







RESEARCH ARTICLE

Open Access



Estimating life expectancy adjusted by self-rated health status in the United States: national health interview survey linked to the mortality

Hyunsoon Cho^{1*} , Zhuoqiao Wang², K. Robin Yabroff³ , Benmei Liu⁴ , Timothy McNeel² ,
Eric J. Feuer⁴  and Angela B. Mariotto⁴ 

Abstract

Background: Life expectancy is increasingly incorporated in evidence-based screening and treatment guidelines to facilitate patient-centered clinical decision-making. However, life expectancy estimates from standard life tables do not account for health status, an important prognostic factor for premature death. This study aims to address this research gap and develop life tables incorporating the health status of adults in the United States.

Methods: Data from the National Health Interview Survey (1986–2004) linked to mortality follow-up through to 2006 (age ≥ 40 , $n = 729,531$) were used to develop life tables. The impact of self-rated health (excellent, very good, good, fair, poor) on survival was quantified in 5-year age groups, incorporating complex survey design and weights. Life expectancies were estimated by extrapolating the modeled survival probabilities.

Results: Life expectancies incorporating health status differed substantially from standard US life tables and by health status. Poor self-rated health more significantly affected the survival of younger compared to older individuals, resulting in substantial decreases in life expectancy. At age 40 years, hazards of dying for white men who reported poor vs. excellent health was 8.5 (95% CI: 7.0, 10.3) times greater, resulting in a 23-year difference in life expectancy (poor vs. excellent: 22 vs. 45), while at age 80 years, the hazards ratio was 2.4 (95% CI: 2.1, 2.8) and life expectancy difference was 5 years (5 vs. 10). Relative to the US general population, life expectancies of adults (age < 65) with poor health were approximately 5–15 years shorter.

Conclusions: Considerable shortage in life expectancy due to poor self-rated health existed. The life table developed can be helpful by including a patient perspective on their health and be used in conjunction with other predictive models in clinical decision making, particularly for younger adults in poor health, for whom life tables including comorbid conditions are limited.

Keywords: Self-rated health status, Life expectancy, Survival, Mortality, health survey, National Health Interview Survey

Background

Life expectancy is the most common summary measure of mortality used to describe the overall health status of a population. It is increasingly incorporated in evidence-based screening and treatment guidelines to ensure that patients with limited life expectancy exposed to the potential harms of screening and treatment live long

*Correspondence: hscho@ncc.re.kr

¹ Department of Cancer Control and Population Health, Graduate School of Cancer Science and Policy, National Cancer Center, 323 Ilsan-ro, Ilsandong-gu, Goyang-si, Gyeonggi-do 10408, Republic of Korea
Full list of author information is available at the end of the article



enough to experience the potential benefits [1–4]. For example, the survival benefit from the early detection and effective treatment of colorectal or breast cancer may only occur 5 years after the screening test [2]. An individual with a shorter life expectancy (i.e., < 5 years) would not live long enough to benefit from early detection, yet such an individual will be exposed to the psychological distress of undergoing a cancer screening and diagnosis, complications from biopsies, and the risk of overtreatment. Thus, providing accurate life expectancy estimates that incorporate an individual's health status is important to identify individuals who may not benefit from screening or treatment and to include discussions on the pros and cons of the intervention in the clinical decision-making process.

Predictions of life expectancies for the average person are usually obtained from national life tables stratified by age, sex, and race [5]. However, the survival experience of the average person does not fully account for other important prognostic factors of premature death, such as health status, comorbidity, and various behavioral and socioeconomic risk factors. Previous studies have developed tools to estimate life expectancy using additional characteristics, such as comorbidities, [6–10] smoking status, [11] and socioeconomic factors [12–14]. However, these tools can have limited applications. For example, Medicare claim data life tables by comorbidity status have only been calculated for the elderly population (65 years and older) based on the comorbidities that were treated and reimbursed. However, such data may not be readily available at the point of care. Moreover, little research has been conducted in the younger population, where individuals with high comorbidity burden or poor health status are likely to have a very different life expectancy than an average individual of the same age.

Studies have shown that self-rated health is a good predictor of mortality, [15–31] even after controlling for various objective health measurements (e.g., disease conditions, health risk behaviors such as smoking and exercise). Self-rated health is a summary measure that represents an individual's evaluation of his/her own health, and captures the full array of illnesses a person has, possible disease symptoms yet to be diagnosed or treated, as well as the severity of current illness. Importantly, self-rated health consists of a single, easily understood and answered, and well-validated question that a physician or other health professional can query a patient, and it does not require comprehensive access to medical records or claim data [29]. However, studies estimating life tables solely by self-rated health in US adults are limited. Moreover, life expectancy by self-rated health status for the US adults compared to the US general population has not been estimated.

This study aims to address this research gap and develop life tables, and life expectancy estimates that incorporate self-rated health in an exceptionally large national sample of adults in the United States.

Methods

Data sources and study population

This study used data from the 1986–2004 public use National Health Interview Survey (NHIS) linked to mortality follow-up through to December 31, 2006 [32]. The NHIS is an annual in-person household survey of the civilian noninstitutionalized population of the United States conducted by the National Center for Health Statistics (NCHS). The civilian noninstitutional population includes adults who do not live in institutions (for example, correctional facilities, long-term care hospitals, and nursing homes) and who are not on active duty in the Armed Forces [33]. Response rates were well over 80% throughout the 19 years of NHIS data included in this study. The NCHS provides files to allow the linkage of NHIS participants to the National Death Index through December 31, 2006 for vital status information. The date of death was available in quarterly months for those who died; in addition, based on NCHS guidance, individuals that were not linked to the National Death Index were presumed to be alive on December 31, 2006. Adults younger than 40 years were excluded because 98.2% of those individuals were alive at the end of follow-up and would be considered censored cases in the survival analysis.

Measures

The self-ratings of health status were based on responses to the question in the NHIS Family Core component: “Would you say your health in general is excellent, very good, good, fair, or poor?” All 5 health status categories were evaluated in this study. Demographic characteristics such as age, sex, and race (White, Black, other/multiple, unknown) were also considered.

Survival estimation

Survival time was measured from the NHIS interview (month) to death (quarter) or end of follow-up in 2006. We estimated age-specific survival curves stratified by sex. We fit discrete time proportional hazards models by sex and 5-year age groups (i.e., 40–44, 45–49, 50–54, 55–59, 60–64, 65–69, 70–74, 75–79, 80–84, 85+). The following covariates were included in the model: health status (i.e., excellent, very good, good, fair, poor), race (i.e., white, black, other), survey year, and age, in a linear term within each age group. As proxy responses were accepted for family members that were not at home during the interview, reporting of self-rated health status

(i.e., self, proxy) was also adjusted in the analysis. The models were fitted separately for each age-group to account for changes in baseline hazards by age and to more accurately characterize discrepancies in mortality caused by health status in each age group. The discrete time complementary log-log model, [34–36] a discrete analog of the Cox proportional hazards model, was employed for the proportional hazards models. All statistical analyses were conducted using SAS 9.3 (SAS Institute, Cary NC, USA) and SAS-callable SUDAAN [37] to account for the complex NHIS survey design and weights using SAS PROC SURVEYLOGISTIC. Details on model formulation and the SAS programs are provided in the supplementary materials. Sensitivity analyses were conducted for different model assumptions, and they too are reported in the supplemental materials. As approximately 36% reports were from family members as opposed to being self-reported, the effects of the proxy reporting of health status on hazards of death was also assessed.

Life expectancy calculation

Life expectancy calculations require extrapolation of survival models beyond observed data. Although we had long-term follow-up, with the maximum follow-up at 21 years (Median: 9.8, IQR: 5.5, 14.7), annual survival probabilities beyond the follow-up period need to be estimated. Similar to other studies, the best matching US life table approach was used [6, 7] to extrapolate survival post-follow-up. For example, the life table estimated from the survival model for individuals aged 60 years stops at age 81 years. To extrapolate survival beyond age 81, we compared the modeled survival probabilities of a 60-year-old white man with excellent health to the age-specific US life tables of white men between ages 40 and 70, and identified the US life table of a 51-year-old white man as the best matching survival experience. Then, we defined a 51-year old as health adjusted age for a 60-year-old white man with excellent health. This best-matching life table was then used to calculate life expectancy (i.e., life expectancy at age 51 from the US life table is the health status adjusted life expectancy for a 60-year-old white man with excellent health). In this study, we used decennial 2000 US life tables by sex and race because they correspond closely to the study period [38, 39]. More details on these methods are provided in the supplemental materials (Supplemental Fig. 1).

Results

Demographics and health status

The study population consisted of 729,531 US adult respondents, aged 40 years and older at the time of the survey. The study population was predominately white (85.2%) and approximately half were women (53.4%)

Table 1 Characteristics of study population: US adult respondents, age 40 years and older at the time of the interview, National Health Interview Survey, 1986–2004, with mortality follow-up through 2006

Variable		Sample size, n ^a	Wt ^b %
Age at interview	40–44	136,139	18.9
	45–49	116,846	16.2
	50–54	98,198	13.6
	55–59	83,240	11.4
	60–64	75,184	10.1
	65–69	69,394	9.2
	70–74	58,886	7.9
	75–79	44,336	6.1
	80–84	28,184	3.9
	85+	19,124	2.7
Sex	Male	333,588	46.6
	Female	395,943	53.4
Race	White	601,184	85.2
	Black	90,655	9.9
	Other / multiple	34,368	4.5
	Unknown	3324	0.4
Health Status ^c	Excellent	178,179	24.9
	Very good	198,346	28.0
	Good	211,181	28.6
	Fair	97,056	12.7
	Poor	42,141	5.4
	Unknown	2628	0.3
Survival status	Dead	165,616	20.5
	Alive	563,915	79.5

^a Sample size, total sample size was 729,531

^b Weighted percent

^c Includes health status reported by self (63.4%), proxy (36%) and missing (0.6%) reporting status information

(Table 1). Health status ranged from 24.9% *excellent*, 28% *very good*, 28.6% *good*, 12.7% *fair*, to 5.4% *poor* (Table 1). The distribution of self-rated health status differed by age at the interview, with *excellent* health reported proportionately lower in the older participants, while *fair* to *poor* health status was higher in the older participants (Fig. 1). Health ratings were similar in men and women, though men were slightly more likely to rate their health as “excellent” relative to women in younger participants: for example, the proportion of those who rated their health as *excellent* among men and women at age 40 were: 39.2 and 36.4% in whites and 30.9 and 23.5% in blacks, respectively.

Impact of health status on survival

Individuals reporting poor and fair health had a much higher risk of death compared to individuals reporting

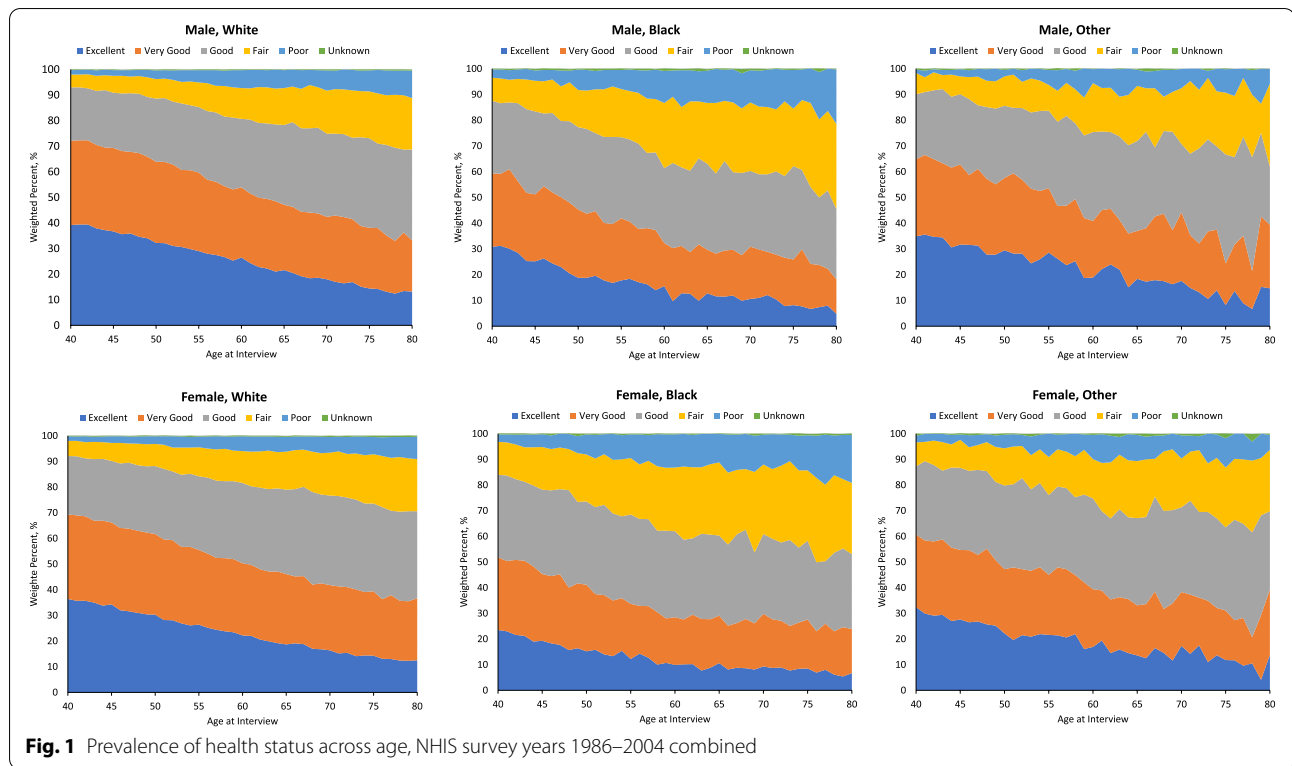


Fig. 1 Prevalence of health status across age, NHIS survey years 1986–2004 combined

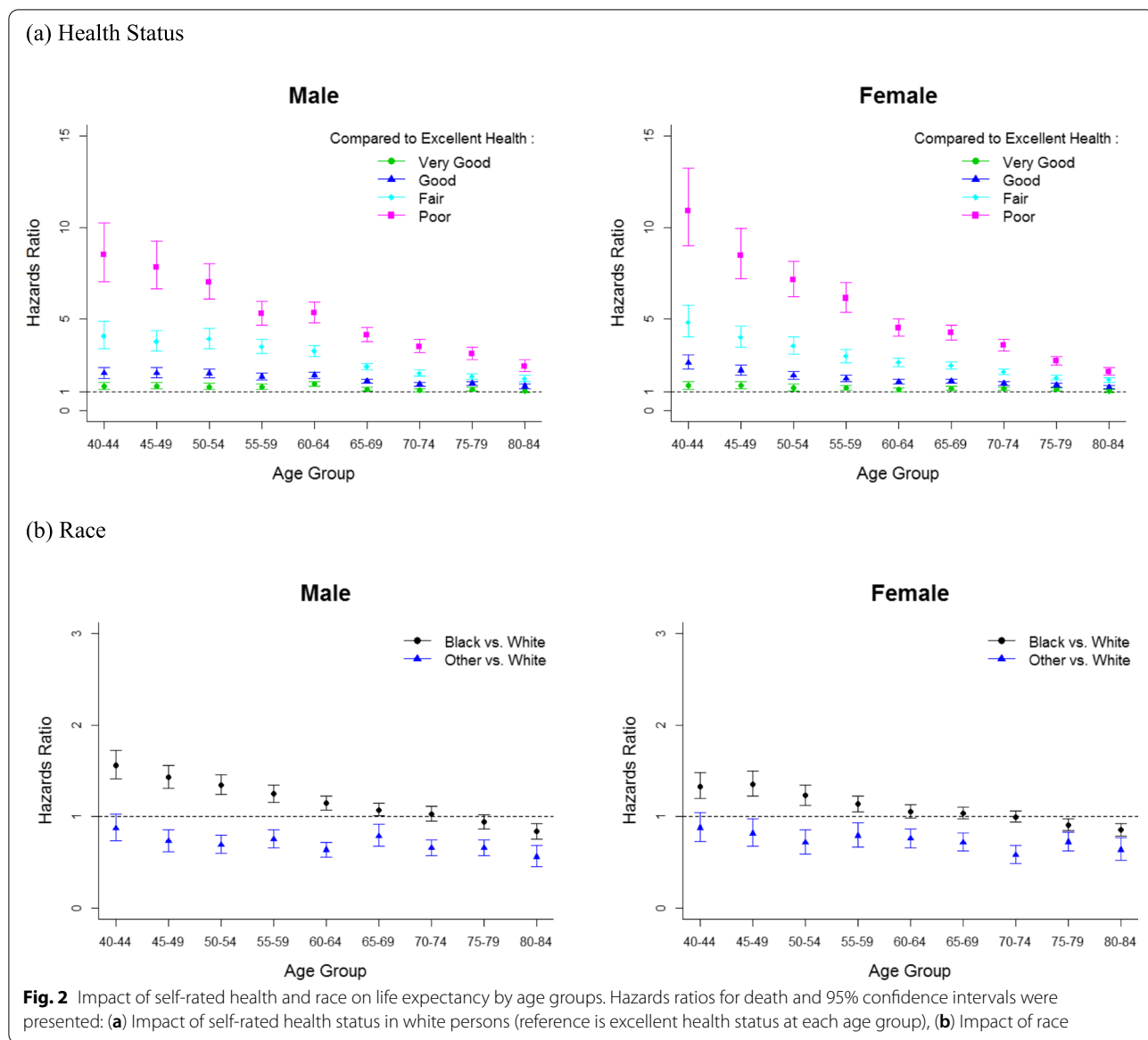
excellent health, and the difference in risk by health status was substantially higher in younger than older people (Fig. 2a). For example, among white men aged 40 years, those reporting *poor* health had 8.5 times (95% CI: 7.0, 10.3) greater risk of death compared with those reporting *excellent* health; whereas among white men aged 80 years, risk of death was 2.4 times higher (95% CI: 2.1, 2.8) for poor health compared to excellent health. In white women, comparisons of *poor* and *excellent* health correspond to 10.9 times (95% CI: 9.0, 13.2) greater risk of death at age 40 years and 2.1 times (95% CI: 1.9, 2.4) at age 80 years. Risk of death for those with *very good* and *good* health, compared to those in *excellent* health at the same chronological age, were similar across age groups. Figure 2b shows the effect of race on risk of death. Relative to white participants, black participants had a greater risk of death at younger ages but had similar or lower risk at older ages. At younger ages, the impact of health status on survival and life expectancy was greater in black compared to whites. While other races had a lower risk of death compared to whites overall.

Estimated survival varied greatly by age and health status among white participants (Fig. 3), while the results were similar for blacks and other races (data not shown). There is almost no or little difference in survival among health status at a younger age. However, the small change in the survival probability will result in a large difference

in life expectancy. The effect of health status on survival is strong for middle to old age. At older age, the effect of health status becomes smaller again. In other words, life expectancy in higher ages has less variability.

The survival experiences of individuals with favorable health status were generally similar (i.e., *excellent*, *very good*, and *good* health in younger participants; *excellent* and *very good* health in the elderly). As expected, survival probabilities were dramatically lower in older participants. Survival was close to 100% for individuals aged 40 to 55 with *excellent* to *good* health. Indeed, the survival estimated from the US life tables is generally close to the survival for people in *good* health. Individuals of all ages who reported *fair* and *poor* health had worse survival estimates relative to the US general population.

Notably, older individuals' survival was higher in the NHIS study population compared to the US general population. This is likely because the NHIS sample did not include institutionalized individuals, who likely have a higher mortality risk. For example, 5-year (and 10-year) survival of the average white woman conditional on surviving at age 80 is 78% (50%) in the NHIS population, while it is 72% (41%) in the US general population. As such, at age 80, the survival experience of the US general population was close to individuals who reported fair health in the NHIS.

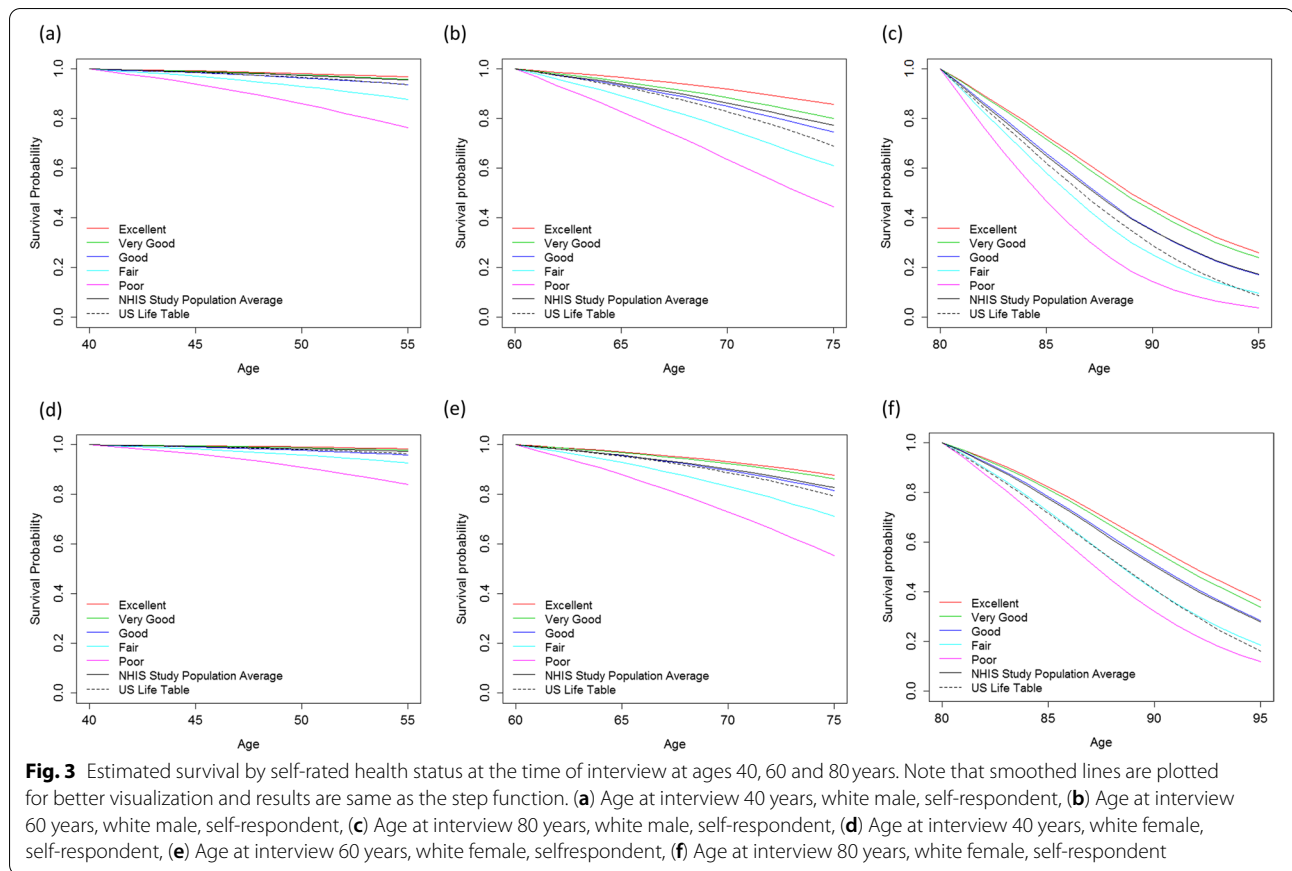


The sensitivity analysis used different models and found that the estimated survival outcomes were similar regardless of model selection (Supplemental Fig. 2). Compared to self-report, family reported (proxy-report) health status showed lower mortality risk (risk of death) in reported excellent health at all ages and higher mortality risk in individuals with poor health at older ages. However, the magnitudes were minor and the effects were not significant in some age groups (Supplemental Fig. 3).

Life expectancy adjusted by health status

Life expectancy varied considerably by self-rated health status. When compared to the US general population

matched by age, sex, and race, the life expectancies of those in *excellent* and *very good* health were longer, while of individuals with *fair* and *poor* health were shorter (Table 2). The remaining life expectancy at 60 males in the US general population was 20 for whites, 17 for blacks, 24 for others. In contrast, when considering health status, the remaining life expectancies varied substantially; those were 27 for *excellent*, 24 for *very good*, 21 for *good*, 17 for *fair*, 14 for *poor* health status in whites; 28, 25, 22, 17, 12 for blacks; 31, 28, 25, 20, 16 for other races. Variabilities between life expectancy by health status and general population average were more substantial in blacks than those in other races. The results were similar in the females; however, there existed a larger variability by



health status. Discrepancies in life expectancy by health status were greater in younger ages. For example, life expectancies of white men in *excellent* health versus *poor* health is 45 versus 22 years (i.e., 23-year difference) at age 40, whereas it is 10 versus 5 years (i.e., 5-year difference) at age 80. In other words, at age 40, life expectancy (the differences in life expectancy relative to the US population average) for *excellent*, *very good*, *good*, *fair*, *poor* were 45 (8), 42 (5), 37 (0), 29 (-8), 22 (-15); their remaining life vary greatly from 22 to 45, and the difference with US average was greater up to 15 years. In contrast, those were 10 (3), 10 (2), 8 (1), 7 (0), 5 (-2) at age 80, showing, at older ages, less variability in life expectancy. Life expectancy highly depends on age, and in older ages, the effect of other factors (such as self-rated health), will become less distinctive.

The life expectancies of the NHIS study population were closer to reported good health status rates in all ages (Table 2). Life expectancies of the average US general population in the elderly are shorter, particularly in blacks, than those estimated from *good* health status in the NHIS population. For example, in blacks, the average life expectancy of the US general population is 4 years shorter at age 65 and 3 years shorter at age 70 (Table 2).

Discussion

This study developed life tables and estimated the life expectancy by self-rated health for men and women in the United States using a large population-based dataset. Life expectancy was found to vary considerably by health status. Consistent with previous studies, [25, 27] self-rated health was found to be a reliable predictor of long-term survival, and individuals who rated their health as excellent had higher survival and life expectancy estimates than individuals with poor or fair health.

More importantly, self-rated health was found to have a greater effect on survival and life expectancy at younger ages compared to older ages. Younger people with poor health had a significantly shorter life expectancy compared to average persons at the same age. The life tables by health status presented here are an improvement compared to the national average and are particularly important for the understudied younger populations because clinical decision-making and patient-centered care may involve an assessment of benefits and harms for a younger person similar to that for an elderly person. Therefore, this study, by providing such data, represents an important contribution

Table 2 Life expectancy adjusted by self-rated health status compared with the life expectancies of the US average population

Male													
Age ^a	US Pop. Average LE ^b	Study Pop. Average LE ^c Diff ^d		Self-rated Health Status									
		LE ^e	Diff ^f	Excellent LE ^e Diff ^f	Very Good LE ^e Diff ^f	Good LE ^e Diff ^f	Fair LE ^e Diff ^f	Poor LE ^e Diff ^f					
White													
40	37	41	4	45 (8)	42 (5)	37 (0)	29 (-8)	22 (-15)					
45	32	36	4	42 (9)	38 (5)	33 (1)	26 (-6)	20 (-13)					
50	28	32	3	37 (9)	33 (5)	29 (1)	22 (-6)	18 (-10)					
55	24	27	3	32 (8)	29 (5)	26 (2)	20 (-4)	16 (-8)					
60	20	22	2	27 (7)	24 (4)	21 (2)	17 (-3)	14 (-6)					
65	16	18	2	22 (6)	20 (4)	18 (1)	14 (-2)	10 (-6)					
70	13	15	2	18 (5)	16 (3)	14 (1)	12 (-1)	9 (-4)					
75	10	12	2	14 (4)	13 (3)	12 (2)	9 (-1)	7 (-3)					
80	7	8	1	10 (3)	10 (2)	8 (1)	7 (0)	5 (-2)					
Black													
40	32	37	4	44 (12)	39 (7)	35 (3)	27 (-5)	20 (-12)					
45	28	33	5	40 (12)	37 (9)	32 (4)	25 (-3)	18 (-10)					
50	24	29	5	36 (12)	33 (9)	28 (4)	21 (-3)	15 (-9)					
55	20	25	4	32 (12)	30 (9)	26 (5)	18 (-2)	15 (-6)					
60	17	20	3	28 (11)	25 (8)	22 (5)	17 (-1)	12 (-5)					
65	14	17	3	23 (9)	20 (6)	18 (4)	14 (0)	10 (-4)					
70	11	14	3	18 (6)	17 (6)	15 (3)	12 (1)	9 (-3)					
75	9	11	2	15 (6)	14 (4)	12 (3)	10 (1)	7 (-2)					
80	7	9	2	11 (4)	11 (4)	10 (2)	8 (1)	6 (-1)					
Other													
40	42	41	-1	46 (5)	43 (1)	37 (-5)	31 (-11)	23 (-19)					
45	37	37	0	43 (6)	40 (3)	35 (-2)	29 (-8)	22 (-15)					
50	32	34	2	40 (7)	37 (5)	32 (-1)	25 (-7)	20 (-13)					
55	28	29	1	34 (6)	32 (4)	28 (0)	22 (-6)	18 (-10)					
60	24	25	1	31 (7)	28 (4)	25 (1)	20 (-4)	16 (-8)					
65	20	20	1	25 (5)	23 (3)	20 (0)	17 (-3)	13 (-6)					
70	16	17	1	20 (5)	19 (3)	17 (1)	15 (-1)	11 (-5)					
75	13	14	1	17 (4)	16 (3)	14 (1)	12 (-1)	9 (-3)					
80	10	11	2	14 (4)	13 (4)	12 (2)	10 (1)	8 (-2)					
Female													
Age ^a	US Pop. Average LE ^b	Study Pop. Average LE ^c Diff ^d		Self-rated Health Status									
		LE ^e	Diff ^f	Excellent LE ^e Diff ^f	Very Good LE ^e Diff ^f	Good LE ^e Diff ^f	Fair LE ^e Diff ^f	Poor LE ^e Diff ^f					
White													
40	41	44	3	51 (10)	47 (6)	39 (-2)	34 (-7)	26 (-16)					
45	37	39	3	45 (9)	41 (5)	36 (-1)	30 (-6)	22 (-14)					
50	32	35	3	40 (8)	38 (6)	33 (1)	28 (-4)	21 (-11)					
55	28	30	3	35 (7)	33 (5)	29 (2)	24 (-3)	18 (-10)					
60	23	25	2	29 (6)	28 (4)	25 (2)	20 (-3)	15 (-8)					
65	19	21	2	25 (6)	23 (4)	21 (2)	17 (-2)	13 (-6)					
70	15	17	1	20 (5)	19 (4)	17 (1)	15 (-1)	11 (-4)					
75	12	14	2	15 (3)	15 (3)	14 (2)	12 (0)	9 (-3)					
80	9	11	2	12 (3)	12 (3)	11 (2)	10 (1)	8 (-2)					
Black													
40	38	42	5	53 (15)	49 (11)	41 (4)	34 (-4)	24 (-13)					
45	33	37	4	47 (14)	43 (10)	38 (4)	30 (-3)	22 (-11)					
50	29	33	4	42 (13)	40 (10)	34 (5)	27 (-2)	20 (-9)					
55	25	28	3	38 (13)	35 (10)	31 (6)	24 (-1)	18 (-7)					
60	21	24	3	32 (10)	30 (9)	26 (5)	20 (-1)	16 (-6)					
65	18	20	3	27 (9)	25 (7)	22 (4)	18 (0)	14 (-4)					
70	14	17	3	22 (8)	20 (6)	18 (4)	15 (1)	11 (-3)					
75	11	14	2	17 (5)	16 (5)	14 (3)	13 (1)	9 (-2)					
80	9	11	3	14 (5)	13 (4)	12 (3)	10 (1)	9 (0)					
Other													
40	46	44	-2	51 (6)	49 (3)	41 (-5)	35 (-10)	27 (-19)					
45	41	39	-2	47 (6)	44 (3)	38 (-3)	32 (-9)	24 (-16)					
50	36	37	1	44 (8)	40 (4)	36 (0)	30 (-6)	23 (-13)					
55	32	32	0	38 (6)	35 (4)	32 (0)	26 (-5)	20 (-12)					
60	27	26	-1	32 (4)	30 (3)	27 (0)	22 (-5)	17 (-10)					
65	23	23	0	28 (5)	26 (3)	24 (1)	20 (-3)	16 (-7)					
70	19	20	2	24 (6)	23 (4)	21 (2)	18 (-1)	14 (-4)					
75	15	17	1	19 (4)	18 (3)	17 (1)	14 (-1)	12 (-3)					
80	12	13	1	15 (3)	15 (3)	14 (2)	12 (1)	10 (-2)					

Table 2 (continued)

For an easy representation of the results, differences in life expectancies from 3 to 5 is shaded in light colors, 6 to 10 is shaded in medium colors, and more than 10 is shaded in dark colors. Blue colors represent greater and orange colors represent shorter life expectancies relative to the US average population

^a Chronological age

^b Life expectancy obtained from 2000 US decennial life table

^c Average life expectancy of study population estimated from the model

^d Differences in life expectancies calculated as the average life expectancy in the study population minus life expectancy in the average US population (2000 US life table)

^e Life expectancy adjusted by self-rated health status estimated from the model

^f Differences in life expectancies calculated as life expectancy adjusted by health status minus life expectancy in the average US population (2000 US life table)

to the scarce literature on individuals younger than 66 years old.

This study builds upon a growing body of research addressing health status and life expectancy for use in clinical decision-making [6–9, 24, 25, 40]. Some studies incorporated utility scores, a composite index, for example, the Health Utility Index in Canada [38, 39], for the health state classifications or health-related quality of life in life expectancy evaluations. Other studies estimated healthy life expectancy [41–43], defined as the number of years that a person at a given age (often calculated at birth) can expect to live in good health (or full health), and utilized as a summary metric for mortality. In contrast, our life expectancy estimates adjusted for self-reported health status provide the average remaining life expectancy at specific health status reported. The measurement is based on a single question regarding health status, which is easily asked and answered, and essential demographic information. Meanwhile, prior works that have created life tables taking into account comorbidity status have been estimated from Medicare claims linked to the Surveillance Epidemiology and End Results (SEER) data for individuals aged 66 years and older [6, 7, 9]. However, these life tables are limited by the fact that they are unavailable for ages below 66, based on comorbidities measured from claims data, which only identify treated comorbidities, and they do not account for other factors that may affect health and mortality. Similarly, even prognostic indices and prediction models utilizing health status are often tailored to the elderly population and require additional information (i.e., existing illnesses) other than general health status to formulate a prognostic index to predict mortality. This study provides an extensive set of life tables by race, sex, and self-reported health status; it included participants aged between 40 and 64. Moreover, self-rated health has been shown in some studies to be a more accurate measure of health status than objective measures, [30, 44] as it reflects an individual's perception of health and likely captures the full array of illness in a particular person.

We used an age-specific discrete time proportional hazards model to minimize assumptions and capture changes in the baseline hazards of each age group. The

models were stratified by 5 years age groups because we did not want to assume proportionally in the baseline hazard age for different age groups, e.g., 40–44 and 60–64. However, the proportionality assumption was only used for ages within an age group. The proportional hazards assumption for other covariates was checked. This model allowed us to account for variation in both baseline hazards and the effect of health status on survival by age group. Therefore, the model enabled us to characterize the impact of health status in each age group and compare life expectancy adjusted for health status with the US life table estimates of the same age. Other models could be considered depending on the assumptions regarding the baseline hazards and the effect of health status across age groups. An overview of our sensitivity analyses using different models is provided in the Supplemental Material. Estimated survival outcomes were similar across various approaches.

There were several limitations to this study. The average survival and life expectancy of the NHIS study population was close to those who reported *good* health. However, as the NHIS population consists only of noninstitutionalized individuals, our average survival and life expectancy estimates were slightly higher than those in the US general population, especially among the elderly population. Although the results of this study should only be applied to noninstitutionalized individuals, as the issues associated with clinical decision-making in the screening and treatment of the institutionalized population may differ from the noninstitutionalized population in other ways, it is possible that the individuals who respond to surveys, such as the NHIS, may be healthier on an average than the general population. In this study, we adjusted for proxy-reporting in the survival models and reported survival and life expectancies for self-reported health status. Life expectancies estimated from proxy information were similar to the self-reported data. For an individual who is using this calculator is their health status at that time. Our life tables were constructed using the health status at the survey and baseline. In the survey, we do not have information on their future health status and the individual. Incorporating possible changes and trajectories of the individual's future health status in the life expectancy

remained as future research. Lastly, This study aimed to evaluate and construct a life table that can be used as a simple tool to aid clinical decision-making; thus, we only included health status and essential demographic variables in the model. This study did not include specific risk factors that may interact with health status and impact life expectancy estimates, such as smoking status and/or functional limitations. As such, this will be an important area for future research. Despite the limitations mentioned above, this study is a large population-based study using a single question about health status, which is well-validated and easily obtained by physicians or other health care providers. Moreover, self-rated health has been used in numerous population-level studies. It is also a conventional way to address a topic of health status—when is it to be covered in physician-patient communications [30]. Thus, these life tables can provide health care providers a simple tool for assessing life expectancy while adjusting for the health status of their patients.

The National Cancer Institute Surveillance Research Program is developing an online system, the SEER Cancer Survival Calculator, to individually predict cancer patients' risk of death from cancer and other causes in the US [45, 46]. However, individuals diagnosed with specific cancers are not necessarily representative of the US population at large; thus, using US life tables to represent their general health profile has limitations [47]. The SEER Cancer Survival Calculator tool uses life tables with comorbidities prior to formulating a cancer diagnosis to estimate risk of death because of other causes [45, 46]. However, these life tables are not available for individuals below the age of 65 years. To this end, the life tables using self-reported health developed in this study will be used to estimate the risk of death because of other causes and to provide more accurate estimates of other causes of mortality than general US life tables based solely on demographic characteristics.

Conclusions

Although US life tables represent the life expectancy of the average person in the general population, they do not fully capture the variation in survival by health status. As shown in this study, health status is strongly associated with life expectancy, especially for younger adults. The life table developed can be helpful by including a patient perspective on their health and be used in conjunction with other predictive models in clinical decision making, particularly for younger adults in poor health, for whom life tables including comorbid conditions are limited.

Abbreviations

NHIS: National Health Interview Survey; NCHS: National Center for Health Statistics; SEER: Surveillance Epidemiology and End Results.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12889-021-12332-0>.

Additional file 1.

Acknowledgments

We thank Dr. Barry Graubard at the National Cancer Institute, USA, for his helpful comments and discussions on the survey data analysis.

Authors' contributions

HC, RY, EF and AM contributed to the conceptualization and design. HC, ZW, BL, TM and AM contributed to the data analysis. ZW and TM conducted data curation. All coauthors contributed to the interpretation of the results. HC, RY and AM drafted the article. All coauthors contributed to the critical revision and editing of the manuscript. All coauthors approved the final version of the manuscript.

Funding

H. Cho received funding from the National Cancer Institute, USA (HHSN261201400609P), and the National Research Foundation of Korea, funded by the Korean Ministry of Science and ICT (Grant No. NRF-2020R1A2C1A01011584). The funding sources had no role in the design of the study and collection, analysis, and interpretation of data, and in writing the manuscript.

Availability of data and materials

The datasets analyzed during the current study are publicly available from the National Center for Health Statistics, USA (<https://www.cdc.gov/nchs/data-linkage/mortality.htm>).

Declarations

Ethics approval and consent to participate

Not applicable. This study is the secondary analysis of publicly available de-identified data, the National Health Interview Survey (NHIS). The use of the NHIS public use files does not require IRB review nor an exempt determination. Informed consent is not applicable.

Consent for publication

Not applicable.

Competing interests

None. There are no conflicts of interest to declare.

Author details

¹Department of Cancer Control and Population Health, Graduate School of Cancer Science and Policy, National Cancer Center, 323 Ilsan-ro, Ilsan-dong-gu, Goyang-si, Gyeonggi-do 10408, Republic of Korea. ²Information Management Services Inc., Calverton, MD, USA. ³Surveillance and Health Equity Science, American Cancer Society, Atlanta, GA, USA. ⁴Division of Cancer Control and Population Sciences, National Cancer Institute, Bethesda, MD, USA.

Received: 27 October 2020 Accepted: 28 November 2021

Published online: 20 January 2022

References

- Hoffman KE, Nguyen PL, Ng AK, D'Amico AV. Prostate cancer screening in men 75 years old or older: an assessment of self-reported health status and life expectancy. *J Urol*. 2010;183(5):1798–802.

2. Lee SJ, Boscardin WJ, Stijacic-Cenzer I, Conell-Price J, O'Brien S, Walter LC. Time lag to benefit after screening for breast and colorectal cancer: meta-analysis of survival data from the United States, Sweden, United Kingdom, and Denmark. *BMJ* : British Medical Journal. 2013;346.
3. Walter LC, Covinsky KE. Cancer screening in elderly patients. *JAMA*. 2001;285(21):2750–6.
4. Walter LC, Lindquist K, Covinsky KE. Relationship between health status and use of screening mammography and Papanicolaou smears among women older than 70 years of age. *Ann Intern Med*. 2004;140(9):681–8.
5. Wei R, Anderson RN, Curtin LR, Arias E. U.S. decennial life tables for 1999–2001: State life tables. *Natl Vital Stat Rep*. 2012;60(9):1–66.
6. Mariotto AB, Wang Z, Klabunde CN, Cho H, Das B, Feuer EJ. Life tables adjusted for comorbidity more accurately estimate noncancer survival for recently diagnosed cancer patients. *J Clin Epidemiol*. 2013;66(12):1376–85.
7. Cho H, Klabunde CN, Yabroff KR, Wang Z, Meekins A, Lansdorp-Vogelaar I, et al. Comorbidity-adjusted life expectancy: a new tool to inform recommendations for optimal screening strategies. *Ann Intern Med*. 2013;159(10):667–76.
8. Lee SJ, Lindquist K, Segal MR, Covinsky KE. Development and validation of a prognostic index for 4-year mortality in older adults. *JAMA*. 2006;295(7):801–8.
9. Tan A, Kuo Y-F, Goodwin JS. Predicting life expectancy for community-dwelling older adults from Medicare claims data. *Am J Epidemiol*. 2013;178(6):974–83.
10. Carey EC, Covinsky KE, Lui L-Y, Eng C, Sands LP, Walter LC. Prediction of mortality in community-living frail elderly people with long-term care needs. *J Am Geriatr Soc*. 2008;56(1):68–75.
11. Woloshin S, Schwartz LM, Welch HG. The risk of death by age, sex, and smoking status in the United States: putting health risks in context. *J Natl Cancer Inst*. 2008;100(12):845–53.
12. Chetty R, Stepner M, Abraham S, et al. The association between income and life expectancy in the United States, 2001–2014. *JAMA*. 2016;315(16):1750–66.
13. Lochner K, Pamuk E, Makuc D, Kennedy BP, Kawachi I. State-level income inequality and individual mortality risk: a prospective, multilevel study. *Am J Public Health*. 2001;91(3):385–91.
14. Sorlie PD, Backlund E, Keller JB. US mortality by economic, demographic, and social characteristics: the National Longitudinal Mortality Study. *Am J Public Health*. 1995;85(7):949–56.
15. McCallum J, Shadbolt B, Wang D. Self-rated health and survival: a 7-year follow-up study of Australian elderly. *Am J Public Health*. 1994;84(7):1100–5.
16. Yu ESH, Kean YM, Slymen DJ, Liu WT, Zhang M, Katzman R. Self-perceived health and 5-year mortality risks among the elderly in Shanghai, China. *Am J Epidemiol*. 1998;147(9):880–90.
17. Helmer C, Barberger-Gateau P, Letenneur L, Dartigues J-F. Subjective health and mortality in French elderly women and men. *J Gerontol Ser B Psychol Sci Soc Sci*. 1999;54B(2):S84–92.
18. McGee DL, Liao Y, Cao G, Cooper RS. Self-reported health status and mortality in a multiethnic US cohort. *Am J Epidemiol*. 1999;149(1):41–6.
19. Idler EL, Russell LB, Davis D. Survival, functional limitations, and self-rated health in the NHANES I epidemiologic follow-up study, 1992. *Am J Epidemiol*. 2000;152(9):874–83.
20. Saliba D, Elliott M, Rubenstein LZ, Solomon DH, Young RT, Kamberg CJ, et al. The vulnerable elders survey: a tool for identifying vulnerable older people in the community. *J Am Geriatr Soc*. 2001;49(12):1691–9.
21. Heidrich J, Liese AD, Löwel H, Keil U. Self-rated health and its relation to all-cause and cardiovascular mortality in southern Germany. Results from the MONICA Augsburg cohort study 1984–1995. *Ann Epidemiol*. 2002;12(5):338–45.
22. Kawada T. Self-rated health and life prognosis. *Arch Med Res*. 2003;34(4):343–7.
23. Benjamins MR, Hummer RA, Eberstein IW, Nam CB. Self-reported health and adult mortality risk: an analysis of cause-specific mortality. *Social Science & Medicine* 2004, 59(6):1297–1306.
24. Schonberg MA, Davis RB, McCarthy EP, Marcantonio ER. Index to predict 5-year mortality of community-dwelling adults aged 65 and older using data from the National Health Interview Survey. *J Gen Intern Med*. 2009;24(10):1115.
25. Schonberg MA, Davis RB, McCarthy EP, Marcantonio ER. External validation of an index to predict up to 9-year mortality of community-dwelling adults aged 65 and older. *J Am Geriatr Soc*. 2011;59(8):1444–51.
26. Deshpande AD, McQueen A, Coups EJ. Different effects of multiple health status indicators on breast and colorectal cancer screening in a nationally representative US sample. *Cancer Epidemiol*. 2012;36(3):270–5.
27. Hoffman RM, Koyama T, Albertsen PC, Barry MJ, Daskivich TJ, Goodman M, et al. Self-reported health status predicts other-cause mortality in men with localized prostate Cancer: results from the prostate Cancer outcomes study. *J Gen Intern Med*. 2015;30(7):924–34.
28. DeSalvo KB, Bloser N, Reynolds K, He J, Muntner P. Mortality prediction with a single general self-rated health question. *J Gen Intern Med*. 2006;21(3):267–75.
29. DeSalvo KB, Fan VS, McDonnell MB, Fihn SD. Predicting mortality and health-care utilization with a single question. *Health Serv Res*. 2005;40(4):1234–46.
30. Idler EL, Benyamini Y. Self-rated health and mortality: a review of twenty-seven community studies. *J Health Soc Behav*. 1997;38(1):21–37.
31. Jylhä M. What is self-rated health and why does it predict mortality? Towards a unified conceptual model. *Soc Sci Med*. 2009;69(3):307–16.
32. National Center for Health Statistics. Office of Analysis and Epidemiology, The National Health Interview Survey (1986–2004) Linked Mortality Files, mortality follow-up through 2006: Matching Methodology, May 2009. Hyattsville, Maryland. (Available at the following address: http://www.cdc.gov/nchs/data/datalinkage/matching_methodology_nhis_final.pdf), Accessed July 2021.
33. US bureau of labor statistics: <https://www.bls.gov/bls/glossary.htm>, accessed July 2021.
34. Allison PD. Discrete-time methods for the analysis of event histories. *Sociol Methodol*. 1982;13:61–98.
35. Prentice RL, Gloeckler LA. Regression analysis of grouped survival data with application to breast Cancer data. *Biometrics*. 1978;34(1):57–67.
36. Singer JD, Willett JB. It's about time: using discrete-time survival analysis to study duration and the timing of events. *J Educ Behav Stat*. 1993;18(2):155–95.
37. SUDAAN Language manual, release 11.0. In. Research Triangle Park, NC, USA: Research Triangle Institute; 2012.
38. U.S. Annual Life Tables, <https://seer.cancer.gov/expsurvival/>, accessed July 2021.
39. Arias E, Curtin LR, Wei R, Anderson RN: U.S. decennial life tables for 1999–2001, United States life tables. *Natl Vital Stat Rep* 2008, 57(1):1–36.
40. Pollack CE, Blackford AL, Schoenborn NL, Boyd CM, Pears KS, DuGoff EH. Comparing prognostic tools for Cancer screening: considerations for clinical practice and performance assessment. *J Am Geriatr Soc*. 2016;64(5):1032–8.
41. World Health Organization: THE GLOBAL HEALTH OBSERVATORY, Healthy life expectancy (HALE) at birth, <https://www.who.int/data/gho/indicator-metadata-registry/imr-details/66>, accessed July 2021.
42. The United Nations: Life Expectancy at Birth, https://www.un.org/esa/sustdev/natlinfo/indicators/methodology_sheets/health/health_life_expectancy.pdf, accessed July 2021.
43. GBD 2019 Demographics Collaborators: Global age-sex-specific fertility, mortality, healthy life expectancy (HALE), and population estimates in 204 countries and territories, 1950–2019: a comprehensive demographic analysis for the Global burden of disease study 2019. *Lancet* 2020, 396(10258):1160–1203.
44. Idler E, Leventhal H, McLaughlin J, Leventhal E. In sickness but not in health: self-ratings, identity, and mortality. *J Health Soc Behav*. 2004;45(3):336–56.
45. Feuer EJ, Lee M, Mariotto AB, Cronin KA, Scoppa S, Penson DF, et al. The cancer survival query system: making survival estimates from the surveillance, epidemiology, and end results program more timely and relevant for recently diagnosed patients. *Cancer*. 2012;118(22):5652–62.
46. Feuer EJ, Rabin BA, Zou Z, Wang Z, Xiong X, Ellis JL, et al. The surveillance, epidemiology, and end results Cancer survival calculator SEER*CS: validation in a managed care setting. *JNCI Monographs*. 2014;2014(49):265–74.
47. Cho H, Mariotto AB, Bhupinder MS, Klabunde CN, Feuer EJ: Assessing Non-Cancer Health Status of US Cancer Patients: Other-Cause Survival and Comorbidity. *American Journal of Epidemiology*, doi: 101093/aje/kws580 2013.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.