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

Rachel Koehler, Zach Blundell, Zhaner Mou

Introduction

Motivation

The recent decline in life expectancy in the United States raises critical questions about the health and well-being of the population. Understanding the causes of this decline is essential for designing effective public health interventions and policies to reverse this declining trend.

Detailed Background

The paper examines the recent decline in life expectancy in the United States, focusing on the period from 2015 to 2017, when life expectancy dropped for three consecutive years. The decline has affected nearly all race/ethnic and gender groups, and is attributed to several factors, including a rise in opioid overdose deaths, suicide, homicide, and Alzheimer's disease, as well as a slowdown in the decline of cardiovascular disease mortality. These trends indicate the importance of addressing the underlying causes, including socioeconomic disparities and geographic inequalities.

Problem Statement

This article addresses the problem of declining life expectancy in the U.S. by examining the contributing factors across demographic groups, causes of death, and geographic regions.

Summary of Main Results

This article shows that the recent decline in U.S. life expectancy is primarily driven by increases in opioid overdose deaths, homicide, and Alzheimer's disease, with significant contributions from socioeconomic and geographic disparities, while also highlighting the role of stagnating declines in cardiovascular disease mortality.

Approach

Data

The study uses mortality data from the Centers for Disease Control and Prevention (CDC) WONDER database, which includes detailed cause-of-death information by age, gender, race/ethnicity, and geographic location from 1999 to 2018.

Methods

The authors employ descriptive analyses to examine trends in life expectancy and mortality rates across demographic groups (race/ethnicity, gender, and age). They provide a detailed breakdown of mortality trends by specific causes of death, such as unintentional poisonings, suicide, homicide, CVD, and Alzheimer's disease, and use age-adjusted death rates to compare trends across demographic groups, revealing significant disparities in how these causes affect different populations.

Results

FIGURE 1 PLOT AND CODE
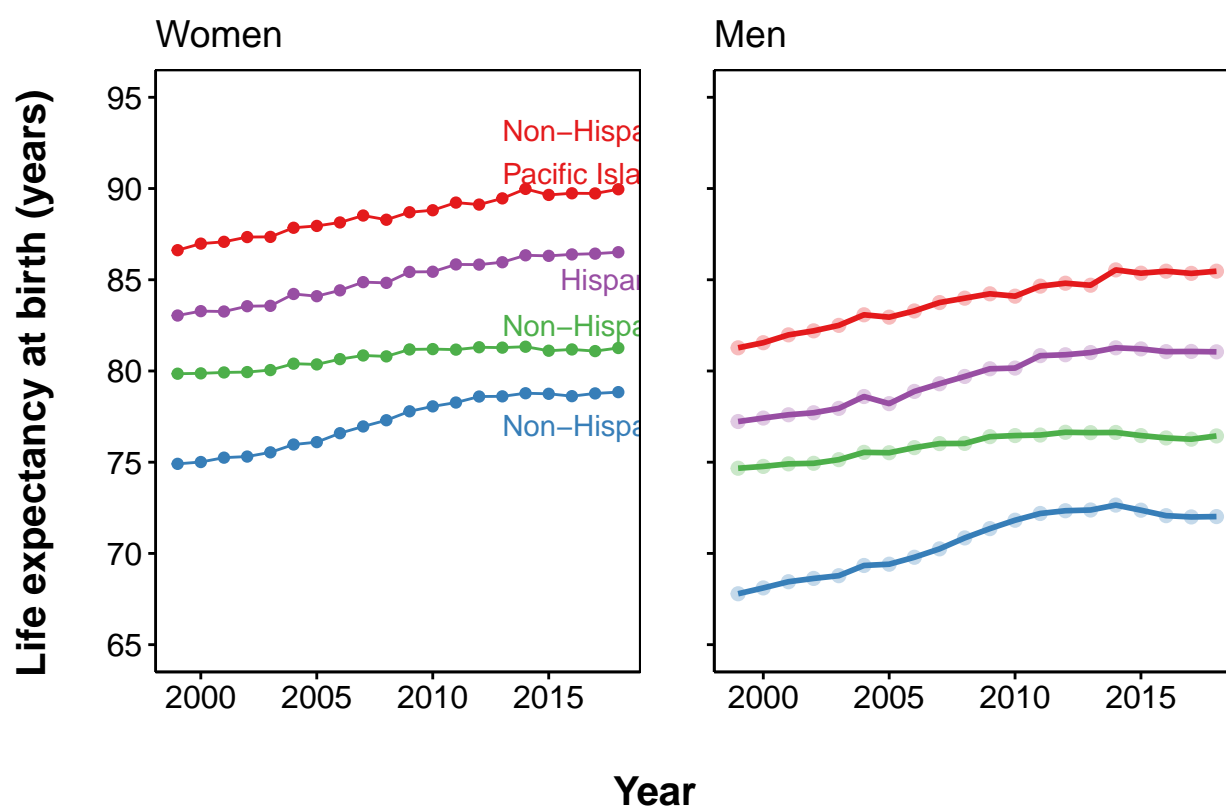
Life expectancy at birth in the United States, by gender and race/ethnicity, 1999–2018. Author's calculations of data from the Centers for Disease Control and Prevention (CDC) WONDER (13; data are from the Multiple Cause of Death Files, 1999-2018).

This plot shows the general trend that for every race/eth category given in the data, there is a general increase in life expectancy from 2000-2017. However, there is still a persistent gap between men and women for every race/eth category
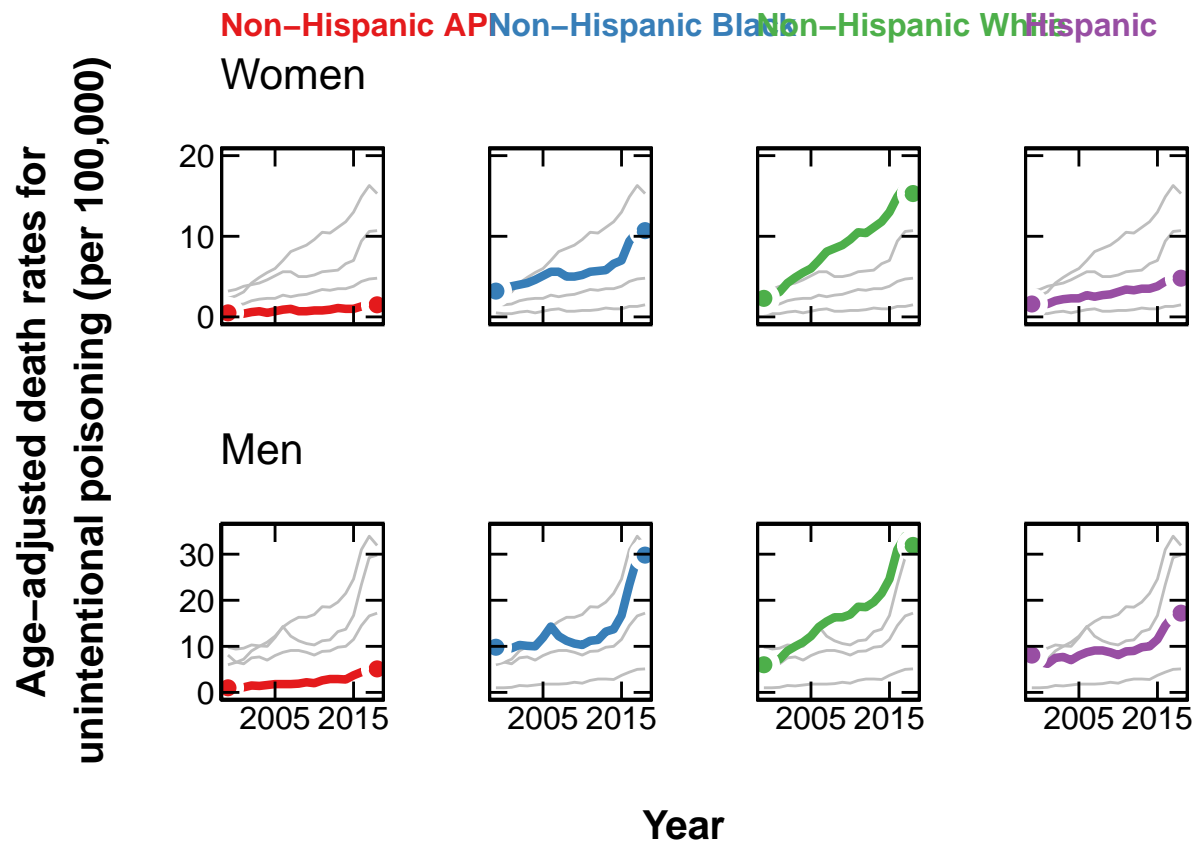
FIGURE 2 PLOT AND CODE
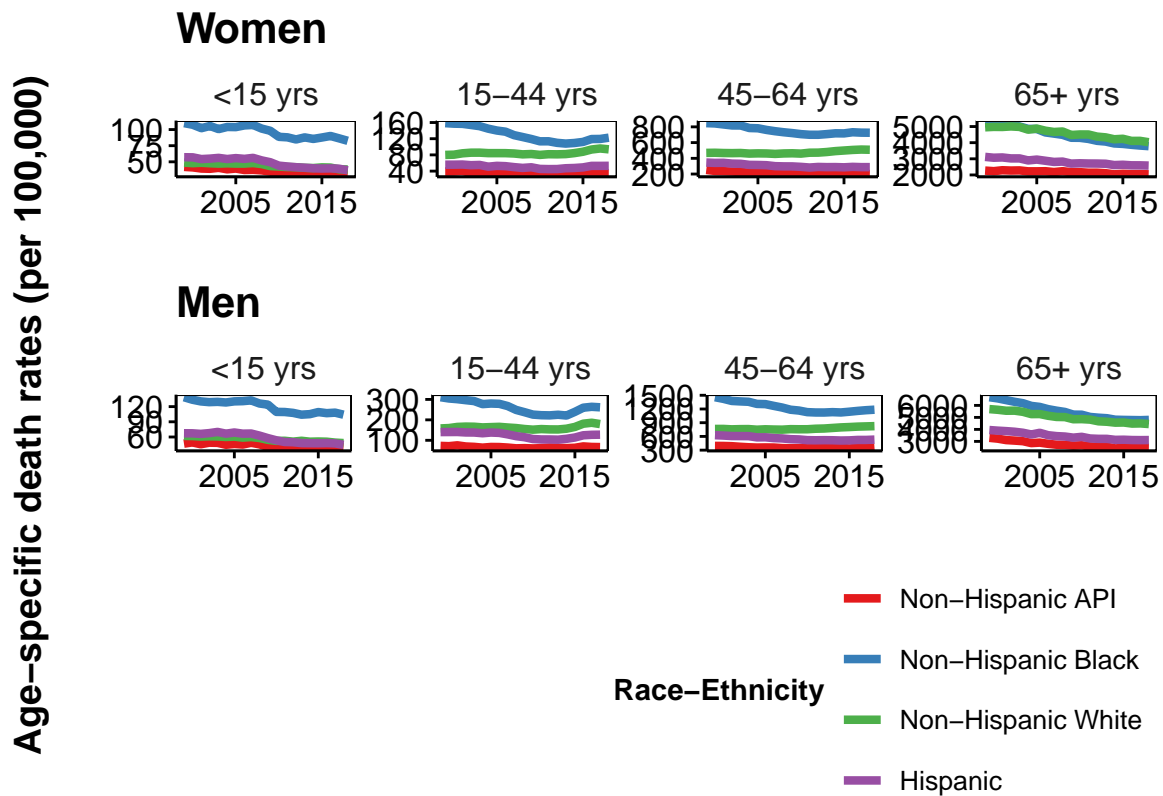
Age-specific mortality rates in the United States, by gender and race/ethnicity, 1999–2018. Data from CDC WONDER (13; data are from the Multiple Cause of Death Files, 1999-2018).

This plot shows how the age specific death rates vary across age ranges. From this we still see an overall decrease from the beginning of the time frame to the end. However, there a few groups that see an increase over the time frame particularly the 15-44 age group. The reasoning behind this will be explored through future figures.

FIGURE 4 PLOT AND CODE (passed over 3 due to needing STATA)

Trends in age-adjusted unintentional poisoning death rates, 1999–2018, by gender and race/ethnicity (ICD-10 codes X40–X49). Authors' calculations of data from CDC WONDER (13; data are from the Multiple Cause of Death Files, 1999–2018).

In this figure, we see that across every race/eth group, there is an increase in the age-adjusted death rates for unintentional poisoning over the time period (refering to mainly overdoses from medication/drugs). We can also see how some groups (Non-hispanic Black/White people) are the most effected and this effect does not vary much through gender, though Non-hispanic Black and Hispanic men see a greater increase than their respective female counterparts.

FIGURE 5 PLOT AND CODE
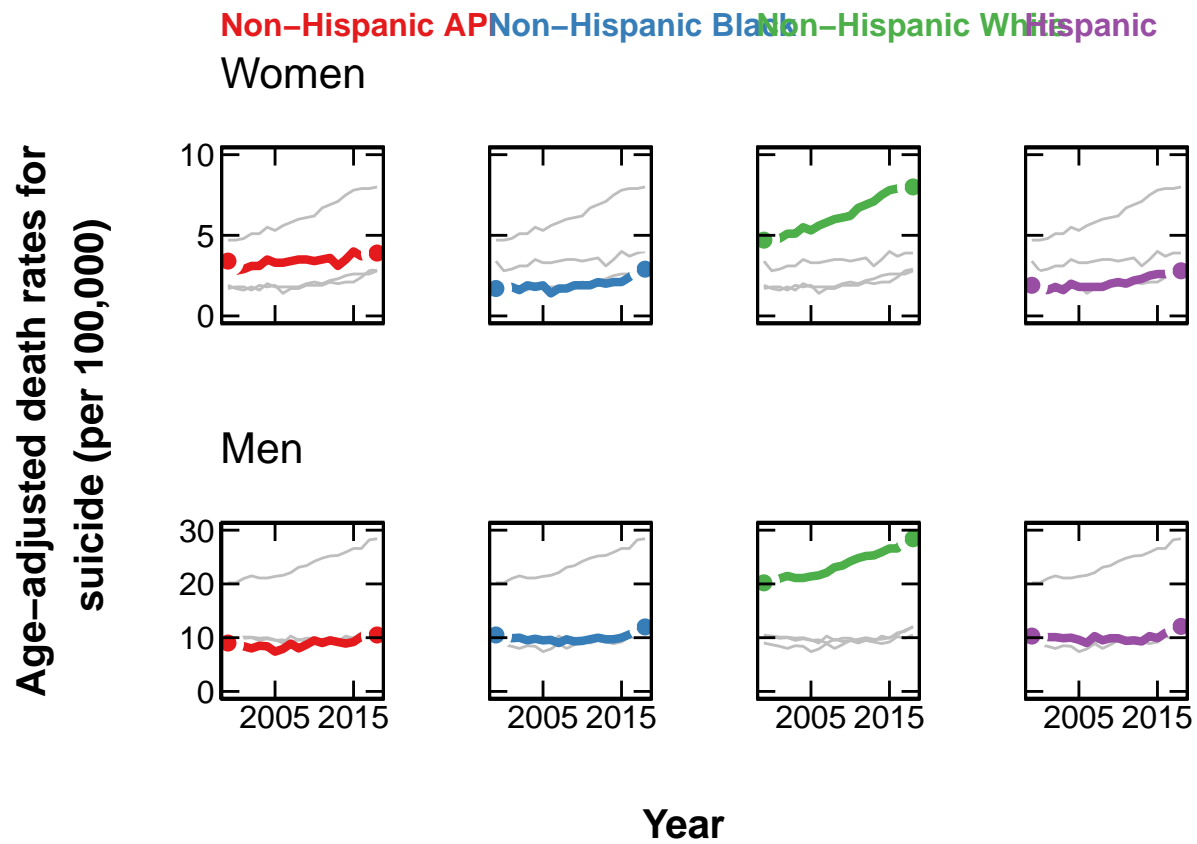
Trends in age-adjusted suicide death rates, 1999–2018, by gender and race/ethnicity (ICD-10 codes X60–X84). Authors' calculations of data from CDC WONDER (13; data are from the Multiple Cause of Death Files, 1999–2018).

For this plot we see that for age-adjusted suicide death rates, Non-hispanic White people see the greatest increase in the rate while having the highest baseline level in both men and women. In fact, we also see that for ever other race/eth category, the rate stays fairly level for the entire time period.

FIGURE 6 PLOT AND CODE

Women

Race−Ethnicity
- Non−Hispanic White
- Non−Hispanic Black
- Non−Hispanic API
- Hispanic
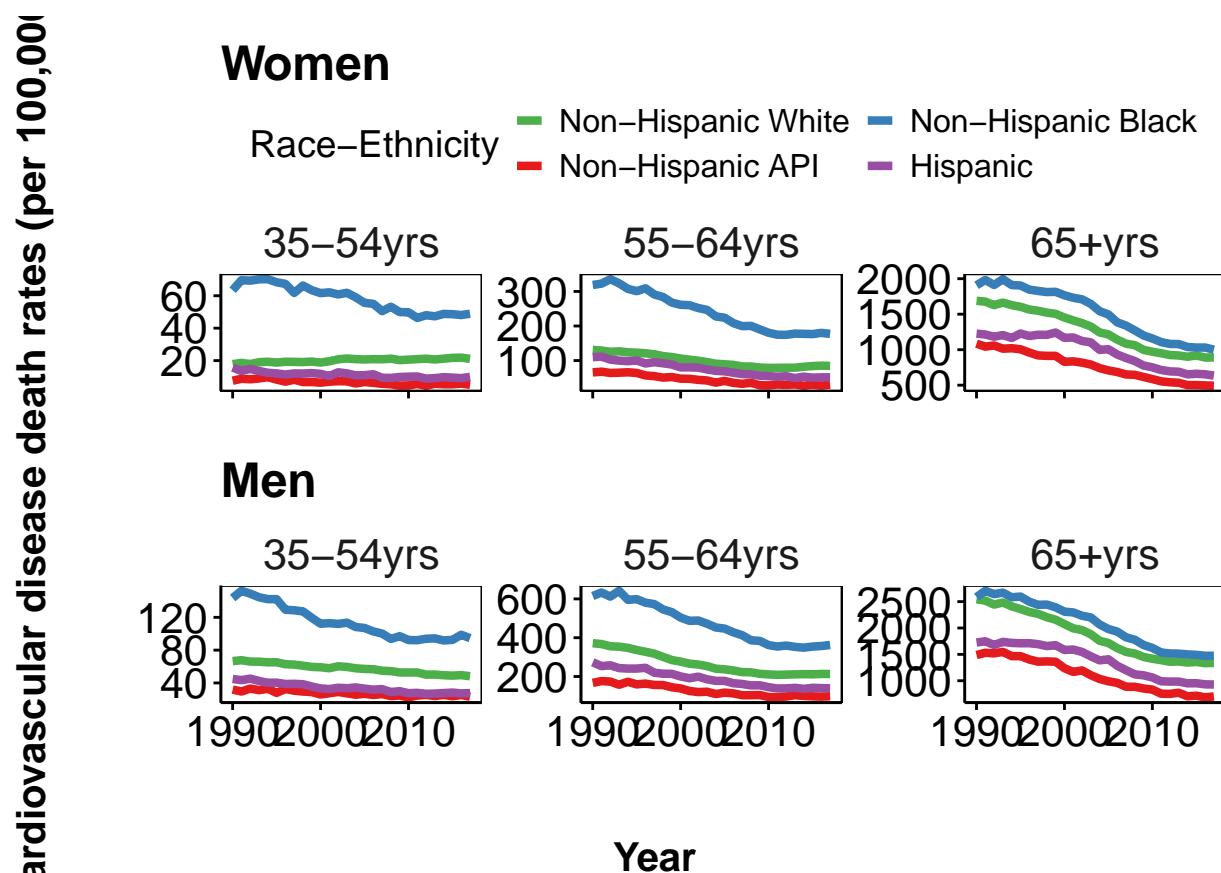
Age-adjusted cardiovascular disease death rates per 100,000, by age group, gender, and race/ethnicity, 1990–2017. Author's calculations of data from the National Center for Health Statistics (98).

For this plot, we can see the reason why there was a relatively large decrease in death rates (and thus increase in life expectancy) for every race/eth group as the rate of CVD deaths drops for pretty almost every race/eth group(with some exceptions) across age groups. We see the most notable difference in the oldest age group (they see higher incidence of CVD deaths so this is understandable). It's interesting to see the increasing rates across age groups purely from how the y-axis frame changes.

FIGURE 7 PLOT AND CODE
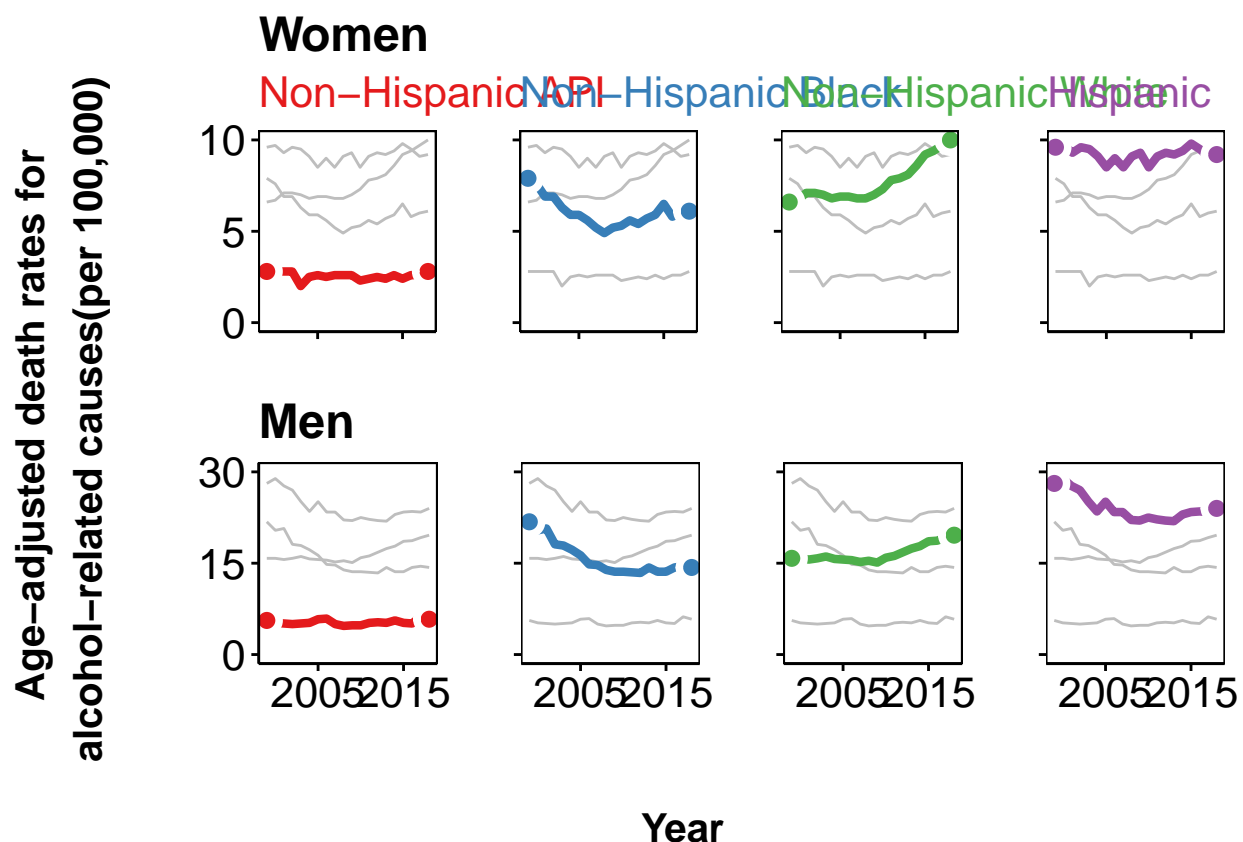
Age-adjusted alcohol-related death rates per 100,000, by age group, gender, and race/ethnicity, 1999–2018 [ICD-10 codes K70, K73, K47 (chronic liver disease and cirrhosis), F47 (alcohol use disorders)]. Authors' calculations of data from CDC WONDER (13; data are from the Multiple Cause of Death Files, 1999–2018).

Finally, we see in these graphs that death rates for alcohol related causes vary across race/eth groups, especially when looking at Non-hispanic Black and White populations. For Non-Hispanic Black people, we can actually see a decrease in rates as opposed to non-hispanic Whites who see an increase in rates over the time period.

Discussion

From the paper, we see that there is a general trend in increasing life expectancy and decreasing death rates, though there is a need to look deeper into the reasoning why this occurs and what factors are at play. There is a clear decrease in CVD related deaths across almost all age groups and ethnicity, but there has also been an increase in opiod related deaths, meaning the full picture is more complex than simply "people are living longer" and interventions should be targeted to reflect these complexities.

For one, the clear increase in the death rates from unintentional poisonings (opiods) across all groups is incredibly concerning and should be targeted by both policy and practitioners. Interventions can include needle exchange programs, increasing naloxone availability, increase people and provider awareness, increasing access to treatment, and de-stigmatizing drugs and their related treatments. Though there has been incredible progress in some areas, such as CVD, progress can still be made trough reducing cost of treatment, increasing access, improving lifestyle interventions, and reducing inequalities across groups. This same

attitude should be help across all factors for all groups: though great progress has been made in some areas, we can work to improve the health and livelihood of people and should take steps to do so in all areas.

Article Figures

```
## Rows: 1232 Columns: 13
## -- Column specification ----------------------------------------------------
## Delimiter: ","
## dbl (13): sex, race, age, cod, de, ie, te, mx2014, mx2017, pctgap, tdeaths20...
##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.
```

oution of Leading Causes of Death to Life Expectancy Decline

uring this period was for non–Hispanic Black men, primarily due to unintentio
which together accounted for more than 75% of their decline in life expecta
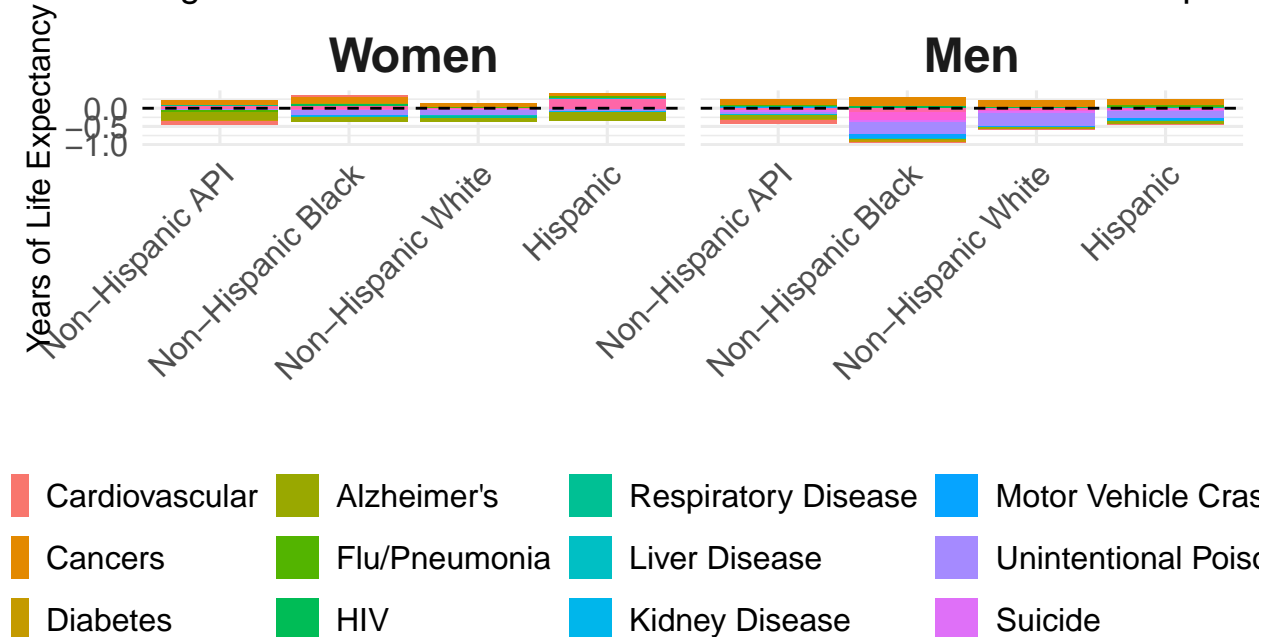
a Source: CDC Mortality Statistics (2014–2017). Negative values indicate a decline in life

We can see in this stacked bar chart of causes for all available race/eth groups in both men and women, Black men see the largest negative decrease in life expectancy from 2014-2017. From this graph, it's clear the main factor for this, and as covered in the paper is unintentional poisoning (opiods). The other large factors contributing to this is homicides and crashses. We analyze the death rate of this group further in the next plot

8

Death Rate Trends for Non–Hispanic Black Males (2000–2018

100,000 for Non–Hispanic Black males, showing a disproportionately high mortality burd
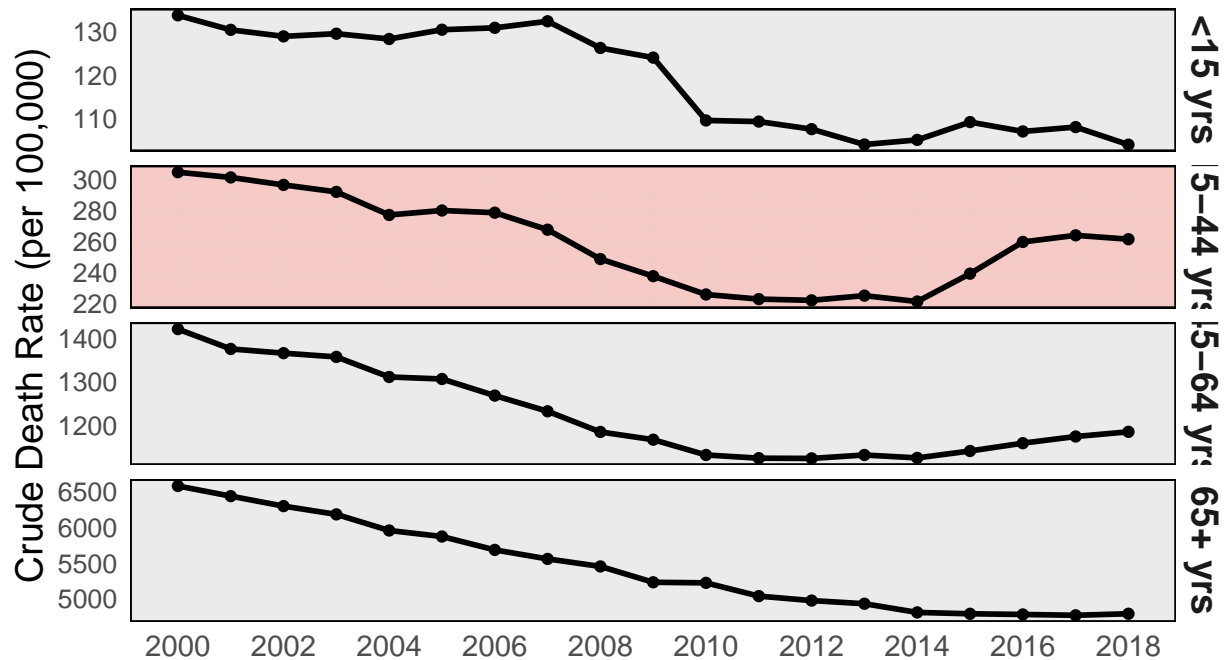
Source: Centers for Disease Control and Prevention (CDC) WONDER

From this graph, we can see that across all age groups from 2014-2017, which is what our previous graph was analyzing, there is either a stagnation in crude death rate OR a rather substantial increase, as seen in the 15-44 age group. This is rather concerning because we can see a rather steady decrease in the death rate in all age groups until 2014. This, along with the previous graph and the paper, really highlights how big of a concern opiods are in this community and that interventions should maybe prioritize this group considering how heavily they have been affected

Appendix: All code for this report

```
library(tidyverse)
library(plotly)
library(here)
library(ggrepel)
library(patchwork)
library(readr)
library(ggtext)
library(ggplot2)
library(gridExtra)
library(grid)
library(cowplot)
library(knitr)
########## FIGURE 1 ##########

figure1data <- read.delim("jp-le-age-sex-race.txt")
```

```r
# age as factor, split into different age ranges depending on the data coding from CDC WONDER
figure1data$age <- recode_factor(figure1data$age3, `1`= "At birth",
  `2`= "At age 25", `3`= "At age 65")

# race-ethnicity as factor, set based on race/eth code from 1 to 5
figure1data$raceeth <- recode_factor(figure1data$race, `1`= "Non-Hispanic AIAN",
  `2`= "Non-Hispanic API", `3`= "Non-Hispanic Black",
  `4`= "Non-Hispanic White", `5`= "Hispanic")

# gender-group, bring together the coding of the gender of the participant
figure1data$gender <- recode_factor(figure1data$sex, `1`= "Women", `2`= "Men")

#theme that matches paper, given by the code in paper, sometimes varies depending on the plots/figure n
stheme <- theme_classic() +
  theme(
    # Title and subtitle (smaller than before)
    plot.title = element_text(size = 16, face = "bold"),
    plot.subtitle = element_text(size = 14),

    # Axis text and titles (smaller sizes)
    axis.text.x = element_text(size = 12, colour = "black"),
    axis.text.y = element_text(size = 12, colour = "black"),
    axis.title.x = element_text(face = "bold", size = 12),   # bold x-axis title
    axis.title.y = element_text(face = "bold", size = 12),

    # Remove grid and add a border around the panel
    panel.grid = element_blank(),
    panel.border = element_rect(color = "black", fill = NA),

    # Show tick marks on both axes
    axis.ticks = element_line(color = "black"),

    # Strip elements (for facets)
    strip.text = element_text(size = 12),
    strip.background = element_rect(colour = "white"),

    # Remove the legend
    legend.position = "none"
  )


# life expectancy at birth - women
#Begin with Women, based off figure in paper,

womenfig1 <- subset(figure1data, age3==1 & sex==1 & race!=1)

w <- ggplot(womenfig1, aes(x = year0, y = ex, color = raceeth)) +
  geom_point() +
  geom_line() +
  labs(y = " ", x = "") +
  scale_y_continuous(
    limits = c(65, 95),
    breaks = seq(65, 95, by = 5)  # Explicit breaks every 5 years, again set by paper
```

```
  ) +
  scale_color_manual(values=c("#e41a1c", "#377eb8", "#4daf4a", "#984ea3")) +
 ggtitle(" ", subtitle="Women") +
  annotate("text", label = "Non-Hispanic Asian/\nPacific Islander",
           x = 2013, y = 92, size = 4, colour = "#e41a1c", hjust=0) +
  annotate("text", label = "Hispanic",
           x = 2015.5, y = 85, size = 4, colour = "#984ea3", hjust=0) +
  annotate("text", label = "Non-Hispanic White",
           x = 2013, y = 82.5, size = 4, colour = "#4daf4a", hjust=0) +
  annotate("text", label = "Non-Hispanic Black",
           x = 2013, y = 77, size = 4, colour = "#377eb8", hjust=0) +
  theme(
   axis.title.y = element_text(face = "bold", size = 15)
  ) +
  stheme
###Specific Race/ETH annotations to match how the paper distinguishes them on the plot

# life expectancy at birth - men

menfig1 <- subset(figure1data, age3==1 & sex==2 & race!=1)

m <- ggplot(menfig1, aes(x = year0, y = ex, color = raceeth)) +
  geom_point(alpha = 0.3, size = 2) +
  geom_line(size = 1) +
  labs(y = "", x = "") +
  scale_y_continuous(
    limits = c(65, 95),           # Match upper limit to women's plot
    breaks = seq(65, 95, by = 5)  # Same breaks as women's plot
  ) +
    scale_color_manual(values=c("#e41a1c", "#377eb8", "#4daf4a", "#984ea3")) +
 ggtitle(" ", subtitle="Men") +
  stheme +
  theme(
    axis.text.y = element_blank(),  # Remove y-axis numbers
  )

###Race/Eth coding from women plot not included (only needed for women since colors match)
#The limits and breaks for life expectency are set so the plots are alligned
#Only need numbers on y-axis of women plot since they line up

#combine genders to make Fig. 1
fig1_final_plot <- grid.arrange(
  w, m, ncol = 2,
  bottom = textGrob(
    "Year",
    gp = gpar(fontface = "bold", fontsize = 15)
  ),
  left = textGrob(
    "Life expectancy at birth (years)",
    gp = gpar(fontface = "bold", fontsize = 15),
    rot = 90
  )
)
```

```
########## FIGURE 2 ##########

figure2nhisp <- read.delim("asmr-sex-race-hisp-1999-2018.txt")
figure2hisp <- read.delim("asmr-sex-race-hisp-1999-2018.txt")

# Read and sort Non-Hispanic data. Keep demographics, deaths, population, and create crude rates

nhisp <- read_tsv(
  "asmr-sex-race-nhisp-1999-2018.txt",
  skip = 1,
  col_names = c("notes", "gender", "gcode", "race", "rcode", "age", "acode",
                "year", "ycode", "deaths", "pop", "crate"),
  n_max = 1760,
  col_types = "cccfccfddddd"
)

##Add in per 100,000 rates
nhisp2 <- nhisp %>%
  select(gender, race, acode, year, deaths, pop) %>%
  mutate(racen = as.numeric(race),
         rate = deaths / pop * 100000)

# read and sort Hispanic data. Keep demographics, deaths, population, and create crude rates
hisp <- read_tsv(
  "asmr-sex-race-hisp-1999-2018.txt",
  skip = 1,
  col_names = c("notes", "gender", "gcode", "age", "acode", "year", "ycode",
                "deaths", "pop", "crate"),
  n_max = 440,
  col_types = "ccccfddddd"
)

##Add in per 100,000 rates
hisp2 <- hisp %>%
  select(gender, acode, year, deaths, pop) %>%
  mutate(racen = 5,    # Hispanic group identifier
         rate = deaths / pop * 100000)

# Combine Hispanics and non-Hispanics
fig2_data_combined <- bind_rows(nhisp2, hisp2) %>%
  select(-race, -rate) %>%  # Remove race and preliminary rate
  mutate(
    racen = recode_factor(as.character(racen),
                          `1` = "Non-Hispanic AIAN",
                          `2` = "Non-Hispanic API",
                          `3` = "Non-Hispanic Black",
                          `4` = "Non-Hispanic White",
                          `5` = "Hispanic"),
    age4 = recode_factor(as.character(acode),
                          `1` = "<15 yrs",
                          `1-4` = "<15 yrs",
                          `5-14` = "<15 yrs",
                          `15-24` = "15-44 yrs",
```

```r
                              `25-34` = "15-44 yrs",
                              `35-44` = "15-44 yrs",
                              `45-54` = "45-64 yrs",
                              `55-64` = "45-64 yrs",
                              `65-74` = "65+ yrs",
                              `75-84` = "65+ yrs",
                              `85+` = "65+ yrs")
  ) %>%
  group_by(gender, racen, age4, year) %>%
  select(-acode) %>%  # acode is no longer needed after recoding
  summarise(across(c(deaths, pop), sum), .groups = "drop") %>%  # sum over the groups
  mutate(rate = deaths / pop * 100000) %>%
  filter(racen != "Non-Hispanic AIAN")  # Exclude AIAN if needed

# Define a shared theme with the legend positioned at the bottom
theme_with_legend <- theme_classic() +
  theme(
    plot.title = element_text(size = 16, face = "bold"),
    plot.subtitle = element_text(size = 14),
    axis.text.x = element_text(size = 12, colour = "black"),
    axis.text.y = element_text(size = 12, colour = "black"),
    axis.title.x = element_text(face = "bold", size = 12),
    axis.title.y = element_text(face = "bold", size = 14),
    panel.grid = element_blank(),
    panel.border = element_rect(color = "black", fill = NA),
    axis.ticks = element_line(color = "black"),
    strip.text = element_text(size = 12),
    strip.background = element_rect(colour = "white"),
    legend.position = "bottomright",  # Set legend to bottom
    legend.justification = c(1, 0),  # Align legend to bottom right
    legend.title = element_text(face = "bold"),
    legend.text = element_text(size = 10),
    legend.background = element_blank(),  # Remove legend box
    legend.key = element_blank()  # Remove legend key background
  )

# Women plot (w2)
F2w2 <- ggplot(subset(fig2_data_combined, gender == "Female"),
            aes(x = year, y = rate, colour = racen)) +
  geom_line(size = 1.5) +
  facet_wrap(~ age4, nrow = 1, scales = "free") +
  scale_color_manual(name = "Race-Ethnicity",
                     values = c("#e41a1c", "#377eb8", "#4daf4a", "#984ea3")) +
  guides(color = guide_legend(ncol = 1)) +  # Stack legend vertically
  labs(y = "", x = " ") +
  ggtitle("Women") +
  scale_x_continuous(breaks = c(2005, 2015)) +
  theme_with_legend +
  theme(strip.text = element_text(face = "plain"))

# Men plot (m2)
F2m2 <- ggplot(subset(fig2_data_combined, gender == "Male"),
            aes(x = year, y = rate, colour = racen)) +
```

```r
  geom_line(size = 1.5) +
  facet_wrap(~ age4, nrow = 1, scales = "free") +
  scale_color_manual(name = "Race-Ethnicity",
                     values = c("#e41a1c", "#377eb8", "#4daf4a", "#984ea3")) +
  guides(color = guide_legend(ncol = 1)) +  # Stack legend vertically
  labs(y = "", x = " ") +
  ggtitle("Men") +
  scale_x_continuous(breaks = c(2005, 2015)) +
  theme_with_legend +
  theme(strip.text = element_text(face = "plain"))

# Combine the plots and correctly collect the legend at the bottom
fig2_final_plot <- ggdraw() +
  draw_plot(
    (F2w2 / F2m2) +
      plot_layout(guides = "collect") &
      theme(legend.position = "bottom", legend.justification = "right"),  # Align legend to the right w
    x = 0.04, y = 0, width = 0.92, height = 1
  ) +
  draw_label(
    "Age-specific death rates (per 100,000)",
    x = 0.01, y = 0.5, angle = 90,
    vjust = 1.5, fontface = "bold", size = 14
  )

# Display the final plot
plot(fig2_final_plot)
########## FIGURE 4 ##########

### STEP1: data processing

# upload datasets: non-Hispanics & Hispanics
up <- read_tsv("upoison-race-aadr-1999-2018.txt", skip=1,
               col_names=c("notes", "gender", "gcode", "race", "rcode", "year", "ycode",
                           "deaths", "pop", "crate", "aadr"), n_max=160,
               col_types = "cccfcdcdddd")
uph <- read_tsv("upoison-hisp-aadr-1999-2018.txt", skip=1,
                col_names=c("notes", "gender", "gcode", "year", "ycode",
                            "deaths", "pop", "crate", "aadr"), n_max=40,
                col_types = "cccdcdddd")

# keep demographics, deaths, population, create rates
up2 <- select(up, gender, race, year, deaths, pop, aadr) %>%
  mutate(racen = as.numeric(race), rate = deaths / pop * 100000)
uph2 <- select(uph, gender, year, deaths, pop, aadr) %>%
  mutate(racen = 5, rate = deaths / pop * 100000)

# put Hispanics and non-Hispanics together
ur <- bind_rows(up2, uph2) %>%
  select(-race) %>%
  mutate(racen = recode_factor(racen, `1`= "Non-Hispanic\nAIAN",
                               `2`= "Non-Hispanic\nAPI", `3`= "Non-Hispanic\nBlack",
                               `4`= "Non-Hispanic\nWhite", `5`= "Hispanic")) %>%
```

```r
    filter(racen != "Non-Hispanic\nAIAN")

### STEP2: plot

# modify plot characteristics
stheme <- theme_classic() +
  theme(plot.title = element_text(size = 12, face = "bold"),
        plot.subtitle = element_text(size=12)) +
  theme(axis.text.x = element_text(size = 12, colour = "black"),
        axis.title.y = element_text(size=12, angle=90, colour="black", face="bold"),
        axis.text.y = element_text(size = 12, colour="black"),
        legend.position="none",
        panel.grid.major = element_blank(),
        panel.grid.minor = element_blank(),
        panel.border = element_rect(colour = "black", fill = NA, linewidth=1)) +
  theme(axis.line.x=element_line(colour="black"),
        axis.line.y=element_line(colour="black"),
        axis.ticks = element_line(colour="black"),
        strip.text = element_text(size = 16),
        strip.background = element_rect(colour="white")) +
  theme(axis.ticks.length=unit(-0.25, "cm"))


# 1) plot for men

# Non-Hispanic API menx
nhac <- "#e41a1c"
amp <- ggplot(subset(ur, gender=="Male"),
              aes(x=year, y=aadr, group=racen))  +
  geom_line(show.legend=F, colour="grey") +
  geom_line(data=subset(ur, gender=="Male" & racen=="Non-Hispanic\nAPI"),
            colour=nhac, size=1.5) +
  geom_point(data=subset(ur, gender=="Male" & (year==1999 | year==2018) &
                         racen=="Non-Hispanic\nAPI"), shape=21, colour="white",
            fill=nhac, size=3, stroke=2) +
  scale_y_continuous(limits = c(0, 35), breaks = c(0, 10, 20, 30), sec.axis=dup_axis()) +
  scale_x_continuous(breaks=c(2005,2015), sec.axis=dup_axis()) +
  xlab("") + ylab("") + stheme +
  ggtitle("", subtitle="Men") +
  theme(plot.subtitle = element_text(size=16),
        axis.line.y.right = element_blank(),
        axis.text.y.right = element_blank(),
        axis.line.x.top = element_line(colour = "black"),
        axis.ticks.x.top = element_line(colour = "black"),
        axis.text.x.top = element_blank())

# Non-Hispanic Black men
nhbc <- "#377eb8"
bmp <- ggplot(subset(ur, gender=="Male"),
              aes(x=year, y=aadr, group=racen))  +
  geom_line(show.legend=F, colour="grey") +
  geom_line(data=subset(ur, gender=="Male" & racen=="Non-Hispanic\nBlack"),
            colour=nhbc, size=1.5) +
```

```r
  geom_point(data=subset(ur, gender=="Male" & (year==1999 | year==2018) &
                          racen=="Non-Hispanic\nBlack"), shape=21, colour="white",
          fill=nhbc, size=3, stroke=2) +
  scale_y_continuous(limits = c(0, 35), breaks = c(0, 10, 20, 30), sec.axis=dup_axis()) +
  scale_x_continuous(breaks=c(2005,2015), sec.axis=dup_axis()) +
  xlab("") + ylab("") + stheme +
  theme(axis.line.y = element_blank(),
        axis.text.y = element_blank(),
        axis.line.x.top = element_line(colour = "black"),
        axis.ticks.x.top = element_line(colour = "black"),
        axis.text.x.top = element_blank())

# Non-Hispanic White men
nhwc <- "#4daf4a"
wmp <- ggplot(subset(ur, gender=="Male"),
              aes(x=year, y=aadr, group=racen))  +
  geom_line(show.legend=F, colour="grey") +
  geom_line(data=subset(ur, gender=="Male" & racen=="Non-Hispanic\nWhite"),
          colour=nhwc, size=1.5) +
  geom_point(data=subset(ur, gender=="Male" & (year==1999 | year==2018) &
                          racen=="Non-Hispanic\nWhite"), shape=21, colour="white",
          fill=nhwc, size=3, stroke=2) +
  scale_y_continuous(limits = c(0, 35), breaks = c(0, 10, 20, 30), sec.axis=dup_axis()) +
  scale_x_continuous(breaks=c(2005,2015), sec.axis=dup_axis()) +
  xlab("") + ylab("") + stheme +
  theme(axis.line.y = element_blank(),
        axis.text.y = element_blank(),
        axis.line.x.top = element_line(colour = "black"),
        axis.ticks.x.top = element_line(colour = "black"),
        axis.text.x.top = element_blank())

# Hispanic men
hc <- "#984ea3"
hmp <- ggplot(subset(ur, gender=="Male"),
              aes(x=year, y=aadr, group=racen))  +
  geom_line(show.legend=F, colour="grey") +
  geom_line(data=subset(ur, gender=="Male" & racen=="Hispanic"),
          colour=hc, size=1.5) +
  geom_point(data=subset(ur, gender=="Male" & (year==1999 | year==2018) &
                          racen=="Hispanic"), shape=21, colour="white",
          fill=hc, size=3, stroke=2) +
  scale_y_continuous(limits = c(0, 35), breaks = c(0, 10, 20, 30), sec.axis=dup_axis()) +
  scale_x_continuous(breaks=c(2005,2015), sec.axis=dup_axis()) +
  xlab("") + ylab("") + stheme +
  theme(axis.line.y = element_blank(),
        axis.text.y = element_blank(),
        axis.line.x.top = element_line(colour = "black"),
        axis.ticks.x.top = element_line(colour = "black"),
        axis.text.x.top = element_blank())

# plots for men together
mp <- (amp | bmp | wmp | hmp)
```

```r
# 2) plot for women

# Non-Hispanic API women
nhac <- "#e41a1c"
awp <- ggplot(subset(ur, gender=="Female"),
              aes(x=year, y=aadr, group=racen))  +
  geom_line(show.legend=F, colour="grey") +
  geom_line(data=subset(ur, gender=="Female" & racen=="Non-Hispanic\nAPI"),
            colour=nhac, size=1.5) +
  geom_point(data=subset(ur, gender=="Female" & (year==1999 | year==2018) &
                           racen=="Non-Hispanic\nAPI"), shape=21, colour="white",
             fill=nhac, size=3, stroke=2) +
  scale_y_continuous(limits=c(0,20), breaks=c(0,10,20), sec.axis=dup_axis()) +
  scale_x_continuous(breaks=c(2005,2015), sec.axis=dup_axis()) +
  xlab("") + ylab("") + stheme +
  ggtitle("Non-Hispanic API", subtitle="Women") +
  theme(plot.title = element_text(colour=nhac),
        plot.subtitle = element_text(size=16),
        axis.line.y = element_blank(),
        axis.text.x = element_blank(),
        axis.text.y.right = element_blank())

# Non-Hispanic Black women
nhbc <- "#377eb8"
bwp <- ggplot(subset(ur, gender=="Female"),
              aes(x=year, y=aadr, group=racen))  +
  geom_line(show.legend=F, colour="grey") +
  geom_line(data=subset(ur, gender=="Female" & racen=="Non-Hispanic\nBlack"),
            colour=nhbc, size=1.5) +
  geom_point(data=subset(ur, gender=="Female" & (year==1999 | year==2018) &
                           racen=="Non-Hispanic\nBlack"), shape=21, colour="white",
             fill=nhbc, size=3, stroke=2) +
  scale_y_continuous(limits=c(0,20), breaks=c(0,10,20), sec.axis=dup_axis()) +
  scale_x_continuous(breaks=c(2005,2015), sec.axis=dup_axis()) +
  xlab("") + ylab("") + stheme +
  ggtitle("Non-Hispanic Black", subtitle="") +
  theme(plot.title = element_text(colour=nhbc),
        axis.text.x = element_blank(),
        axis.line.y = element_blank(),
        axis.text.y = element_blank())

# Non-Hispanic White women
nhwc <- "#4daf4a"
wwp <- ggplot(subset(ur, gender=="Female"),
              aes(x=year, y=aadr, group=racen))  +
  geom_line(show.legend=F, colour="grey") +
  geom_line(data=subset(ur, gender=="Female" & racen=="Non-Hispanic\nWhite"),
            colour=nhwc, size=1.5) +
  geom_point(data=subset(ur, gender=="Female" & (year==1999 | year==2018) &
                           racen=="Non-Hispanic\nWhite"), shape=21, colour="white",
             fill=nhwc, size=3, stroke=2) +
  scale_y_continuous(limits=c(0,20), breaks=c(0,10,20), sec.axis=dup_axis()) +
  scale_x_continuous(breaks=c(2005,2015), sec.axis=dup_axis()) +
```

```r
  xlab("") + ylab("") + stheme +
  ggtitle("Non-Hispanic White", subtitle="") +
  theme(plot.title = element_text(colour=nhwc),
        axis.text.x = element_blank(),
        axis.line.y = element_blank(),
        axis.text.y = element_blank())

# Hispanic women
hc <- "#984ea3"
hwp <- ggplot(subset(ur, gender=="Female"),
              aes(x=year, y=aadr, group=racen))  +
  geom_line(show.legend=F, colour="grey") +
  geom_line(data=subset(ur, gender=="Female" & racen=="Hispanic"),
            colour=hc, size=1.5) +
  geom_point(data=subset(ur, gender=="Female" & (year==1999 | year==2018) &
                              racen=="Hispanic"), shape=21, colour="white",
             fill=hc, size=3, stroke=2) +
  scale_y_continuous(limits=c(0,20), breaks=c(0,10,20), sec.axis=dup_axis()) +
  scale_x_continuous(breaks=c(2005,2015), sec.axis=dup_axis()) +
  xlab("") + ylab("") + stheme +
  ggtitle("Hispanic", subtitle="") +
  theme(plot.title = element_text(colour=hc),
        axis.text.x = element_blank(),
        axis.line.y = element_blank(),
        axis.text.y = element_blank())

# put plots for women together
wp <- (awp | bwp | wwp | hwp)

# now put men and women's plots together and label
F4_p <- wp / mp + plot_layout(widths = rep(10, 4), heights = rep(4, 2))
y.grob <- textGrob("Age-adjusted death rates for \n unintentional poisoning (per 100,000)",
                   gp=gpar(fontface="bold", fontsize=15), rot=90)
x.grob <- textGrob("Year",
                   gp=gpar(fontface="bold",  fontsize=15))
grid.arrange(patchworkGrob(F4_p), left = y.grob, bottom = x.grob)
########## FIGURE 5 ##########

### STEP1: data processing

# upload datasets: (age-adjusted) suicide rates from CDC WONDER
s <- read_tsv("suic-race-aadr-1999-2018.txt", skip=1,
  col_names=c("notes", "gender", "gcode", "race", "rcode", "year", "ycode",
            "deaths", "pop", "crate", "aadr"), n_max=160,
  col_types = "cccfcdcddddd")
sh <- read_tsv("suic-hisp-aadr-1999-2018.txt", skip=1,
              col_names=c("notes", "gender", "gcode", "year", "ycode",
                        "deaths", "pop", "crate", "aadr"), n_max=40,
              col_types = "cccdcdddd")

# keep demographics, deaths, population, create rates
s2 <- select(s, gender, race, year, deaths, pop, aadr) %>%
  mutate(racen = as.numeric(race), rate = deaths / pop * 100000)
```

```r
sh2 <- select(sh, gender, year, deaths, pop, aadr) %>%
  mutate(racen = 5, rate = deaths / pop * 100000)

# put Hispanics and non-Hispanics together
sr <- bind_rows(s2, sh2) %>%
  select(-race) %>%
  mutate(racen = recode_factor(racen, `1`= "Non-Hispanic\nAIAN",
        `2`= "Non-Hispanic\nAPI", `3`= "Non-Hispanic\nBlack",
        `4`= "Non-Hispanic\nWhite", `5`= "Hispanic")) %>%
  filter(racen != "Non-Hispanic\nAIAN")



### STEP2: plot

# modify plot characteristics

##### 2  #####
##### set some plot characteristics
stheme <- theme_classic() +
  theme(plot.title = element_text(size = 12, face = "bold"),
        plot.subtitle = element_text(size=12)) +
  theme(axis.text.x = element_text(size = 12, colour = "black"),
        axis.title.y = element_text(size=12, angle=90, colour="black", face="bold"),
        axis.text.y = element_text(size = 12, colour="black"),
        legend.position="none",
        panel.grid.major = element_blank(),
        panel.grid.minor = element_blank(),
        panel.border = element_rect(colour = "black", fill = NA, linewidth=1)) +
  theme(axis.line.x=element_line(colour="black"),
        axis.line.y=element_line(colour="black"),
        axis.ticks = element_line(colour="black"),
        strip.text = element_text(size = 16),
        strip.background = element_rect(colour="white")) +
  theme(axis.ticks.length=unit(-0.25, "cm"))

# 1) plot for men
###FOR ALL MEN, ENSURE THAT THE Y-AXIS MARKS LINE UP (They have the same limits and breaks)
##For each race/eth of men/women, ensure they line up for years (same breaks)
# Non-Hispanic API men
nhac <- "#e41a1c"
amp <- ggplot(subset(sr, gender=="Male"),
              aes(x=year, y=aadr, group=racen))  +
  geom_line(show.legend=F, colour="grey") +
  geom_line(data=subset(sr, gender=="Male" & racen=="Non-Hispanic\nAPI"),
            colour=nhac, size=1.5) +
  geom_point(data=subset(sr, gender=="Male" & (year==1999 | year==2018) &
                            racen=="Non-Hispanic\nAPI"), shape=21,
             colour="white", fill=nhac, size=3, stroke=2) +
  scale_y_continuous(limits=c(0,30), breaks=c(0,10,20,30), sec.axis=dup_axis()) +
  scale_x_continuous(breaks=c(2005,2015), sec.axis=dup_axis()) +
  xlab("") + ylab("") + stheme +
  ggtitle("", subtitle="Men") +
  theme(plot.subtitle = element_text(size=16),
```

```r
        axis.line.y.right = element_blank(),
        axis.text.y.right = element_blank(),
        axis.line.x.top = element_line(colour = "black"),
        axis.ticks.x.top = element_line(colour = "black"),
        axis.text.x.top = element_blank())


# Non-Hispanic Black men
nhbc <- "#377eb8"
bmp <- ggplot(subset(sr, gender=="Male"),
      aes(x=year, y=aadr, group=racen))  +
  geom_line(show.legend=F, colour="grey") +
  geom_line(data=subset(sr, gender=="Male" & racen=="Non-Hispanic\nBlack"),
          colour=nhbc, size=1.5) +
  geom_point(data=subset(sr, gender=="Male" & (year==1999 | year==2018) &
                           racen=="Non-Hispanic\nBlack"), shape=21,
          colour="white", fill=nhbc, size=3, stroke=2) +
  scale_y_continuous(limits=c(0,30), breaks=c(0,10,20,30), sec.axis=dup_axis()) +
  scale_x_continuous(breaks=c(2005,2015), sec.axis=dup_axis()) +
  xlab("") + ylab("") + stheme +
  theme(axis.line.y = element_blank(),
        axis.text.y = element_blank(),
        axis.line.x.top = element_line(colour = "black"),
        axis.ticks.x.top = element_line(colour = "black"),
        axis.text.x.top = element_blank())


# Non-Hispanic White men
nhwc <- "#4daf4a"
wmp <- ggplot(subset(sr, gender=="Male"),
      aes(x=year, y=aadr, group=racen))  +
  geom_line(show.legend=F, colour="grey") +
  geom_line(data=subset(sr, gender=="Male" & racen=="Non-Hispanic\nWhite"),
          colour=nhwc, size=1.5) +
  geom_point(data=subset(sr, gender=="Male" & (year==1999 | year==2018) &
                           racen=="Non-Hispanic\nWhite"), shape=21,
          colour="white", fill=nhwc, size=3, stroke=2) +
  scale_y_continuous(limits=c(0,30), breaks=c(0,10,20,30), sec.axis=dup_axis()) +
  scale_x_continuous(breaks=c(2005,2015), sec.axis=dup_axis()) +
  xlab("") + ylab("") + stheme +
  theme(axis.line.y = element_blank(),
        axis.text.y = element_blank(),
        axis.line.x.top = element_line(colour = "black"),
        axis.ticks.x.top = element_line(colour = "black"),
        axis.text.x.top = element_blank())


# Hispanic men
hc <- "#984ea3"
hmp <- ggplot(subset(sr, gender=="Male"),
      aes(x=year, y=aadr, group=racen))  +
  geom_line(show.legend=F, colour="grey") +
  geom_line(data=subset(sr, gender=="Male" & racen=="Hispanic"),
          colour=hc, size=1.5) +
  geom_point(data=subset(sr, gender=="Male" & (year==1999 | year==2018) &
                           racen=="Hispanic"), shape=21,
```

```
                  colour="white", fill=hc, size=3, stroke=2) +
      scale_y_continuous(limits=c(0,30), breaks=c(0,10,20,30), sec.axis=dup_axis()) +
      scale_x_continuous(breaks=c(2005,2015), sec.axis=dup_axis()) +
      xlab("") + ylab("") + stheme +
      theme(axis.line.y = element_blank(),
            axis.text.y = element_blank(),
            axis.line.x.top = element_line(colour = "black"),
            axis.ticks.x.top = element_line(colour = "black"),
            axis.text.x.top = element_blank())

# put plots for men together
mp <- (amp | bmp | wmp | hmp)

# 2) plot for women
###FOR ALL WOMEN, ENSURE Y-AXIS LIMITS AND BREAKS ARE THE SAME, ONLY NEED LABELING ON LEFT MOST PLOT
# Non-Hispanic API women
nhac <- "#e41a1c"
awp <- ggplot(subset(sr, gender=="Female"),
         aes(x=year, y=aadr, group=racen)) +
    geom_line(show.legend=F, colour="grey") +
    geom_line(data=subset(sr, gender=="Female" & racen=="Non-Hispanic\nAPI"),
              colour=nhac, size=1.5) +
    geom_point(data=subset(sr, gender=="Female" & (year==1999 | year==2018) &
                            racen=="Non-Hispanic\nAPI"), shape=21,
               colour="white", fill=nhac, size=3, stroke=2) +
    scale_y_continuous(limits=c(0,10), breaks=c(0,5,10), sec.axis=dup_axis()) +
    scale_x_continuous(breaks=c(2005,2015), sec.axis=dup_axis()) +
    xlab("") + ylab("") + stheme +
    ggtitle("Non-Hispanic API", subtitle="Women") +
    theme(plot.title = element_text(colour=nhac),
          plot.subtitle = element_text(size=16),
          axis.line.y = element_blank(),
          axis.text.x = element_blank(),
          axis.text.y.right = element_blank())

# Non-Hispanic Black women
nhbc <- "#377eb8"
bwp <- ggplot(subset(sr, gender=="Female"),
         aes(x=year, y=aadr, group=racen)) +
    geom_line(show.legend=F, colour="grey") +
    geom_line(data=subset(sr, gender=="Female" & racen=="Non-Hispanic\nBlack"),
              colour=nhbc, size=1.5) +
    geom_point(data=subset(sr, gender=="Female" & (year==1999 | year==2018) &
                            racen=="Non-Hispanic\nBlack"), shape=21,
               colour="white", fill=nhbc, size=3, stroke=2) +
    scale_y_continuous(limits=c(0,10), breaks=c(0,5,10), sec.axis=dup_axis()) +
    scale_x_continuous(breaks=c(2005,2015), sec.axis=dup_axis()) +
    xlab("") + ylab("") + stheme +
    ggtitle("Non-Hispanic Black", subtitle="") +
    theme(plot.title = element_text(colour=nhbc),
          axis.text.x = element_blank(),
          axis.line.y = element_blank(),
          axis.text.y = element_blank())
```

```r
# Non-Hispanic White women
nhwc <- "#4daf4a"
wwp <- ggplot(subset(sr, gender=="Female"),
       aes(x=year, y=aadr, group=racen))  +
  geom_line(show.legend=F, colour="grey") +
  geom_line(data=subset(sr, gender=="Female" & racen=="Non-Hispanic\nWhite"),
            colour=nhwc, size=1.5) +
  geom_point(data=subset(sr, gender=="Female" & (year==1999 | year==2018) &
                          racen=="Non-Hispanic\nWhite"), shape=21,
             colour="white", fill=nhwc, size=3, stroke=2) +
  scale_y_continuous(limits=c(0,10), breaks=c(0,5,10), sec.axis=dup_axis()) +
  scale_x_continuous(breaks=c(2005,2015), sec.axis=dup_axis()) +
  xlab("") + ylab("") + stheme +
  ggtitle("Non-Hispanic White", subtitle="") +
  theme(plot.title = element_text(colour=nhwc),
        axis.text.x = element_blank(),
        axis.line.y = element_blank(),
        axis.text.y = element_blank())

# Hispanic women
hc <- "#984ea3"
hwp <- ggplot(subset(sr, gender=="Female"),
       aes(x=year, y=aadr, group=racen))  +
  geom_line(show.legend=F, colour="grey") +
  geom_line(data=subset(sr, gender=="Female" & racen=="Hispanic"),
            colour=hc, size=1.5) +
  geom_point(data=subset(sr, gender=="Female" & (year==1999 | year==2018) &
                          racen=="Hispanic"), shape=21,
             colour="white", fill=hc, size=3, stroke=2) +
  scale_y_continuous(limits=c(0,10), breaks=c(0,5,10), sec.axis=dup_axis()) +
  scale_x_continuous(breaks=c(2005,2015), sec.axis=dup_axis()) +
  xlab("") + ylab("") + stheme +
  ggtitle("Hispanic", subtitle="") +
  theme(plot.title = element_text(colour=hc),
        axis.text.x = element_blank(),
        axis.line.y = element_blank(),
        axis.text.y = element_blank())

# put plots for women together
wp <- (awp | bwp | wwp | hwp)

##Combine together the plots and add in labels for x and y axis
F5_p <- wp / mp + plot_layout(widths = rep(40, 4), heights = rep(4, 2))
y.grob <- textGrob("Age-adjusted death rates for \nsuicide (per 100,000)",
                   gp=gpar(fontface="bold", fontsize=15), rot=90)
x.grob <- textGrob("Year",
                   gp=gpar(fontface="bold",  fontsize=15))
grid.arrange(patchworkGrob(F5_p), left = y.grob, bottom = x.grob)
########## FIGURE 6 ##########

#### Deaths from CVD (age-adjusted within broad age groups)
##### from SEER*Stat
```

```r
# Raw data, read out demographics,year, and the cvd outcomes alpong with count/pop

raw <- read_tsv("aaasdr-cvd-1990-2017.txt",
                col_names=c("sex", "raceeth", "age4", "year", "aadr", "count", "pop"), col_types = "dd

# rescale year (since it beings at 0 instead of 1990)
raw$year <- raw$year + 1990

# race-ethnicity as factor, delineated by 0-4 in the CDC WONDER data
raw$raceethf <- recode_factor(raw$raceeth, `0`= "Non-Hispanic White", `1`= "Non-Hispanic Black", `2`= "

# age-group, 4 seperate groups, final plot only uses 3
raw$age4f <- recode_factor(raw$age4, `0`= "15-34yrs", `1`= "35-54yrs", `2`= "55-64yrs", `3`= "65+yrs")

###Set custom theme of plots that matches paper (given in code)
stheme <- theme_classic() +
  theme(
    plot.title = element_text(size = 18, face = "bold"),
    plot.subtitle = element_text(size=16),
    axis.text.x = element_text(size = 16, colour = "black"),
    axis.title.y = element_text(size=16, angle=90, colour="black"),
    axis.text.y = element_text(size = 16, colour="black"),
    legend.position="none",
    panel.grid = element_blank(),  # Remove all grid lines
    axis.line = element_line(colour = "black"),  # Visible axis lines
    axis.ticks = element_line(colour = "black"),  # Ensure ticks are visible
    strip.text = element_text(size = 16),
    strip.background = element_rect(colour = "white"),
    panel.border = element_rect(colour = "black", fill = NA, size = 0.5)  # Border around each panel
  )

# Plot for Men (Bottom - Keep x-axis), only keep age groups that are >35
m <- ggplot(subset(raw, sex == 0 & age4>0 & raceeth!=2),
       aes(x = year, y = aadr, colour = as.factor(raceethf))) +
  geom_line(size=1.5) +
  facet_wrap(~ age4f, nrow=1, scales = "free_y") +

  scale_color_manual(name="Race-Ethnicity",
    values=c("#4daf4a","#377eb8", "#e41a1c", "#984ea3")) +
  labs(y = "",  # Y-axis label
       x = "") +  # X-axis label
  ggtitle("Men") +
  stheme +
  theme(
    axis.title.y = element_text(size = 18, hjust = -0.01, colour = "black"),  # Centered and larger
    axis.title.x = element_text(size = 16, colour = "black")  # Increased size
  )

# Plot for Women (Top - Remove x-axis elements), only keep age groups that are >35
w <- ggplot(subset(raw, sex == 1 & age4>0 & raceeth!=2),
       aes(x = year, y = aadr, colour = as.factor(raceethf))) +
  geom_line(size=1.5) +
  facet_wrap(~ age4f, nrow=1, scales = "free_y") +
```

```r
    scale_color_manual(name="Race-Ethnicity",
      values=c("#4daf4a","#377eb8", "#e41a1c", "#984ea3")) +
    labs(y = "", x = "") +
    ggtitle("Women") +
    stheme +
    theme(
      axis.text.x = element_blank(),
      axis.line.x = element_blank(),
      legend.position = "top",  # Position legend at the top
      legend.title = element_text(size=14),  # Adjust legend title size
      legend.text = element_text(size=12),   # Adjust legend text size
      legend.key.size = unit(0.8, "lines"),  # Adjust legend key size
      legend.margin = margin(t=0, unit="cm") # Reduce top margin
    ) +
    guides(color = guide_legend(nrow=2, byrow=TRUE))  # Arrange legend in two rows

###Combine the plots and add in titles of each respective axis
f6p <- w / m + plot_layout(widths = rep(40, 4), heights = rep(4, 2))
y.grob <- textGrob("Cardiovascular disease death rates (per 100,000)",
                   gp=gpar(fontface="bold", fontsize=15), rot=90)
x.grob <- textGrob("Year",
                   gp=gpar(fontface="bold",  fontsize=15))
grid.arrange(patchworkGrob(f6p), left = y.grob, bottom = x.grob)
########## FIGURE 7 ##########

#Pull out raw data

# age-adjusted rates for non-Hispanics
# cirrhosis and chronic liver dx, pull out the year, crude deaths, code, and demographic
e <- read_tsv("liver-race-aadr-1999-2018.txt", skip=1,
  col_names=c("notes", "gender", "gcode", "race", "rcode", "year", "ycode",
              "deaths", "pop", "crate", "aadr"), n_max=160,
  col_types = "cccfcdcdddd")

# keep demographics, deaths, population, create crude rates
e2 <- select(e, gender, race, year, deaths, pop, aadr) %>%
  mutate(racen = as.numeric(race), rate = deaths / pop * 100000)

# age-adjusted rates for Hispanics, same process as above
eh <- read_tsv("liver-hisp-aadr-1999-2018.txt", skip=1,
  col_names=c("notes", "gender", "gcode", "year", "ycode",
              "deaths", "pop", "crate", "aadr"), n_max=40,
  col_types = "cccdcdddd")

# keep demographics, deaths, population, create crude rates
eh2 <- select(eh, gender, year, deaths, pop, aadr) %>%
  mutate(racen = 5, rate = deaths / pop * 100000)

# put Hispanics and non-Hispanics together
er <- bind_rows(e2, eh2) %>%
  select(-race) %>%
  mutate(racen = recode_factor(racen, `1`= "Non-Hispanic\nAIAN",
        `2`= "Non-Hispanic\nAPI", `3`= "Non-Hispanic\nBlack",
```

```r
        `4`= "Non-Hispanic\nWhite", `5`= "Hispanic")) %>%
  filter(racen != "Non-Hispanic\nAIAN")



##Theme of paper taken from published code
stheme <- theme_classic() +
  theme(
    plot.title = element_text(size = 18, face = "bold"),
    plot.subtitle = element_text(size=16),
    axis.text.x = element_text(size = 16, colour = "black"),
    axis.title.y = element_text(size=16, angle=90, colour="black"),
    axis.text.y = element_text(size = 16, colour="black"),
    legend.position="none",
    panel.grid = element_blank(),  # Remove all grid lines
    axis.line = element_line(colour = "black"),  # Visible axis lines
    axis.ticks = element_line(colour = "black"),  # Ensure ticks are visible
    strip.text = element_text(size = 16),
    strip.background = element_rect(colour = "white"),
    panel.border = element_rect(colour = "black", fill = NA, size = 0.5)  # Border around each panel
  )


# Non-Hispanic API men
nhac <- "#e41a1c"
amp <- ggplot(subset(er, gender=="Male"),
       aes(x=year, y=aadr, group=racen))  +
  geom_line(show.legend=F, colour="grey") +
  geom_line(data=subset(er, gender=="Male" & racen=="Non-Hispanic\nAPI"),
            colour=nhac, size=1.5) +
  geom_point(data=subset(er, gender=="Male" & (year==1999 | year==2018) &
    racen=="Non-Hispanic\nAPI"), shape=21, colour="white", fill=nhac, size=3, stroke=2) +
  scale_y_continuous(limits=c(0,30), breaks=c(0,15,30)) +
  scale_x_continuous(breaks=c(2005,2015)) +
  xlab("") + ylab("") + stheme +
  ggtitle("Men") +  # Removed subtitle
  theme(plot.subtitle = element_blank(),
        axis.title.x = element_text(size = 20, hjust = 4, colour = "black"),
        axis.title.y = element_text(size = 20, hjust = -0.2, colour = "black"))

# Non-Hispanic Black man
nhbc <- "#377eb8"
bmp <- ggplot(subset(er, gender=="Male"),
       aes(x=year, y=aadr, group=racen))  +
  geom_line(show.legend=F, colour="grey") +
  geom_line(data=subset(er, gender=="Male" & racen=="Non-Hispanic\nBlack"),
            colour=nhbc, size=1.5) +
  geom_point(data=subset(er, gender=="Male" & (year==1999 | year==2018) &
    racen=="Non-Hispanic\nBlack"), shape=21, colour="white", fill=nhbc, size=3, stroke=2) +
  scale_y_continuous(limits=c(0,30), breaks=c(0,15,30)) +
  scale_x_continuous(breaks=c(2005,2015)) +
  xlab("") + ylab("") + stheme +
  ggtitle("") +  # Removed subtitle
  theme(plot.subtitle = element_blank(),
        axis.line.y = element_blank(), axis.text.y = element_blank())
```

```r
# Non-Hispanic White men
nhwc <- "#4daf4a"
wmp <- ggplot(subset(er, gender=="Male"),
       aes(x=year, y=aadr, group=racen))  +
  geom_line(show.legend=F, colour="grey") +
  geom_line(data=subset(er, gender=="Male" & racen=="Non-Hispanic\nWhite"),
            colour=nhwc, size=1.5) +
  geom_point(data=subset(er, gender=="Male" & (year==1999 | year==2018) &
    racen=="Non-Hispanic\nWhite"), shape=21, colour="white", fill=nhwc, size=3, stroke=2) +
  scale_y_continuous(limits=c(0,30), breaks=c(0,15,30)) +
  scale_x_continuous(breaks=c(2005,2015)) +
  xlab("") + ylab("") + stheme +
  ggtitle("") +
  theme(plot.subtitle = element_text(),
        axis.line.y = element_blank(), axis.text.y = element_blank())

# Hispanic men
hc <- "#984ea3"
hmp <- ggplot(subset(er, gender=="Male"),
       aes(x=year, y=aadr, group=racen))  +
  geom_line(show.legend=F, colour="grey") +
  geom_line(data=subset(er, gender=="Male" & racen=="Hispanic"),
            colour=hc, size=1.5) +
  geom_point(data=subset(er, gender=="Male" & (year==1999 | year==2018) &
    racen=="Hispanic"), shape=21, colour="white", fill=hc, size=3, stroke=2) +
  scale_y_continuous(limits=c(0,30), breaks=c(0,15,30)) +
  scale_x_continuous(breaks=c(2005,2015)) +
  xlab("") + ylab("") + stheme +
  ggtitle("") +
  theme(plot.subtitle = element_text(),
        axis.line.y = element_blank(), axis.text.y = element_blank())

# put the plots together
mp <- (amp | bmp | wmp | hmp) + plot_annotation(subtitle = 'Men')

##### Plot for women

# Non-Hispanic API women
nhac <- "#e41a1c"
awp <- ggplot(subset(er, gender=="Female"),
       aes(x=year, y=aadr, group=racen))  +
  geom_line(show.legend=F, colour="grey") +
  geom_line(data=subset(er, gender=="Female" & racen=="Non-Hispanic\nAPI"),
            colour=nhac, size=1.5) +
  geom_point(data=subset(er, gender=="Female" & (year==1999 | year==2018) &
    racen=="Non-Hispanic\nAPI"), shape=21, colour="white", fill=nhac, size=3, stroke=2) +
  scale_y_continuous(limits=c(0,10), breaks=c(0,5,10)) +
  scale_x_continuous(breaks=c(2005,2015)) +
  xlab("") + ylab("") + stheme +
  ggtitle("Women", subtitle="Non-Hispanic API") +
  theme(plot.subtitle = element_text(colour=nhac),
        axis.text.x = element_blank(),
          axis.line.y = element_blank())
```

```r
# Non-Hispanic Black women
nhbc <- "#377eb8"
bwp <- ggplot(subset(er, gender=="Female"),
       aes(x=year, y=aadr, group=racen)) +
  geom_line(show.legend=F, colour="grey") +
  geom_line(data=subset(er, gender=="Female" & racen=="Non-Hispanic\nBlack"),
            colour=nhbc, size=1.5) +
  geom_point(data=subset(er, gender=="Female" & (year==1999 | year==2018) &
    racen=="Non-Hispanic\nBlack"), shape=21, colour="white", fill=nhbc, size=3, stroke=2) +
  scale_y_continuous(limits=c(0,10), breaks=c(0,5,10)) +
  scale_x_continuous(breaks=c(2005,2015)) +
  xlab("") + ylab("") + stheme +
  ggtitle("", subtitle="Non-Hispanic Black") +
  theme(plot.subtitle = element_text(colour=nhbc),
        axis.text.x = element_blank(),
        axis.line.y = element_blank(), axis.text.y = element_blank())


# Non-Hispanic White women
nhwc <- "#4daf4a"
wwp <- ggplot(subset(er, gender=="Female"),
       aes(x=year, y=aadr, group=racen)) +
  geom_line(show.legend=F, colour="grey") +
  geom_line(data=subset(er, gender=="Female" & racen=="Non-Hispanic\nWhite"),
            colour=nhwc, size=1.5) +
  geom_point(data=subset(er, gender=="Female" & (year==1999 | year==2018) &
    racen=="Non-Hispanic\nWhite"), shape=21, colour="white", fill=nhwc, size=3, stroke=2) +
  scale_y_continuous(limits=c(0,10), breaks=c(0,5,10)) +
  scale_x_continuous(breaks=c(2005,2015)) +
  xlab("") + ylab("") + stheme +
  ggtitle("", subtitle="Non-Hispanic White") +
  theme(plot.subtitle = element_text(colour=nhwc), axis.text.x = element_blank(),
        axis.line.y = element_blank(), axis.text.y = element_blank())

# Hispanic women
hc <- "#984ea3"
hwp <- ggplot(subset(er, gender=="Female"),
       aes(x=year, y=aadr, group=racen)) +
  geom_line(show.legend=F, colour="grey") +
  geom_line(data=subset(er, gender=="Female" & racen=="Hispanic"),
            colour=hc, size=1.5) +
  geom_point(data=subset(er, gender=="Female" & (year==1999 | year==2018) &
    racen=="Hispanic"), shape=21, colour="white", fill=hc, size=3, stroke=2) +
  scale_y_continuous(limits=c(0,10), breaks=c(0,5,10)) +
  scale_x_continuous(breaks=c(2005,2015)) +
  xlab("") + ylab("") + stheme +
  ggtitle("", subtitle="Hispanic") +
  theme(plot.subtitle = element_text(colour=hc),
        axis.title.y = element_text(size = 18, hjust = -0.5, colour = "black"),
        axis.text.x = element_blank(),
        axis.line.y = element_blank(), axis.text.y = element_blank())
```

```r
# put the plots together
wp <- (awp | bwp | wwp | hwp)

F7p <- wp / mp + plot_layout(widths = rep(40, 4), heights = rep(4, 2))
y.grob <- textGrob("Age-adjusted death rates for \nalcohol-related causes(per 100,000)",
                   gp=gpar(fontface="bold", fontsize=15), rot=90)
x.grob <- textGrob("Year",
                   gp=gpar(fontface="bold", fontsize=15))
grid.arrange(patchworkGrob(F7p), left = y.grob, bottom = x.grob)
# Read data
rawd <- read_csv("le-age-cause-decomp-2014-2017.csv")

# Summarize by sex, race, and cause of death
rawt <- rawd %>%
  group_by(sex, race, cod) %>%
  summarise(total = sum(cont) * -1, .groups = 'drop')

# Create grouping for all causes ("Total change")
tot <- aggregate(rawt$total, list(sex = rawt$sex, race = rawt$race), FUN = sum)
tot$cod <- 15
colnames(tot)[3] <- "total"

# Bind together
raw <- rawt %>%
  bind_rows(tot)

# Recode gender as a factor
raw$gender <- recode_factor(raw$sex, `1` = "Women", `2` = "Men")

# Recode race-ethnicity as a factor
raw$raceeth <- recode_factor(raw$race,
  `1` = "Non-Hispanic API",
  `2` = "Non-Hispanic Black",
  `3` = "Non-Hispanic White",
  `4` = "Hispanic"
)

# Recode causes of death
raw$codf <- recode_factor(raw$cod,
  `1` = "Cardiovascular", `2` = "Cancers",
  `3` = "Diabetes", `4` = "Alzheimer's", `5` = "Flu/Pneumonia",
  `6` = "HIV", `7` = "Respiratory Disease", `8` = "Liver Disease",
  `9` = "Kidney Disease", `10` = "Motor Vehicle Crashes",
  `11` = "Unintentional Poisoning", `12` = "Suicide",
  `13` = "Homicide", `14` = "All Other Causes",
  `15` = "Total Change"
)

# Summarize by gender, race, and cause
raw15 <- raw %>%
  group_by(gender, raceeth, codf) %>%
  summarise(total = sum(total), .groups = 'drop')
```

```r
# Remove "Total Change"
raw15_filtered <- raw15 %>%
  filter(codf != "Total Change")

  # Vertical stacked bar chart
ggplot(raw15_filtered, aes(x = raceeth, y = total, fill = codf)) +
  geom_bar(stat = "identity", position = "stack") +
  geom_hline(yintercept = 0, linetype = "dashed", color = "black") +  # Horizontal line at y=0
  facet_wrap(~gender, nrow = 1) +  # Side-by-side comparison of Women and Men
  theme_minimal(base_size = 12) +  # Improve text readability
  labs(
    title = "Contribution of Leading Causes of Death to Life Expectancy Decline (2014-2017)",
    subtitle = "The largest decline during this period was for non-Hispanic Black men, primarily due to
    x = " ",
    y = "Years of Life Expectancy Change",
    fill = "Cause of Death",
    caption = "Data Source: CDC Mortality Statistics (2014-2017). Negative values indicate a decline in
  ) +
  theme(
    plot.title = element_text(face = "bold", size = 16, hjust = 0.5),  # Center title
    plot.subtitle = element_text(size = 14, hjust = 0.5),  # Center subtitle
    plot.caption = element_text(size = 12, hjust = 0.5),  # Center caption
    strip.text = element_text(size = 18, face = "bold"),  # Larger facet labels (Women/Men)
    axis.text.x = element_text(size = 12, angle = 45, hjust = 1),  # Rotate x-axis labels
    axis.text.y = element_text(size = 12),  # Increase readability of y-axis
    legend.text = element_text(size = 12),  # Larger legend text
    legend.position = "bottom",  # Move legend to bottom
    legend.title = element_text(size = 12, face = "bold")  # Make legend title larger
  )




black_men_data <- fig2_data_combined %>%
  filter(gender == "Male", racen == "Non-Hispanic Black", year >= 2000) %>%
  mutate(
    line_color = ifelse(age4 == "15-44 yrs", "highlight", "neutral"),  # Color lines
    bg_color = ifelse(age4 == "15-44 yrs", "#F5CAC7", "#EAEBEB")  # Background color for each facet
  )

# Plot with facet-specific background colors
ggplot(black_men_data, aes(x = year, y = rate, group = age4, color = line_color)) +
  geom_rect(aes(xmin = -Inf, xmax = Inf, ymin = -Inf, ymax = Inf, fill = bg_color), alpha = 0.3) +  # B
  geom_line(size = 1) +
  geom_point(size = 1.5) +
  facet_wrap(~age4, ncol = 1, scales = "free_y", strip.position = "right") +  # Stack vertically, label
  scale_color_manual(values = c("highlight" = "black", "neutral" = "black")) +  # Custom line colors
  scale_fill_identity() +  # Use colors from data without needing a legend
  labs(
    title = "Death Rate Trends for Non-Hispanic Black Males (2000-2018)",
    subtitle = "Crude death rates per 100,000 for Non-Hispanic Black males, showing a disproportionately
    caption = "Source: Centers for Disease Control and Prevention (CDC) WONDER",
    x = " ",
    y = "Crude Death Rate (per 100,000)"
```

```
) +
theme_minimal() +
theme(
  strip.text.y = element_text(size = 14, face = "bold"),  # Strip text formatting
  text = element_text(size = 14),
  strip.background = element_blank(),  # Remove default strip background
  axis.text.y = element_text(size = 10),  # Make y-axis labels smaller
  plot.title = element_text(size = 16, face = "bold", hjust = 0.5),  # Bold title
  plot.subtitle = element_text(size = 12, color = "gray40", hjust = 0.5),  # Subtitle formatting
  plot.caption = element_text(size = 10, color = "gray50"),  # Caption formatting
  panel.background = element_rect(fill = "white", color = NA),  # Overall plot background white
  legend.position = "none"  # Hide legend for cleaner visualization
) +
scale_x_continuous(limits = c(2000, 2018), breaks = seq(2000, 2018, 2))
```