# Provisional Life Expectancy Estimates for 2021

**Elizabeth Arias, Ph.D., Betzaida Tejada-Vera, M.S., Kenneth D. Kochanek, M.A., and Farida B. Ahmad, M.P.H.**

## Introduction

The National Center for Health Statistics (NCHS) collects and disseminates the nation's official vital statistics through the National Vital Statistics System. NCHS uses provisional vital statistics data for conducting public health surveillance and final data for producing annual national natality and mortality statistics. NCHS publishes annual and decennial national life tables based on final vital statistics data. To assess the effects of excess mortality related to the COVID-19 pandemic on life expectancy, NCHS published the first ever provisional life expectancy estimates for the year 2020 [1,2]. Life expectancy estimates presented in this report are based on provisional mortality data for 2021 and final data for 2019 and 2020. Provisional data are early estimates based on death certificates received, processed, and coded but not finalized by NCHS. These estimates are considered provisional because death certificate information may be revised, and additional death certificates may be received until approximately 6 months after the end of the year.

This report presents life expectancy estimates calculated using complete period life tables based on provisional death counts for 2021 by sex and for the total, Hispanic, non-Hispanic American Indian or Alaska Native (AIAN), non-Hispanic Asian, non-Hispanic Black, and non-Hispanic White populations. Estimates for the Native Hawaiian or Other Pacific Islander population were not produced because data needed to evaluate race and ethnicity misclassification on death certificates for this population are not currently available [3]. There are two types of life tables: the cohort (or generation) and the period (or current) life table. The cohort life table presents the mortality experience of a particular birth cohort from the moment of birth through consecutive ages in successive calendar years. The period life table does not represent the mortality experience of an actual birth cohort but rather presents what would happen to a hypothetical cohort if it experienced throughout its entire life the mortality conditions of a particular period. This report also presents contributions of causes of death to the changes in life expectancy using a life table partitioning technique (Technical Notes).

**Keywords**: Hispanic origin • race • cause of death • National Vital Statistics System

## Data and Methods

Provisional life expectancy estimates were calculated using complete period life tables based on provisional death counts for 2021 from death records received and processed by NCHS as of April 24, 2022; provisional numbers of births for the same period based on birth records received and processed by NCHS as of May 3, 2022; and July 1, 2021, postcensal population estimates based on the 2010 decennial census. Provisional death rates are typically computed using death data after a 3-month lag, as completeness and timeliness of provisional death data can vary by many factors, including cause of death, month of the year, and age of the decedent [4,5]. Mortality data used in this report include over 99% of the deaths that occurred in 2021, but certain jurisdictions and age groups may be underrepresented for later months [5]. Deaths requiring investigation, including infant deaths and those from external injuries and drug overdose may be underestimated [6]. See Technical Notes for more information about the calculation of the complete period life tables and life table partitioning by cause of death. Provisional 2021 life expectancy estimates are compared with final estimates for years 2019 and 2020 to describe changes in life expectancy in the United States since the start of the COVID-19 pandemic.

## Results

### Life expectancy in the United States

The Table summarizes life expectancy by age, race and Hispanic origin, and sex. Life expectancy at birth represents the average number of years a group of infants would live if they were to experience throughout life the age-specific death rates prevailing during a period. In 2021, life expectancy at birth was 76.1 years, declining by 0.9 year from 77.0 in 2020 [3]. Life expectancy at birth for males in 2021 was 73.2 years, representing a decline of 1.0 year from 74.2 years in 2020. For females, life expectancy declined to 79.1 years, decreasing 0.8 year from 79.9 years in 2020 (Figure 1). Excess deaths due to COVID-19 and other causes in 2020 and 2021 led to an overall decline in life expectancy between 2019 and 2021 of 2.7 years for the total population, 3.1 years for males, and 2.3 years for females (Figure 1) [7].

Table. Provisional life expectancy, by age, race and Hispanic origin, and sex: United States, 2021

| Age (years) | All races and origins | | | Hispanic | | | Non-Hispanic American Indian or Alaska Native | | | Non-Hispanic Asian | | | Non-Hispanic Black | | | Non-Hispanic White | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Total | Male | Female | Total | Male | Female | Total | Male | Female | Total | Male | Female | Total | Male | Female | Total | Male | Female |
| 0. . . . . . . . . . . . . . . | 76.1 | 73.2 | 79.1 | 77.7 | 74.4 | 81.0 | 65.2 | 61.5 | 69.2 | 83.5 | 81.2 | 85.6 | 70.8 | 66.7 | 74.8 | 76.4 | 73.7 | 79.2 |
| 1. . . . . . . . . . . . . . . | 75.6 | 72.6 | 78.5 | 77.1 | 73.8 | 80.4 | 64.7 | 61.0 | 68.7 | 82.7 | 80.4 | 84.8 | 70.6 | 66.5 | 74.5 | 75.7 | 73.0 | 78.5 |
| 5. . . . . . . . . . . . . . . | 71.6 | 68.7 | 74.6 | 73.1 | 69.8 | 76.4 | 60.9 | 57.1 | 64.8 | 78.8 | 76.5 | 80.8 | 66.7 | 62.6 | 70.7 | 71.8 | 69.1 | 74.6 |
| 10. . . . . . . . . . . . . . | 66.7 | 63.8 | 69.7 | 68.2 | 64.9 | 71.5 | 55.9 | 52.2 | 59.9 | 73.8 | 71.5 | 75.9 | 61.8 | 57.7 | 65.7 | 66.8 | 64.1 | 69.6 |
| 15. . . . . . . . . . . . . . | 61.7 | 58.8 | 64.7 | 63.2 | 59.9 | 66.5 | 51.0 | 47.3 | 55.0 | 68.8 | 66.6 | 70.9 | 56.9 | 52.8 | 60.8 | 61.9 | 59.2 | 64.7 |
| 20. . . . . . . . . . . . . . | 56.9 | 54.1 | 59.8 | 58.4 | 55.1 | 61.6 | 46.4 | 42.7 | 50.3 | 63.9 | 61.7 | 65.9 | 52.2 | 48.3 | 56.0 | 57.0 | 54.4 | 59.8 |
| 25. . . . . . . . . . . . . . | 52.2 | 49.5 | 55.0 | 53.7 | 50.6 | 56.8 | 42.1 | 38.6 | 45.8 | 59.1 | 56.9 | 61.0 | 47.8 | 44.2 | 51.3 | 52.3 | 49.8 | 54.9 |
| 30. . . . . . . . . . . . . . | 47.6 | 45.1 | 50.2 | 49.1 | 46.1 | 52.0 | 38.0 | 34.7 | 41.5 | 54.3 | 52.1 | 56.1 | 43.5 | 40.0 | 46.7 | 47.7 | 45.3 | 50.2 |
| 35. . . . . . . . . . . . . . | 43.1 | 40.7 | 45.5 | 44.5 | 41.7 | 47.2 | 34.3 | 31.2 | 37.4 | 49.4 | 47.3 | 51.2 | 39.1 | 35.9 | 42.1 | 43.1 | 40.9 | 45.5 |
| 40. . . . . . . . . . . . . . | 38.6 | 36.4 | 40.9 | 39.9 | 37.3 | 42.5 | 30.8 | 28.0 | 33.8 | 44.6 | 42.5 | 46.3 | 35.0 | 32.0 | 37.7 | 38.7 | 36.5 | 40.8 |
| 45. . . . . . . . . . . . . . | 34.2 | 32.1 | 36.4 | 35.5 | 33.0 | 37.8 | 27.4 | 24.8 | 30.0 | 39.9 | 37.9 | 41.5 | 30.9 | 28.1 | 33.4 | 34.3 | 32.3 | 36.3 |
| 50. . . . . . . . . . . . . . | 30.0 | 28.0 | 31.9 | 31.1 | 28.8 | 33.3 | 24.4 | 22.1 | 26.7 | 35.2 | 33.3 | 36.7 | 26.9 | 24.4 | 29.2 | 30.0 | 28.1 | 31.9 |
| 55. . . . . . . . . . . . . . | 25.9 | 24.0 | 27.6 | 26.9 | 24.8 | 28.8 | 21.5 | 19.5 | 23.5 | 30.6 | 28.9 | 32.0 | 23.2 | 20.9 | 25.2 | 25.9 | 24.1 | 27.6 |
| 60. . . . . . . . . . . . . . | 22.0 | 20.4 | 23.5 | 23.0 | 21.1 | 24.6 | 18.9 | 17.2 | 20.4 | 26.1 | 24.6 | 27.4 | 19.7 | 17.6 | 21.5 | 21.9 | 20.4 | 23.4 |
| 65. . . . . . . . . . . . . . | 18.3 | 16.9 | 19.6 | 19.3 | 17.6 | 20.6 | 16.3 | 15.1 | 17.4 | 21.9 | 20.5 | 22.9 | 16.5 | 14.8 | 18.0 | 18.3 | 16.9 | 19.5 |
| 70. . . . . . . . . . . . . . | 14.8 | 13.7 | 15.8 | 15.7 | 14.4 | 16.7 | 13.7 | 12.7 | 14.5 | 17.8 | 16.7 | 18.6 | 13.6 | 12.2 | 14.7 | 14.7 | 13.6 | 15.7 |
| 75. . . . . . . . . . . . . . | 11.5 | 10.6 | 12.3 | 12.4 | 11.3 | 13.1 | 11.2 | 10.5 | 11.8 | 14.0 | 13.1 | 14.5 | 10.9 | 9.7 | 11.7 | 11.4 | 10.5 | 12.1 |
| 80. . . . . . . . . . . . . . | 8.6 | 7.9 | 9.1 | 9.3 | 8.5 | 9.7 | 9.1 | 8.6 | 9.3 | 10.4 | 9.8 | 10.7 | 8.4 | 7.5 | 8.9 | 8.4 | 7.8 | 8.9 |
| 85. . . . . . . . . . . . . . | 6.1 | 5.6 | 6.4 | 6.7 | 6.1 | 6.9 | 7.2 | 6.9 | 7.2 | 7.3 | 6.9 | 7.4 | 6.2 | 5.6 | 6.5 | 5.9 | 5.5 | 6.2 |
| 90. . . . . . . . . . . . . . | 4.1 | 3.9 | 4.3 | 4.6 | 4.3 | 4.6 | 5.6 | 5.5 | 5.4 | 4.8 | 4.7 | 4.8 | 4.5 | 4.1 | 4.6 | 4.0 | 3.7 | 4.1 |
| 95. . . . . . . . . . . . . . | 2.8 | 2.7 | 2.9 | 3.2 | 3.0 | 3.1 | 4.4 | 4.4 | 4.1 | 3.1 | 3.1 | 3.0 | 3.2 | 3.0 | 3.3 | 2.7 | 2.6 | 2.7 |
| 100. . . . . . . . . . . . . | 2.0 | 2.0 | 2.0 | 2.3 | 2.2 | 2.1 | 3.5 | 3.6 | 3.3 | 2.1 | 2.2 | 2.0 | 2.4 | 2.3 | 2.3 | 1.9 | 1.8 | 1.9 |

NOTES: Life tables by race and Hispanic origin have been adjusted for race and ethnicity misclassification on death certificates; see Technical Notes in this report. Estimates are based on provisional data for 2021. Provisional data are subject to change as additional data are received.
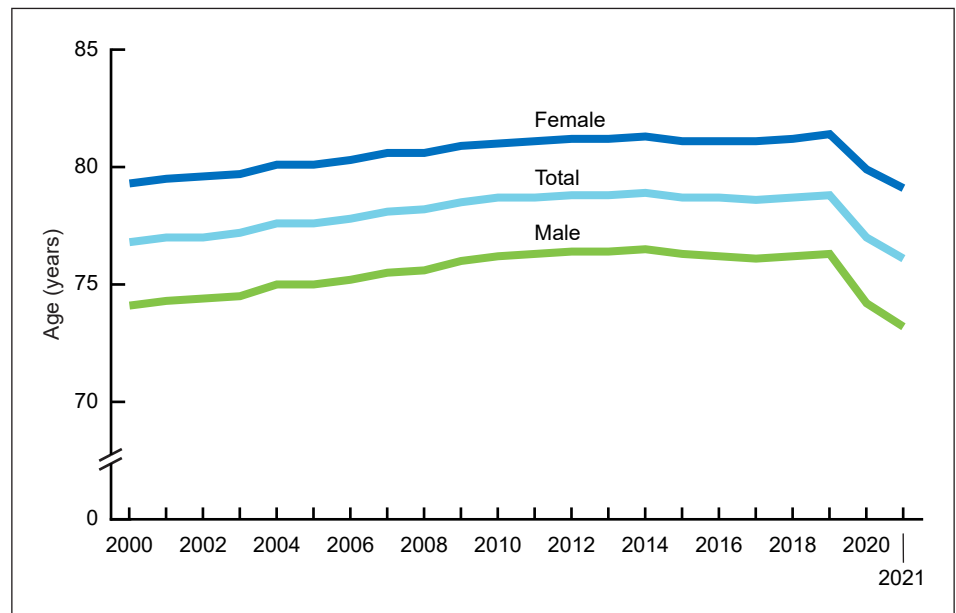
SOURCE: National Center for Health Statistics, National Vital Statistics System, Mortality.

U.S. Department of Health and Human Services • Centers for Disease Control and Prevention • National Center for Health Statistics • National Vital Statistics System

2

The difference in life expectancy between the sexes was 5.9 years in 2021, increasing from 5.7 in 2020. Between 2000 and 2010, the difference in life expectancy between the sexes narrowed from 5.2 years to its lowest level of 4.8 years, but then increased in 2020 and 2021 to levels not seen since 1996 (when the difference was 6 years) (Figure 1).

## Life expectancy by Hispanic origin and race

Between 2020 and 2021, life expectancy decreased by 1.9 years for the non-Hispanic AIAN population (67.1 to 65.2) (Figure 2). It decreased by 1.0 year for the non-Hispanic White population (77.4 to 76.4), by 0.7 year for the non-Hispanic Black population (71.5 to 70.8), by 0.2 year for the Hispanic population (77.9 to 77.7), and by 0.1 year for the non-Hispanic Asian population (83.6 to 83.5). Increases in excess deaths led to a decline in life expectancy between 2019 and 2021 of 6.6 years for the non-Hispanic AIAN population, 4.2 years for the Hispanic population, 4.0 years for the non-Hispanic Black population, 2.4 years for the non-Hispanic White population, and 2.1 years for the non-Hispanic Asian population.

Figure 1. Life expectancy at birth, by sex: United States, 2000–2021
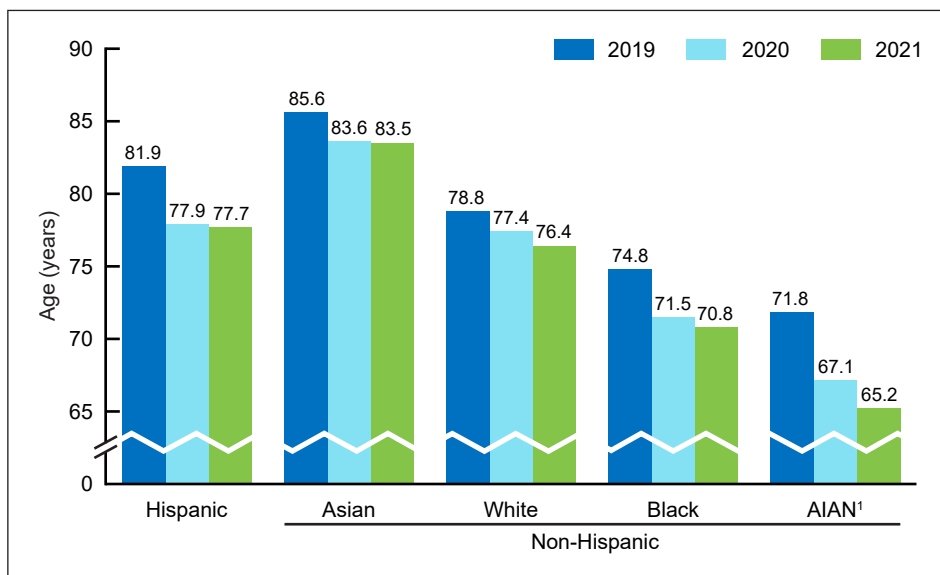


NOTES: Estimates are based on provisional data for 2021. Provisional data are subject to change as additional data are received. Estimates for 2000–2020 are based on final data.
SOURCE: National Center for Health Statistics, National Vital Statistics System, Mortality.

Among the 10 Hispanic-origin and race-sex groups (Figure 3), the decrease in life expectancy between 2020 and 2021 was greatest for non-Hispanic AIAN males, whose life expectancy declined by 2.3 years (63.8 to 61.5), followed by non-Hispanic AIAN females with a decline of 1.5 years (70.7 to 69.2), non-Hispanic Black and non-Hispanic White males with a decline of 1.1 years each (67.8 to 66.7) and (74.8 to 73.7), respectively, non-Hispanic White females with a decline of 0.9 year (80.1 to 79.2), non-Hispanic Black females with a decline of 0.6 year (75.4 to 74.8), Hispanic and non-Hispanic Asian females with a decline of 0.3 year each (81.3 to 81.0) and (85.9 to 85.6), respectively, and Hispanic males with a decline of 0.2 year (74.6 to 74.4). Non-Hispanic Asian males experienced an increase in life expectancy of 0.1 year (81.1 to 81.2) between 2020 and 2021.

Overall, increases in excess deaths during 2020 and 2021 led to decreases in life expectancy at birth of 7.1 years for non-Hispanic AIAN males, 5.8 years for non-Hispanic AIAN females, 4.7 years for Hispanic males, 4.6 years for non-Hispanic Black males, 3.4 years for Hispanic females, 3.3 years for non-Hispanic Black females, 2.6 years for non-Hispanic White males, 2.3 years for non-Hispanic Asian males, 2.1 years for non-Hispanic White females, and 1.8 years for non-Hispanic Asian females.

Figure 2. Life expectancy at birth, by Hispanic origin and race: United States, 2019–2021



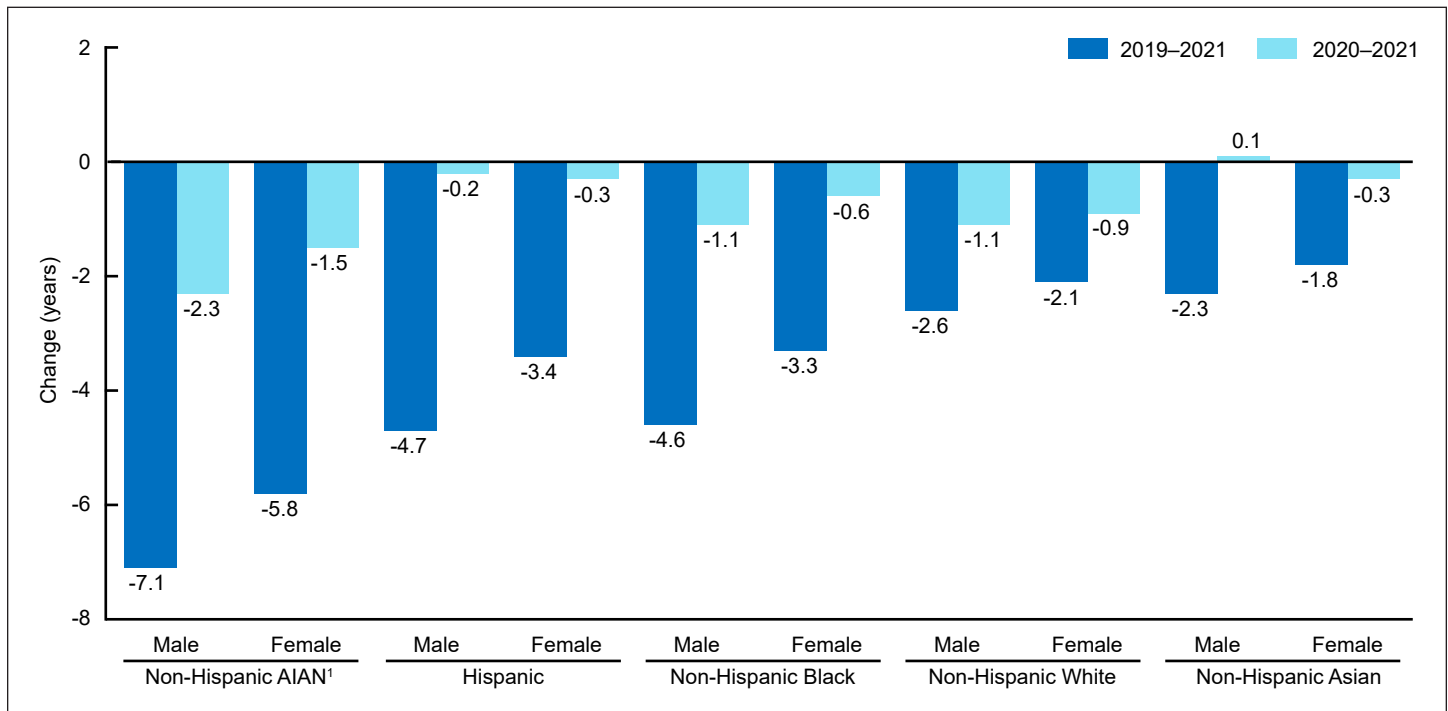[1]American Indian or Alaska Native.
NOTES: Estimates are based on provisional data for 2021. Provisional data are subject to change as additional data are received. Estimates for 2019 and 2020 are based on final data. Life tables by race and Hispanic origin are based on death rates that have been adjusted for race and Hispanic-origin misclassification on death certificates; see Technical Notes in this report.
SOURCE: National Center for Health Statistics, National Vital Statistics System, Mortality.

U.S. Department of Health and Human Services • Centers for Disease Control and Prevention • National Center for Health Statistics • National Vital Statistics System

3

Figure 3. Change in life expectancy at birth, by Hispanic origin and race: United States, 2019–2021 and 2020–2021



[1] American Indian or Alaska Native.
NOTES: Estimates are based on provisional data for 2021. Provisional data are subject to change as additional data are received. Estimates for 2019 and 2020 are based on final data. Life tables by race and Hispanic origin are based on death rates that have been adjusted for race and Hispanic-origin misclassification on death certificates; see Technical Notes in this report.
SOURCE: National Center for Health Statistics, National Vital Statistics System, Mortality.

## Effect on life expectancy of changes in cause-specific mortality

Increases or decreases in life expectancy represent the sum of positive and negative contributions of cause-specific death rates. Declines in cause-specific mortality contribute to increases in life expectancy, while increases contribute to decreases in life expectancy. If the negative contributions (increases in cause-specific death rates) are greater than the positive contributions (decreases in cause-specific deaths rates), then the result is a decline in life expectancy. If negative and positive contributions offset each other, then there would be no change in life expectancy (see Technical Notes for a description of the partitioning method).

The decline of 0.9 year in life expectancy between 2020 and 2021 was primarily due to increases in mortality due to COVID-19 (50.0% of the negative contribution), unintentional injuries (15.9%), heart disease (4.1%), chronic

liver disease and cirrhosis (3.0%), and suicide (2.1%) (Figure 4). The decline in life expectancy would have been even greater were it not for the offsetting effects of decreases in mortality due to influenza and pneumonia (38.5%), chronic lower respiratory diseases (28.8%), Alzheimer disease (18.3%), perinatal conditions (6.3%), and Parkinson disease (2.3%).

For the male population, the 1.0-year decline in life expectancy was mostly due to increases in mortality due to COVID-19 (49.5%), unintentional injuries (19.1%), suicide (3.6%), chronic liver disease and cirrhosis (3.4%), and homicide (2.5%). The decline in life expectancy was offset by decreases in mortality due to influenza and pneumonia (29.5%), chronic lower respiratory diseases (26.2%), cancer (12.0%), Alzheimer disease (11.4%), and perinatal conditions (8.3%).

For females, the decline in life expectancy of 0.8 year was primarily due to increases in mortality due to COVID-19 (51.2%), unintentional injuries (14.8%), heart disease (5.7%), stroke (3.5%), and
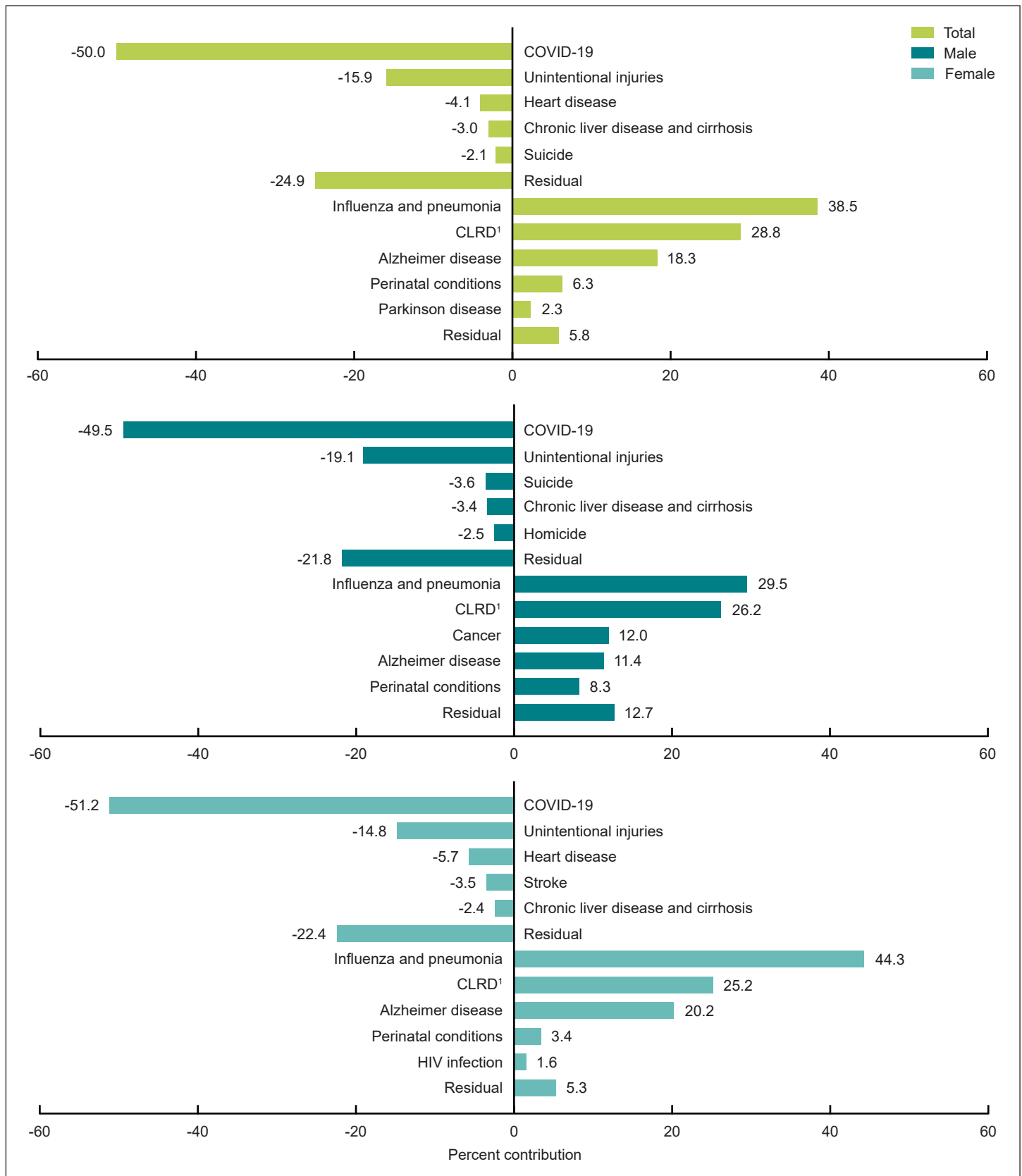
chronic liver disease and cirrhosis (2.4%). The decline in life expectancy was offset by decreases in mortality due to influenza and pneumonia (44.3%), chronic lower respiratory diseases (25.2%), Alzheimer disease (20.2%), perinatal conditions (3.4%), and HIV infection (1.6%).

The non-Hispanic AIAN population experienced the greatest decline in life expectancy (1.9 years) between 2020 and 2021. The decline was due primarily to increases in mortality due to COVID-19 (21.4%), unintentional injuries (21.3%), chronic liver disease and cirrhosis (18.6), suicide (5.4%), and heart disease (3.4%). The decline in life expectancy would have been greater if not for the offsetting declines in mortality due to homicide (23.0%), influenza and pneumonia (21.2%), congenital malformations (12.4%), perinatal conditions (9.4%), and benign neoplasms (5.6%) (Figure 5).

The second greatest decline in life expectancy between 2020 and 2021 was in the non-Hispanic White population (1.0 year). The decline was primarily due to increases in mortality due to

U.S. Department of Health and Human Services • Centers for Disease Control and Prevention • National Center for Health Statistics • National Vital Statistics System

4

Figure 4. Contribution of leading causes of death to the change in life expectancy, by sex and total population: United States, 2020–2021



Legend:
- Total
- Male
- Female

**Total**

| Cause | Percent contribution |
|---|---|
| COVID-19 | -50.0 |
| Unintentional injuries | -15.9 |
| Heart disease | -4.1 |
| Chronic liver disease and cirrhosis | -3.0 |
| Suicide | -2.1 |
| Residual | -24.9 |
| Influenza and pneumonia | 38.5 |
| CLRD[1] | 28.8 |
| Alzheimer disease | 18.3 |
| Perinatal conditions | 6.3 |
| Parkinson disease | 2.3 |
| Residual | 5.8 |

**Male**

| Cause | Percent contribution |
|---|---|
| COVID-19 | -49.5 |
| Unintentional injuries | -19.1 |
| Suicide | -3.6 |
| Chronic liver disease and cirrhosis | -3.4 |
| Homicide | -2.5 |
| Residual | -21.8 |
| Influenza and pneumonia | 29.5 |
| CLRD[1] | 26.2 |
| Cancer | 12.0 |
| Alzheimer disease | 11.4 |
| Perinatal conditions | 8.3 |
| Residual | 12.7 |

**Female**

| Cause | Percent contribution |
|---|---|
| COVID-19 | -51.2 |
| Unintentional injuries | -14.8 |
| Heart disease | -5.7 |
| Stroke | -3.5 |
| Chronic liver disease and cirrhosis | -2.4 |
| Residual | -22.4 |
| Influenza and pneumonia | 44.3 |
| CLRD[1] | 25.2 |
| Alzheimer disease | 20.2 |
| Perinatal conditions | 3.4 |
| HIV infection | 1.6 |
| Residual | 5.3 |

Percent contribution

[1]Chronic lower respiratory diseases.
NOTES: Estimates are based on provisional data for 2021. Provisional data are subject to change as additional data are received. Estimates for 2020 are based on final data. Life tables by race and Hispanic origin are based on death rates that have been adjusted for race and Hispanic-origin misclassification on death certificates; see Technical Notes in this report.
SOURCE: National Center for Health Statistics, National Vital Statistics System, Mortality.

U.S. Department of Health and Human Services • Centers for Disease Control and Prevention • National Center for Health Statistics • National Vital Statistics System

5

Figure 5. Contribution of leading causes of death to change in life expectancy, by Hispanic origin and race: Non-Hispanic American Indian or Alaska Native and non-Hispanic White populations, 2020–2021
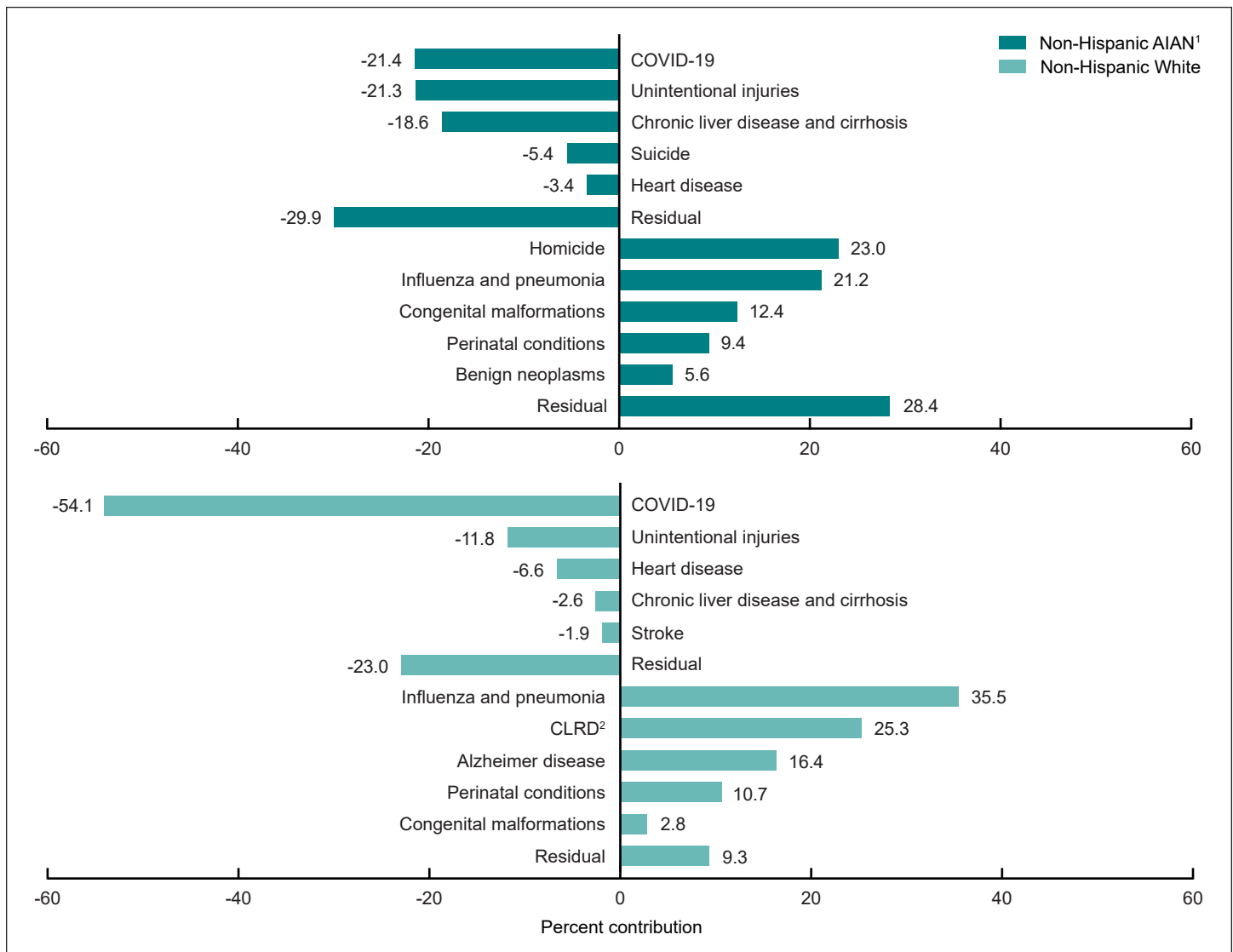


[1]American Indian or Alaska Native.
[2]Chronic lower respiratory diseases.
NOTES: Estimates are based on provisional data for 2021. Provisional data are subject to change as additional data are received. Estimates for 2020 are based on final data. Life tables by race and Hispanic origin are based on death rates that have been adjusted for race and Hispanic-origin misclassification on death certificates; see Technical Notes in this report.
SOURCE: National Center for Health Statistics, National Vital Statistics System, Mortality.

COVID-19 (54.1%), unintentional injuries (11.8%), heart disease (6.6%), chronic liver disease and cirrhosis (2.6%), and stroke (1.9%). The negative effects of these causes were offset by decreases in mortality due to influenza and pneumonia (35.5%), chronic lower respiratory diseases (25.3%), Alzheimer disease (16.4%), perinatal conditions (10.7%), and congenital malformations (2.8%) (Figure 5).

The non-Hispanic Black population had the third greatest decline in life expectancy (0.7 year). The decline was due primarily to increases in mortality due to COVID-19 (35.0%), unintentional injuries (22.7%), stroke (4.8%), suicide (2.7%), and chronic liver disease and cirrhosis (2.5%). The decrease in life expectancy was offset by decreases in mortality due to heart disease (20.8%), influenza and pneumonia (20.3%), perinatal conditions (14.1%), chronic lower respiratory diseases (13.6%), and cancer (10.2%) (Figure 6).

The Hispanic population had the fourth largest decline in life expectancy between 2020 and 2021 (0.2 year). This decrease was primarily due to increases in mortality due to unintentional injuries (31.2%), COVID-19 (25.5%), chronic liver disease and cirrhosis (5.1%), homicide (4.8%), and suicide (2.8%). The decline in life expectancy would have been greater were it not for the offsetting effects of decreases in mortality due to heart disease (36.8%), influenza and pneumonia (22.8%), diabetes (12.4%), Alzheimer disease (12.2%), and chronic lower respiratory diseases (9.8%) (Figure 6).

The non-Hispanic Asian population experienced the smallest decline in life

U.S. Department of Health and Human Services • Centers for Disease Control and Prevention • National Center for Health Statistics • National Vital Statistics System

6

Figure 6. Contribution of leading causes of death to change in life expectancy, by Hispanic origin and race: Non-Hispanic Black and Hispanic populations, 2020–2021
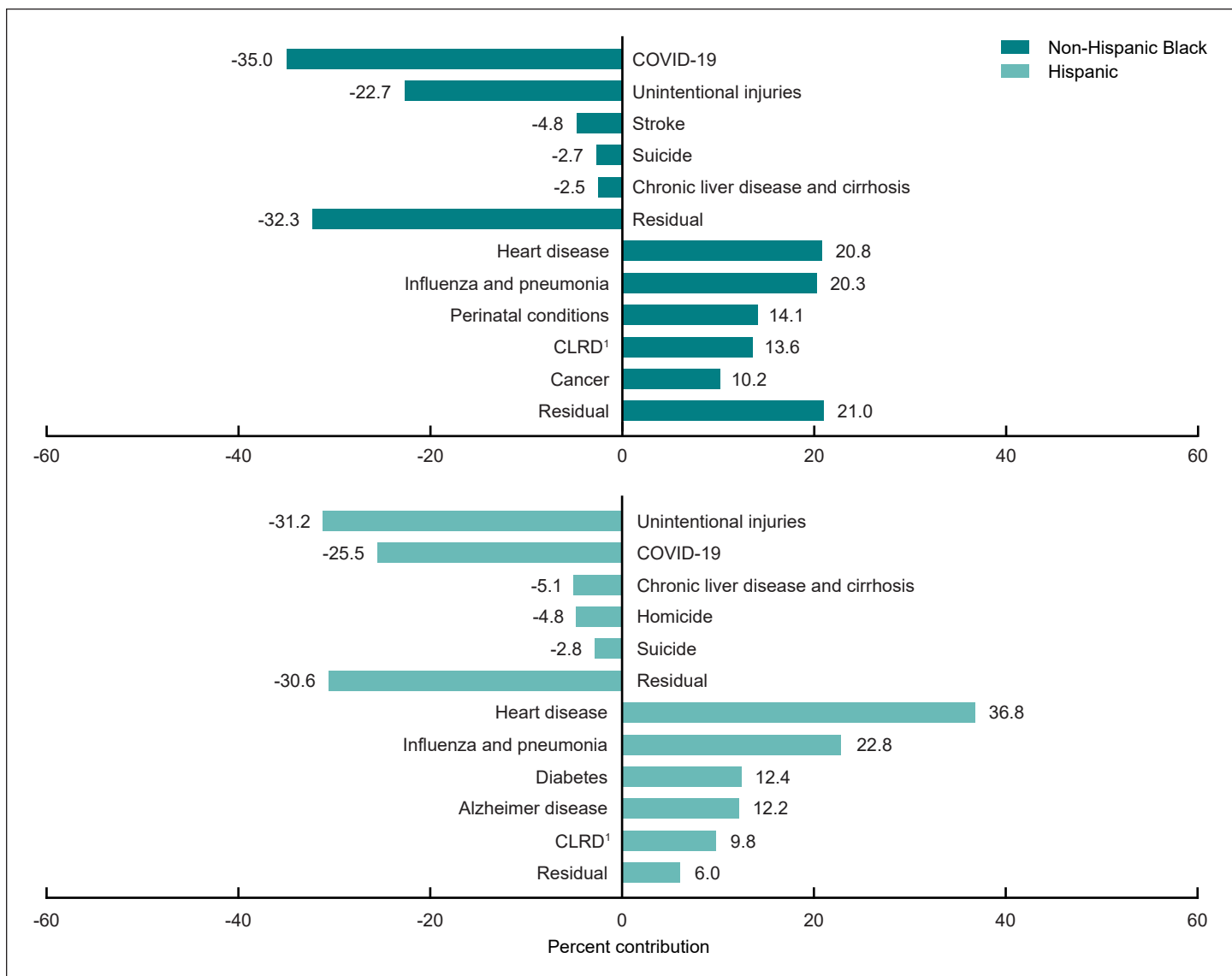


[1]Chronic lower respiratory diseases.
NOTES: Estimates are based on provisional data for 2021. Provisional data are subject to change as additional data are received. Estimates for 2020 are based on final data. Life tables by race and Hispanic origin are based on death rates that have been adjusted for race and Hispanic-origin misclassification on death certificates; see Technical Notes in this report.
SOURCE: National Center for Health Statistics, National Vital Statistics System, Mortality.

expectancy (0.1 year), primarily due to increases in mortality due to cancer (21.4%), COVID-19 (16.6%), unintentional injuries (15.0%), stroke (7.9%), and suicide (6.6%). The decline in life expectancy was offset by decreases in mortality due to influenza and pneumonia (35.7%), congenital malformations (19.4%), chronic lower respiratory diseases (16.6%), diabetes (10.6%), and heart disease (9.3%) (Figure 7).

# Discussion and Conclusions

U.S. life expectancy at birth for 2021, based on nearly final data, was 76.1 years, the lowest it has been since 1996. Male life expectancy (73.2) and female life expectancy (79.1) also declined to levels not seen since 1996. The non-Hispanic AIAN population experienced the largest decline in life expectancy, from 67.1 in 2020 to 65.2 years in 2021, the same life expectancy of the total U.S. population in 1944 (8). The non-Hispanic White

population had the second greatest decline in life expectancy (77.4 to 76.4) and was the lowest seen since 1995 for the White population (regardless of Hispanic origin). Life expectancy for the non-Hispanic Black population declined from 71.5 to 70.8 years, a level last seen in 1996 for the Black population (regardless of Hispanic origin). Life expectancy for the Hispanic population declined to 77.7 years, a level lower than in 2006 (80.3), the first year for which life expectancy estimates by Hispanic origin were produced (9). The

U.S. Department of Health and Human Services • Centers for Disease Control and Prevention • National Center for Health Statistics • National Vital Statistics System

7

Figure 7. Contribution of leading causes of death to change in life expectancy, by Hispanic origin and race: Non-Hispanic Asian population, 2020–2021
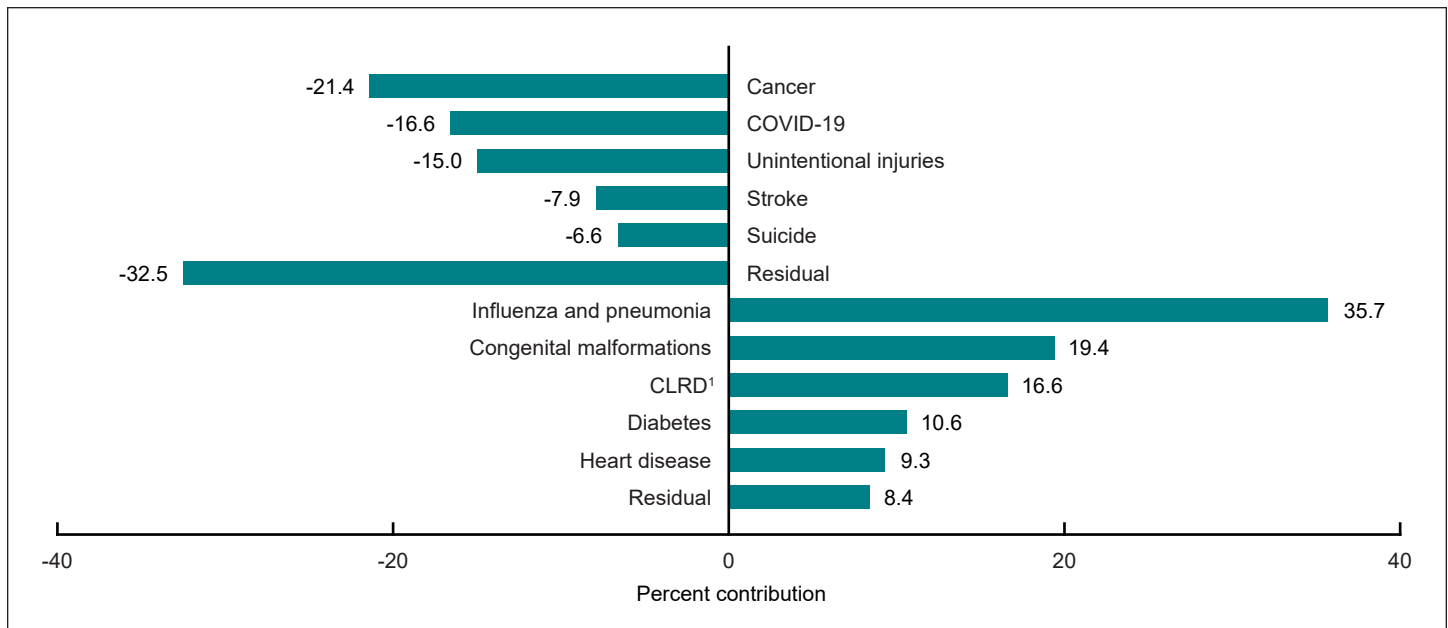


[1] Chronic lower respiratory diseases.
NOTES: Estimates are based on provisional data for 2021. Provisional data are subject to change as additional data are received. Estimates for 2020 are based on final data. Life tables by race and Hispanic origin are based on death rates that have been adjusted for race and Hispanic-origin misclassification on death certificates; see Technical Notes in this report.
SOURCE: National Center for Health Statistics, National Vital Statistics System, Mortality.

non-Hispanic Asian population had the smallest decline in life expectancy (83.6 to 83.5) and maintained its status as the population with the highest life expectancy in the United States.

Between 2020 and 2021, racial and ethnic disparities in life expectancy increased in some cases and declined in others. For example, the non-Hispanic White advantage over the non-Hispanic AIAN population increased by 8.7% between 2020 (10.3 years) and 2021 (11.2). The non-Hispanic White advantage over the non-Hispanic Black population declined by 5.1% between 2020 (5.9) and 2021 (5.6). Life expectancy for the Black population has consistently been lower than that of the White population, but the gap had been narrowing during the past three decades, from 7.1 years in 1993 to 4.0 in 2019 (10).

The Hispanic life expectancy advantage over the non-Hispanic White population increased by 1.6% between 2020 (0.5 year) and 2021 (1.3). Between 2019 and 2020, the Hispanic population lost most of the life expectancy advantage (3.1 to 0.5) it had experienced relative to the non-Hispanic White population since first recorded in 2006 (8,10). The non-Hispanic

Asian life expectancy advantage over the non-Hispanic White population increased by 14.5% between 2020 (6.2) and 2021 (7.1).

COVID-19 was the leading cause contributing negatively to the change in life expectancy for the total population and for three of the five Hispanic-origin and race groups shown in this report. Mortality due to COVID-19 contributed 54.1%, 35.0%, and 21.4% to the decline in life expectancy for the non-Hispanic White, non-Hispanic Black, and non-Hispanic AIAN populations, respectively. For the Hispanic and non-Hispanic Asian populations, COVID-19 was the second leading cause contributing to the decline in life expectancy by 25.5% and 16.6%, respectively. Unintentional injuries was the leading cause contributing to the decline in life expectancy for the Hispanic population (31.2%) and the second leading cause for the non-Hispanic Black, non-Hispanic AIAN, and non-Hispanic White populations, contributing to the decline by 22.7%, 21.3%, and 11.8%, respectively. It had the third largest effect on the decline in life expectancy for the non-Hispanic Asian population (15.0%). Increases in unintentional injury deaths

in 2021 were largely driven by drug overdose deaths.

Between 2019 and 2021, life expectancy in the United States declined 2.7 years, with most of the decline (66.7%) occurring the first year of the COVID-19 pandemic. During the 2-year period, large disparities were seen in loss of life expectancy by Hispanic origin and race. The non-Hispanic AIAN population lost 6.6 years, followed by the Hispanic population with a loss of 4.2 years, the non-Hispanic Black population with a loss of 4.0 years, the non-Hispanic White population with a loss of 2.4 years, and the non-Hispanic Asian population with a loss of 2.1 years. The Hispanic and non-Hispanic Asian populations experienced over 95% of their respective declines during the first year of the pandemic. The non-Hispanic Black and non-Hispanic AIAN populations experienced 82.5% and 71.2% of their declines during the first year. Unlike these populations, the non-Hispanic White population experienced close to one-half of their decline (41.7%) during the second year of the pandemic.

The provisional mortality data on which the life tables are based have several

U.S. Department of Health and Human Services • Centers for Disease Control and Prevention • National Center for Health Statistics • National Vital Statistics System

8

limitations. First, the timeliness of death certificate data varies by jurisdiction and time. Some jurisdictions have historically taken longer to submit death certificates because of paper records, staffing shortages, or other local issues. The effects of recent changes in timeliness will not be clear until data are finalized. Another limitation is the variation in timeliness due to age and cause of death. Certain age groups, particularly under 5 years, may be underrepresented (5). Death certificates for deaths requiring investigation, including infant deaths, deaths involving external injuries, and drug overdose deaths take longer to complete; these deaths may be underreported in the 3–6 months after the deaths occurred. For example, the increase in drug overdose deaths in 2021 caused substantial delays in the completion of death certificates for some jurisdictions. Lastly, the timeliness of death certificate data by race and ethnicity has not been studied. Differences in timeliness by these factors may result in underestimation of deaths for specific groups. The underestimation of infant deaths, for example, will have a disproportionate effect on life expectancy at birth, given the latter's sensitivity to infant mortality, which is generally higher than mortality at all other ages up to the mid-50s or so.

# References

1. Arias E, Tejada-Vera B, Ahmad F. Provisional life expectancy estimates for January through June, 2020. Vital Statistics Rapid Release; no 10. Hyattsville, MD: National Center for Health Statistics. February 2021. DOI: https://dx.doi.org/10.15620/cdc:100392.

2. Arias E, Tejada-Vera B, Ahmad F, Kochanek KD. Provisional life expectancy estimates for 2020. Vital Statistics Rapid Release; no 15. Hyattsville, MD: National Center for Health Statistics. July 2021. DOI: https://dx.doi.org/10.15620/cdc:107201.

3. Arias E, Xu JQ. United States life tables, 2020. National Vital Statistics Reports; vol 71 no 1. Hyattsville, MD: National Center for Health Statistics. 2022. DOI: https://dx.doi.org/10.15620/cdc:118055.

4. Ahmad FB, Cisewski JA. Quarterly provisional estimates for selected indicators of mortality, 2020–Quarter 4, 2021. National Center for Health Statistics. National Vital Statistics System, Vital Statistics Rapid Release Program. 2022. Available from: https://www.cdc.gov/nchs/nvss/vsrr/mortality-dashboard.htm.

5. Ahmad FB, Dokpesi P, Escobedo L, Rossen L. Timeliness of death certificate data by sex, age, and geography. Vital Statistics Rapid Release; no 9. Hyattsville, MD: National Center for Health Statistics. June 2020. Available from: https://www.cdc.gov/nchs/data/vsrr/vsrr009-508.pdf.

6. Rossen LM, Ahmad FB, Spencer MR, Warner M, Sutton P. Methods to adjust provisional counts of drug overdose deaths for underreporting. Vital Statistics Rapid Release; no 6. Hyattsville, MD: National Center for Health Statistics. August 2018. Available from: https://www.cdc.gov/nchs/data/vsrr/report006.pdf.

7. Rossen LM, Ahmad FB, Anderson RN, Branum AM, Du C, Krumholz HM, et al. Disparities in excess mortality associated with COVID-19 – United States, 2020. MMWR Morb Mortal Wkly Rep 70(33):1114–9. 2021.

8. Arias E, Xu JQ. United States life tables, 2017. National Vital Statistics Reports; vol 68 no 7. Hyattsville, MD: National Center for Health Statistics. 2019. Available from: https://www.cdc.gov/nchs/data/nvsr/nvsr68/nvsr68_07-508.pdf.

9. Arias E. United States life tables by Hispanic origin. National Center for Health Statistics. Vital Health Stat 2(152). 2010. Available from: https://www.cdc.gov/nchs/data/series/sr_02/sr02_152.pdf.

10. Markides KS, Coreil J. The health of Hispanics in the southwestern United States: An epidemiologic paradox. Public Health Rep 101(3):253–65. 1986.

11. Office of Management and Budget. Revisions to the standards for the classification of federal data on race and ethnicity. Fed Register 62(210):58782–90. 1997. Available from: https://www.govinfo.gov/content/pkg/FR-1997-10-30/pdf/97-28653.pdf.

12. Office of Management and Budget. Race and ethnic standards for federal statistics and administrative reporting. Statistical Policy Directive 15. 1977. Available from: https://wonder.cdc.gov/wonder/help/populations/bridged-race/directive15.html.

13. Arias E, Heron M, Hakes JK. The validity of race and Hispanic-origin reporting on death certificates in the United States: An update. National Center for Health Statistics. Vital Health Stat 2(172). 2016. Available from: https://www.cdc.gov/nchs/data/series/sr_02/sr02_172.pdf.

14. Arias E, Xu JQ, Curtin S, Bastian B, Tejada-Vera B. Mortality profile of the non-Hispanic American Indian or Alaska Native population, 2019. National Vital Statistics Reports; vol 70 no 12. Hyattsville, MD: National Center for Health Statistics. 2021. DOI: https://dx.doi.org/10.15620/cdc:110370.

15. Arias E, Curtin SC, Tejada-Vera B. U.S. decennial life tables for 2009–2011: Methodology of the United States life tables. National Vital Statistics Reports; vol 69 no 10. Hyattsville, MD: National Center for Health Statistics. 2020.

16. Preston SH, Heuveline P, Guillot M. Demography: Measuring and modeling population processes. Oxford, England: Blackwell Publishers. 2001.

17. Thatcher AR, Kannisto V, Vaupel JW. The force of mortality at ages 80 to 120. Odense, Denmark: Odense University Press. 1998.

U.S. Department of Health and Human Services • Centers for Disease Control and Prevention • National Center for Health Statistics • National Vital Statistics System

9

18. Andreev KF, Bourbeau RR. Frailty modeling of Canadian and Swedish mortality at adult and advanced ages. Population Association of America Annual Meeting. 2007.

19. Elo IT, Turra CM, Kestenbaum B, Fergusson BR. Mortality among elderly Hispanics in the United States: Past evidence and new results. Demography 41(1):109–28. 2004.

20. Turra CM, Elo IT. The impact of salmon bias on the Hispanic mortality advantage: New evidence from Social Security data. Popul Res Policy Rev 27(5):515–30. 2008.

21. Brass W. On the scale of mortality. In: Brass W, editor. Biological aspects of demography. London, England: Taylor and Francis. 99–110. 1971.

22. Himes CL, Preston SH, Condran GA. A relational model of mortality at older ages in low mortality countries. Popul Stud 48(2):269–91. 1994.

23. Arias E, Xu J, Jim MA. Period life tables for the non-Hispanic American Indian and Alaska Native Population, 2007–2009. Am J Public Health 104(Suppl 3):S312–9. 2014.

24. Chiang CL. The life table and its applications. Malabar, FL: Krieger Publishing. 1984.

25. Wei R, Curtin LR, Arias E, Anderson RN. United States decennial life tables for 1999–2001, methodology of the United States life tables. National Vital Statistics Reports; vol 57 no 4. Hyattsville, MD: National Center for Health Statistics. 2008.

26. Silcocks PB, Jenner DA, Reza R. Life expectancy as a summary of mortality in a population: Statistical considerations and suitability for use by health authorities. J Epidemiol Community Health 55(1):38–43. 2001.

27. Arriaga EE. Changing trends in mortality decline during the last decades. In: Ruzicka L, Wunsch G, Kane P, editors. Differential mortality: Methodological issues and biosocial factors. Oxford, England: Clarendon Press. 1989.

28. Arriaga EE. Measuring and explaining the change in life expectancies. Demography 21(1):83–96. 1984.

29. Kochanek KD, Maurer JD, Rosenberg HM. Causes of death contributing to changes in life expectancy: United States, 1984–89. National Center for Health Statistics. Vital Health Stat 20(23). 1994. Available from: https://www.cdc.gov/nchs/data/series/sr_20/sr20_023.pdf.

30. Kochanek KD, Murphy SL, Xu J, Arias E, Tejada-Vera B. Deaths: Final data for 2020. National Vital Statistics Reports. Hyattsville, MD: National Center for Health Statistics. 2022. [Forthcoming].

U.S. Department of Health and Human Services • Centers for Disease Control and Prevention • National Center for Health Statistics • National Vital Statistics System

10

# Technical Notes

The methodology used to estimate provisional 2021 complete period life tables (Internet Tables I1–I21) on which the life expectancy estimates presented in this report are based is the same as that used to produce annual U.S. life tables, with some modifications (3). First, the life tables presented in this report are based on provisional death counts rather than on final death counts. Second, they are based on monthly population estimates based on the 2010 decennial census rather than on annual midyear population estimates based on the 2020 decennial census. Third, an adjustment factor based on 2020 differences between vital statistics and Medicare data is used to estimate mortality for ages 65–99 for the total, non-Hispanic White, and non-Hispanic Black populations. The main reason for the differences in methodology is data availability. Final death counts for the year 2021 will not be available until late in 2022. Similarly, postcensal 2021 midyear population estimates based on the 2020 decennial census are not yet available. A correction factor is used to adjust death rates at the oldest ages because Medicare data, used to supplement vital statistics data at older ages, is not yet available. Another difference is the use of provisional birth counts rather than final birth counts and linked birth and infant death data used for life tables by Hispanic origin and race, as these data are not yet available (Internet Tables I1–I18).

Standard errors of the two most important functions, the probability of dying and life expectancy (Internet Tables I19–I20), are estimated under the assumption that the data are only affected by random error because over 99% of deaths that occurred in 2021 are included. However, the possibility that certain jurisdictions and age groups may be underrepresented for later months could potentially lead to biases not accounted for by the estimated standard errors. Other possible errors, including age and Hispanic origin and race misreporting on death certificates are also not considered in the calculation of the variances or standard errors of the life table functions.

## Data for calculating life table functions

### Vital statistics data

Mortality data used to estimate the life tables presented in this report include over 99% of the deaths that occurred in 2021, although certain jurisdictions and age groups may be underrepresented for later months. Death data are typically over 99% complete 3 months after the date of death, but this can vary by jurisdiction, age of the decedent, and cause of death. Most jurisdictions submit over 90% of death data by 3 months after the date of death, but some jurisdictions take longer to submit death records. Death data for decedents aged under 5 years are 90% complete 3 months after the date of death, and 95% complete by 6 months. Provisional estimates of infant mortality are typically presented with a 9-month lag, as infant deaths require additional investigation and death certificates take longer to complete. Timeliness also varies by cause of death, with deaths due to external causes taking additional time to investigate and complete death certificates. Provisional estimates for most external causes of death (for example, falls, suicides, and unintentional injuries) are presented with a 6-month lag, while drug overdose deaths are presented with a 9-month lag.

Beginning with the 2018 data year, all 50 states and the District of Columbia reported deaths based on the 2003 revision of the U.S. Standard Certificate of Death for the entire year (3). The revision is based on the 1997 Office of Management and Budget (OMB) standards (11), which allow individuals to report more than one race and increased the race choices from four to five by separating the Asian and Pacific Islander groups. The Hispanic category did not change, remaining consistent with previous reports.

The Hispanic-origin and race groups in this report follow the 1997 standards and differ from the race categories used in reports for data years before 2018. From 2003 through 2017, not all deaths were reported using the 2003 certificate revision that allowed the reporting of more than one race based on the 1997 OMB race standard (11). During those years, multiple-race data were bridged to the 1977 standard single-race categories (12). Use of the bridged-race process was discontinued for the reporting of mortality statistics in 2018 when all states collected data on race according to the 1997 OMB guidelines for the full data year.

### Census population data

The population data used to estimate the life tables shown in this report are July 1, 2021, postcensal population estimates based on the 2010 decennial census and are available from the U.S. Census website at: https://www.census.gov/programs-surveys/popest/technical-documentation/research/evaluation-estimates/2020-evaluation-estimates/2010s-national-detail.html.

## Preliminary adjustment of the data

### Adjustment for unknown age

An adjustment is made to account for the small proportion of deaths for which age is not reported on the death certificate. The number of deaths in each age category is adjusted proportionally to account for those with not-stated ages. The following factor ($F$) is used to make the adjustment. $F$ is calculated for the total and for each sex group within a racial and ethnic population for which life tables are constructed:

$$F = \frac{D}{D^a} \qquad [1]$$

where $D$ is the total number of deaths and $D^a$ is the total number of deaths for which age is stated. $F$ is then applied by multiplying it by the number of deaths in each age group.

U.S. Department of Health and Human Services • Centers for Disease Control and Prevention • National Center for Health Statistics • National Vital Statistics System

11

## Adjustment for misclassification of Hispanic origin and race on death certificates

Two data sources were used to adjust for race and ethnicity misclassification on death certificates. For the Hispanic, non-Hispanic Asian, non-Hispanic Black, and non-Hispanic White populations, the National Longitudinal Mortality Study (NLMS) was used to produce classification ratios (or correction factors) to adjust observed sex- and age-specific death rates for misclassification on death certificates (13). NLMS consists of a series of Current Population Surveys (CPS) (1979–2011) linked to vital statistics mortality data through the National Death Index (NDI) (14). For the non-Hispanic American Indian or Alaska Native (AIAN) population, an extract of the 2010 Census Edited File (CEF)–Census Unedited File (CUF) Match File containing records for people classified as AIAN alone or in combination with another race in the 2010 decennial census was linked to NDI to identify decedents for the period April 1, 2010, to December 31, 2011. The resulting 34,366 CEF–CUF Match AIAN Extract–Mortality Linked Data decedent records were used to estimate classification ratios to correct for race and ethnicity misclassification on death certificates for the AIAN population (14).

The classification ratios consist of a comparison of self-reported Hispanic origin and race on CPS or the decennial census with Hispanic origin and race reported on the death certificates of the samples of decedents in NLMS who died during the period 1999–2011 and decedents in the CEF–CUF Match AIAN Extract who died between April 1, 2010, and December 31, 2011 (13,14). Linked records are used to estimate sex- age-specific ratios of survey or census Hispanic origin and race counts to death certificate counts (13,14).

The survey or census death certificate ratio (or classification ratio) is the ratio of the count (weighted in the case of CPS) of self-reported race and ethnicity on the survey or census to the count (weighted in the case of CPS) of the same

racial or ethnic category on the death certificates of the sample of NLMS (CEF–CUF Match AIAN Extract) decedents described previously. It can be interpreted as the net difference in assignment of a specific Hispanic-origin and race category between the two classification systems and can be used as a correction factor for race and ethnicity misclassification (7,9). It is assumed that the race and ethnicity reported by a survey or census respondent is more reliable than proxy reporting of race and ethnicity by a funeral director who has little personal knowledge of the decedent. Also, the 1997 OMB standards mandate that self-identification be the standard used for the collection and recording of race and ethnicity information (11).

Classification ratios discussed previously are used to adjust the age-specific number of deaths for ages 1–95 and over for the total Hispanic, non-Hispanic AIAN, non-Hispanic Asian, non-Hispanic Black, and non-Hispanic White populations, and by sex for each group, as follows:

$$_nD_x = {_nD_x^F} \cdot {_nCR_x} \qquad [2]$$

where $_nD_x^F$ is the age-specific number of deaths adjusted for unknown age as described above, $_nCR_x$ are the sex- and age-specific classification ratios used to correct for the misclassification of Hispanic origin and race on death certificates, and $_nD_x$ are the final age-specific counts of death adjusted for age and race and ethnicity misclassification.

Classification ratios for infant deaths are unreliable due to small sample sizes. Corrections for racial and ethnic misclassification of infant deaths are addressed by using infant death counts and live birth counts from linked birth and infant death data files rather than the traditional birth and death data files (3). In the linked file, each infant death record is linked to its corresponding birth record so that the race and ethnicity of the mother reported on the birth record can be ascribed to the infant death record. Due to the unavailability of linked birth and infant death or final

birth data at this time, provisional birth and death data were used to estimate infant mortality for 2021 provisional life tables. Typically, infant mortality rates based on linked birth and infant death data show that using the traditional files overestimates the infant mortality rate by 3.3% for Hispanic, 1.8% for non-Hispanic AIAN, and 4.7% for non-Hispanic Black infants; and underestimates the infant mortality rate by 22.2% for non-Hispanic Asian and 2.7% for non-Hispanic White infants (3).

## Interpolation of $P_x$ and $D_x$

Anomalies—both random and those associated with reporting age at death—can be problematic when using vital statistics and census data by single years of age to estimate the probability of death (1,3). Graduation techniques are often used to eliminate these anomalies and to derive a smooth curve by age. Beers' ordinary minimized fifth difference formula is used to obtain smoothed values of population counts ($P_x$) and death counts ($D_x$) from 5-year age groupings of $_nP_x$ from age 0 to 99 and $_nD_x$ from age 5 to 99, and where $_nD_x$ has first been adjusted for not-reported age and race and ethnicity misclassification on the death certificate (15).

## Calculation of the probability of dying ($q_x$)

The first step in the calculation of a complete period life table is the estimation of the age-specific probability of dying, $q_x$, which is derived from the age-specific death rate, $m_x$ (3,16). In the life table cohort,

$$m_x = \frac{d_x}{L_x}$$

where $d_x$ is the number of deaths occurring between ages $x$ and $x + 1$, and $L_x$ is the number of person-years lived by the life table cohort between ages $x$ and $x + 1$. The conversion of the age-specific

U.S. Department of Health and Human Services • Centers for Disease Control and Prevention • National Center for Health Statistics • National Vital Statistics System

12

death rate, $m_x$, to the age-specific probability of death, $q_x$, is as follows:

$$q_x = \frac{m_x}{1+(1-a_x)m_x} \quad [3]$$

where $a_x$ is the fraction of the number of person-years lived in the age interval by members of the life table cohort who died in the interval. When the age interval is 1 year, except at infancy, $a_x = 1/2$; in other words, deaths occur on average midway through the age interval. As a result,

$$q_x = \frac{m_x}{1+\frac{1}{2}m_x} \quad [4]$$

Because the complete period life table is based on the age-specific death rates of a current population observed for a specific calendar year, the life table death rate is equivalent to the observed death rates of the current population:

$$m_x = \frac{d_x}{L_x} = M_x = \frac{D_x}{P_x}$$

where $D_x$ is the Beers' smoothed number of deaths adjusted for not-stated age and race and ethnicity misclassification on the death certificate (for the Hispanic, non-Hispanic AIAN, non-Hispanic Asian, non-Hispanic Black, and non-Hispanic White populations), and $P_x$ is the Beers' smoothed population at risk of dying between ages $x$ and $x + 1$. Then,

$$q_x = \frac{M_x}{1+\frac{1}{2}M_x} = \frac{D_x}{P_x+\frac{1}{2}D_x} \quad [5]$$

This procedure is used to estimate vital statistics age-specific probabilities of death for ages 1–99.

## Calculation of $q_x$ at age 0

The higher mortality observed in infancy is associated with a high concentration of deaths occurring at the beginning of the age interval rather than in the middle. As a result, whenever possible, it is best to assign deaths to the appropriate birth cohorts. Consequently, the probability of death at birth, $q_0$, is calculated using a birth cohort method that uses a separation factor ($f$) defined as the proportion of infant deaths in

year $t$ occurring in infants born in the previous year ($t - 1$). The value $f$ is estimated by categorizing infant deaths by date of birth. The probability of death is then calculated as:

$$q_0 = \frac{D_0(1-f)}{B^t} + \frac{D_0(f)}{B^{t-1}} \quad [6]$$

where $D_0$ is the number of infant deaths adjusted for not-stated age in 2021, $B^t$ is the number of live births in 2021, and $B^{t-1}$ is the number of live births in 2020.

## Probabilities of dying at the oldest ages for the total, non-Hispanic Black, and non-Hispanic White populations

To estimate the annual U.S. life tables, Medicare data are used to supplement vital statistics data for the estimation of $q_x$ at the oldest ages because these data are more accurate as proof of age is required for enrollment in the Medicare program. Medicare data are used to estimate the probability of dying for ages 66 and over for the total, non-Hispanic Black, and non-Hispanic White populations.

The method consists of the following steps. First, vital statistics and Medicare death rates are blended in the age range 66–99. Second, a logistic model is used to smooth the blended death rates in the age range 85–99 and predict death rates for ages 100–120. Third, final resulting death rates, $M_x$, are converted to $q_x$.

For ages 66–94, vital statistics death rates, $M_x^V$, and Medicare death rates, $M_x^M$, are blended with a weighting process that gives gradually declining weight to vital statistics data and gradually increasing weight to Medicare data. For ages 95–99, $M_x^M$ is used exclusively. Blended $M_x$ is thus obtained as follows:

$$M_x = \frac{1}{30}\left[(95-x)M_x^V + (x-65)M_x^M\right] \quad [7]$$

when $x = 66,...,94$

and $M_x = M_x^M$

when $x = 95,...,99$.

$M_x^M$ is estimated as follows:

$$M_x^M = \frac{D_x^M}{P_x^M}$$

where $D_x^M$ is the age-specific Medicare death count, and $P_x^M$ is the age-specific Medicare midyear population count.

Because Medicare data for 2021 are not yet available, a modification of the method described above is used to estimate blended vital and Medicare death rates, $M_x$. The difference between vital statistics death rates, $M_x^V$, and Medicare death rates, $M_x^M$, for ages 66–99 does not vary significantly from year to year. As a result, it is possible to use the previous year's information to generate an adjustment factor to approximate blended vital and Medicare death rates, $M_x$, for the year in which Medicare data are unavailable.

An adjustment factor, $R_x$, is used to approximate blended vital statistics and Medicare death rates, $M_x$, for the total, non-Hispanic Black, and non-Hispanic White populations by sex in 2021 as follows:

$$M_x = M_x^V \cdot R_x \quad [8]$$

where $M_x^V$ are the observed vital statistics death rates in 2021 and $R_x$ is calculated as follows:

$$R_x = \frac{{}^{2020}M_x}{{}^{2020}M_x^V}$$

where ${}^{2020}M_x$ is the blended vital statistics and Medicare death rate and ${}^{2020}M_x^V$ is the vital statistics death rate for ages 66 to 99 in 2020.

A logistic model proposed by Kannisto is then used to smooth $M_x$ in the age range 85–99 and predict $M_x$ in the age range 100–120 (17). The start of the modeled age range varies by race- and ethnicity-specific population because it is a function of the age at which the rate of change in the age-specific death rates peaks. Currently, the rate of change in the age-specific death rate rises steadily up to approximately ages 80–85 or so and then begins to decline. As a result,

U.S. Department of Health and Human Services • Centers for Disease Control and Prevention • National Center for Health Statistics • National Vital Statistics System

13

it is difficult to model a large age span, such as 65–100, with one simple model without over-smoothing and as a result changing the underlying mortality pattern seen in the population of interest (18). Further, the observed data for the age range 65–85 or so is reliable and robust, as indicated by the very close similarity between vital statistics and Medicare death rates, so it is unnecessary to model (smooth) the entire age span (65–100).

The Kannisto model is a simple form of a logistic model in which the logit of $u_x$ (or the natural log of the odds of $u_x$) is a linear function of age, $x$ (17). It is expressed as:

$$\ln\left(\frac{u_x}{1-u_x}\right) = \ln(\alpha) + \beta x \qquad [9]$$

where $u_x$, the force of mortality (or instantaneous death rate), is defined as:

$$u_x = \frac{\alpha e^{\beta x}}{1+\alpha e^{\beta x}}$$

Because $u_x$ is not directly observed but is closely approximated by $m_x$, and $m_x = M_x$, then the logit of $M_x$ is modeled instead. A maximum-likelihood generalized linear model estimation procedure is used to fit the following model in the age range 85–99 years:

$$\ln\left(\frac{M_x}{1-M_x}\right) = \ln(\alpha) + \beta x \qquad [10]$$

Then, the estimated parameters are used to predict $\bar{M}_x$ as follows:

$$\bar{M}_x = \frac{e^a e^{bx}}{1+e^a e^{bx}} \text{ or, equivalently,}$$

$$\bar{M}_x = \frac{e^{a+bx}}{1+e^{a+bx}} \qquad [11]$$

where $a$ and $b$ are the predicted values of parameters $\ln(\alpha)$ and $\beta$, respectively, given by fitting model 10.

Finally, the predicted probability of death, $\bar{q}_x$, for ages 85–120 is estimated by converting $\bar{M}_x$ as follows:

$$\bar{q}_x = \frac{\bar{M}_x}{1+\frac{1}{2}\bar{M}_x} \qquad [12]$$

The probability of death is extrapolated to age 120 to estimate the life table population until no survivors remain. This information is then used to estimate $L_x$ for ages 100–120, which is used to close the table with the age category 100 and over, combined (discussed in the following section).

## Probabilities of dying at the oldest ages for the Hispanic, non-Hispanic AIAN, and non-Hispanic Asian populations

As noted previously, Medicare data are unreliable for the Hispanic, non-Hispanic AIAN, and non-Hispanic Asian populations due to inconsistencies in the Medicare race and ethnicity classification system. As a result, other methods were used to estimate mortality at the oldest ages for these populations. Beyond age 80, mortality estimates based strictly on vital statistics data for the Hispanic, non-Hispanic AIAN, and non-Hispanic Asian populations are too low, despite correction for ethnic misclassification on the death certificate.

A consistent finding across diverse studies has been that Hispanic mortality in the adult and advanced ages varies between approximately 80% and 89% relative to that of the non-Hispanic White population (9,19,20). The Brass relational logit model takes advantage of the relationship between Hispanic and non-Hispanic White mortality previously identified and has been widely and successfully used to predict the mortality of one population relative to another at the older ages (21–23). Using the age-specific mortality pattern of the non-Hispanic White population as the standard, the Brass relational logit model is used to predict Hispanic mortality in the older ages. The standard is fit to Hispanic data in the age interval 45–80, and the predicted parameters are used to estimate the probabilities of death for ages 76–100. This method allows the relationship between the two populations in the younger ages to be extended to the older ages (9,22,23).

Although similar information is not available for the non-Hispanic AIAN and non-Hispanic Asian populations, with a slight modification, the Brass relational logit model was successfully used to produce reliable complete period life tables for the non-Hispanic AIAN population in Indian Health Service Contract Health Service Delivery Area counties (23). The choice of the non-Hispanic White population as the standard population is based on several factors. First, it is the most widely used comparison population in the study of racial and ethnic disparities given its social and economic privilege. Second, it is the largest population in the United States and has the most reliable mortality data. Third, the relationship between the age-specific mortality patterns of the non-Hispanic AIAN and non-Hispanic Asian populations and the non-Hispanic White population remains constant throughout the age span 45–80 (45–84 for the non-Hispanic AIAN population). Assuming that this pattern continues to the oldest ages is reasonable because the final results are consistent with expected age-specific mortality patterns at the oldest ages.

The Brass relational logit model expresses the age-specific mortality pattern of a population of interest as a function of the age-specific mortality pattern of a standard population and is expressed as:

$$\bar{Y}_x = \alpha + \beta Y_x^S \qquad [13]$$

where $\bar{Y}_x$ is the predicted logit of the probability of death, $q_x$, in the population of interest, that is,

$$logit[q_x] = \ln\left[\frac{q_x}{1-q_x}\right]$$

$Y_x^S$ is the logit of the probability of death in the standard population, $q_x^S$, that is,

$$logit[q_x^S] = \ln\left[\frac{q_x^S}{1-q_x^S}\right]$$

$\alpha$ is the predicted parameter that measures the level of mortality of the population of interest relative to the standard population, and $\beta$ is the

U.S. Department of Health and Human Services • Centers for Disease Control and Prevention • National Center for Health Statistics • National Vital Statistics System

14

predicted parameter that measures the slope of the mortality function of the population of interest relative to the standard population (9,22,23).

A maximum-likelihood generalized linear model estimation procedure was used to fit equation 13 in the age range 45–80 (45–84 for the non-Hispanic AIAN population). The resulting predicted parameters $\alpha$ and $\beta$ were then used to estimate the predicted probability of death for ages 76–120 (80–120 for the non-Hispanic AIAN population). The value $q_x$ was predicted to age 120 to estimate the life table population until no survivors remain, as was done for the other population groups. This information was then used to estimate $L_x$ for ages 100–120, which was used to close the table with the age category 100 and over, combined (discussed in the next section).

Predicted $\bar{q}_x$ is estimated by transforming its logit, $\bar{Y}_x$ back as follows:

$$\bar{q}_x = \frac{\exp\left[\bar{Y}_x\right]}{\left[1+\exp\bar{Y}_x\right]} = \frac{\exp\left[\alpha+\beta Y_x^S\right]}{1+\exp\left[\alpha+\beta Y_x^S\right]} \quad [14]$$

To ensure a smooth transition from vital $q_x^V$ and predicted $\bar{q}_x$, the two were blended for ages 76–80 (80–84 for the non-Hispanic AIAN population) with a graduating process as follows:

$$q_x = \frac{1}{6}\left[(81-x)q_x^V + (x-75)\bar{q}_x\right] \quad [15]$$

when $x = 76,...,80$, and

$$q_x = \frac{1}{6}\left[(85-x)q_x^V + (x-79)\bar{q}_x\right]$$

when $x = 80,...,84$.

Finally, to close the table at age 100 and over (combined), $_\infty q_{100}$ is set equal to 1.0 because all survivors to this age will die at some point in the open-ended age interval. Once $q_x$ is obtained for each single year of age, the other life table functions are easily calculated.

## Calculation of remaining life table functions for all groups

### Survivor function ($l_x$)

The life table radix, $l_0$, is set at 100,000. For ages greater than 0, the number of survivors remaining at exact age $x$ is calculated as

$$l_x = l_{x-1}(1-q_{x-1}) \quad [16]$$

### Decrement function ($d_x$)

The number of deaths occurring between ages $x$ and $x + 1$ is calculated from the survivor function:

$$d_x = l_x - l_{x+1} = l_x\, q_x \quad [17]$$

Note that $_\infty d_{100} = {}_\infty l_{100}$ because $_\infty q_{100} = 1.0$.

### Person-years lived ($L_x$)

Person-years lived for ages 1–99 are calculated assuming that the survivor function declines linearly between ages $x$ and $x + 1$. This gives the formula:

$$L_x = \frac{1}{2}(l_x + l_{x+1}) = l_x - \frac{1}{2}d_x \quad [18]$$

For $x = 0$, the separation factor $f$ is used to calculate $L_0$:

$$L_0 = fl_0 + (1-f)l_1 \quad [19]$$

Finally, $_\infty L_{100}$ is estimated as the sum of the extrapolated $L_x$ values for ages 100–120.

### Person-years lived at and above age $x$ ($T_x$)

$T_x$ is calculated by summing $L_x$ values at and above age $x$:

$$T_x = \sum_{x=0}^{\infty} L_x \quad [20]$$

### Life expectancy at age $x$ ($e_x$)

Life expectancy at exact age $x$ is calculated as:

$$e_x = \frac{T_x}{l_x} \quad [21]$$

## Variances and standard errors of the probability of dying and life expectancy

Variances are estimated under the assumption that the mortality data on which the life tables are based are not affected by sampling error and are subject only to random variation. However, although over 99% of deaths that occurred from January through December 2021 are included, the data may be biased by the possibility that certain jurisdictions and age groups may be underrepresented for later months. These errors as well as those resulting from age and Hispanic origin and race misreporting on death certificates are not considered in the calculation of the variances or standard errors of the life table functions.

The methods used to estimate the variances of $q_x$ and $e_x$ are based on Chiang (24), with some necessary modifications due to the use of Medicare data and statistical modeling for smoothing and prediction of old age death rates. Based on the assumption that deaths are binomially distributed, Chiang proposed the following equation for the variance of $q_x$:

$$Var(q_x) = \frac{q_x^2(1-q_x)}{D_x} \quad [22]$$

where $D_x$ is the age-specific number of deaths. For the total, non-Hispanic White, and non-Hispanic Black populations, this equation is used to estimate $Var(q_x)$ throughout the age span, with a modification where for ages less than 66 years, $D_x$ is the deaths from vital statistics data, smoothed by interpolation and adjusted for the number of deaths with age not stated and for the non-Hispanic White and non-Hispanic Black populations adjusted for race and ethnicity misclassification. For ages 66 and above, $D_x$ was obtained by treating the population as a cohort population and calculated from $q_x$ (16,25):

$$P_x = \frac{(P_{x-1} - 0.5D_{x-1}) \bullet (2 - q_x)}{2} \quad [23]$$

U.S. Department of Health and Human Services • Centers for Disease Control and Prevention • National Center for Health Statistics • National Vital Statistics System

15

$$D_x = \frac{q_x P_x}{1 - 0.5 q_x}$$

where $D_x$ is the number of deaths at age $x$ and $P_x$ is the mid-year population at age $x$.

This method is used under the assumption that the application of the adjustment factor $R_x$ is equivalent to using blended vital statistics and Medicare data.

For the Hispanic, non-Hispanic AIAN, and non-Hispanic Asian populations, equation 20 was used for ages 0–75 (0–79 for non-Hispanic AIAN). Because $q_x$ was predicted based on the Brass relational logit model for ages 76–120 (80–120 for non-Hispanic AIAN), the Delta method was used to approximate its variance for these ages as follows:

$$Var(\overline{q}_x) = \left( \frac{\exp\left(\alpha + \beta Y_x^S\right)}{\left[\exp\left(\alpha + \beta Y_x^S\right) + 1\right]^2} \right)^2 \bullet Var\left(\alpha + \beta Y_x^S\right) \qquad [24]$$

For ages 76–80 (80–84 for non-Hispanic AIAN), the variance of $q_x$ is calculated as:

$$\frac{1}{6}\left[(81 - x) \bullet Var(q_x) + (x - 75) \bullet Var(\overline{q}_x)\right],$$

when $x = 76,\ldots 80$, and

$$\frac{1}{6}\left[(85 - x) \bullet Var(q_x) + (x - 79) \bullet Var(\overline{q}_x)\right],$$

when $x = 80,\ldots 84$.

### Standard error of $q_x$

$$SE(q_x) = \sqrt{Var(q_x)} \qquad [25]$$

For all groups, variances of the life expectancies for ages 0–99 years are estimated using Chiang's equation:

$$Var(e_x) = \frac{\sum_{x=0}^{x=99} l_x^2 \bullet \left[(1 - 0.5) + e_{x+n}\right]^2 \bullet Var\left(q_x\right)}{l_x^2} \qquad [26]$$

Chiang assumed that because $q_{100+} = 1.00$, then $Var(q_{100+}) = 0$, and therefore, $Var(e_{100+}) = 0$. Silcocks et al proposed that in the final age group, life expectancy is dependent on the mean length of survival and not on the probability of survival, and therefore, the assumption of no variance is incorrect, and $Var(e_{100+})$ can be approximated as (26):

$$Var(e_{100+}) \approx \frac{\frac{l_{100+}^2}{M_{100+}^4} \bullet Var\left(M_{100+}\right)}{l_{100+}^2}$$

### Standard error of $e_x$

$$SE(e_x) = \sqrt{Var(e_x)} \qquad [27]$$

## Causes of death contributing to changes in life expectancy

To measure changes in mortality, a discrete method developed by Arriaga (27,28) was used to estimate the contribution of mortality change by causes of death based on changes in life expectancy, which is described as a procedure that "estimates the number of years added to or removed from life expectation because of the decrease or increase (respectively) of the central mortality rates of life tables" (27). With this method, one can partition the change in life expectancy over time or between two separate groups of populations. In this report, Arriaga's technique is used to partition changes in life expectancy at birth in the United States from 2020 to 2021 by cause of death.

The method partitions changes into component additive parts and identifies the causes of death having the greatest influence, positive or negative, on changes in life expectancy based on rankable causes of death (29). This is the same method used by the National Center for Health Statistics annually to analyze changes in life expectancy (30).

## Acknowledgments

**Suggested citation**

Arias E, Tejada-Vera B, Kochanek KD, Ahmad FB. Provisional life expectancy estimates for 2021. Vital Statistics Rapid Release; no 23. Hyattsville, MD: National Center for Health Statistics. August 2022. DOI: https://dx.doi.org/10.15620/cdc:118999.

**Copyright information**

**National Center for Health Statistics**

Brian C. Moyer, Ph.D., *Director*

Amy M. Branum, Ph.D., *Associate Director for Science*

**Division of Vital Statistics**

Steven Schwartz, Ph.D., *Director*

Andrés A. Berruti, Ph.D., M.A., *Associate Director for Science*

CS332641

U.S. Department of Health and Human Services • Centers for Disease Control and Prevention • National Center for Health Statistics • National Vital Statistics System

16