linked pooled cohort database ($n = 40,291$ cancer patients) to examine disparities in patient survival by a broad range of social determinants, including race/ethnicity, nativity, educational attainment, income/poverty level, occupation, housing tenure, physical and mental health status, smoking, physical activity, body mass index, and alcohol consumption. We used Cox proportional hazards models to estimate mortality hazard ratios and cause-specific 1-year, 5-year, and 10-year survival rates for all-cancer and lung cancer. During 1997-2015, the 10-year age-adjusted (all-cause) survival rate for cancer patients with professional and managerial occupations was 89.66%, significantly higher than the survival rate of 83.17% for laborers or 83.66% for the unemployed. Cancer patients with renting house had significantly lower age-adjusted survival rates than those owning house (82.65% vs. 85.80%). The 10-year age-adjusted survival rates were significantly greater among cancer patients with regular physical activity than those without regular physical activity (90.18% vs. 83.24%). Age-adjusted survival rates were significantly reduced for cancer patients with lower income and education, poor health, and serious psychological distress, and among current and former smokers. The gap in survival narrowed with additional sociodemographic, health, or behavioral adjustment. Similarly large differentials were found in lung cancer survival. Marked disparities in all-cancer and lung cancer survival were found by a wide range of sociodemographic and health characteristics.

Key Words Cancer, Survival, Social class, Health behaviors, Lung neoplasm

INTRODUCTION

In 2019, there were more than 16.9 million cancer survivors in the United States [1]. The 5-year relative survival rates increased from 49% during 1975-1977 to 67% during 2010-2016, due to advances in treatment and earlier diagnosis for some cancers [1,2]. Although survival of cancer patients has improved, 1.9 million new cancer cases were expected to be diagnosed in 2021 and 608,570 cancer patients were estimated to die [1,3]. Particularly, lung and bronchus, as

the most common cause of cancer death for both men and women, accounted for 22% of total cancer deaths in 2021, which is two times higher than the number of prostate cancer deaths and 1.47 times higher than the number of breast cancer deaths [3]. The 5-year relative survival rate for lung cancer patients was only 21% during 2010-2016, increasing from 12% during 1975-1977 [1,2].

Common risk factors of cancer incidence [4-7], mortality [4,5,7-9], and survival [4,5,8,10-14] are well-documented, including sex, race/ethnicity, income, education level, occupa-

Received January 6, 2022, Revised April 4, 2022, Accepted May 23, 2022

Correspondence to Hyunjung Lee, E-mail: hyunjung.lee0001@gmail.com, https://orcid.org/0000-0003-0783-5484



This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License, which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Copyright © 2022 Korean Society of Cancer Prevention

tion, housing tenure, cigarette smoking, excess body weight, alcohol consumption, physical activity, and psychological distress [4-6,12,14,15]. Besides these factors, use of screening tests has decreased mortality rates [4,8], and stage at diagnosis and treatment affect cancer mortality or survival rates [8]. Non-Hispanic Blacks have higher death rates for all cancers combined, compared with other racial/ethnic groups [5,8,16]. Low socioeconomic position in education, occupation, and income has been shown to increase risk of lung cancer incidence by 61%, 48%, and 37%, compared with the highest socioeconomic status (SES) levels [6], while one meta-analysis found that lung cancer survival was associated with income but not consistently with education [10]. Health behaviors are important factors, given that 28.8% of cancer deaths are attributable to cigarette smoking, followed by obesity (6.5%), alcohol intake (4.0%), fruit and vegetable consumption (2.7%), and physical inactivity (2.2%) [7]. Cancer mortality has been shown to be 33% higher for adults with serious psychological distress, compared to those without psychological distress [9].

Existing registry-based studies of survival have focused on disparities in survival of cancer patients by sex and race/ethnicity, limited measures of SES, comorbidity, stage, or area-level SES. Patient survival rates are found to be significantly lower for males and non-Hispanic blacks, compared with females and non-Hispanic whites [1,2,5,8]. During 2010-2016, the 5-year relative survival rate for all cancers combined for females was 68.5%, higher than the rate of 66.4% for males [2]. The 5-year relative survival rate for whites was 68%, compared with 63% for blacks [1,2]. The 5-year survival for lung cancer female patients aged 66-74 years with no comorbidity was 60.3% at localized disease stage, 27.8% at regional disease stage, and 3.7% at distant disease stage, while those with severe comorbidity presented 36.3%, 15.9%, and 1.8% [17]. Differences in cancer survival have also been found by socioeconomic deprivation. During 1988-1999, the 10-year survival rate for cancer patients was 41% in the most-deprived neighborhoods, compared with 60.4% for those in the least-deprived neighborhoods [5]. However, most research on cancer patient survival uses registry-based (e.g., Surveillance, Epidemiology, and End Results [SEER]) incidence and survival data that have limited SES, nativity/immigrant status, and health-risk information.

To address this gap in research, given the highest rate of lung cancer incidence in the United States, our study examines disparities in the predicted all-cause and cancer survival rates among all cancer patients and lung cancer patients in the United States by sociodemographic, health, and behavioral characteristics, using a nationally representative longitudinal dataset with a 19-years of mortality follow-up.

MATERIALS AND METHODS

Data

The data for this study are derived from the 1997-2014 cross-sectional National Health Interview Survey (NHIS) linked to the 1997-2015 cumulative National Death Index (NDI) [18]. As a nationally representative, annual cross-sectional household interview survey, NHIS provides demographic, socioeconomic, and health characteristics of the civilian, non-institutionalized population in the United States [19]. The NHIS uses geographically clustered sampling techniques to select the sample of dwelling units [19]. The sample is designed each month's sample to be nationally representative and collected continuously from January to December each year [19]. We used the NHIS Sample Adults file including more specific and detailed information such as cancer diagnosis or cancer type [20]. NHIS sample size varies from year to year. In the Sample Adult files, for example, the 1997 NHIS contains 36,116 persons with an 80.4% response rate [21], and the 2014 NHIS contains 36,697 persons with a 58.9% response rate [20]. The National Center for Health Statistics (NCHS) developed public-use versions of the NHIS linked with death certificate records from the NDI. For this study, we used the 1997-2015 pooled NHIS linked mortality file containing 19 years of longitudinal mortality follow-up data from the date of survey participation through December 31, 2015 [22,23].

Sample

The study sample was restricted to cancer patients aged 18 and older from the years of 1997 to 2014 NHIS sample adult files ($n = 42,767$, Fig. 1). Cancer patients were identified by asking respondents whether they were ever diagnosed by a doctor or other health professional as having cancer. Then, cancer patients without mortality status ($n = 1,230$), those without information on age at first diagnosis for cancer ($n = 626$), or sample with death month being earlier than interview month ($n = 4$), were eliminated from the analysis. The final pooled eligible sample size excluding missing values in covariates was 40,291 for all cancer patients and 1,287 lung cancer patients. Missing values, accounting for less than 1% of the study sample, for race/ethnicity (0.01%), nativity (0.10%), education (0.43%), marital status (0.14%), housing tenure (0.12%), self-reported physical health status (0.16%), and smoking (0.66%) were excluded from the analyses.

For missing values accounting for more than 1% of the total observations, poverty status (17.52%), psychological distress (1.94%), regular physical activity (2.96%), body mass index (BMI) (2.99%), and alcohol consumption (1.45%), we created missing categories so as to not lose too many observations in the analysis. Although a knowledgeable proxy can respond for sample adults, if they are physically or mentally unable to respond to survey, there might be a potential selection bias since survey respondents were healthy enough

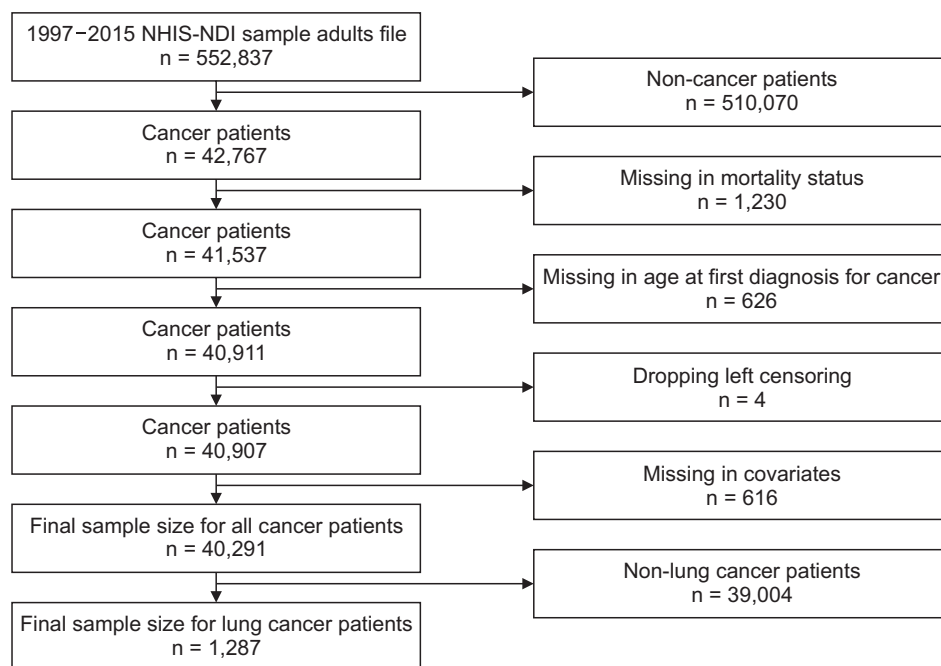


Figure 1. Flow chart of final sample size for all cancer patients and lung cancer patients in 1997-2015 NHIS-NDI. NHIS-NDI, National Health Interview Survey-National Death Index.

to participate in the survey and survived a substantial period after the cancer diagnosis.

Measurement

We examined all-cause mortality and cancer mortality (ICD-10 codes: C00-C97)-related outcomes: age-adjusted and covariate-adjusted predicted cancer-specific survival rates for all cancer patients and for lung cancer patients.

Follow-up time for individuals who died during the study period was estimated by the number of months from age first diagnosed with cancer to the month/year of death. For individuals who were alive was estimated by the number of months from age first diagnosed with cancer to December 31, 2015. For cancer patients with more than one cancer, we calculated the mean of years since first diagnosis on multiple cancers. Since the NHIS-NDI database provides only the quarter of death, we assumed that death occurred in the middle of the quarter, February, May, August, or November [24].

Sociodemographic, health, and behavioral characteristics

Based on the previous literature, we selected the following determinants of health for model estimation of cancer mortality: age, sex, race/ethnicity, education, poverty status, occupation, housing tenure, marital status, nativity/immigrant status, region of residence, self-reported physical health status, psychological distress, regular physical activity, BMI, smoking status, alcohol consumption, and survey years [4-10]. Definition of measures is provided in Table S1.

Analytic approach

Weighted proportions of all-cause death and cancer death by

each covariate were calculated for all cancer patients and for lung cancer patients, using survey weight. For the lung cancer model, binary measure of the US-born/foreign-born variable was used and alcohol consumption was dropped. Cox proportional hazards models were used to derive predicted all-cause and cancer-specific survival rates among cancer patients, controlling for individual characteristics and year-fixed effects.

First, we examined the Cox models controlling for a different set of covariates: 1) age and survey year (age-adjusted model); 2) age, survey year, sex, race/ethnicity, education, poverty status, occupation, housing tenure, marital status, nativity/immigrant status, and region (SES-adjusted model); and 3) all covariates in the SES-adjusted model plus self-reported health status, psychological distress, regular physical activity, BMI, smoking status, and alcohol consumption (fully-adjusted model). All Cox regression analyses with hazard ratios were reported in Table S2, 3.

Second, the predicted survival rates were estimated for each of the sociodemographic, health, and behavioral characteristics using Stata post-estimation command, *survci*, at the mean age and reference group for other characteristics. This approach provides survival rates for prototypical individuals, which is different from the average survival rates of the sample. Since Cox regression is a semi-parametric model, Stata first estimates the baseline survivor function, when all covariates are equal to zero [25,26]. Once Stata has an estimate of the baseline survivor function, a covariate-adjusted survival rate was predicted at the specified characteristics [26]. These are the characteristics of the prototypical individual for this study: age = 63.52 years for all cancer patients (68.11 years for lung cancer patients); male; non-Hispanic

white; master's degree or higher; income \geq 400% of federal poverty level (FPL); professional and managerial occupation; homeowner; married; US-born; West region; excellent/very good/good health; no serious psychological distress; regular physical activity; BMI < 25; never smoker; lifetime abstainer; and year = 1997.

We separately estimated and reported survival rates for each individual characteristic by all-cause and cancer specific death, by 1-year, 5-year, and 10 year-follow-up, and by age-, SES-, and fully-adjusted models for both all cancer patients and lung cancer patients [27]. Cancer specific survival rate and SES- and fully-adjusted model analyses were reported in Tables S4-S7.

Our approach is different from previous studies using SEER data, in which expected survival rates are derived from life tables by SES, geography, and race developed by the SEER program [2,27,28]. For cancer death, individuals surviving beyond the follow-up period and those dying from causes other than cancer were treated as right-censored observations. The model assumes that hazard rates are a log-linear function of parameters representing the effects of covariates [29,30]. We checked the hazards proportionality assumption [30] by inspecting the plots of [log(-log) survival function] against survival time t for the various covariate categories including those for all covariates in the fully-adjusted model [31]. These plots were found to be approximately parallel and hence the proportionality assumption was taken to be satisfied by the data. All analyses were conducted by Stata 17 (StataCorp LLC, College Station, TX, USA), and the Cox model was fitted using `stcox` (Stata) procedure [32]. The study was exempt from Institutional Review Board approval as it utilized a de-identified public use dataset.

RESULTS

Descriptive statistics

Our study sample of cancer patients were more likely to be female (57.46%), non-Hispanic white (87.95%), those with a high school diploma (30.85%), those at or above 400% of the poverty level (34.35%), those with professional and managerial occupations (27.33%), homeowners (80.57%), currently married (61.61%), the US-born (93.21%), residents of the South (37.51%), those with excellent/very good/good health status (73.82%), those without psychological distress (42.31%), those with no regular physical activity (42.50%), those with BMI < 25 (36.62%), never smokers (45.29%), and current drinkers (57.40%) (Table 1). Lung cancer patients showed a similar pattern, except that they were more likely to be male (51.98%), unemployed or outside the labor force (31.82%), to report fair/poor health status (52.96%), and former smokers (64.05%). The total number of all-cause deaths and cancer deaths during the 19-year follow-up were 11,840 and 4,229 for all cancers combined, and 798 and 499 for lung cancer patients.

All-sites combined

1) Survival rates by sociodemographic characteristics

Lower survival was associated with racial minority and lower SES such as education, income, or occupation among cancer patients. The estimated difference in the 10-year age-adjusted all-cause survival rate between non-Hispanic black and non-Hispanic white cancer patients was 5.92 percentage points (79.52% vs. 85.44%). The estimated 10-year age-adjusted all-cause survival rate for cancer patients with less than high school was 81.58% and 88.38% for cancer patients with master's degree or higher (Table 2). The estimated difference in the 10-year age-adjusted all-cause survival rate between cancer patients below the poverty level and those with income \geq 400% of the FPL was 7.38 percentage points (80.63% vs. 88.01%). The 10-year age-adjusted all-cause survival rate for cancer patients with professional and managerial occupations was 89.66%, significantly higher than the survival rate for cancer patients with other occupations. The estimated 10-year age-adjusted all-cause survival rate for cancer patients renting house was 82.65%, significantly lower than the survival rate of 85.80% for cancer patients owning house. The 10-year age-adjusted all-cause survival rate for cancer patients currently married was 85.42%, higher than the survival rate for cancer patients divorced/separated (84.35%), or never married (82.75%). The estimated 10-year age-adjusted all-cause survival rate for US-born cancer patients was 84.98%, lower than the survival rate for foreign-born cancer patients residing in the US for 15 years or longer (85.94%). The gap in cancer survival narrowed with additional sociodemographic, health, or behavioral adjustment (Table S4).

2) Survival rates by health status and health behaviors

Lower survival was consistently associated with poor physical and mental health status. The estimated 10-year age-adjusted all-cause survival rate for cancer patients with fair or poor health status was 77.92%, significantly lower than the survival rate of 87.99% for cancer patients with excellent/very good/good health status. The estimated 10-year age-adjusted all-cause survival rate for cancer patients with serious psychological distress was 79.57%, significantly lower than the survival rate of 86.71% for cancer patients with no psychological distress.

The estimated 10-year age-adjusted all-cause survival rate for cancer patients with regular physical activity was 90.18%, significantly higher than the survival rate for cancer patients without regular physical activity (83.24%) or those with less than regular physical activity (87.92%). The 10-year age-adjusted all-cause survival rate for cancer patients with BMI < 25 was 84.49%, significantly higher than the survival rate for cancer patients with severe obesity (BMI \geq 40 [82.37%]) but lower than the survival rate for cancer patients with BMI 25 to 29 (85.93%), BMI 30 to 39 (84.83%). Among cancer

Table 1. Weighted proportion (%) of socio-demographic, health, and behavioral characteristics of US cancer patients and lung cancer patients aged ≥ 18 years, 1997-2015

Characteristics	All cancers patients (n = 40,291)		Lung cancer patients (n = 1,287)	
	Patients	Weighted proportion	Patients	Weighted proportion
All-cause deaths	11,840	26.89 (0.28)	798	62.37 (1.62)
Cancer deaths	4,229	9.87 (0.18)	499	39.62 (1.70)
Sex				
Male	15,640	42.54 (0.31)	626	51.98 (1.69)
Female	24,651	57.46 (0.31)	661	48.02 (1.69)
Race/ethnicity				
Non-Hispanic white	33,543	87.95 (0.21)	1,026	85.46 (1.08)
Non-Hispanic black	3,312	5.60 (0.15)	160	8.43 (0.87)
Hispanic/Latino	2,519	4.38 (0.13)	75	4.38 (0.58)
American Indian/Alaska native	212	0.50 (0.05)	9	0.71 (0.29)
Asian/Pacific islander	620	1.40 (0.07)	15	0.91 (0.28)
Non-Hispanic other race	85	0.17 (0.03)	2	0.10 (0.07)
Education				
Less than high school	6,666	14.31 (0.24)	320	22.98 (1.34)
High school	12,326	30.85 (0.29)	467	38.58 (1.60)
Some college	11,065	27.37 (0.30)	287	21.24 (1.44)
Bachelor's degree	6,034	16.31 (0.25)	127	10.41 (1.09)
\geq Master's degree	4,200	11.16 (0.22)	86	6.79 (0.83)
Poverty status (%), ratio of family income to FPL				
< 100	4,349	7.76 (0.16)	184	10.91 (0.99)
≥ 100 to < 200	7,103	14.91 (0.22)	282	19.59 (1.32)
≥ 200 to < 300	6,033	14.93 (0.23)	212	17.79 (1.31)
≥ 300 to < 400	4,047	10.57 (0.19)	122	10.13 (1.03)
≥ 400	11,857	34.35 (0.36)	251	22.74 (1.29)
Missing	6,902	17.48 (0.27)	236	18.84 (1.24)
Occupation				
Professional and managerial	10,018	27.33 (0.31)	210	17.88 (1.39)
Sales/clerical and tech support	8,827	22.44 (0.26)	219	17.63 (1.31)
Service	4,376	10.46 (0.19)	121	8.79 (0.91)
Craft and repair	4,549	11.84 (0.23)	204	16.01 (1.15)
Laborer	1,448	3.71 (0.13)	61	6.05 (0.89)
Others	595	1.50 (0.07)	25	1.81 (0.42)
Unemployed/outside labor force	10,478	22.72 (0.30)	447	31.82 (1.46)
Housing tenure (home ownership)				
Renter	10,178	19.43 (0.30)	373	22.90 (1.40)
Owner	30,113	80.57 (0.30)	914	77.10 (1.40)
Marital status				
Currently married	19,172	61.61 (0.31)	566	57.07 (1.52)
Widowed	9,122	15.62 (0.21)	353	20.18 (1.10)
Divorced/ separated	8,205	15.14 (0.21)	269	16.04 (1.06)
Never married	3,792	7.63 (0.16)	99	6.71 (0.97)
Years in the US ^a				
US-born	37,269	93.21 (0.16)	1,188	92.30 (0.95)
< 15 years	374	0.88 (0.06)	99	7.70 (0.95)
≥ 15 years	2,648	5.91 (0.15)		N/A
Region				
Northeast	6,877	18.10 (0.34)	236	19.63 (1.39)
Midwest	9,522	24.27 (0.37)	324	25.75 (1.38)
South	14,895	37.51 (0.48)	505	40.55 (1.61)
West	8,997	20.12 (0.37)	222	14.07 (1.22)

Table 1. Continued

Characteristics	All cancers patients (n = 40,291)		Lung cancer patients (n = 1,287)	
	Patients	Weighted proportion	Patients	Weighted proportion
Self-reported health status				
Excellent/very good/good	28,980	73.82 (0.29)	588	47.04 (1.72)
Fair/poor	11,311	26.18 (0.29)	699	52.96 (1.72)
K-6 distress scale				
0	16,689	42.31 (0.33)	465	36.73 (1.61)
1-2	8,362	21.05 (0.24)	228	17.74 (1.27)
3-5	6,552	16.24 (0.22)	207	15.28 (1.17)
6-12	6,001	14.21 (0.22)	255	20.42 (1.34)
≥ 13	2,059	4.57 (0.13)	99	7.01 (0.84)
Missing	628	1.61 (0.10)	33	2.81 (0.57)
PA				
No regular PA	18,246	42.50 (0.34)	827	63.34 (1.56)
Less than regular PA	10,704	27.38 (0.29)	266	21.05 (1.44)
Regular PA	10,388	27.55 (0.29)	175	13.74 (1.09)
Missing	953	2.57 (0.11)	19	1.87 (0.48)
BMI				
< 25	14,945	36.62 (0.28)	591	44.60 (1.71)
≥ 25 to < 30	14,043	35.30 (0.28)	420	33.73 (1.65)
≥ 30 to < 40	8,923	22.26 (0.26)	226	18.29 (1.24)
≥ 40	1,412	3.31 (0.11)	29	2.10 (0.46)
Missing	968	2.51 (0.09)	21	1.28 (0.32)
Smoking status				
Never smoker	18,348	45.29 (0.32)	203	15.60 (1.25)
Former smoker	14,910	37.75 (0.31)	804	64.05 (1.61)
Current smoker	7,033	16.96 (0.24)	280	20.34 (1.35)
Alcohol consumption				
Lifetime abstainer	8,178	18.67 (0.27)		
Former drinker	9,759	23.14 (0.28)		
Current drinker	22,023	57.40 (0.34)		
Unknown	331	0.79 (0.05)		

Values for each characteristic are presented as numbers or weighted proportion and standard errors. PA, physical activity; FPL, Federal Poverty Level; K-6, Kessler 6; BMI, body mass index. *For the lung cancer model, binary measure of the US-born/foreign-born variable was used and alcohol consumption was excluded.

patients, the 10-year age-adjusted all-cause survival rate for non-smokers was 88.31%, substantially higher than the survival rate for former smokers (84.48%) or current smokers (78.92%). Among cancer patients, the 10-year age-adjusted all-cause survival rate for lifetime abstainers was 85.10%, significantly higher than the survival rate for former drinkers (81.78%) but lower than the survival rate for current drinkers (86.49%). Similar to all-cause survival rates, the lower cancer-specific survival rates were associated with lower SES, poor health status, and health-risk behaviors (Table S6).

Lung cancer

1) Survival rates by sociodemographic characteristics

The covariate-adjusted survival rates for lung cancer were much lower than the survival rates for all cancers combined, but the pattern of association between individual characteristics and survivorship was consistent with all-cancer survival

rates (Table 3, Table S5). The estimated 10-year age-adjusted all-cause survival rate for lung cancer patients with less than high school education was 46.67%, substantially lower than the survival rate of 57.51% for lung cancer patients with master's degree or more (Table 3). The 10-year age-adjusted all-cause survival rate for lung cancer patients below the poverty level was 50.19%, significantly lower than the survival rate of 56.37% for lung cancer patients with income at or above 400% FPL. The 10-year age-adjusted all-cause survival rate for lung cancer patients with professional and managerial occupations was 59.22%, significantly higher than the survival rate of 51.43% for the unemployed. Differences in all-cause survival rates for lung cancer patients by housing tenure, marital status, and Nativity status were not statistically significant.

2) Survival rates by health status and health behaviors

The estimated 10-year age-adjusted all-cause survival rate

Table 2. Age-adjusted all-cause survival rates by socio-demographic, health, and behavioral characteristics among US cancer patients aged ≥ 18 years, 1997-2015 (n = 40,291)

Characteristics	1 year follow-up	5 year follow-up	10 year follow-up
Sex			
Male	98.79 (98.62, 98.94)	91.52 (90.91, 92.09)	80.94 (79.81, 82.02)
Female	99.23 (99.12, 99.32)	94.52 (94.13, 94.89)	87.42 (86.65, 88.15)
Race/ethnicity			
Non-Hispanic white	99.10 (98.97, 99.21)	93.61 (93.16, 94.03)	85.44 (84.59, 86.25)
Non-Hispanic black	98.69 (98.49, 98.86)	90.83 (90.00, 91.59)	79.52 (77.92, 81.02)
Hispanic	99.07 (98.92, 99.20)	93.44 (92.72, 94.09)	85.07 (83.6, 86.42)
American Indian/Alaska native	98.82 (98.39, 99.14)	91.74 (89.04, 93.79)	81.43 (75.89, 85.81)
Asian/Pacific islander	99.21 (99.01, 99.37)	94.40 (93.21, 95.39)	87.17 (84.61, 89.33)
Non-Hispanic other race	98.87 (98.14, 99.32)	92.09 (87.35, 95.10)	82.18 (72.49, 88.71)
Education			
Less than high school	98.83 (98.67, 98.98)	91.82 (91.20, 92.39)	81.58 (80.42, 82.68)
High school	99.04 (98.90, 99.16)	93.20 (92.69, 93.67)	84.53 (83.56, 85.45)
Some college	99.20 (99.08, 99.30)	94.31 (93.86, 94.73)	86.97 (86.08, 87.8)
Bachelor's degree	99.23 (99.11, 99.33)	94.50 (94.01, 94.95)	87.37 (86.38, 88.29)
\geq Master's degree	99.29 (99.18, 99.39)	94.95 (94.46, 95.41)	88.38 (87.36, 89.33)
Poverty status (%), ratio of family income to FPL			
< 100	98.77 (98.58, 98.93)	91.36 (90.63, 92.04)	80.63 (79.23, 81.95)
≥ 100 to < 200	98.95 (98.80, 99.08)	92.63 (92.06, 93.16)	83.32 (82.23, 84.36)
≥ 200 to < 300	99.02 (98.88, 99.15)	93.10 (92.54, 93.62)	84.34 (83.24, 85.37)
≥ 300 to < 400	99.09 (98.95, 99.21)	93.57 (92.99, 94.11)	85.36 (84.19, 86.45)
≥ 400	99.27 (99.16, 99.36)	94.78 (94.36, 95.17)	88.01 (87.17, 88.79)
Missing	99.11 (98.98, 99.22)	93.68 (93.18, 94.15)	85.59 (84.61, 86.51)
Occupation			
Professional and managerial	99.37 (99.28, 99.46)	95.53 (95.10, 95.91)	89.66 (88.78, 90.48)
Sales/clerical and tech support	99.33 (99.23, 99.42)	95.24 (94.79, 95.66)	89.03 (88.10, 89.89)
Service	99.24 (99.11, 99.34)	94.57 (93.99, 95.10)	87.55 (86.34, 88.65)
Craft and repair	98.99 (98.83, 99.13)	92.86 (92.14, 93.51)	83.80 (82.37, 85.13)
Laborer	98.94 (98.75, 99.11)	92.56 (91.58, 93.43)	83.17 (81.16, 84.99)
Others	99.18 (98.99, 99.34)	94.20 (93.08, 95.13)	86.71 (84.34, 88.75)
Unemployed/ outside labor force	98.98 (98.83, 99.10)	92.79 (92.28, 93.27)	83.66 (82.69, 84.58)
Housing tenure (home ownership)			
Renter	98.91 (98.75, 99.04)	92.31 (91.73, 92.85)	82.65 (81.55, 83.69)
Owner	99.12 (99.00, 99.23)	93.77 (93.33, 94.19)	85.80 (84.96, 86.60)
Marital status			
Currently married	99.09 (98.97, 99.21)	93.59 (93.13, 94.03)	85.42 (84.53, 86.27)
Widowed	99.10 (98.97, 99.21)	93.62 (93.13, 94.08)	85.48 (84.52, 86.38)
Divorced/ separated	99.02 (98.88, 99.15)	93.10 (92.53, 93.62)	84.35 (83.25, 85.38)
Never married	98.91 (98.74, 99.06)	92.35 (91.60, 93.05)	82.75 (81.24, 84.15)
Years in the US			
US-born	99.06 (98.93, 99.18)	93.39 (92.93, 93.82)	84.98 (84.10, 85.81)
< 15 years	99.00 (98.66, 99.25)	92.93 (90.82, 94.57)	84.00 (79.56, 87.54)
≥ 15 years	99.13 (98.99, 99.25)	93.83 (93.19, 94.41)	85.94 (84.63, 87.14)
Region			
Northeast	99.03 (98.89, 99.15)	93.14 (92.59, 93.65)	84.43 (83.36, 85.43)
Midwest	99.06 (98.93, 99.18)	93.39 (92.88, 93.86)	84.97 (83.98, 85.91)
South	99.02 (98.89, 99.14)	93.11 (92.61, 93.59)	84.38 (83.41, 85.30)
West	99.17 (99.05, 99.27)	94.11 (93.65, 94.55)	86.55 (85.64, 87.42)
Self-reported health status			
Excellent/very good/good	99.27 (99.17, 99.36)	94.80 (94.43, 95.16)	87.99 (87.26, 88.69)
Fair/poor	98.59 (98.39, 98.76)	90.12 (89.41, 90.78)	77.92 (76.63, 79.15)
K-6 distress scale			
0	99.18 (99.07, 99.28)	94.20 (93.76, 94.60)	86.71 (85.87, 87.51)
1-2	99.17 (99.05, 99.27)	94.08 (93.61, 94.51)	86.46 (85.53, 87.32)
3-5	99.04 (98.90, 99.16)	93.22 (92.67, 93.73)	84.59 (83.52, 85.60)
6-12	98.88 (98.72, 99.03)	92.16 (91.53, 92.74)	82.30 (81.09, 83.44)
≥ 13	98.69 (98.48, 98.87)	90.86 (89.93, 91.71)	79.57 (77.75, 81.26)

Table 2. Continued

Characteristics	1 year follow-up	5 year follow-up	10 year follow-up
PA			
No regular PA	98.95 (98.81, 99.08)	92.61 (92.09, 93.09)	83.24 (82.27, 84.17)
Less than regular PA	99.26 (99.16, 99.36)	94.75 (94.33, 95.15)	87.92 (87.07, 88.72)
Regular PA	99.41 (99.32, 99.48)	95.76 (95.40, 96.10)	90.18 (89.44, 90.88)
BMI			
< 25	99.03 (98.90, 99.15)	93.16 (92.67, 93.63)	84.49 (83.54, 85.39)
≥ 25 to < 30	99.13 (99.01, 99.24)	93.83 (93.37, 94.26)	85.93 (85.04, 86.78)
≥ 30 to < 40	99.05 (98.92, 99.17)	93.33 (92.79, 93.82)	84.83 (83.79, 85.81)
≥ 40	98.89 (98.68, 99.06)	92.18 (91.13, 93.11)	82.37 (80.23, 84.30)
Smoking status			
Never smoker	99.29 (99.19, 99.38)	94.92 (94.54, 95.28)	88.31 (87.56, 89.02)
Former smoker	99.03 (98.90, 99.15)	93.17 (92.67, 93.64)	84.48 (83.51, 85.39)
Current smoker	98.65 (98.45, 98.82)	90.55 (89.81, 91.23)	78.92 (77.54, 80.22)
Alcohol consumption			
Lifetime abstainer	99.07 (98.94, 99.19)	93.45 (92.95, 93.93)	85.10 (84.11, 86.03)
Former drinker	98.85 (98.68, 98.99)	91.91 (91.31, 92.47)	81.78 (80.65, 82.86)
Current drinker	99.17 (99.05, 99.27)	94.09 (93.66, 94.49)	86.49 (85.66, 87.27)
Unknown	98.82 (98.51, 99.06)	91.70 (89.95, 93.16)	81.34 (77.75, 84.42)

Values for each characteristic are presented as survival rates and 95% confidence interval. All cancer patients were defined by individuals who were ever diagnosed with cancer regardless of cancer type. The total number of all-cause deaths was 11,840 (272 deaths for 1-year follow-up; 1,907 deaths for 5-year follow-up; 4,207 deaths for 10-year follow-up). Age-adjusted survival rates were estimated after estimating Cox proportional hazards model adjusted for age and survey year. PA, physical activity; FPL, Federal Poverty Level; K-6, Kessler 6; BMI, body mass index.

for lung cancer patients with fair or poor health status was 42.38%, significantly lower than the survival rate of 63.26% for lung cancer patients with excellent/very good/good health status. The estimated 1-year age-adjusted all-cause survival rate for lung cancer patients with serious psychological distress was 40.96%, significantly lower than the survival rate of 58.93% for lung cancer patients with no psychological distress.

The estimated 10-year age-adjusted all-cause survival rate for lung cancer patients with regular physical activity was 67.10%, significantly higher than the survival rate for lung cancer patients without regular physical activity (48.36%) or those with less than regular physical activity (61.57%). The estimated 10-year age-adjusted all-cause survival rate for lung cancer patients with BMI < 25 was 49.88%, lower than the survival rate for lung cancer patients with BMI 25 to 29 (54.45%), BMI 30 to 39 (59.75%), or BMI ≥ 40 (64.33%). Among lung cancer patients, the estimated 10-year age-adjusted all-cause survival rate for non-smokers was 65.00%, significantly higher than the survival rate for former smokers (53.52%) or current smokers (44.79%). We found a similar pattern of association between cancer-specific survival rates and sociodemographic, health, and behavioral characteristics for lung cancer patients (Table S7).

DISCUSSION

In this large prospective study of 40,291 US cancer patients using a relatively long mortality follow-up of 19 years, we

found marked and consistent gradients in age-adjusted all-cause and cancer-specific survival rates by sociodemographic, health, and behavioral characteristics. Even after controlling for socioeconomic and demographic covariates, health-risk behaviors, and health status, significant disparities in survival rates existed but narrowed for the cancer patients. Computation of predicted survival rates for each characteristic, adjusting for other sociodemographic, health, and behavioral characteristics, is a particularly novel feature of our study.

Our findings on survival gap by sex, race/ethnicity, education, income was consistent with previous studies in terms of higher survival rates for females, whites, individuals with high education level, high income, or insurance, married individuals, home owners, or those in Western states, compared with their counterparts [1,2,11-14]. However, in terms of the magnitude, we found higher survival rates since we estimated survival rates for the prototypical individuals who are most likely to have higher survival rates than others given their better SES and health-risk profiles. Our estimates are not directly comparable with the average survival rates for cancer patients, but we recommend interpreting the results by comparing across covariate categories and focusing on the gap in survival, such as between homeowners and renters.

Our study contributes to the existing literature on social determinants of health and cancer health disparities by estimating survival rates by various individual-level sociodemographic, behavioral, and health status factors among cancer patients, which have not been well studied in prior studies.

Table 3. Age-adjusted all-cause survival rates by socio-demographic, health, and behavioral characteristics among US lung cancer patients aged ≥18 years, 1997-2015 (n = 1,287)

Characteristics	1 year follow-up	5 year follow-up	10 year follow-up
Sex			
Male	93.07 (90.45, 94.98)	66.50 (58.86, 73.04)	49.25 (40.14, 57.73)
Female	94.38 (92.21, 95.95)	72.00 (65.17, 77.72)	56.54 (47.83, 64.35)
Race/ethnicity			
Non-Hispanic white	93.90 (91.59, 95.59)	69.95 (62.93, 75.91)	53.76 (45.02, 61.72)
Non-Hispanic black	92.97 (89.96, 95.10)	66.14 (56.92, 73.83)	48.77 (37.80, 58.85)
Hispanic	94.14 (91.01, 96.20)	70.98 (60.21, 79.34)	55.14 (41.61, 66.76)
American Indian/Alaska native	97.62 (90.43, 99.43)	87.24 (57.05, 96.73)	78.88 (37.77, 94.38)
Asian/Pacific islander	93.39 (87.28, 96.62)	67.85 (47.39, 81.75)	50.98 (27.44, 70.40)
Non-Hispanic other race	98.60 (90.14, 99.81)	92.34 (55.90, 98.91)	87.07 (36.44, 98.12)
Education			
Less than high school	92.58 (89.71, 94.67)	64.51 (56.24, 71.62)	46.67 (37.03, 55.74)
High school	94.09 (91.79, 95.76)	70.72 (63.54, 76.74)	54.75 (45.73, 62.89)
Some college	94.94 (92.78, 96.47)	74.45 (67.10, 80.40)	59.87 (50.22, 68.24)
Bachelor's degree	93.26 (90.12, 95.42)	67.25 (57.27, 75.40)	50.16 (38.16, 61.02)
≥ Master's degree	94.56 (91.82, 96.39)	72.75 (63.22, 80.19)	57.51 (45.26, 67.98)
Poverty status (%), ratio of family income to FPL			
< 100	93.24 (90.29, 95.32)	67.23 (58.01, 74.87)	50.19 (39.07, 60.31)
≥ 100 to < 200	93.21 (90.47, 95.18)	67.09 (58.80, 74.07)	50.00 (40.04, 59.16)
≥ 200 to < 300	93.73 (91.16, 95.58)	69.28 (61.20, 76.00)	52.87 (42.87, 61.90)
≥ 300 to < 400	93.66 (90.70, 95.70)	68.96 (59.34, 76.74)	52.44 (40.62, 62.97)
≥ 400	94.35 (91.99, 96.03)	71.89 (64.21, 78.20)	56.37 (46.59, 65.04)
Missing	93.95 (91.44, 95.74)	70.19 (62.19, 76.81)	54.08 (44.08, 63.05)
Occupation			
Professional and managerial	94.85 (92.30, 96.57)	74.04 (65.03, 81.06)	59.22 (47.45, 69.20)
Sales/clerical and tech support	96.23 (94.32, 97.51)	80.40 (73.06, 85.93)	68.37 (58.07, 76.64)
Service	93.63 (90.20, 95.88)	68.78 (57.32, 77.75)	52.09 (38.08, 64.37)
Craft and repair	94.62 (91.89, 96.45)	73.05 (63.46, 80.50)	57.85 (45.48, 68.37)
Laborer	94.27 (90.64, 96.52)	71.51 (58.60, 81.03)	55.75 (39.57, 69.19)
Others	94.46 (89.61, 97.08)	72.34 (54.76, 84.02)	56.87 (35.13, 73.75)
Unemployed/ outside labor force	93.50 (91.08, 95.29)	68.28 (61.09, 74.42)	51.43 (42.65, 59.51)
Housing tenure (home ownership)			
Renter	93.01 (90.22, 95.02)	66.29 (57.99, 73.33)	48.96 (39.06, 58.14)
Owner	93.90 (91.62, 95.58)	69.99 (63.07, 75.86)	53.80 (45.20, 61.65)
Marital status			
Currently married	93.57 (91.09, 95.38)	68.63 (61.14, 74.98)	52.04 (42.86, 60.44)
Widowed	93.59 (91.03, 95.44)	68.73 (60.83, 75.35)	52.16 (42.47, 60.98)
Divorced/ separated	93.85 (91.34, 95.65)	69.81 (61.91, 76.39)	53.60 (43.76, 62.46)
Never married	94.30 (91.35, 96.26)	71.71 (61.51, 79.64)	56.15 (43.22, 67.23)
Nativity status			
US-born	93.63 (91.26, 95.37)	68.87 (61.81, 74.90)	52.34 (43.67, 60.3)
Foreign-born	94.55 (91.81, 96.39)	72.79 (63.24, 80.24)	57.61 (45.34, 68.09)
Region			
Northeast	93.35 (90.65, 95.29)	67.68 (59.44, 74.61)	50.73 (40.73, 59.88)
Midwest	94.33 (92.00, 96.00)	71.82 (64.24, 78.06)	56.24 (46.59, 64.81)
South	93.07 (90.41, 95.02)	66.55 (58.71, 73.24)	49.27 (39.89, 57.96)
West	94.62 (92.38, 96.22)	73.08 (65.64, 79.16)	57.96 (48.35, 66.42)
Self-reported health status			
Excellent/very good/good	95.58 (93.83, 96.83)	77.09 (71.14, 81.97)	63.26 (55.21, 70.26)
Fair/poor	91.87 (88.85, 94.09)	61.40 (53.16, 68.62)	42.38 (33.16, 51.30)
K-6 distress scale			
0	94.82 (92.77, 96.30)	73.83 (67.13, 79.39)	58.93 (50.19, 66.65)
1-2	94.29 (91.86, 96.02)	71.54 (63.51, 78.10)	55.77 (45.55, 64.81)
3-5	93.48 (90.80, 95.40)	68.10 (59.77, 75.07)	51.19 (41.03, 60.44)
6-12	92.55 (89.56, 94.71)	64.31 (55.55, 71.78)	46.32 (36.13, 55.90)
≥ 13	91.41 (87.44, 94.17)	59.93 (48.55, 69.58)	40.96 (28.55, 52.96)

Table 3. Continued

Characteristics	1 year follow-up	5 year follow-up	10 year follow-up
PA			
No regular PA	92.98 (90.38, 94.90)	66.00 (58.49, 72.46)	48.36 (39.44, 56.70)
Less than regular PA	95.26 (93.21, 96.70)	75.77 (68.63, 81.51)	61.57 (52.04, 69.75)
Regular PA	96.08 (94.27, 97.33)	79.60 (72.85, 84.84)	67.10 (57.70, 74.87)
BMI			
< 25	93.21 (90.65, 95.09)	67.04 (59.50, 73.49)	49.88 (40.81, 58.29)
≥ 25 to < 30	94.04 (91.69, 95.74)	70.50 (63.14, 76.67)	54.45 (45.21, 62.78)
≥ 30 to < 40	94.93 (92.70, 96.49)	74.37 (66.75, 80.50)	59.75 (49.75, 68.39)
≥ 40	95.64 (91.37, 97.82)	77.60 (60.83, 87.86)	64.33 (42.25, 79.79)
Smoking status			
Never smoker	95.74 (93.79, 97.09)	78.08 (71.00, 83.64)	65.00 (55.30, 73.10)
Former smoker	93.88 (91.56, 95.57)	69.84 (62.81, 75.79)	53.52 (44.78, 61.48)
Current smoker	92.20 (89.10, 94.45)	63.05 (54.18, 70.67)	44.79 (34.65, 54.41)

Values for each characteristic are presented as survival rates and 95% confidence interval. The total number of all-cause deaths was 798 (70 deaths for 1-year follow-up; 345 deaths for 5-year follow-up; 519 deaths for 10-year follow-up). Age-adjusted survival rates were estimated after estimating Cox proportional hazards model adjusted for age and survey year. PA, physical activity; FPL, Federal Poverty Level; K-6, Kessler 6; BMI, body mass index.

For example, estimating cancer survivorship associated with self-assessed health status, psychological distress, smoking, and BMI is especially new and shows the significance of these factors in determining cancer survival. Showing the large gap in all-cause survival between non-smokers and current smokers and between cancer patients with regular physical activity and those with no regular physical activity, our study findings provided the evidence of smoking status and physical activity as important factors in all-cause survival for cancer patients. Public health programs and policy interventions such as physical activity and tobacco control/cessation campaigns among cancer patients might be effective measures to improve survival rates.

Persistence of marked disparities in cancer survivorship among racial/ethnic and SES groups might reflect healthcare access and treatment disparities and shows the potential for further improvements in cancer survival [5,33]. Future research is needed to examine differentials in survival rates by neighborhood SES or built environmental factors, levels of urbanization/rurality [10,33,34], social supports [35], or quality of health care [36-38]. Although we focused on individual-level SES factors due to data availability, area-based SES has been strongly associated with cancer survival [10,33,34]. Cancer patients in neighborhoods with lower SES have markedly lower survival rates than those in higher SES neighborhoods [10,33,34]. Furthermore, it would be important to determine if similar individual patient-level socio-behavioral inequalities in survival exist for other major cancer sites such as colorectal, prostate, breast, and cervical cancer. Social isolation has been shown to increase risk of all-cause mortality by 66% and breast cancer mortality by 114% [35]. Considering that our study found significant disparities in survival by psychological distress among cancer patients, it may be worthwhile to examine the effect of any mediators

such as social network or supports, alleviating the association between psychological distress and cancer survival. High quality cancer care is also important in improving survivorship among cancer patients. Stratifying follow-up care by cancer patients' risk, such as a supported self-management for patients at low risk, a shared care for those at moderate risk, and complex case management for those at high risk, will help improve quality of care [37,38]. Using patient-reported outcome measures integrated into eHealth platforms will help stratify risk among patients [37,38]. Based on risk-stratification, various healthcare professionals, including survivorship clinics, psychosocial and medical experts, and consultants, need to collaborate at all points of the cancer care pathways from cancer diagnosis through end of life care [36-38].

This study has limitations. First, our study only contains the NHIS sample eligible for linkage to the NDI. Excluding samples ineligible for linkage may lead to biased mortality or survival estimates. To address this bias, we used the adjusted original sampling weight to account for the NHIS-NDI mismatches [23]. Second, our findings may be affected by the omitted-variable bias. While our Cox regression models were controlled for self-reported BMI, fair/poor health status, and psychological distress, there could be other potential confounders of cancer-related information, such as site, stage at diagnosis, and first course treatment, all of which are not available in the NHIS. Third, since the NHIS excludes the institutionalized population, cancer survivors participating in NHIS might be likely diagnosed with earlier stage, more treatable, and less fatal cancers than cancer survivors as a whole, which would have overestimated the survival time. Our survey analyses are not generalizable to the total cancer patients due to this potential selection bias. Finally, all the covariates in the NHIS-NDI database were time-fixed at the baseline as of the survey date. Several of the covariates

such as SES, health status, behavioral risk factors, and psychological distress could have varied over the long mortality follow-up period of 19 years, which would have influenced their estimated impact on cancer mortality and survivorship. In particular, mental health-related indicators and subjective health status can be severely affected by reverse causation as the more advanced the stage, the worse these indicators can be. Future studies need to examine the association using longitudinal data on individual sociodemographic and health-related characteristics.

In a nationally representative study, we found marked and consistent gradients in age-adjusted all-cause and cancer survival rates by socioeconomic, demographic, health, and behavioral characteristics. The disparities in survival rates for the cancer patients narrowed with the additional covariate adjustment. Smoking status and physical activity were the most important risk factors influencing all-cause survival among cancer patients, while self-reported health status and psychological distress showed the largest differentials in cancer-specific survival.

FUNDING

None.

CONFLICTS OF INTEREST

No potential conflicts of interest were disclosed.

SUPPLEMENTARY MATERIALS

Supplementary materials can be found via <https://doi.org/10.15430/JCP.2022.27.2.89>.

ORCID

Hyunjung Lee, <https://orcid.org/0000-0003-0783-5484>
Gopal K. Singh, <https://orcid.org/0000-0002-7820-0323>

REFERENCES

1. American Cancer Society. Cancer facts & figures 2021. Atlanta, American Cancer Society, 2021.
2. Howlader N, Noone AM, Krapcho M, Miller D, Brest A, Yu M, et al. SEER Cancer Statistics Review, 1975-2017, National Cancer Institute. https://seer.cancer.gov/csr/1975_2017/. Accessed Dec 1, 2021.
3. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2021. *CA Cancer J Clin* 2021;71:7-33. Erratum in: *CA Cancer J Clin* 2021;71:359.
4. American Cancer Society. Cancer prevention & early detection facts & figures 2019-2020. Atlanta, American Cancer Society, 2019.
5. Singh GK, Jemal A. Socioeconomic and racial/ethnic disparities in cancer mortality, incidence, and survival in the United States, 1950-2014: over six decades of changing patterns and widening inequalities. *J Environ Public Health* 2017;2017:2819372.
6. Sidorchuk A, Agardh EE, Aremu O, Hallqvist J, Allebeck P, Moradi T. Socioeconomic differences in lung cancer incidence: a systematic review and meta-analysis. *Cancer Causes Control* 2009;20:459-71.
7. Islami F, Goding Sauer A, Miller KD, Siegel RL, Fedewa SA, Jacobs EJ, et al. Proportion and number of cancer cases and deaths attributable to potentially modifiable risk factors in the United States. *CA Cancer J Clin* 2018;68:31-54.
8. Ward E, Jemal A, Cokkinides V, Singh GK, Cardinez C, Ghafoor A, et al. Cancer disparities by race/ethnicity and socioeconomic status. *CA Cancer J Clin* 2004;54:78-93.
9. Lee H, Singh GK. The association between psychological distress and cancer mortality in the United States: results from the 1997-2014 NHIS-NDI record linkage study. *Ann Behav Med* 2021;55:621-40.
10. Finke I, Behrens G, Weisser L, Brenner H, Jansen L. Socioeconomic differences and lung cancer survival-systematic review and meta-analysis. *Front Oncol* 2018;8:536.
11. Stubbs VC, Rajasekaran K, Cannady SB, Newman JG, Ibrahim SA, Brant JA. Social determinants of health and survivorship in parotid cancer: an analysis of the National Cancer Database. *Am J Otolaryngol* 2020;41:102307.
12. Niu X, Roche LM, Pawlish KS, Henry KA. Cancer survival disparities by health insurance status. *Cancer Med* 2013;2:403-11.
13. Ellington TD, Henley SJ, Wilson RJ, Miller JW. Breast cancer survival among males by race, ethnicity, age, geographic region, and stage - United States, 2007-2016. *MMWR Morb Mortal Wkly Rep* 2020;69:1481-4.
14. Kogevinas M, Marmot MG, Fox AJ, Goldblatt PO. Socioeconomic differences in cancer survival. *J Epidemiol Community Health* 1991;45:216-9.
15. Zavala VA, Bracci PM, Carethers JM, Carvajal-Carmona L, Coggins NB, Cruz-Correa MR, et al. Cancer health disparities in racial/ethnic minorities in the United States. *Br J Cancer* 2021;124:315-32.
16. Ward EM, Sherman RL, Henley SJ, Jemal A, Siegel DA, Feuer EJ, et al. Annual report to the nation on the status of cancer, featuring cancer in men and women age 20-49 years. *J Natl Cancer Inst* 2019;111:1279-97.
17. Edwards BK, Noone AM, Mariotto AB, Simard EP, Boscoe FP, Henley SJ, et al. Annual report to the nation on the status of cancer, 1975-2010, featuring prevalence of comorbidity and impact on survival among persons with lung, colorectal, breast, or prostate cancer. *Cancer* 2014;120:1290-314.
18. Blewett L, Drew J, King M, Williams K. IPUMS health surveys: national health interview survey, version 6.4. <https://doi.org/10.18128/D070.V6.4>. Accessed December 1, 2021.
19. National Center for Health Statistics. National Health Interview Survey. https://www.cdc.gov/nchs/nhis/about_nhis.htm. Accessed December 1, 2021.

20. National Center for Health Statistics (NCHS). Survey description, National Health Interview Survey, 2014. Hyattsville, National Center for Health Statistics, 2015.
21. National Center for Health Statistics. Survey description, National Health Interview Survey, 1997. Hyattsville, National Center for Health Statistics, 2000.
22. National Center for Health Statistics (NCHS). Public-use Linked Mortality File. Hyattsville, National Center for Health Statistics, 2019.
23. National Center for Health Statistics (NCHS). The Linkage of National Center for Health Statistics survey data to the National Death Index- 2015 Linked Mortality File (LMF): methodology overview and analytic considerations. Hyattsville, National Center for Health Statistics, 2019.
24. Siahpush M, Singh GK, Tibbits M, Pinard CA, Shaikh RA, Yaroch A. It is better to be a fat ex-smoker than a thin smoker: findings from the 1997-2004 National Health Interview Survey-National Death Index linkage study. *Tob Control* 2014;23:395-402.
25. Kalbfleisch JD, Prentice RL. The statistical analysis of failure time data. 2nd ed. Hoboken, John Wiley & Sons, 2011.
26. Cefalu M. Pointwise confidence intervals for the covariate-adjusted survivor function in the Cox. *Stata J* 2011;11:64-81.
27. Mariotto AB, Noone AM, Howlader N, Cho H, Keel GE, Garshell J, et al. Cancer survival: an overview of measures, uses, and interpretation. *J Natl Cancer Inst Monogr* 2014;2014:145-86. Erratum in: *J Natl Cancer Inst Monogr* 2015;2015:97.
28. Mariotto AB, Zou Z, Johnson CJ, Scoppa S, Weir HK, Huang B. Geographical, racial and socio-economic variation in life expectancy in the US and their impact on cancer relative survival. *PLoS One* 2018;13:e0201034.
29. Yamaguchi K. Event history analysis. Newbury Park, SAGE Publications, 1991.
30. Singer JD, Willett JB. Applied longitudinal data analysis: modeling change and event occurrence. New York, Oxford University Press, 2003.
31. Singh GK, Kposowa AJ. A comparative analysis of infant mortality in major Ohio cities: significance of socio-biological factors. *Appl Behav Sci Rev* 1994;2:77-94.
32. StataCorp. Stata statistical software: release 15. College Station, StataCorp LLC, 2017.
33. Singh GK, Williams SD, Siahpush M, Mulhollen A. Socio-economic, rural-urban, and racial inequalities in US cancer mortality: part I-all cancers and lung cancer and part II-colorectal, prostate, breast, and cervical cancers. *J Cancer Epidemiol* 2011;2011:107497.
34. Castañeda-Avila MA, Jesdale BM, Beccia A, Bey GS, Epstein MM. Differences in survival among multiple myeloma patients in the United States SEER population by neighborhood socioeconomic status and race/ethnicity. *Cancer Causes Control* 2021;32:1021-8.
35. Kroenke CH, Kubzansky LD, Schernhammer ES, Holmes MD, Kawachi I. Social networks, social support, and survival after breast cancer diagnosis. *J Clin Oncol* 2006;24:1105-11.
36. Loonen JJ, Blijlevens NM, Prins J, Dona DJ, Den Hartogh J, Senden T, et al. Cancer survivorship care: person centered care in a multidisciplinary shared care model. *Int J Integr Care* 2018;18:4.
37. Jacobs LA, Shulman LN. Follow-up care of cancer survivors: challenges and solutions. *Lancet Oncol* 2017;18:e19-29.
38. Warrington L, Absolom K, Velikova G. Integrated care pathways for cancer survivors - a role for patient-reported outcome measures and health informatics. *Acta Oncol* 2015;54:600-8.