

# Supplemental Material: Declining Life Expectancy in the United States: Missing the Trees for the Forest

Sam Harper<sup>1,2,3</sup>

Corinne A. Riddell<sup>4</sup>

Nicholas B. King<sup>1,2,5</sup>

<sup>1</sup> Department of Epidemiology, Biostatistics & Occupational Health, McGill University, Montreal Canada

<sup>2</sup> Institute for Health and Social Policy, McGill University, Montreal, Canada

<sup>3</sup> Department of Public Health, Erasmus Medical Center, Rotterdam, The Netherlands

<sup>4</sup> Division of Epidemiology & Biostatistics, School of Public Health, UC Berkeley, Berkeley, USA

<sup>5</sup> Biomedical Ethics Unit, McGill University, Montreal, Canada

## 1 DATA SOURCES AND METHODS

### 1.1 Data

We used a number of data sources to construct plots and to analyze life expectancy trends. Data on US life expectancy and life expectancy for other Organization for Economic Cooperation and Development countries were abstracted using the World Bank's World Development Indicators online database (The World Bank Group 2020), specifically the series on life expectancy at birth. For estimates within the US, the underlying data source was the US National Vital Statistics System, maintained by the [National Center for Health Statistics](#), which collects information on all deaths occurring in the United States each year.

Although NCHS collects data on race and gender, accurate mortality statistics for the "American Indian and Alaska Native" group are difficult to produce because of misclassification of Indigenous individuals as white on death certificates. Thirty percent of deaths to Indigenous individuals are estimated to be misclassified, hindering abilities to calculate life expectancy or calculate accurate, cause-specific mortality rates. In 2008, Arias and colleagues (Arias et al. 2008) adjusted for misreporting of Indigenous ethnicity and estimated life expectancy for this group to be 71.1, 2.8 years lower than the next-lowest life expectancy estimate for Non-Hispanic Blacks. The age-adjusted mortality rate from opioid overdoses (not adjusted for ethnicity misreporting) was estimated to be between the mortality rates for non-Hispanic whites and non-Hispanic blacks during 2017 (Woolf and Schoomaker 2019). Even with under-reporting Indigenous Americans have the highest reported rates of suicide, and the second-highest reported rate of homicide (Kochanek et al. 2019). A more recent update by Arias and colleagues (Arias et al. 2016) found that the quality of reporting for the American Indian and Alaska Native population remains poor. For these reasons we were unable to estimate life expectancy and cause-specific mortality for this population.

#### 1.1.1 CDC WONDER

We extracted mortality data by age, gender, and cause of death for various analyses using the CDC's Wide-ranging Online Data for Epidemiologic Research, called [WONDER](#) (Centers for Dis-

ease Control and Prevention and National Center for Health Statistics 2020). We used the “Underlying Cause of Death 1999-2018” [database](#) to generate mortality trends by age, gender, race, ethnicity, and cause-of-death, as well as geography for some figures.

### 1.1.2 SEER Stat

In addition to CDC WONDER we also used the SEER\*Stat interface (National Cancer Institute 2019) to produce some mortality rates that were more difficult to do using CDC WONDER, such as age-adjusted rates within broad age categories, as well as longer term trends for some causes of death. However, the underlying data for SEER\*Stat are also maintained and produced by NCHS. See the mortality documentation at the SEER\*Stat [website](#).

## 1.2 Estimation of life expectancy and decomposition

We used standard methods for creating abridged life tables (Preston, Heuveline, and Guillot 2001), with 11 10-year age groups (0-1, 1-4, 5-14, … , 85+). We calculated the contribution of each age group and cause-of-death to the difference in life expectancy between the years 2014 and 2017 using the methods developed by Arriaga (Arriaga 1984). Briefly, the contribution of a particular age group is both a direct function of the difference in age-specific mortality rates at that age plus an additional contribution resulting from the fact that mortality differences at a given age will produce additional survivors at older ages (Arriaga 1984), given by the formula below:

$$n\Delta_x = \left[ \underbrace{l_x^{2017}/l_0^{2017}}_{\text{fraction of survivors}} \times \underbrace{\left( \frac{nL_x^{2014}}{l_x^{2014}} - \frac{nL_x^{2017}}{l_x^{2017}} \right)}_{\text{direct effect}} \right] + \underbrace{\left[ \frac{T_{x+n}^{2014}}{l_{x+n}^{2014}} \times \frac{\frac{l_x^{2017}l_{x+n}^{2014}}{l_x^{2014}} - l_{x+n}^{2017}}{l_0^{2017}} \right]}_{\text{indirect effect + interaction}}$$

where  $n\Delta_x$  is the total contribution for a given age interval between  $x$  and  $x+n$ ,  $l_x$  is number alive at age  $x$ ,  $L_x$  is the person-years lived in the interval, and  $T_x$  is the person-time lived beyond age  $x$ .

The contribution  $n\Delta_x^i$  of each cause of death  $i$  within a given age group is a function of the difference between the two time periods in the proportion of deaths due to a given cause (Arriaga 1989):

$$n\Delta_x^i = n\Delta_x \times \frac{\overbrace{\left( np_x^{i,2014} \times nr_x^{2014} \right) - \left( np_x^{i,2017} \times nr_x^{2017} \right)}^{\text{difference in share of deaths for cause } i}}{\underbrace{nr_x^{2014} - nr_x^{2017}}_{\text{overall mortality rate difference}}}$$

where  $n\Delta_x$  is the total contribution for an age group,  $np_x^i$  is the proportion of deaths within age group  $x$  due to cause  $i$ , and  $nr_x$  is the overall age-specific death rate. The total difference in life expectancy is the net sum of the age-cause components:

$$\sum_i n\Delta_x^i = n\Delta_x, \text{ and } e_0^{2014} - e_0^{2017} = \sum_x n\Delta_x = \sum_x \sum_i n\Delta_x^i$$

The total difference in life expectancy at birth between 2017 and 2014 is the sum of the age-cause specific components. The total contribution of a given age group to the difference in life expectancy

is equal to the sum of its contributions across all causes of death. Likewise, the contribution of a particular cause of death is the sum of its contributions across age groups.

### **1.3 Joinpoint analysis of life expectancy trends**

We also used joinpoint regression software (National Cancer Institute 2020) to examine when trends in life expectancy and mortality may have changed in recent years. This was used to produce the estimates of changes in US life expectancy relative to OECD countries, as well as for the estimates of US life expectancy trends by gender and race-ethnicity. This program uses segmented weighted least squares regression and iteratively adds up to 5 segments and uses a permutation test (Kim et al. 2000) to select the model that best fits the observed data.

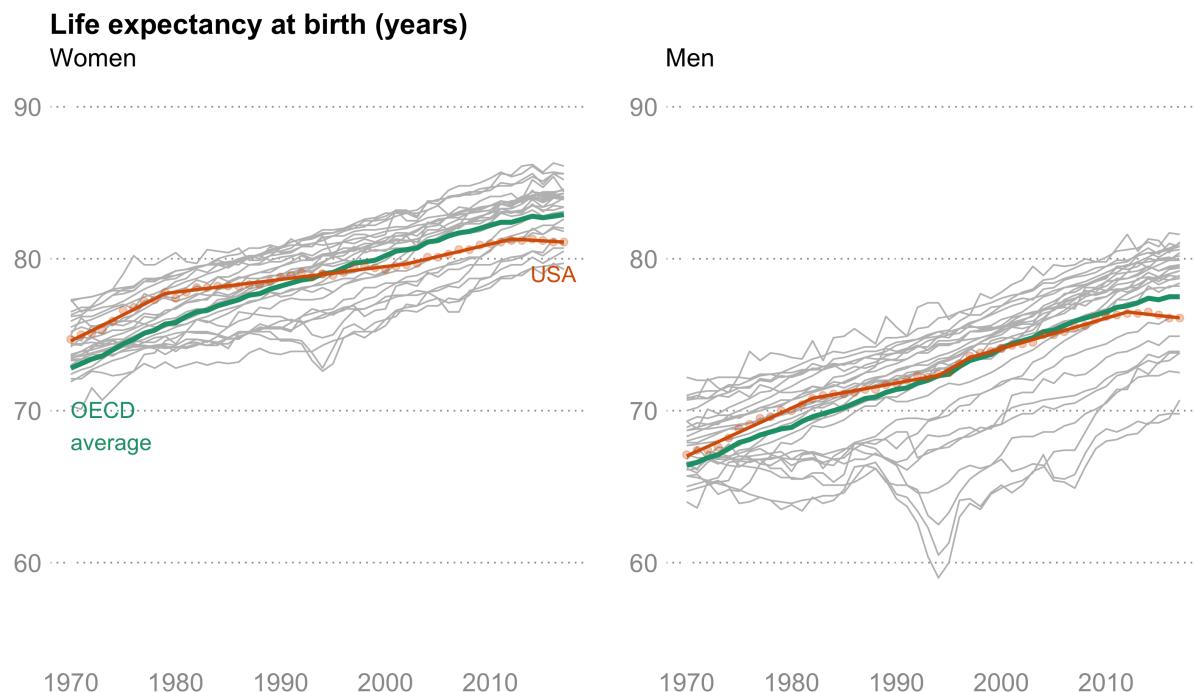
## **2 SUPPLEMENTAL TABLES AND FIGURES**

The data and code to reproduce these tables and figures are provided via the links in each caption, as well as in a public repository at the Open Science Foundation: <https://osf.io/ypwfh/>.

**Supplemental Table 1.** Life expectancy by gender and race-ethnicity, 1999-2018. Source: Authors' calculations of data from CDC WONDER (Centers for Disease Control and Prevention and National Center for Health Statistics 2020). Data: <https://osf.io/y4fzx/>, <https://osf.io/wsdvb/> Code: <https://osf.io/ewxm7/>

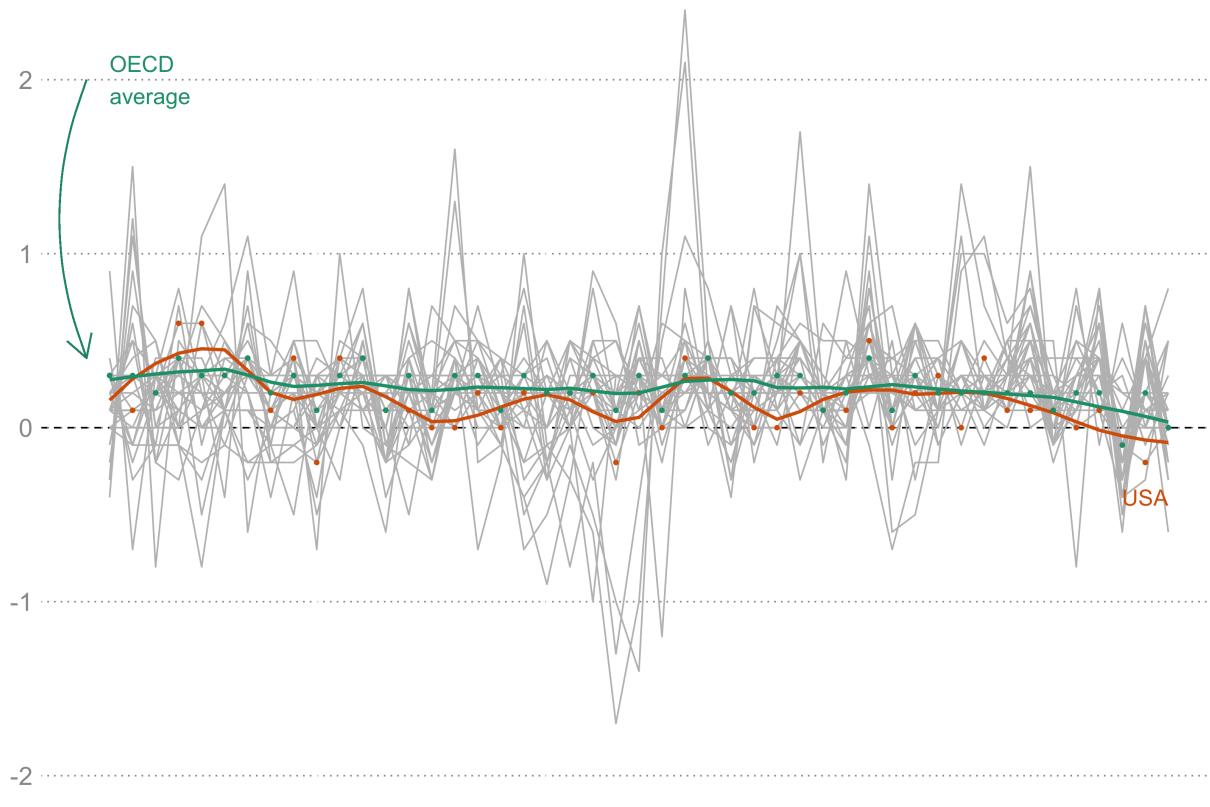
Year	Non-Hispanic API		Non-Hispanic Black		Non-Hispanic White		Hispanic	
	Women	Men	Women	Men	Women	Men	Women	Men
1999	86.6	81.3	74.9	67.8	79.8	74.7	83.0	77.2
2000	87.0	81.6	75.0	68.1	79.9	74.8	83.3	77.4
2001	87.1	82.0	75.3	68.4	79.9	74.9	83.3	77.6
2002	87.3	82.2	75.3	68.6	79.9	74.9	83.5	77.7
2003	87.4	82.5	75.5	68.8	80.0	75.1	83.6	77.9
2004	87.8	83.1	76.0	69.3	80.4	75.5	84.2	78.6
2005	87.9	82.9	76.1	69.4	80.4	75.5	84.1	78.2
2006	88.1	83.3	76.6	69.8	80.6	75.8	84.4	78.9
2007	88.5	83.8	77.0	70.2	80.9	76.0	84.9	79.3
2008	88.3	84.0	77.3	70.9	80.8	76.0	84.8	79.7
2009	88.7	84.2	77.8	71.4	81.2	76.4	85.4	80.1
2010	88.8	84.1	78.1	71.8	81.2	76.5	85.4	80.2
2011	89.2	84.7	78.3	72.2	81.2	76.5	85.8	80.8
2012	89.1	84.8	78.6	72.3	81.3	76.6	85.8	80.9
2013	89.5	84.7	78.6	72.4	81.3	76.6	86.0	81.0
2014	90.0	85.5	78.8	72.7	81.3	76.6	86.3	81.3
2015	89.7	85.4	78.8	72.4	81.1	76.5	86.3	81.2
2016	89.7	85.5	78.6	72.1	81.2	76.3	86.4	81.1
2017	89.7	85.3	78.8	72.0	81.1	76.3	86.4	81.1
2018	90.0	85.5	78.8	72.0	81.3	76.4	86.5	81.0
Changes								
2014-2017	-0.3	-0.2	0.0	-0.7	-0.2	-0.3	0.1	-0.2
2010-2018	1.2	1.4	0.7	0.2	0.1	-0.1	1.1	0.8

**Supplemental Figure 1.** Life expectancy at birth in the United States and 29 other high-income countries, 1969-2017. Source: World Development Indicators (The World Bank Group 2020). Data: <https://osf.io/d2b7c/>, <https://osf.io/n4mj6/> Code: <https://osf.io/muyrk/>



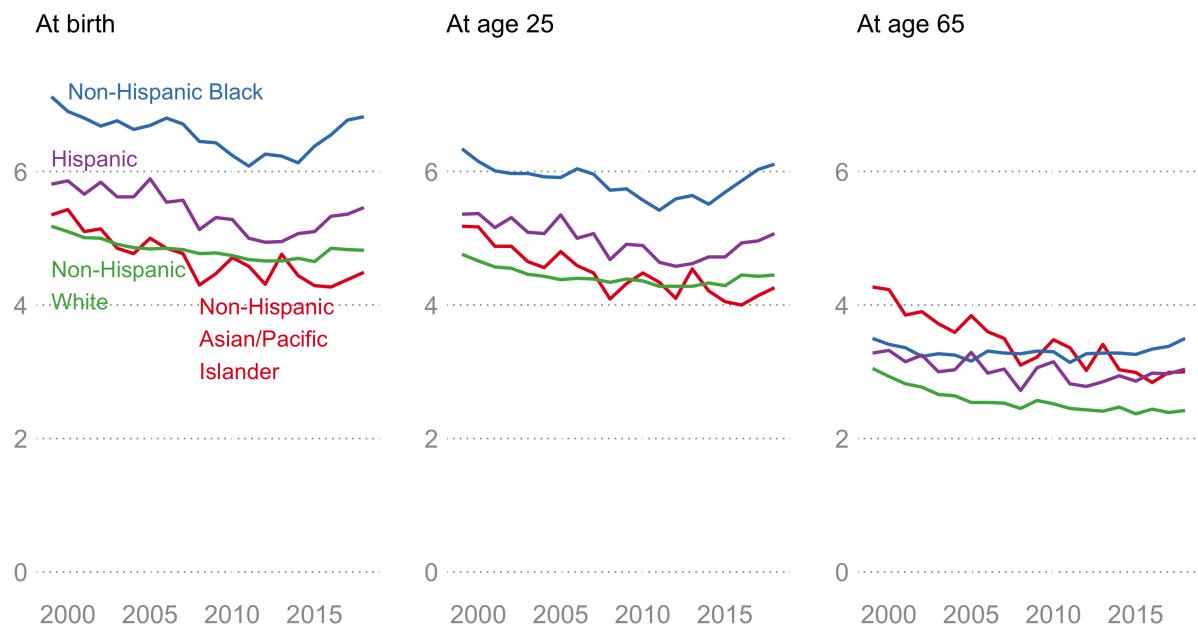
**Supplemental Figure 2.** Annual year-on-year change in life expectancy, USA and average of 29 other OECD countries. Source: World Development Indicators (The World Bank Group 2020). Data: <https://osf.io/d2b7c/>, <https://osf.io/n4mj6/> Code: <https://osf.io/muyrk/>

**Annual change (years)**

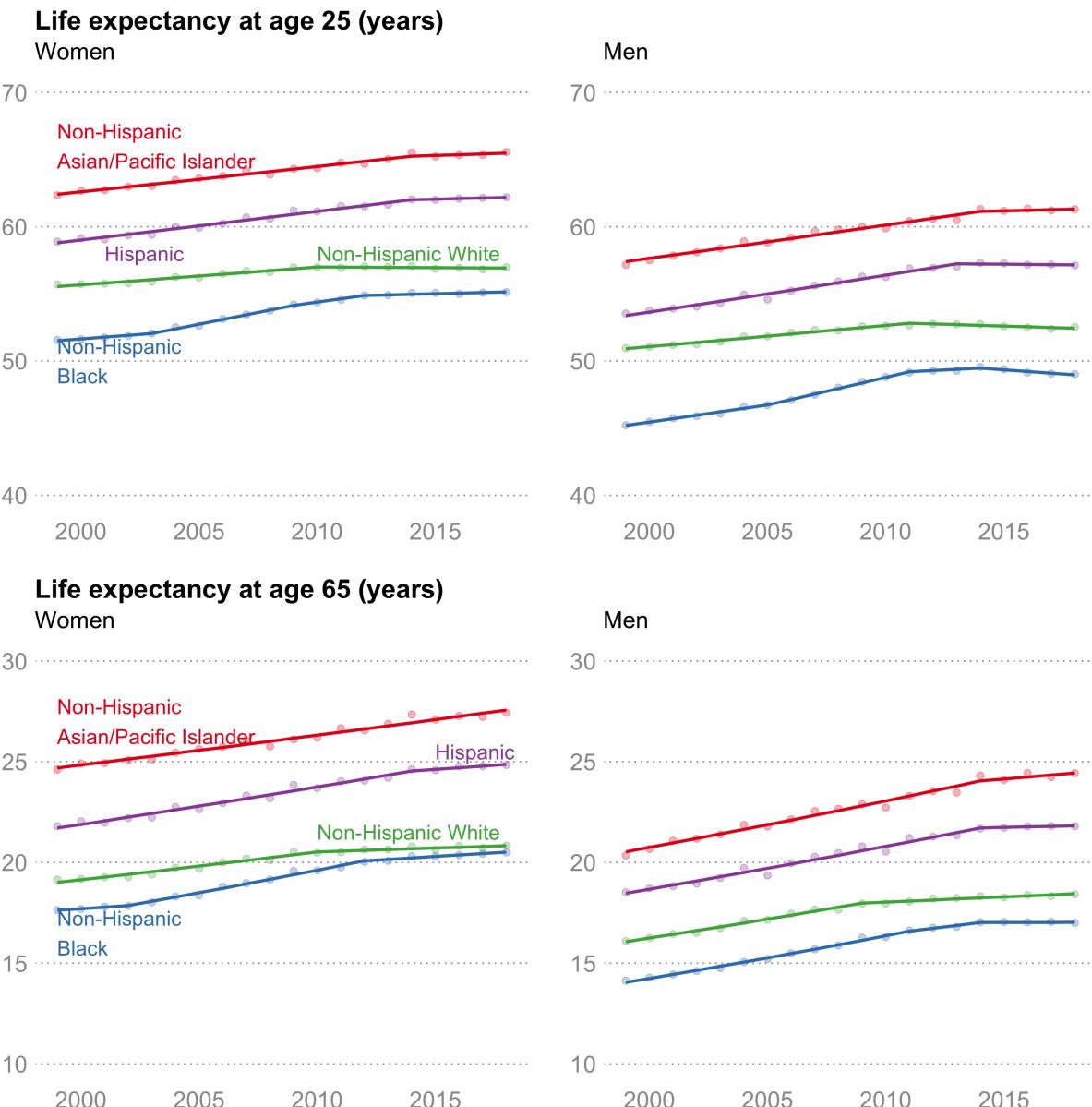


**Supplemental Figure 3.** Trends the gender gap in life expectancy by age and race-ethnicity, 1999-2018. Source: Authors' calculations of data from CDC WONDER (Centers for Disease Control and Prevention and National Center for Health Statistics 2020). Data: <https://osf.io/4s2rz/> Code: <https://osf.io/5xywp/>

### Gender gap (women - men) in life expectancy (years)

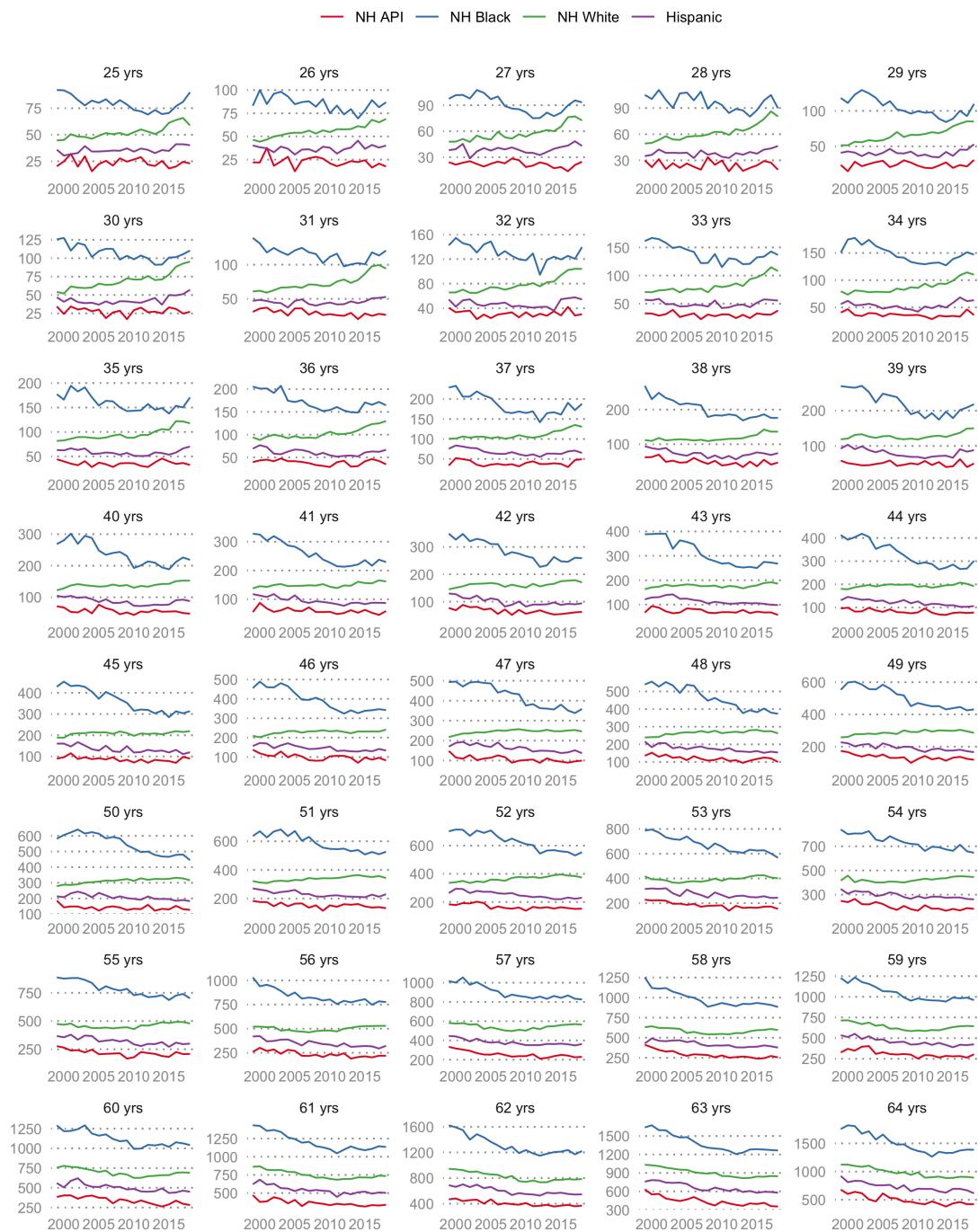


**Supplemental Figure 4.** Life expectancy at age 25 and age 65 in the United States, by gender and race-ethnicity, 1999-2018. Source: Authors' calculations of data from CDC WONDER (Centers for Disease Control and Prevention and National Center for Health Statistics 2020). Data: <https://osf.io/hz864/> <https://osf.io/4s2rz/> Code: <https://osf.io/hz864/>



**Supplemental Figure 5.** Mortality rates by single years of age by race-ethnicity for women, ages 25-64. Source: Authors' calculations of data from CDC WONDER (Centers for Disease Control and Prevention and National Center for Health Statistics 2020). Data: <https://osf.io/uvyx4/>, <https://osf.io/pytev/> Code: <https://osf.io/vb28p/>

### Death rates per 100,000 by single years of age for women, 1999-2018

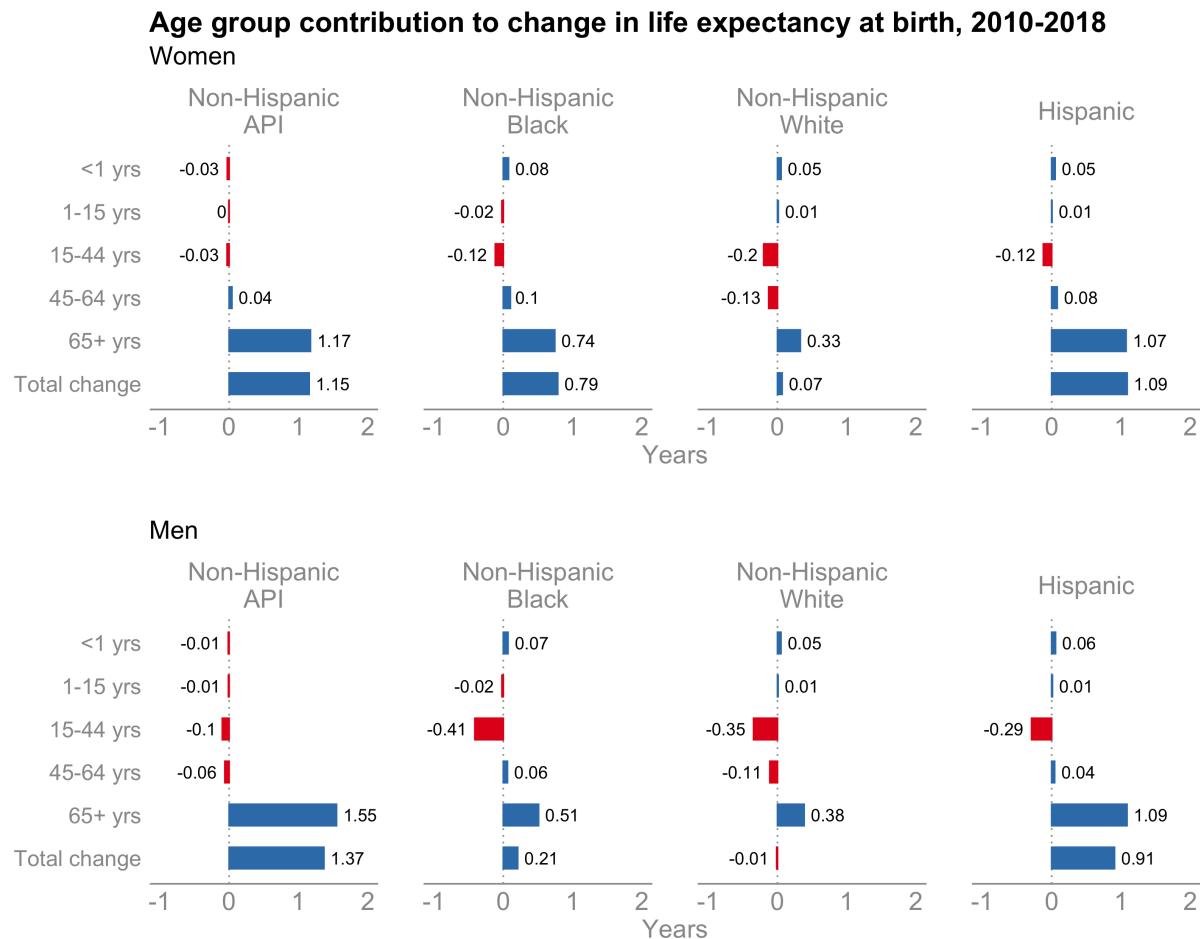


**Supplemental Figure 6.** Mortality rates by single years of age by race-ethnicity for men, ages 25-64. Source: Authors' calculations of data from CDC WONDER (Centers for Disease Control and Prevention and National Center for Health Statistics 2020). Data: <https://osf.io/uvyx4/>, <https://osf.io/pytev/> Code: <https://osf.io/vb28p/>

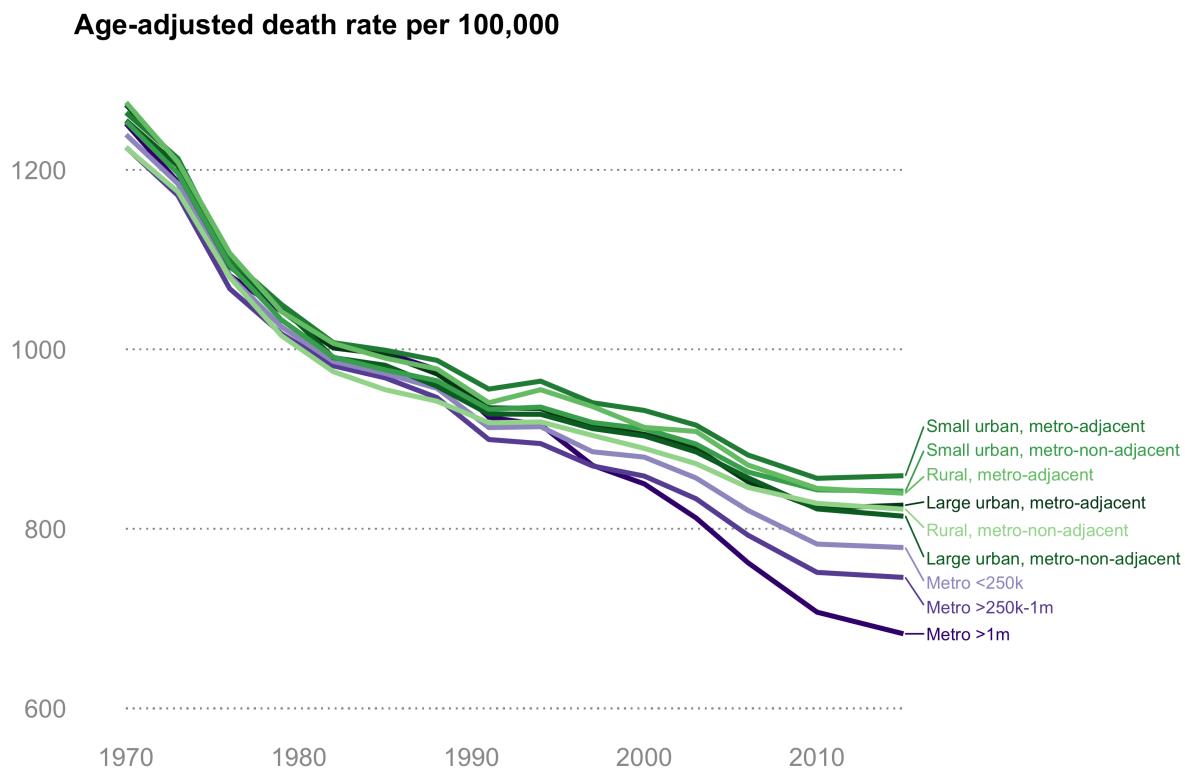
#### Death rates per 100,000 by single years of age for men, 1999-2018



**Supplemental Figure 7.** Age groups contributing to the change in life expectancy at birth in the United States between 2010 and 2018, by gender and race-ethnicity. Red color indicates age groups contributing a decline, blue color indicates age groups contributing an increase. Source: Authors' calculations of data from CDC WONDER (Centers for Disease Control and Prevention and National Center for Health Statistics 2020). Data: <https://osf.io/tk8q3/>, <https://osf.io/mctx3/>, <https://osf.io/utdnv/> Code: <https://osf.io/g9mp2/>, <https://osf.io/qd5w4/>

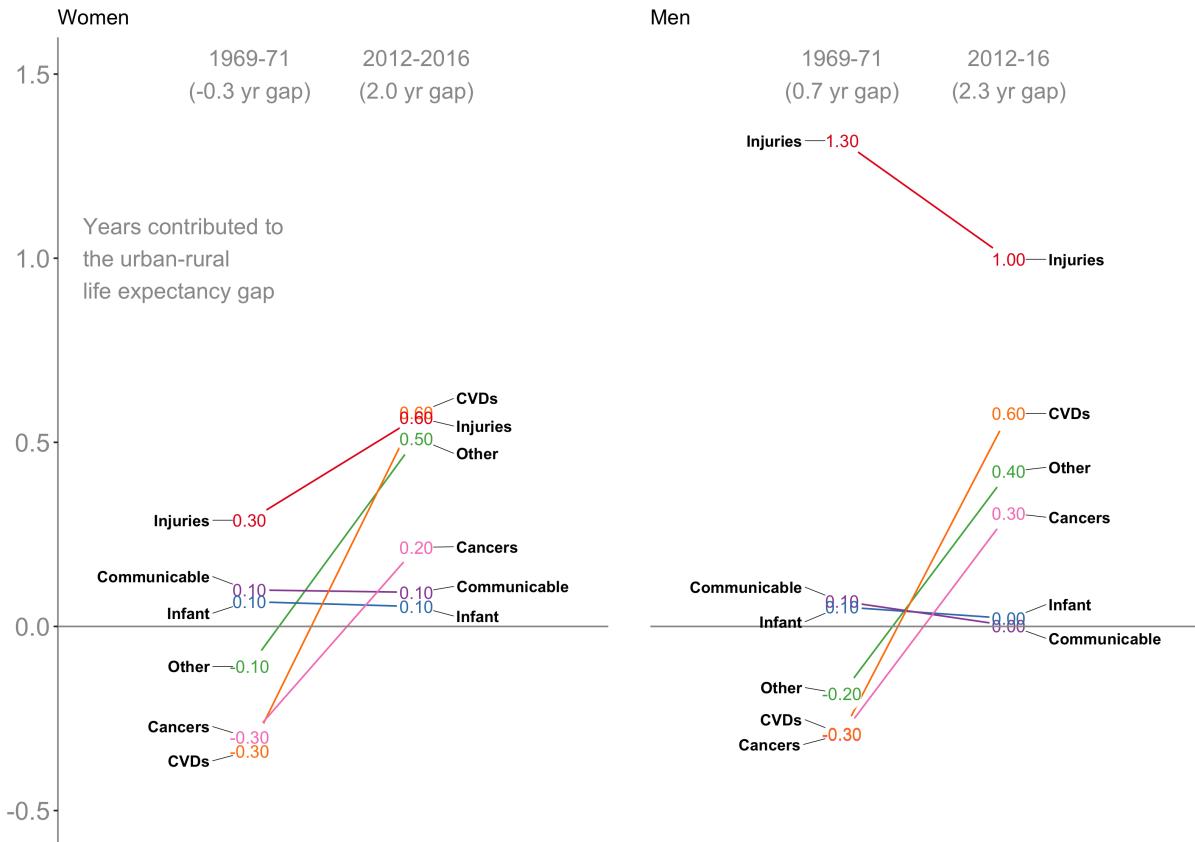


**Supplemental Figure 8.** Age-adjusted death rates per 100,000 by urban (purple) or rural (green) categorization, 1969–2016. Source: Authors' calculations of data from SEER\*Stat (National Cancer Institute 2019). Data: <https://osf.io/rj83e/> Code: <https://osf.io/b856t/>



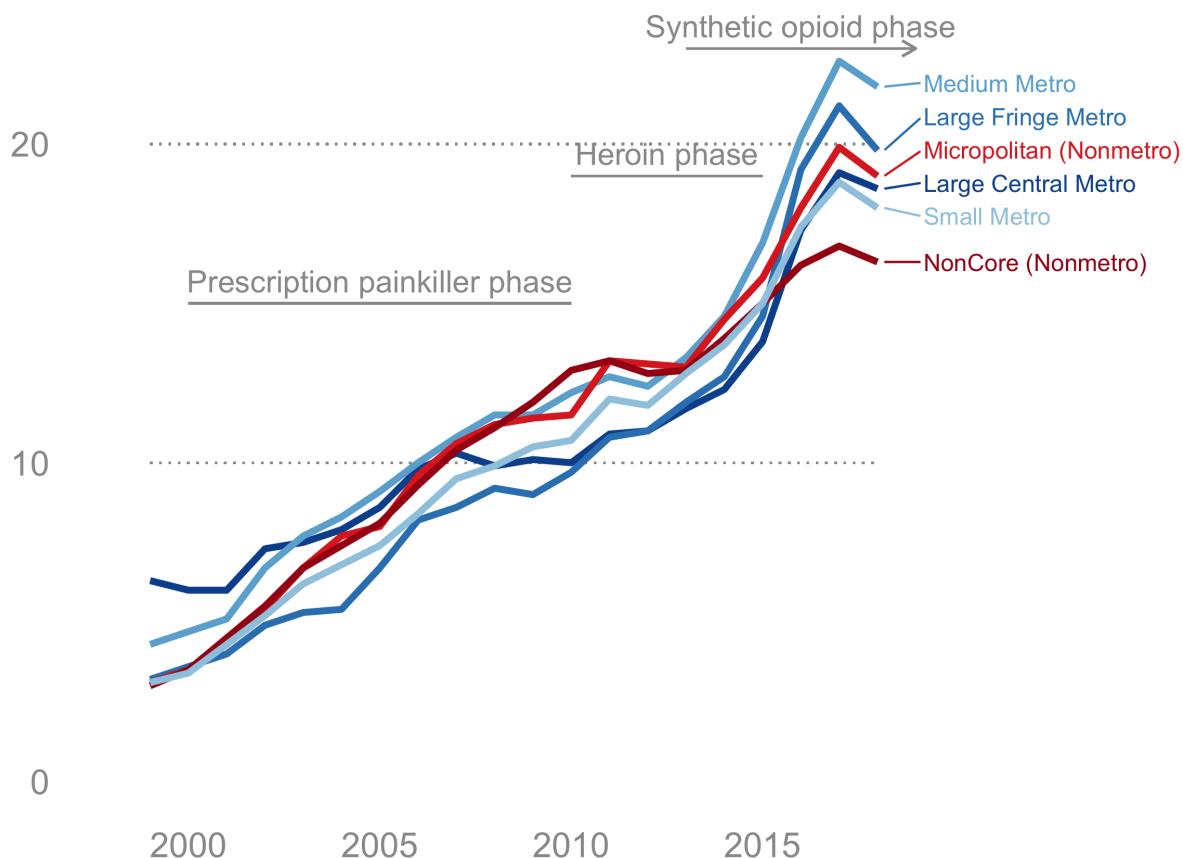
**Supplemental Figure 9.** Decomposition of the urban-rural life expectancy gap by cause-of-death, 1969-71 and 2012-16. Source: Authors' calculations of data from SEER\*Stat (National Cancer Institute 2019). Data: <https://osf.io/rj83e/>, <https://osf.io/rwc68/>, Code: <https://osf.io/4nezp/>, <https://osf.io/d23fz/>

#### Major causes of death contributing to urban-rural differences in life expectancy at birth



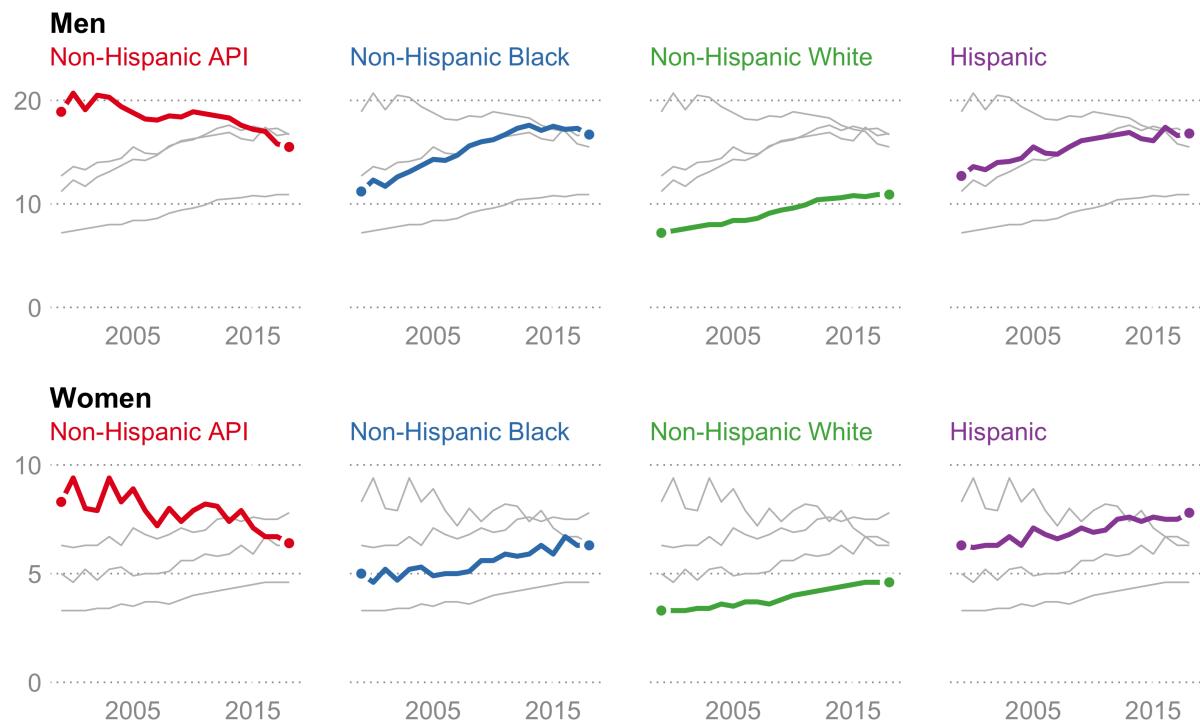
**Supplemental Figure 10.** Trends in age-adjusted unintentional poisoning death rates by urban-rural classification scheme, 1999-2018. Source: Authors' calculations of data from SEER\*Stat (National Cancer Institute 2019). Data: <https://osf.io/zsbfw/> Code: <https://osf.io/m7k2u/>

### Age-adjusted unintentional poisoning death rate per 100,000

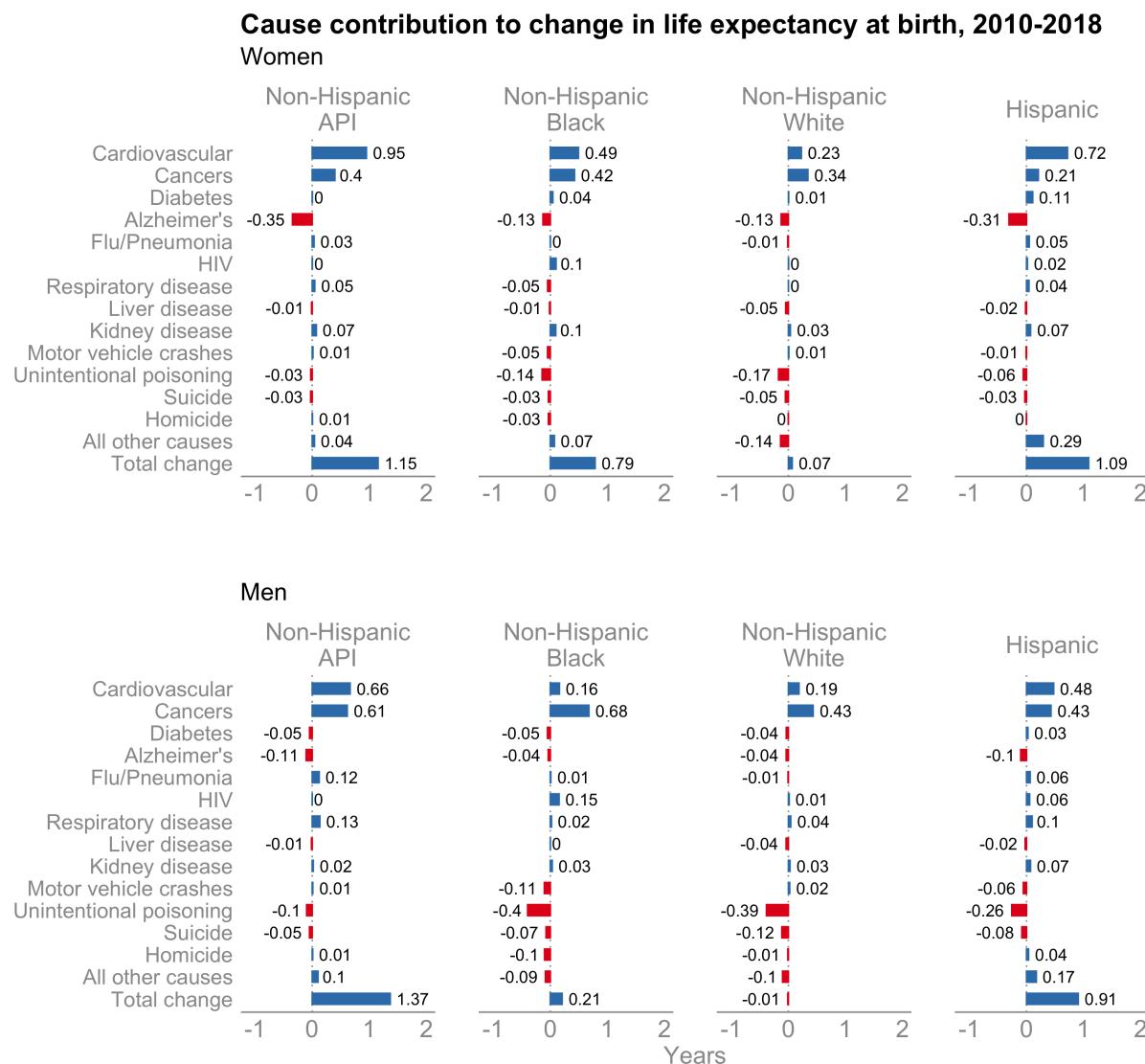


**Supplemental Figure 11.** Trends in age-adjusted liver cancer death rates by race-ethnicity and gender, 1999-2018. Source: Authors' calculations of data from CDC WONDER (Centers for Disease Control and Prevention and National Center for Health Statistics 2020). Data: <https://osf.io/stj6a/>, <https://osf.io/d5jkh/> Code: <https://osf.io/ybdmx/>

Age-adjusted death rates per 100,000 population for liver cancer



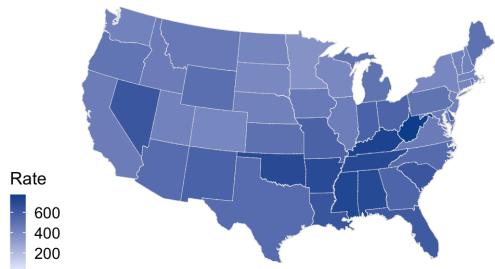
**Supplemental Figure 12.** Causes of death contributing to the change in life expectancy at birth in the United States between 2010 and 2018, by gender and race-ethnicity. Red color indicates causes contributing a decline, blue color indicates causes contributing an increase. Source: Authors' calculations of data from CDC WONDER (Centers for Disease Control and Prevention and National Center for Health Statistics 2020). Data: <https://osf.io/tk8q3/>, <https://osf.io/mctx3/>, <https://osf.io/utdnv/> Code: <https://osf.io/g9mp2/>, <https://osf.io/qd5w4/>



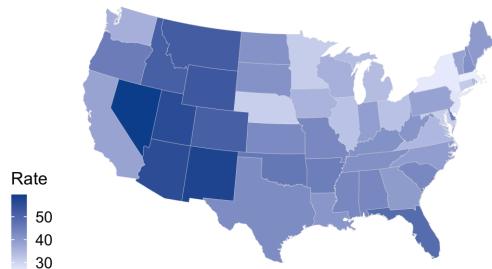
**Supplemental Figure 13.** Age-adjusted death rates per 100,000 for non-Hispanic white men aged 45-54, by state, 2010-2018. Source: Authors' calculations of data from CDC WONDER (Centers for Disease Control and Prevention and National Center for Health Statistics 2020). Data: <https://osf.io/jn5ts/>, <https://osf.io/vdjbp/>, <https://osf.io/xpz6g/>, <https://osf.io/w9ady/> Code: <https://osf.io/va4u7/>

Age-adjusted death rates per 100,000 for non-Hispanic white men aged 45-54, 2010-2018

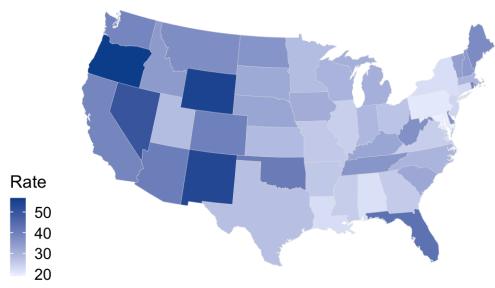
All causes



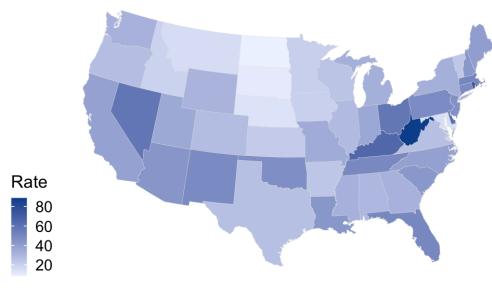
Suicides



Alcohol-related



Unintentional poisoning



## LITERATURE CITED

- Arias, Elizabeth, Melonie Heron, National Center for Health Statistics, Jahn Hakes, and US Census Bureau. 2016. “The Validity of Race and Hispanic-Origin Reporting on Death Certificates in the United States: An Update.” *Vital and Health Statistics. Series 2, Data Evaluation and Methods Research*, no. 172 (January): 1–21.
- Arias, Elizabeth, William S. Schauman, Karl Eschbach, Paul D. Sorlie, and Eric Backlund. 2008. “The Validity of Race and Hispanic Origin Reporting on Death Certificates in the United States.” *Vital and Health Statistics. Series 2, Data Evaluation and Methods Research*, no. 148 (October): 1–23.
- Arriaga, Eduardo E. 1989. “Changing Trends in Mortality Decline During the Last Decades.” In *Differential Mortality: Methodological Issues and Biosocial Factors*, edited by LT Ruzicka, GJ Wunsch, and P Kane, 105–29. Oxford, England: Clarendon Press.
- . 1984. “Measuring and Explaining the Change in Life Expectancies.” *Demography* 21 (1): 83–96. <https://doi.org/10.2307/2061029>.
- Centers for Disease Control and Prevention, and National Center for Health Statistics. 2020. “Underlying Cause of Death, 1999–2018 on CDC WONDER Online Database, Released in 2020. Data Are from the Multiple Cause of Death Files, 1999–2018.” <https://wonder.cdc.gov/ucd-icd10.html>.
- Kim, H. J., M. P. Fay, E. J. Feuer, and D. N. Midthune. 2000. “Permutation Tests for Joinpoint Regression with Applications to Cancer Rates.” *Statistics in Medicine* 19 (3): 335–51. [https://doi.org/10.1002/\(sici\)1097-0258\(20000215\)19:3%3C335:aid-sim336%3E3.0.co;2-z](https://doi.org/10.1002/(sici)1097-0258(20000215)19:3%3C335:aid-sim336%3E3.0.co;2-z).
- Kochanek, Kenneth D., Sherry L. Murphy, Jiaquan Xu, and Elizabeth Arias. 2019. “Deaths: Final Data for 2017.” *National Vital Statistics Reports: From the Centers for Disease Control and Prevention, National Center for Health Statistics, National Vital Statistics System* 68 (9): 1–77.
- National Cancer Institute. 2019. “SEER\*Stat Version 8.3.6. SEER\*Stat Database: Mortality - All COD, Aggregated with County, Total U.S. (1969–2017) <Katrina/Rita Population Adjustment> - Linked to County Attributes - Total U.S., 1969–2018 Counties.” Hyattsville, Maryland: National Cancer Institute, Surveillance Research Program.
- . 2020. “Joinpoint Regression Program, Version 4.8.0.1.” Hyattsville, Maryland: National Cancer Institute, Surveillance Research Program.
- Preston, Samuel H., Patrick Heuveline, and Michel Guillot. 2001. *Demography: Measuring and Modeling Population Processes*. Malden, Mass.: Blackwell.
- The World Bank Group. 2020. “World Development Indicators | DataBank.” <https://databank.worldbank.org/source/development-indicators>.
- Woolf, Steven H., and Heidi Schoomaker. 2019. “Life Expectancy and Mortality Rates in the United States, 1959–2017.” *JAMA* 322 (20): 1996–2016. <https://doi.org/10.1001/jama.2019.16932>.