Diversity, Equity and Inclusion

Mortality and Longevity

U.S. Violent Manner of Death Mortality by Race and Ethnicity A Mortality Study

September | 2023

U.S. Violent Manner of Death Mortality by Race and Ethnicity

A Mortality Study

AUTHORS Tatjana Miljkovic, Ph.D., MBA,
Associate Professor, Miami University

Ying-Ju Chen, Ph.D.,
Assistant Professor, University of Dayton

SPONSOR Diversity, Equity, and Inclusion Strategic Research Program

Contents

U.S. Violent Manner of Death Mortality by Race and Ethnicity

A Mortality Study

Executive Summary

Conventional wisdom in the United States is that mortality by violent causes of death is disproportionately experienced in underrepresented communities as opposed to white communities. This study aims to quantify this experience in more detail than is commonly reported. To achieve this, the authors examined disparities by race and ethnicity and other variables for three violent manner of deaths using data from 27 U.S. states (accounting for 43% of the total U.S. population) between 2016 and 2020 obtained from the Centers for Disease Control and Prevention (CDC) through the National Violent Death Reporting System (NVDRS).

This study considers mortality rates for each of three manners of violent death: suicide, homicide, and “other.” The “other” category of violent manners of death includes deaths due to legal intervention by police or other authorities, unintentional firearm–self-inflicted, undetermined intent, and unintentional firearm inflicted by either another person or unknown who inflicted. Mortality rates differ from the number of deaths; mortality rates represent the number of deaths as a proportion of a population.

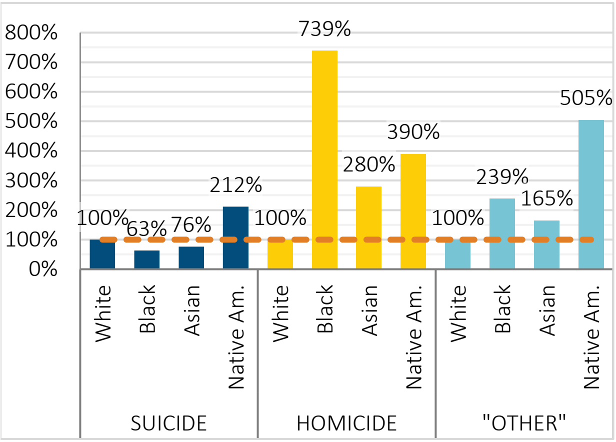
Exploratory data analysis can be a valuable tool to gain an initial understanding of the data, but statistical modeling is necessary to control for specific variables and isolate the effect of specific factors of interest. This study used generalized linear models (GLMs) to isolate the relative risks for race and ethnicity. This type of analysis requires a reference group, which is typically the largest group when one group is not representative of the whole dataset. The largest racial and ethnic groups in the data are whites and non-Hispanic/Latinos, respectively.

The key findings of this mortality rate study can be summarized as follows:

- Suicide (60%) and homicide (30%) are the leading manners of violent death.
- Males are more likely to die than females across all manners of violent death.
- The three 5-year age groups from 20 to34 years are the most susceptible age groups to all violent deaths.
- By race (see Figure 1):
  - Native Americans have the highest risk of dying by suicide and “other” manners of violent death.
  - Blacks/African Americans have the highest risk of dying because of homicide. Blacks are more than 7 times as likely as whites to die by homicide, while Native Americans and Asian/Asian Americans and are 4 and 3 times, respectively, as whites to die by homicide.
- By ethnicity, Hispanics/Latinos have a higher risk than non-Hispanic/Latinos of dying because of any of the manners of violent death (see Figure 2).

Figure 1

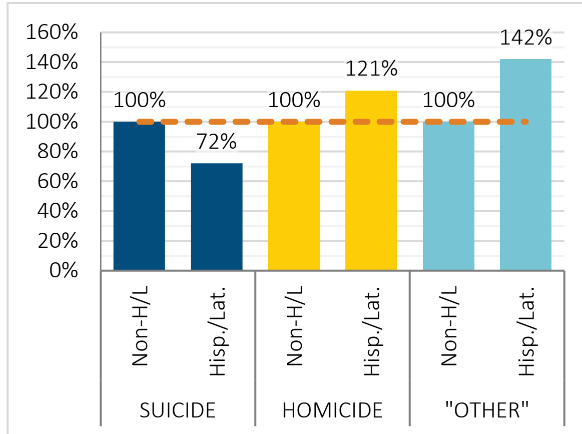
RELATIVE RISKS ISOLATED BY RACE



No other fixed factors are used for suicide; fixed factors for homicide are state, age, and sex; fixed factors for “other" are state, age, sex, and manner.

Figure 2

RELATIVE RISKS ISOLATED BY ETHNICITY



Fixed factor for suicide is sex; fixed factors for homicide are state, age, and sex; fixed factors for “other” are state, age, sex, and manner.

GLM analysis also allowed for finding significant factors of the violent death mortality rates for each manner of death. Race, ethnicity, sex, and age are significant factors in determining rates of each manner of violent death. These findings about mortality rates are summarized as follows:

- Suicide:
  - Race: The highest risk of suicide across all age groups is for Native Americans.
    - A person of Native American descent is more than two times more likely to die by suicide than a white person, regardless of sex and age.
    - Blacks/African Americans and Asians/Asian Americans, on average, have a lower risk of dying because of suicide compared to whites.
    - The risk of suicide for Asians/Asian Americans is, on average, about 25% lower than that for whites (see Figure 1).
  - Ethnicity: The risk of suicide among Hispanics/Latinos is about 30% lower than that for the non-Hispanic/Latino population (see Figure 2).
  - Sex: The risk of suicide for males is about 3.5 times higher than for females.
  - Age: The risk of suicide increases by age group from ages 10–14 to 45–54 years, then slowly decreases toward older age groups. People aged 45–54 are almost eight times more likely to die by suicide than those aged 10–14.
- Homicide:
  - Race: Blacks/African Americans have about a 7.5 times higher risk of dying because of homicide relative to whites. Asians/Asian Americans and Native Americans have about a three and a four times, respectively, higher risk of dying because of homicide relative to whites (Figure 1).
  - Ethnicity: The Hispanic/Latino population has about a 1.2 higher risk of dying because of homicide than their non-Hispanic/Latino counterparts (Figure 2).
  - Sex: Males have more than a three times higher risk of dying because of homicide than females.
  - Age: The highest risk of homicide is found for the age group 25–29 years old.
- “Other” manners of violent death:
  - Race: The highest risk of death due to other manners of death is reported for Native Americans. A person of Native American descent is five times more likely to die because of other manners of death than a white person. Blacks/African Americans have about a 2.3 times higher risk of dying because of

other manners of death relative to whites. Asians/Asian Americans and Native Americans have about a 1.6 times and a five times higher risk of dying because of other manners of death compared to whites (see Figure 1).

- Ethnicity: The Hispanic/Latino population has about a 1.4 higher risk of dying because of other manners of death compared to their non-Hispanic/Latino counterparts (see Figure 2).
- Sex: Males have about a 1.5 times higher risk of dying because of homicide than females.
- Age: The highest risk of dying due to other manners of death is found for the age group 30–34 years old. This risk is more than twice as high as the age group 10–14.

Additional key notes:

- Although calendar year was considered as a factor, it was found not significant and therefore was removed from the model.

- The findings presented in this report are heavily reliant on the NVDRS data set, and negative binomial regression models utilized. However, the absence of applicable socioeconomic variables in the data set precludes their inclusion in the models. Factors such as occupation, income, and education may have a substantial influence on the relative rates of race and ethnicity, but their omission limits the scope of this analysis. Therefore, although this report is valuable for comprehending the link between the factors studied and the mortality rates due to violent deaths, it should not be employed to forecast occurrences of violent deaths.

Section 1: Introduction

In this report the authors examine the violent causes of death in 27 U.S. states between 2016 and 2020. Violent causes of death can take different forms, such as suicide, homicide, and "other" deaths. "Other" deaths include unintentional firearm injury, legal intervention, and deaths of undetermined intent. The data on violent deaths have been provided by the National Violent Death Reporting System (NVDRS) as part of the Centers for Disease Control and Prevention (CDC, 2022).

1.1 BACKGROUND AND LITERATURE REVIEW

Research studies and reports on violent deaths in the U.S. have taken several research pathways. The first includes the studies reported by the CDC based on the surveillance of national violent deaths in the U.S. (see Wilson et al. 2022). This report provides tabulated statistics by sex, age group, race and ethnicity, method of injury, type of location where the injury occurred, circumstances of the injury, and other selected characteristics by looking at a single year across a large sample of participating states. The most recent report, based on data from 2019, reported that among a total of 51,627 violent deaths in the U.S., the majority (64.1%) were suicides, followed by homicides (25.1%), deaths of undetermined intent (8.7%), legal intervention deaths (1.4%) and unintentional firearm deaths (<1.0%). Here the deaths due to legal intervention include those caused by law enforcement and other persons with legal authority to use deadly force acting in the line of duty, excluding legal executions. No cross-sectional studies are in in this research pathway to analyze violent deaths across different causes of death and sociodemographic factors. To the best of the authors' knowledge, no studies look at mortality rates over time for different combinations of factors and their categories (e.g., mortality by age and sex). Most importantly, no studies have been found to show efforts in building statistical models based on the CDC data on violent death. These are the gaps that the current research seeks to address.

The second research pathway includes studies of violent death with a special focus on the individual manner of violent deaths as a function of sociodemographic factors (e.g., suicide by age and race):

- Many studies have been generated on the topic of suicide. For example, Ehlman (2022) reported that suicide was among the 10 leading causes of death in the U.S. in 2020 among persons aged 10–64 years and the second leading cause of death among children and adolescents aged 10–14 and adults aged 25–34. A study on suicide conducted by Joe et al. (2008) reported that Native Americans and Alaskan Indigenous Residents experience significant disparities in their prevalence of suicidal ideation, attempts, and deaths compared to all other racial/ethnic groups in the U.S.

- Regarding homicide, Fridel and Fox (2019) reported that almost three-quarters of homicides in the U.S. involve a male killing another male. Kegler et al. (2022) analyzed the impact of the COVID-19 pandemic in 2020 on homicide rates in the U.S. They found that 79% of homicides involve firearms. The high firearm homicide rates are related to factors such as inequities by race, ethnicity, and poverty level. A study by Barber et al. (2016) analyzed 1,552 police homicides in 16 states during 2005 to 2012 and reported that the annual rate of police homicide was 0.24 per 100,000 people and varied fivefold by state and eightfold by race/ethnicity. Among Hispanics/Latinos, the rate was 0.25 per 100,000 people. Among non-Hispanics/Latinos, rates were 0.48 for Blacks, 0.25 for Native Americans, 0.17 for whites, and 0.06 for Asians/Asian Americans. Rates by state varied more than fivefold, with a low of 0.09 per 100,000 people in Massachusetts to a high of 0.43 and 0.51 in Alaska and New Mexico, respectively. An estimated 731 people were killed by law enforcement officials each year in the U.S. from 2005 to 2012.

The findings of these and similar studies are essential to this study because they will be used to validate some of the results generated in this study (homicides by age and race and ethnicity). However, this study extends current literature by looking at individual manners of violent death driven by a combination of sociodemographic factors such as state, sex, age, race, and ethnicity, aggregated over time.

The third research pathway includes comparative studies comparing the trends in the U.S. to other countries. Some of these research efforts are also led by the World Health Organization (WHO). Grinshteyn and Hemenway (2016) studied the mortality rates provided by the WHO for 2010 due to violence of death among 23 Organisation for Economic Co-operation and Development (OECD) countries. The authors found that homicide rates in the U.S. were seven times higher than in other high-income countries, driven by a gun homicide rate that was about 25 times higher. For the age group 15–24 years old, the gun homicide rate was about 25 times higher. Ninety percent of women, 91% of children aged 0 to 14 years, 92% of youth aged 15 to 24 years, and 82% of all people killed by firearms were from the United States in the sample of 23 OECD countries. Studying homicide, Jaffe (2018) reported that people living in the U.S. are 25 times more likely to die in a gun homicide than in other wealthy countries. The mass shootings that often occur in the U.S. and capture so much attention account for just one-half of 1% of all U.S. gun fatalities annually, based on the report “Gun Policy in America” published by the RAND Corporation (2022). Although this is an important research topic, this study uses data for the U.S. only and will not discuss results for other countries.

1.2 OBJECTIVES

This project investigates the mortality rates by identifying violent manners of death in underrepresented communities relative to the white communities in the U.S. by studying the following:

1. Whether mortality rates for the identified violent manner of death differ between underrepresented communities and white communities in the U.S.
2. Whether sex, race, age and the manner of death are significant factors in determining violent mortality rates in the U.S.
3. The relative risk of specific manners of death, including suicide, homicide, legal intervention by police and other authorities, unintentional firearm–self-inflicted, and unintentional firearm–unknown-inflicted.

1.3 TERMINOLOGY

Krug et al. (2002) indicates the WHO defines a violent death as “a death resulting from the intentional use of physical force or power against oneself, another person, or against a group or community.” A slightly modified definition is used by the National Violent Death Reporting System (NVDRS) of the CDC in the U.S. According to the NVDRS, 2023, a violent death is defined as “a death that results from the intentional use of physical force or power, threatened or actual, against oneself, another person, or a group or community.” NVDRS also specifies that the word “power” includes acts of neglect or omission by one person who has control over another.

The violent deaths include deaths due to suicides, homicides, deaths from legal intervention (a subtype of homicide where the victim is killed by or died as a result of law enforcement action in the line of duty), deaths of undetermined intent, unintentional firearm fatalities, and other with an unknown cause. Deaths of undetermined intent include deaths with some evidence of intent but without enough to definitively classify the death as purposeful. Unintentional firearm injury deaths include some deaths that are, in fact, intentional or of undetermined intent; for example, a person shoots himself or herself when using a gun to frighten, control or harm another person. All these manners of death have been reported as part of the NVDRS data and are analyzed as part of this research study.

The different categories of deaths are labeled by “manner” of death by the NVDRS, and the same term is used in this report. Regional Medical Examiner’s Office of Washoe County, Nevada discussed the difference between the manner of death and the cause of death by reporting that the manner of death is the determination of how the injury or disease leads to death (e.g., natural, suicide, homicide or accident). The manner of death should not be confused with the cause of death (e.g., cancer or heart disease). The cause of death is typically referred to as a specific injury or disease that leads to death.

Each manner of death is further defined as follows:

1. Suicide death

Suicide is a death resulting from the intentional use of force against oneself. A preponderance of evidence should indicate that the use of force was intended.

2. Homicide death

Homicide is defined as a death resulting from the intentional use of force or power, threatened or actual, against another person, group or community. A preponderance of the evidence must indicate that the use of force was intentional. Such deaths resulting from legal intervention are included in a separate category below. Two special scenarios that the National Center for Health Statistics (NCHS) regards as homicides are included in the NVDRS definition: (1) arson with no intent to injure a person and (2) stabbing with intent unspecified.

3. Terrorism Homicide Death

Terrorism deaths are homicides or suicides resulting from events labeled by the Federal Bureau of Investigation as acts of terrorism. Terrorism is a mechanism of death rather than a manner of death. The manner of such death is either homicide or suicide.

4. Unintentional Firearm Injury Death

A death resulting from a penetrating injury or gunshot wound from a weapon that uses a powder charge to fire a projectile when there was a preponderance of evidence that the shooting was not intentionally directed at the victim. Other types of unintentional deaths (e.g., accidental overdose) are not covered in this category. This category has three subcategories:

(1) unintentional firearm–self-inflicted,

(2) unintentional firearm–inflicted by another person, and

(3) unintentional firearm–unknown who inflicted.

5. Legal Intervention Death

A death in which the decedent was killed by or died as a result of a law enforcement officer or another peace officer (persons with specified legal authority to use deadly force), including military law enforcement, acting in the line of duty. The term legal intervention is a classification from the WHO's ICD-10 codes. It does not denote the lawfulness or legality of the circumstances surrounding the death.

These deaths can occur during the course of a law enforcement officer's conducting a random or targeted traffic stop, issuing a citation, making an arrest or in pursuit to apprehend a victim (e.g., victim fleeing or escaping arrest), responding to a call to maintain order, minimizing disturbances and/or ensuring safety (e.g., domestic disturbances or to circumvent a suicide crisis) or other actions as part of law enforcement duties.

6. Death Due to Undetermined Intent

   Undetermined intent includes deaths with evidence of intent but without enough to definitively classify the death as purposeful. Opioid overdose deaths of undetermined intent are included in this category.

This report uses terms consistent with their definition. For example, terms such as "male" and "female" are used for sex, which is a biological concept, whereas the terms "men" and "women" apply to gender, which is a social, behavioral, and legal concept.

1.4 RACE AND ETHNICITY

According to the National Human Genome Research Institute (2023), race is a social construct used to group people. These classifications were developed to identify, distinguish, and marginalize certain groups across nations, regions, and the world. Typically based on physical appearance, social factors, and cultural backgrounds, race has been used to divide human populations into distinct groups.

According to the National Cancer Institute (2023), ethnicity is a term that refers to the social and cultural characteristics, backgrounds or experiences shared by a group of people. This includes language, religion, beliefs, values, and behaviors that are often handed down from one generation to the next.

The terminology for race and ethnicity differs across sources for this study. Two sources of data are considered in this study: the NVDRS and the population data downloaded from the CDC's Wide-ranging Online Data for Epidemiologic Research (WONDER) system. The terminology used by the NVDRS and WONDER is somewhat inconsistent with the Society of Actuaries' preferred language (see Table 1).

Table 1

TERMINOLOGY FOR RACE AND ETHNICITY USED ACROSS SOURCES

Race/Ethnicity	This Report	NVDRS	WONDER
Race	Asian/Asian American	Asian/Pacific Islander	Asian and Pacific Islander
Race	Black/African American	Black or African American	Black or African American
Race	Native American/Alaska Indigenous Resident	American Indian/Alaska Native	American Indian and Alaska Native
Race	White	White	White
Ethnicity	Hispanic/Latino	Hispanic	Hispanic or Latino

1.5 LITERATURE REVIEW BACKGROUND

The research conducted in preparing this report was based on a literature review limited to articles published in 2010 or later. The time periods, source data, methods, study objectives, variables studied, and level of detail varied significantly across the articles.

1.6 REPORT STRUCTURE

This report is organized as follows. Section 2 discusses the data sets in detail, including the definitions of the manners of violent death. Section 3 includes the exploratory data analysis for the crude mortality rates per 100,000 people across different sociodemographic factors and manners of death and data visualization. In Section 4 the authors analyze and build the statistical models for modeling mortality rates due to violent deaths by state, race, ethnicity, sex, age group, and manner of death. This section also provides a summary of the results and the main findings. Section 5 includes the limitations of this study. A discussion and conclusion are provided in Section 6.

Section 2: Data Sets

This section discusses the data sources associated with violent deaths and the corresponding population. Both data sources were obtained from the CDC.

This study is based on data from the National Violent Death Reporting System (NVDRS), which is the only state-based violent death reporting system in the U.S. that helps provide information and context on when, where, and how violent deaths occur and who is affected (Barber et al. 2013). The researchers obtained the NVDRS data from the CDC by signing the data-sharing agreement. Therefore, the data are confidential.

Data are from 27 U.S. states between 2016 and 2020. These states represent 43% of the total U.S. population, the largest available sample for the last five years of data, and are Alaska, Arizona, Colorado, Connecticut, Georgia, Indiana, Iowa, Kansas, Kentucky, Maine, Maryland, Massachusetts, Michigan, Minnesota, New Hampshire, New Jersey, New Mexico, North Carolina, Ohio, Oklahoma, Oregon, Rhode Island, South Carolina, Utah, Vermont, Virginia and Wisconsin.

This study analyzes the violent deaths provided by the NVDRS. The categories of violent deaths include:

(1) suicide,

(2) homicide,

(3) terrorism homicide,

(4) unintentional firearm (inflicted by another person, self-inflicted, and unknown who inflicted,

(5) legal intervention (by police or other authority), and

(6) undetermined intent.

Before summarizing the crude mortality rates, the following caveats need to be considered:

- Hawaii was excluded from data years 2017, 2018, and 2020 because of incomplete case reporting.
- New York was excluded from the data year 2019 because of preliminary case reporting.
- Illinois, Pennsylvania, and Washington: For 2016 and 2017 data, aimed to collect data on violent deaths in a subset of counties representing at least 80% of the violent deaths in their state for those years. For 2018 data, Washington went statewide, but Illinois and Pennsylvania continued to collect data only in a subset of counties representing at least 80% of violent deaths in the state that year. Because less than 100% of violent deaths were targeted for data collection, data from these states may not be fully representative of all violent deaths occurring in their states for these given years.
- California collected data for violent deaths in accordance with requirements under which CDC provided NVDRS funding to the state. Consequently, data was collected for four counties in 2017 (n = 1,866, representing 27.8% of violent deaths that occurred in California in 2017), 21 counties in 2018 (n = 3,659; representing 55.1% of violent deaths that occurred in California in 2018), 30 counties in 2019 (n = 3,645; representing 55.3% of violent deaths that occurred in California in 2019), and 35 counties in 2020 (n = 4,675; representing 68.1% of violent deaths that occurred in California in 2020).
- Texas collected data for violent deaths that occurred in four counties in 2020 (n = 2,741, representing 40.5% of all violent deaths that occurred in Texas in 2020) in accordance with the requirements under which CDC provided NVDRS funding to the state.

Because the statistics from these states may not be entirely representative of all violent deaths that took place in them in those time frames, the crude mortality rate for those states and those years will therefore not be provided as recommended by the CDC. Table 2 summarizes the exclusion of the state data from Table B-1 in Appendix B for each year since 2016. Table B-1 summarizes the availability of state data in each year from 2003 to 2020. One should note that the data for Florida are not available for all years. Also, Puerto Rico is not included in the study.

Table 2

EXCLUSION OF STATE DATA FROM TABLE B-1 EACH YEAR SINCE 2016

Year	States
2016	Illinois, Pennsylvania, Washington
2017	California, Hawaii, Illinois, Pennsylvania, Washington
2018	California, Hawaii, Illinois, Pennsylvania
2019	California, New York
2020	California, Hawaii, Texas

The population data used in this project are obtained from the CDC's WONDER database. WONDER is designed as an integrated information and communication system for public health. One purpose of it is to provide the general public with access to specific and detailed information from the CDC. In this project, the bridged-race population data were queried and downloaded from WONDER. The population estimates were obtained by year, state, race (four categories), ethnicity (two categories), sex (two categories), and age (five-year groups). Race and ethnicity are not mutually exclusive categories, so careful consideration was given when merging the population data with the death counts obtained from the NVDRS.

The methods employed here are standard methods already used in actuarial science and statistics. The data structure reflects the standard for calculating mortality rates (see Boucher and Guillén 2009; Dobson and Barnett 2018).

Section 3: Exploratory Data Analysis

3.1 CRUDE RATE CALCULATION

This subsection documents how the crude mortality rates are calculated based on state, sex, age group, race, and ethnicity.

First, the NVDRS deaths are aggregated by the manners of deaths, year, state, sex, age group, race, and ethnicity. The age group definition is consistent with the age variable (five-year groups) in the CDC WONDER data. Second, the data are merged with the population counts obtained from WONDER. Third, the crude mortality rate per 100,000 people, given a manner of death, year, state, sex, age group, race, and ethnicity, is calculated as follows:

$$crude\ rate = \frac{no.of\ deaths}{population} \times 100{,}000.$$

3.2 EXPLORATORY DATA ANALYSIS OBSERVATIONS

The availability of state data is inconsistent across years from 2003 to 2020, and so this crude rate study focuses on 47 U.S. states that collected statewide data in 2020 and information on 55,903 fatal incidents involving 56,030 deaths.

In 2020, of 56,030 deaths in these 47 states, most were suicide, followed by homicide, deaths of undetermined intent, legal intervention deaths, and unintentional firearm deaths (see Table 3). The mortality rates for all manners of deaths were all higher for males than females, with the most significant gap found in suicide rates. Across all age groups, the suicide rates are the highest among people aged 30–34; the homicide rates are the highest among people 20–24. The mortality rates for deaths of undetermined intent and legal intervention deaths are the highest among people 30–34. The mortality rates for unintentional firearm deaths are the highest among those 15–19. In addition, non-Hispanic/Latino whites had the highest mortality rates for all manners of deaths except homicide deaths among all racial and ethnic groups. In contrast, non-Hispanic/Latino Blacks/African Americans experienced the highest homicide rate among all racial and ethnic groups.

Most violent deaths in 2020 (60%) were suicides, while 30% were homicide and 10% were of undetermined intent, accidental, or legal intervention.

Table 3
DISTRIBUTION OF MANNER OF DEATHS IN 2020 IN NVDRS DATA

Manner of Death	Percentage
Suicide	59.58%
Homicide	29.72%
Deaths of undetermined intent	8.75%
Legal intervention deaths	1.26%
Unintentional firearm deaths	0.69%

3.3 DATA VISUALIZATION

To explore the relationship between mortality rates and several factors, an Exploratory Data Analysis (EDA) was conducted using data visualization tools. The EDA was performed on three different groups because of the small number of deaths in certain manners of death, namely, suicide, homicide, and all “other” remaining manners of death (legal intervention by police and other authorities, unintentional firearm–self-inflicted, unintentional firearm–inflicted by another person, and unintentional firearm–unknown who inflicted). Moreover, the EDA focused on data from 27 U.S. states from 2016 to 2020 because the availability of state data varied across years from 2003 to 2020. The states are Alaska, Arizona, Colorado, Connecticut, Georgia, Indiana, Iowa, Kansas, Kentucky, Maine, Maryland, Massachusetts, Michigan, Minnesota, New Hampshire, New Jersey, New Mexico, North Carolina, Ohio, Oklahoma, Oregon, Rhode Island, South Carolina, Utah, Vermont, Virginia, and Wisconsin. This subset of U.S. states represents the largest available sample for the last five years of data.

3.3.1 SUICIDE DEATHS

First, the authors study the mortality rates due to suicide deaths by age group versus sex, race, and ethnicity individually. Figure 3 shows the mortality rates by age group and sex due to suicide deaths per 100,000 people in the 27 U.S. states from 2016 to 2020. The deaths for the year, state, race, and ethnicity are aggregated by age group and sex. The mortality rates for males are more than double those for females among all age groups. More specifically, the mortality rates for females are about 10 per 100,000 females or less, and the mortality rates for males are about 30 per 100,000 males in a given age group between 20 and 84. This is evidence that sex is one of the most significant variables.

Figure 3

SUICIDE MORTALITY RATES PER 100,000 FROM 2016 TO 2020
BY AGE GROUP AND SEX

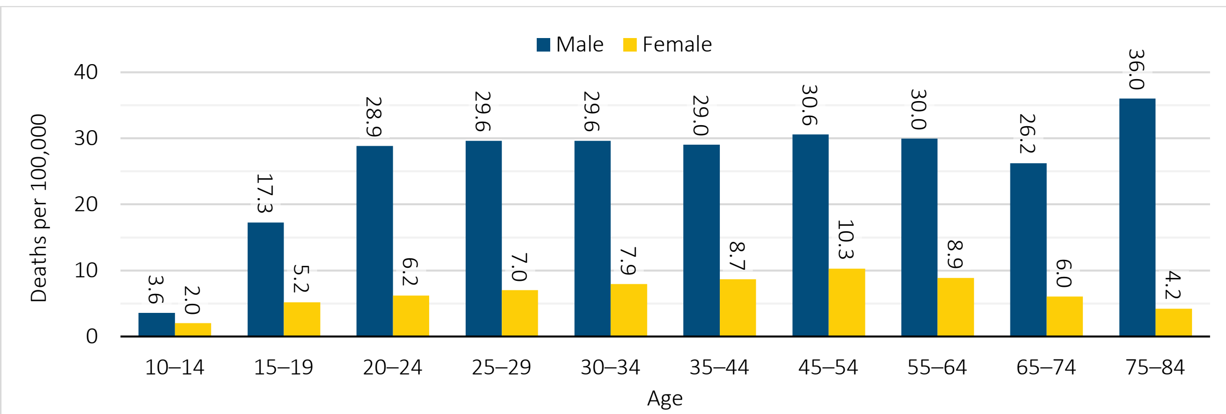


Figure 4 shows the mortality rates by age group and race due to suicide deaths per 100,000 people in the 27 U.S. states from 2016 to 2020. The deaths for the year, state, sex, and ethnicity are aggregated by age group and race. Native Americans have significantly higher mortality rates due to suicide deaths compared to other races, particularly at and before the age range of 30–34. This concerning trend may be attributed to two factors. First, the life expectancy of Native Americans is about seven years lower than that of white Americans (NICOA 2023), which means that other causes of death may have a larger impact on the crude rates in those aged 35 or older. Second, the relatively small population of Native Americans with a relatively high number of deaths further exacerbates the issue. For instance, the average number of deaths of Native Americans aged 25–29 is 66.4 people per year, with an average population of 194,626. However, the average number of deaths of whites in the same age group is 1,474.2 people per year, with an average population of 7,372,767. In addition, the mortality rates for whites due to suicide deaths after the age of 20 are around 20 deaths per 100,000 people. Even with a higher population, white mortality rates are considerably higher than other races at and after age 45. This observation supports the hypothesis that race is a critical variable associated with violent death.

Figure 4

SUICIDE MORTALITY RATES PER 100,000 BY FROM 2016 TO 2020
AGE GROUP AND RACE

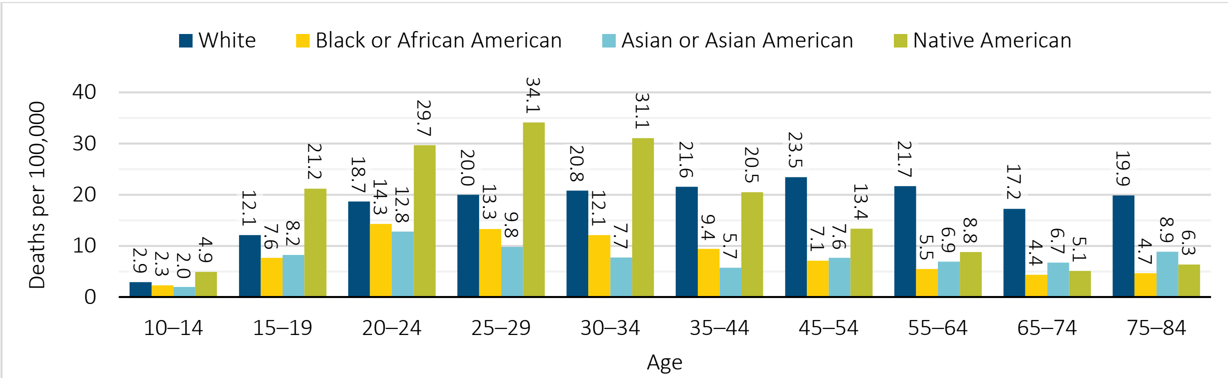
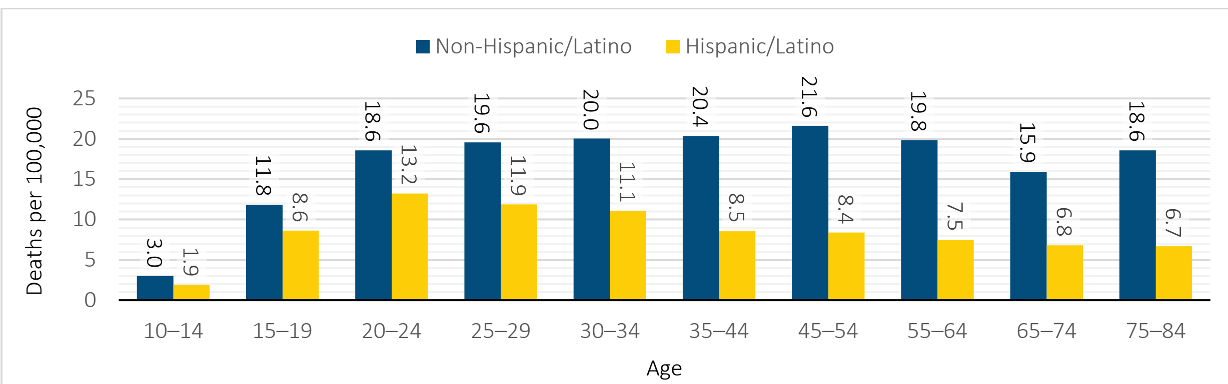


Figure 5 shows the mortality rates by age group and ethnicity due to suicide deaths per 100,000 people in the 27 U.S. states from 2016 to 2020. The deaths for the year, state, sex, and race are aggregated by age group and ethnicity. The observation in the mortality rates shows that it is generally higher for non-Hispanics/Latinos compared to Hispanics/Latinos across all age groups while considering other factors fixed. This indicates that ethnicity should be considered one of the significant variables when studying violent death rates.

Figure 5

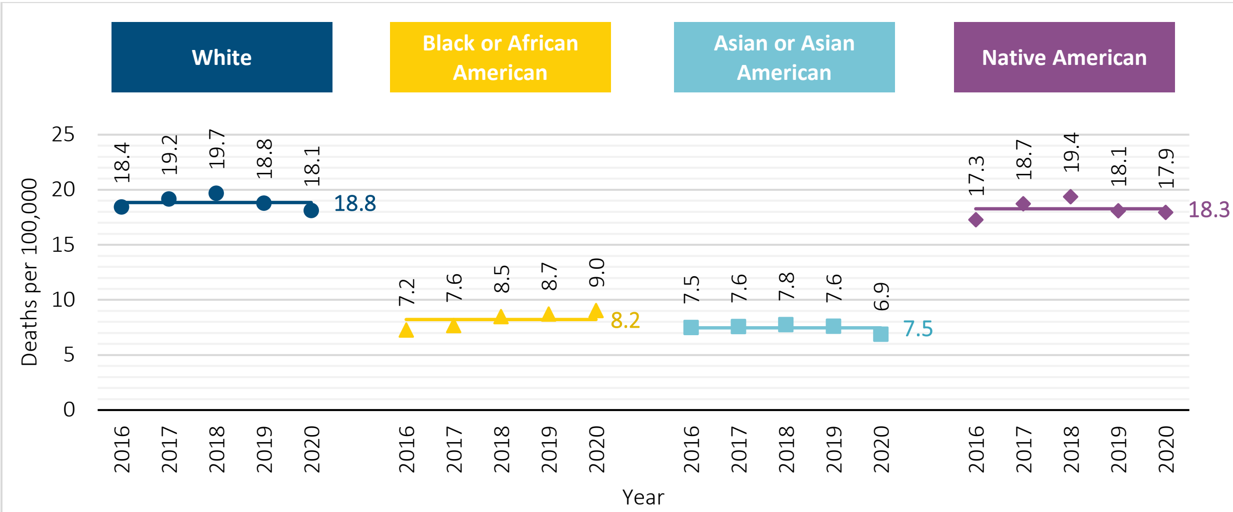
SUICIDE MORTALITY RATES PER 100,000 FROM 2016 TO 2020
BY AGE GROUP AND ETHNICITY



Shifting to examine the effect of race on mortality rates due to suicide deaths, Figure 6 shows the mortality rate by year and race due to suicide deaths per 100,000 people in the 27 U.S. states from 2016 to 2020. The deaths for state, age group, sex, and ethnicity are aggregated by year and race. In Figure 6, the horizontal line for each race indicates the mean of five mortality rates. Without considering other factors, Figure 6 shows that the mortality rate for whites is slightly higher than that for Native Americans. In contrast, the mortality rates for Blacks/African Americans and Asians/Asian Americans are relatively lower. One should note that the mortality rates for Blacks/African Americans show an increasing pattern; however, this phenomenon is not necessarily true over a long time and should be tracked after 2020.

Figure 6

SUICIDE MORTALITY RATES PER 100,000 FROM 2016 TO 2020
BY RACE



Looking at the five-year trend in suicide rates, taking into account race and ethnicity, Figure 7 and Figure 8 show the suicide rates per 100,000 for females and males, respectively, by year, ethnicity, and race in the 27 U.S. states from 2016 to 2020. From these two figures, the authors make three observations:

1. The male mortality rate is overall higher than females across all races, and the male and female differences appear very proportional as the line graphs look similar with the *y*-axis changed.

2. The mortality rate for non-Hispanic/Latinos is higher than for Hispanic/Latinos, except for several Asian or Asian American cases.

3. A significant disparity is seen in mortality rates between Hispanics/Latinos and non-Hispanics/Latinos among whites and Native Americans, with the gap being particularly pronounced among Native Americans.

One should note that although the mortality rates for Hispanic/Latino Asian Americans seem volatile, the difference between the numbers of deaths in years is slight (female: 5, 5, 2, 6, 3; male: 4, 16, 12, 6, 8). However, their corresponding populations in years do not vary much.

Figure 7

FEMALE SUICIDE MORTALITY RATES PER 100,000 FROM 2016 TO 2020
BY YEAR, ETHNICITY AND RACE

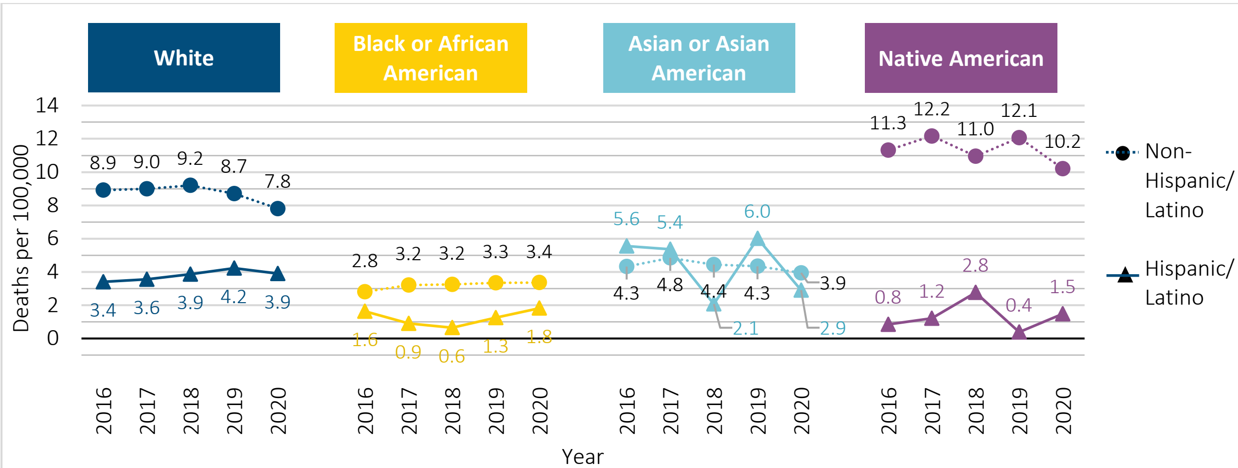
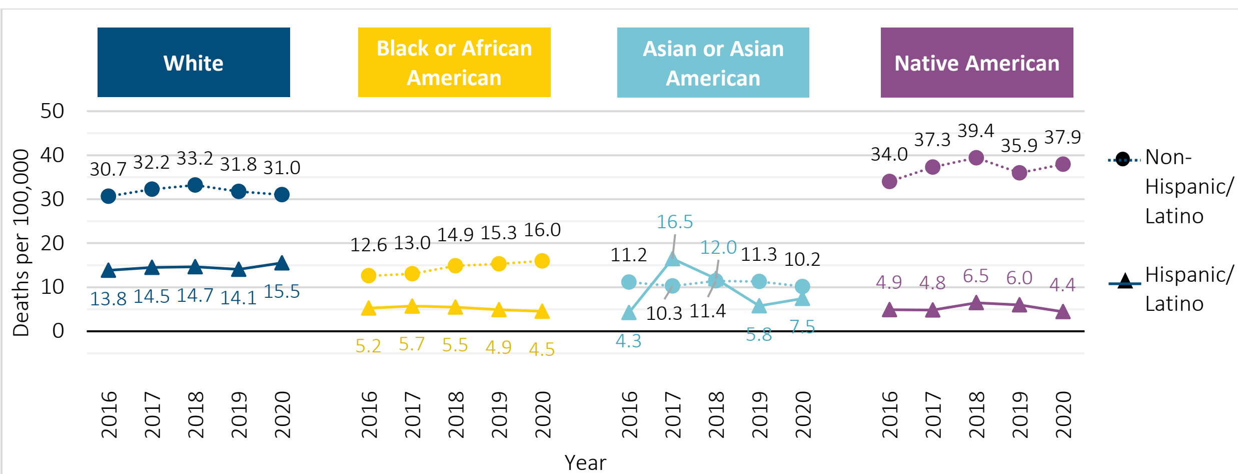


Figure 8

MALE SUICIDE MORTALITY RATES PER 100,000 FROM 2016 TO 2020
BY YEAR, ETHNICITY AND RACE



3.3.2 HOMICIDE DEATHS

This subsection explores the mortality rates due to homicide deaths via different factors. These deaths include homicide and terrorism homicide. Figure 9 shows the mortality rates by age group and sex due to homicide deaths per 100,000 people in the 27 U.S. states from 2016 to 2020. The deaths for the year, state, race, and ethnicity are aggregated by age group and sex. Like suicide cases, the mortality rates for males due to homicide deaths are higher overall than those for females. Unlike suicide, age group seems to have a minor impact on female mortality rates due to homicide. The mortality rate due to homicide for males shows a decreasing pattern after the 25–29 age group. However, this is not the case with the male mortality rates by age group due to suicide. This could indicate that the age group has a higher impact on the study of homicide deaths.

Figure 9

HOMICIDE MORTALITY RATES PER 100,000 FROM 2016 TO 2020
BY AGE GROUP AND SEX

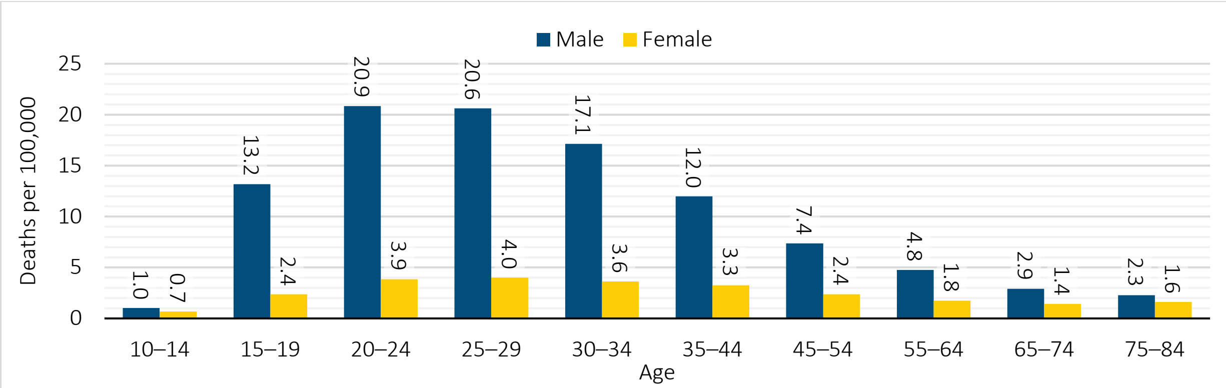


Figure 10 shows the mortality rates by age group and race due to homicide deaths per 100,000 people in the 27 U.S. states from 2016 to 2020. The deaths for the year, state, sex, and ethnicity are aggregated by age group and race. Contrary to the suicide cases, the mortality rate due to homicide for Blacks/African Americans is much higher between age 15 and 64 compared to the mortality rates for the other three populations. In addition, the declining trend of the mortality rates for Blacks/African Americans supports that the age group could play an important role when studying homicide deaths.

Figure 10

HOMICIDE MORTALITY RATES PER 100,000 FROM 2016 TO 2020
BY AGE GROUP AND RACE

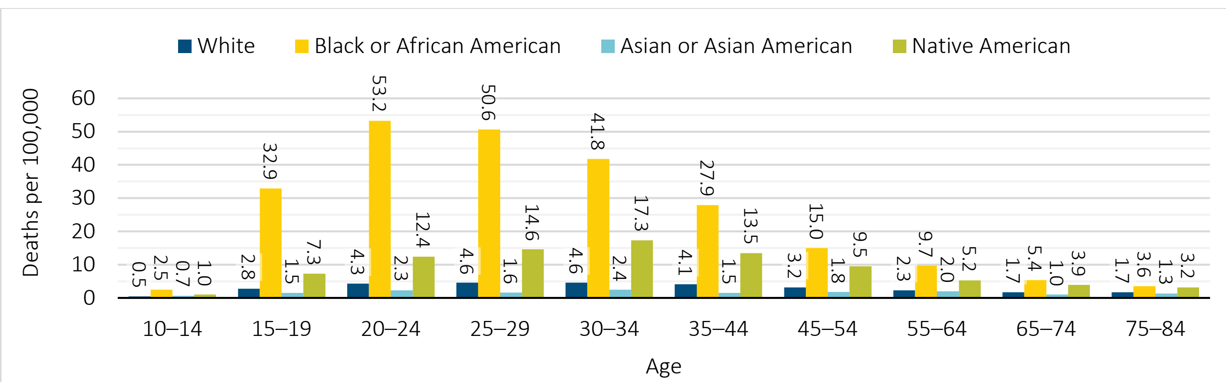
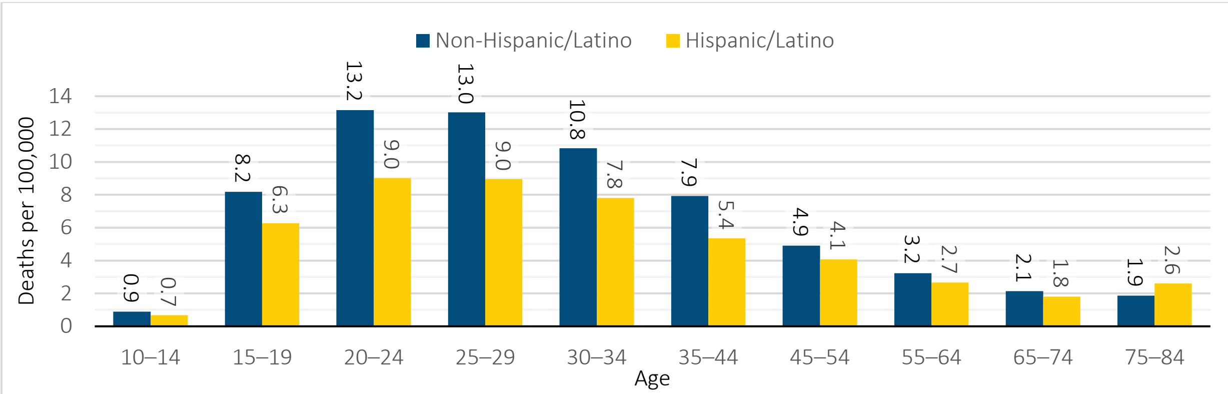


Figure 11 shows the mortality rates by age group and ethnicity due to homicide deaths per 100,000 people in the 27 U.S. states from 2016 to 2020. The deaths for year, state, sex, and race are aggregated by age group and ethnicity. Consistent with the trend due to suicide deaths in Figure 5 the mortality rate for Hispanics/Latinos is lower for all age groups than for non-Hispanics/Latinos.

Figure 11

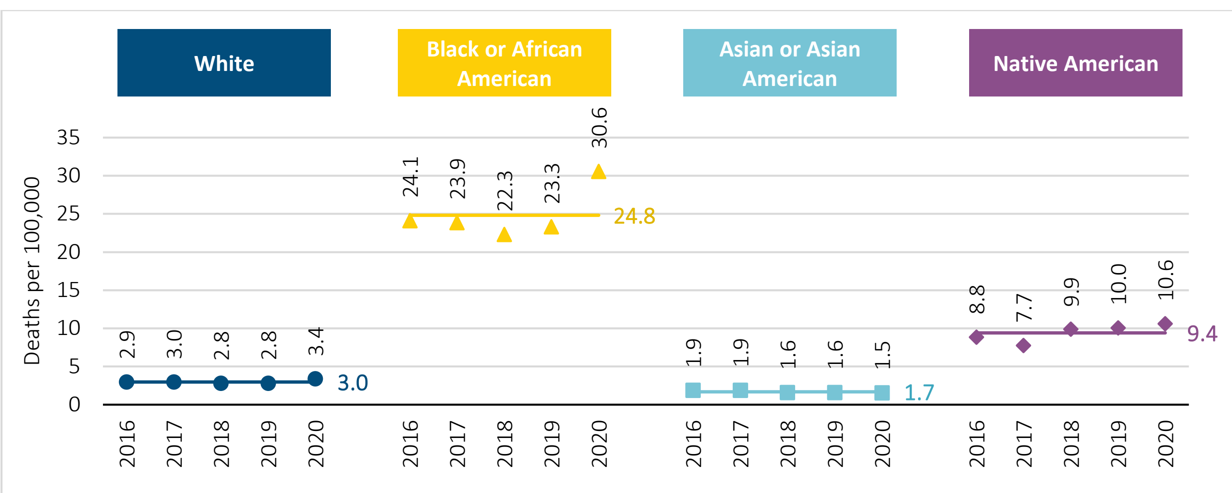
HOMICIDE MORTALITY RATES PER 100,000 FROM 2016 TO 2020
BY AGE GROUP AND ETHNICITY



Next, the impact of race on mortality due to homicide deaths is explored. Figure 12 shows the mortality rate by year and race due to homicide deaths per 100,000 people in the 27 U.S. states from 2016 to 2020. The deaths for state, age group, sex, and ethnicity are aggregated by year and race. Unlike the trend due to suicide seen in Figure 6, Figure 12 indicates that the mortality rates for Blacks/African Americans are significantly higher than those for the other three races when other factors are not considered. When comparing the mortality rates due to suicide and homicide deaths, the mortality rates for Blacks/African Americans due to homicide across the five years are about triple the rates for suicide deaths observed in Figure 6, but the mortality rates for the other three races show the exact opposite trend.

Figure 12

HOMICIDE MORTALITY RATE PER 100,000 FROM 2016 TO 2020
BY RACE



Now consider the five-year trend in average mortality rate, considering race and ethnicity. Figure 13 and Figure 14 show the mortality rates for females and males, respectively, due to homicide deaths per 100,000 people in the 27 U.S. states from 2016 to 2020, broken down by year, ethnicity, and race. The deaths for state and age group are aggregated by year, ethnicity, and race. Consistent with the observations made in Figure 7 and Figure 8, the male mortality rate is overall higher than females across all races, and the mortality rate for non-Hispanics/Latinos is higher than for Hispanics/Latinos, except for several Asian or Asian American cases. Regardless of sex, the big gap in mortality rates occurs in the case of Blacks/African Americans, and the rates of Hispanics/Latinos and non-Hispanics/Latinos are relatively close across the other two races. Note that the data shows reported homicide deaths of female Hispanic/Latino Asians/Asian Americans and female Hispanic/Latino Native Americans were both zero in 2017, but the corresponding populations are 93,345 and 24,136.

Figure 13

FEMALE MORTALITY RATES PER 100,000 FROM 2016 TO 2020
BY YEAR, ETHNICITY AND RACE

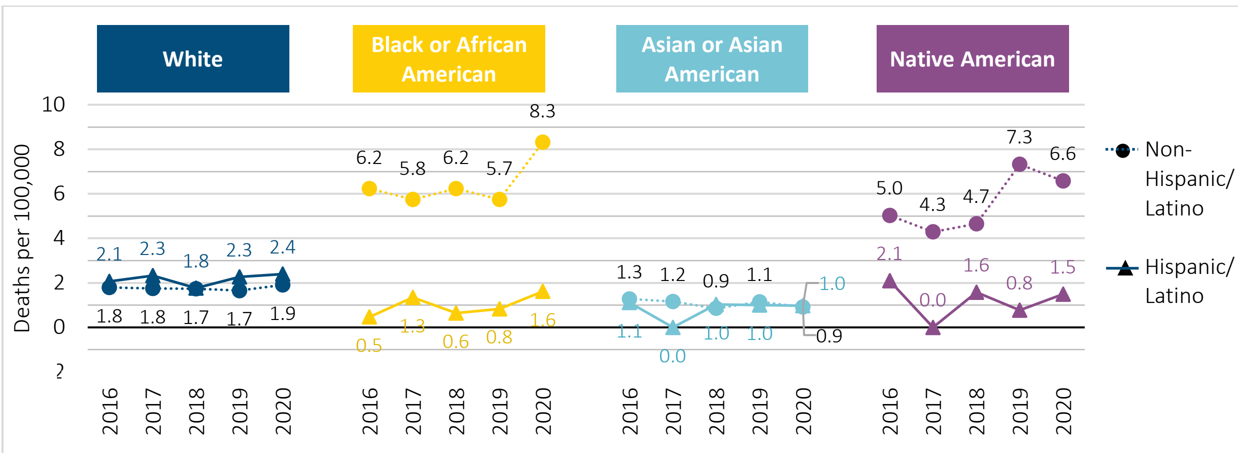
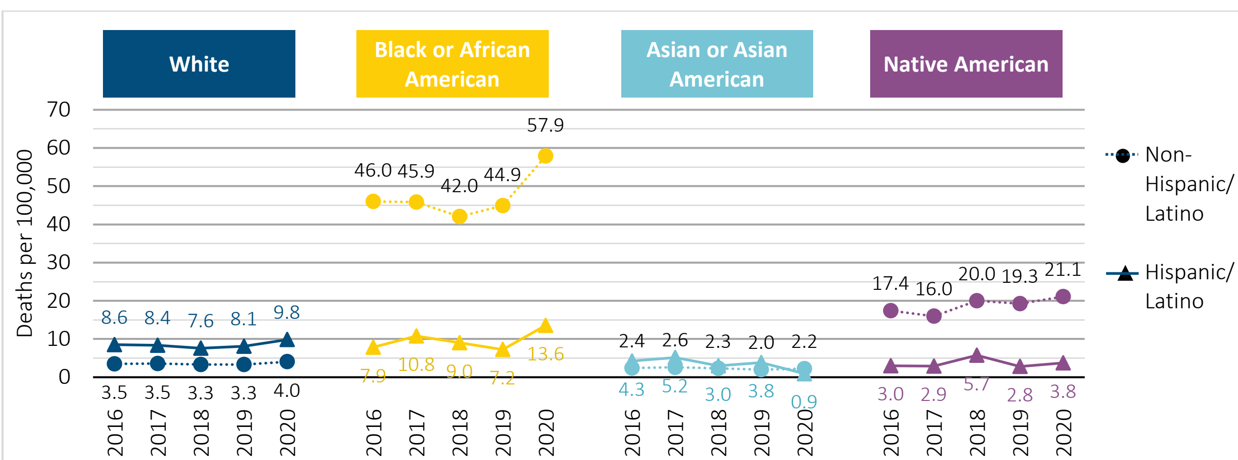


Figure 14

MALE HOMICIDE MORTALITY RATES PER 100,000 FROM 2016 TO 2020
BY YEAR, ETHNICITY AND RACE



3.3.3 "OTHER" MANNERS OF VIOLENT DEATH

In this subsection, the trend in mortality rates due to "other" manners of deaths is studied. These deaths include deaths due to legal intervention, unintentional firearm (self-inflicted, inflicted by others, and unknown who inflicted), and undetermined intent. Figure 15 shows the mortality rates by age group and sex due to "other" manner of deaths per 100,000 people in the 27 U.S. states over 2016–2020. The deaths for year, state, race, and ethnicity are aggregated by age group and sex. Although the trend of the average mortality rate is similar to the trend in suicide and homicide cases, the rates are much smaller.

Figure 15

“OTHER” MANNERS OF DEATH MORTALITY RATES PER 100,000 FROM 2016 TO 2020
BY AGE GROUP AND SEX

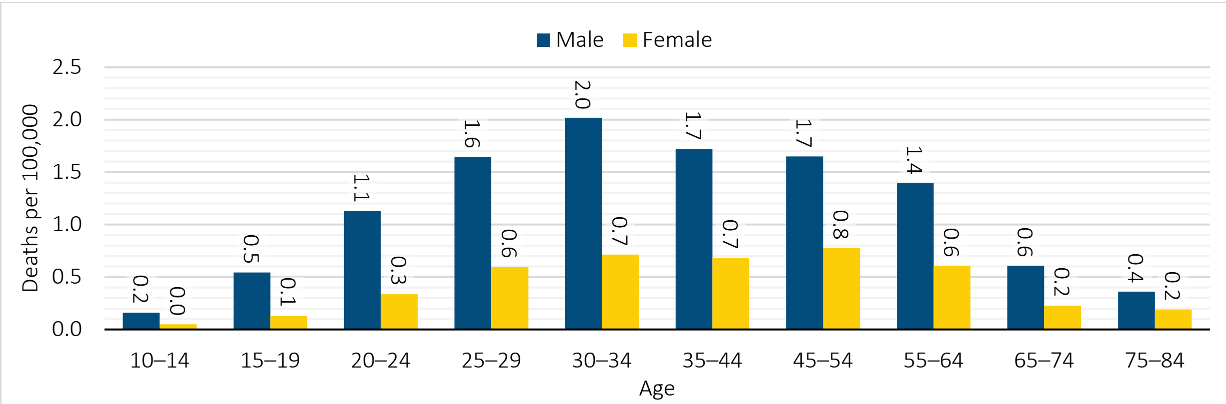


Figure 16 shows the mortality rates by age group and race due to “other” manners of deaths per 100,000 people in the 27 U.S. states from 2016 to 2020. The deaths for the year, state, sex, and ethnicity are aggregated by age group and race. Asians/Asian Americans have consistently lower mortality rates than the other three races across all age groups, and all death rates are much lower than suicide and homicide rates seen in Figure 4 and Figure 10. Although the deviation exists in Figure 16, note that the *y*-axis shows that all crude rates are less than 2.5 deaths per 100,000 people.

Figure 16

“OTHER” MANNERS OF DEATH MORTALITY RATES PER 100,000 FROM 2016 TO 2020
BY AGE GROUP AND RACE

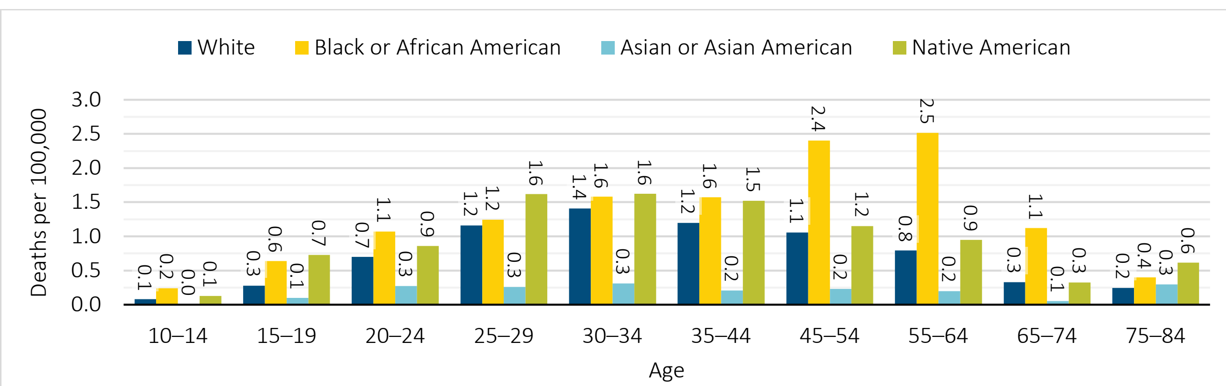
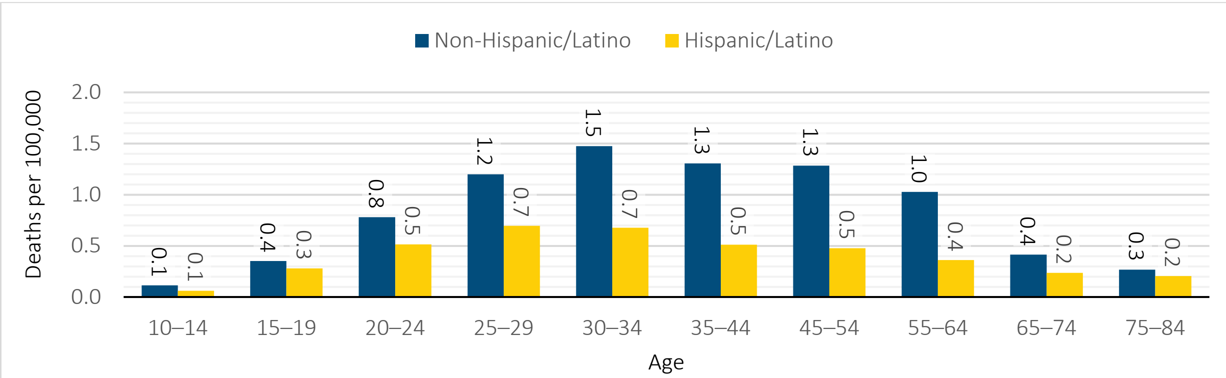


Figure 17 shows the mortality rates by age group and ethnicity due to “other” manners of deaths per 100,000 people in the 27 U.S. states over 2016–2020. The deaths for year, state, sex, and race are aggregated by age group and ethnicity. Like the suicide and homicide cases seen in Figure 2 and Figure 8, the mortality rates for Hispanics/Latinos are lower than those for the non-Hispanic/Latino group.

Figure 17

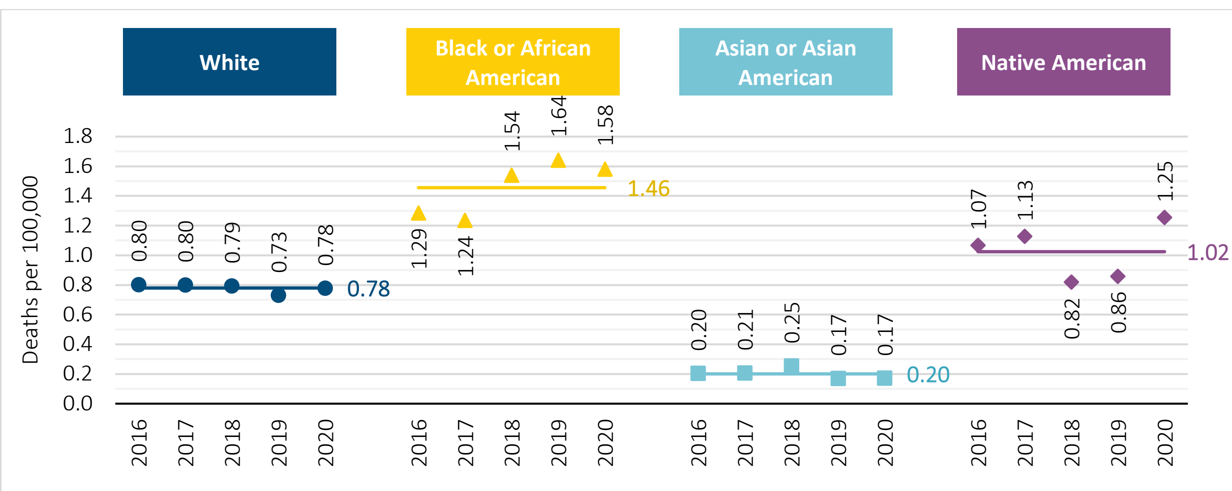
“OTHER” MANNERS OF DEATH MORTALITY RATES PER 100,000 FROM 2016 TO 2020 BY AGE GROUP AND ETHNICITY



Next, the influence of race on mortality due to “other” manners of death is investigated. Figure 18 shows the mortality rate by year and race due to “other” manners of deaths per 100,000 people in the 27 U.S. states over 2016–2020. The deaths for state, age group, sex, and ethnicity are aggregated by year and race. Although the highest mortality rate is about 1.64 death per 100,000 people, which occurred for Blacks/African Americans in 2019, this rate is based on 1,147 deaths in the population of 69,929,580 people. Native Americans in 2020 had about 1.25 deaths per 100,000 people, and this rate is based on 101 deaths in a population of 8,051,460. The standard deviation for Blacks/African Americans and Native Americans is about 1.8 deaths per 100,000 people, and so these values are considered insignificant compared to other rates in the same race category.

Figure 18

“OTHER” MANNERS OF DEATH MORTALITY RATES PER 100,000 FROM 2016 TO 2020 BY RACE



Now look at the five-year trends in mortality, considering race and ethnicity. Figures 19 and 20 show the mortality rates for females and males, respectively, by year, ethnicity, and race due to “other” manners of deaths per 100,000 people in the 27 U.S. states from 2016 to 2020. The deaths for the state and age group are aggregated by year, ethnicity, and race. The mortality rate remains low across all sex, racial, and ethnic groups. Overall, the Hispanic/Latino group has a lower average mortality rate than the non-Hispanic/Latino group, except for a small number of Asian/Asian American cases, but the difference is negligible. In addition, it is important to note that the deaths of female Hispanic/Latino and Asians/Asian Americans from 2017 to 2020 were zero, and this is also true for the corresponding males in 2016 and 2020. The deaths of female Hispanic/Latino Native Americans were also zero in 2017 and 2020. The deaths of female Hispanic/Latino Native Americans were also zero in 2017 and 2020. This is not surprising because these groups have much smaller populations than others.

Figure 19

FEMALE "OTHER" MANNERS OF DEATH MORTALITY RATES PER 100,000 FROM 2016 TO 2020
BY YEAR, ETHNICITY AND RACE

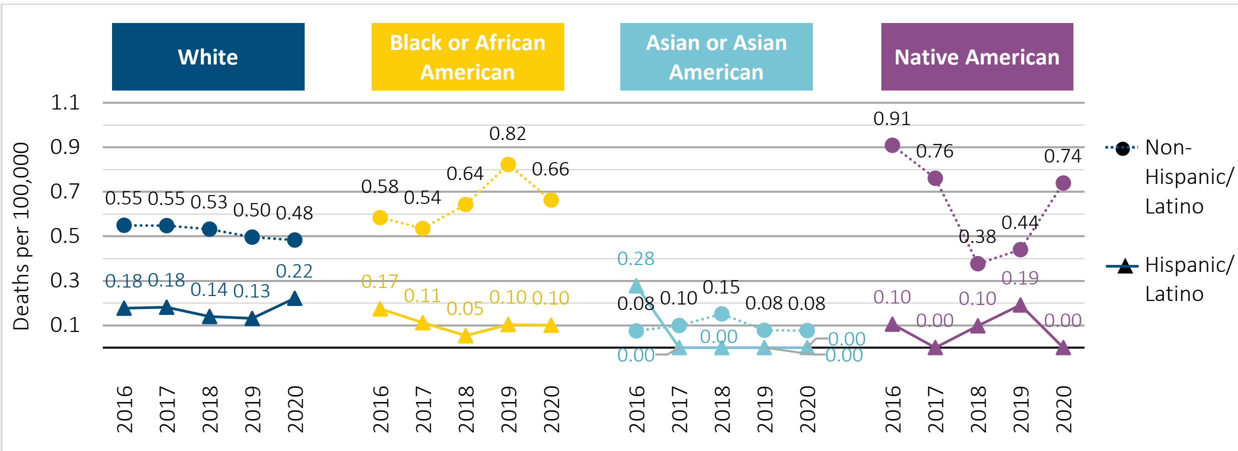
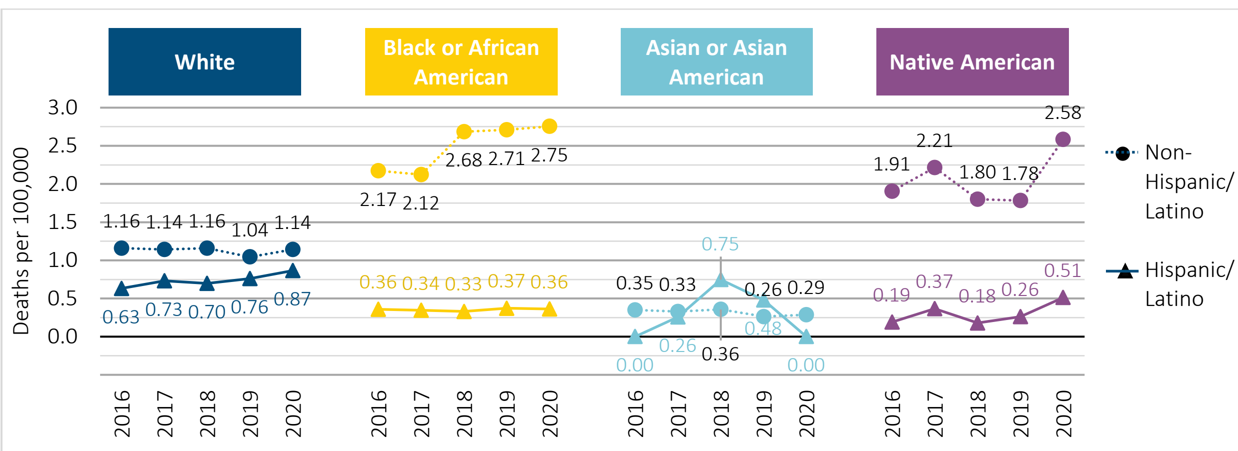


Figure 20

MALE "OTHER" MANNERS OF DEATH MORTALITY RATES PER 100,000 FROM 2016 TO 2020
BY YEAR, ETHNICITY AND RACE



3.4 MOTIVATION TO GO BEYOND EXPLORATORY DATA ANALYSIS

Statistical modeling approaches are essential in studying violent deaths in the U.S. because they offer a more rigorous and sophisticated analysis compared to EDA alone. Although EDA provides an initial understanding of the data and identifies patterns, statistical modeling allows for controlling other variables and isolating the effect of specific factors of interest. For instance, Figure 10 displays mortality rates by age group and race due to homicide deaths while disregarding other variables such as sex, ethnicity, and geographic location. Confounding variables, such as age, sex, and race, can potentially influence the relationship between variables of interest, such as ethnicity and violent deaths. Statistical modeling, such as regression analysis, can control for these confounding variables, enabling researchers to make more accurate conclusions.

Furthermore, statistical modeling approaches enable making predictions based on the model and test hypotheses. For example, one can examine whether the relationship differs between mortality rates for different types of violent deaths, such as homicides and suicides, and age, race, sex or geographic region. This information can help inform policies and interventions to reduce violent deaths.

Utilizing statistical modeling approaches enables systematic and rigorous examination of the relationships between violent deaths and the variables of interest while controlling for potential confounding variables. Although EDA is essential to understand the data, it cannot provide conclusive evidence for complex relationships between mortality rates and violent deaths. Therefore, statistical modeling approaches are necessary to provide reliable results for research and policy decisions.

Section 4: Statistical Analysis

This section discusses statistical modeling and the analysis of the results.

4.1 GLM MODELING

4.1.1 HIERARCHY OF THE MODELS WITH THE NAMING CONVENTION

The relationship between mortality due to violent manner of deaths and the sociodemographic covariates is examined for 27 U.S. states for 2016–2020. Data stratification is done on the basis of the manner of death to deal with the issue of high variability in the data (refer to the overdispersion, discussed in Appendix C). Nine negative binomial GLMs are used for modeling these violent deaths as a function of different combinations of independent variables listed in Table 4.

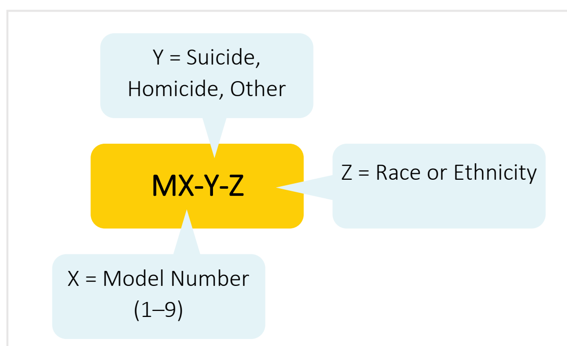
Table 4
SUMMARY OF GLM MODELS WITH THEIR CORRESPONDING FACTORS

MODEL NAME (FIRST 2 CHARACTERS)	FACTORS
M1	Year + State + Sex + Race (or Ethnicity) + Age Group
M2	State + Sex + Race (or Ethnicity) + Age Group
M3	Sex + Race (or Ethnicity) + Age Group
M4	Race (or Ethnicity) + Age Group
M5	Sex + Race (or Ethnicity)
M6	Race (or Ethnicity)
M7	Age Group
M8	Race (or Ethnicity) + Age Group + Age Group:Race (or Ethnicity)
M9	Age Group + Sex

The number of deaths is considered as a response variable, and the population is taken as an offset. The models are presented in order of hierarchy, as shown in Figure 21.

The naming convention for each model is defined as MX-Y-Z, where the first two characters denote the model number (M1–M9), the next character stands for the first initial of the manner of death (suicide, homicide, "other"), and the last character denotes if race or ethnicity is included in the model. For example, the model called M4-S-R would indicate the M4 model for suicide that includes the covariates age group and race, as shown in Table 4. Similarly, the model M4-S-E for suicide would consist of the age group and ethnicity. A list of all 54 model names is provided in Table D-1 in Appendix D.

Figure 21
HIERARCHY OF THE MODELS DEFINED AS MX-Y-Z



The model M8 includes the interaction term (i.e., M8-S-R) between age group and race. One can think about this interaction as that the influence of race (a factor or variable) on the number of violent deaths (response variable) varies based on the age group (another factor). Including the interaction between two factors allows for modeling all different combinations of the levels of these two variables. If the interaction terms are significant, results show evidence of the interaction effect. For example, whites in a specific age group may significantly influence the number of violent deaths reflected in the output. Similarly, including in the model the interaction between ethnicity and age group, checks whether there is a significant influence of ethnicity (factor) on the number of violent deaths (response variable) depending on the age group (another factor). Recall that checking for the interactions between factors is a standard approach in regression modeling.

Race and ethnicity are modeled separately because of data compatibility issues between the two data sources. Death counts and population counts are handled differently between the CDC's NVDRS and WONDER data sets. More details are provided in the next section.

4.1.2 SUMMARY OF THE VARIABLES USED IN THE GLM MODELING

The race and ethnicity variables are handled differently between the NVDRS and WONDER data sets. The NVDRS data include the deaths for race and/or ethnicity combined into eight categories, as shown in Table 5 and Table 6. However, the population estimates for race and ethnicity are mutually exclusive in the WONDER database. Thus, the modeling of race and ethnicity is handled separately.

The categories of the race variable are coded in R and ordered from largest to smallest based on the population size. Refer to Table 5 and Table 6 for additional details.

Table 5

SUMMARY OF NVDRS VARIABLES USED IN GLM INVOLVING RACE

Variable	Description	Variable in R	Coding in R
State	Alaska, Arizona, Colorado, Connecticut, Georgia, Indiana, Iowa, Kansas, Kentucky, Maine, Maryland, Massachusetts, Michigan, Minnesota, New Hampshire, New Jersey, New Mexico, North Carolina, Ohio, Oklahoma, Oregon, Rhode Island, South Carolina, Utah, Vermont, Virginia, Wisconsin	State	Alaska, Arizona, Colorado, Connecticut, Georgia, Indiana, Iowa, Kansas, Kentucky, Maine, Maryland, Massachusetts, Michigan, Minnesota, New Hampshire, New Jersey, New Mexico, North Carolina, Ohio, Oklahoma, Oregon, Rhode Island, South Carolina, Utah, Vermont, Virginia, Wisconsin
Year	2016, 2017, 2018, 2019, 2020	Year	2016, 2017, 2018, 2019, 2020
Sex	Male	Sex	Male
	Female		Female
Race/Ethnicity	White	Race	R1
	Black or African American		R2
	Asian/Pacific Islander		R3
	American Indian/Alaska Native		R4
Cause of Death	Legal intervention (by police or other authority)	Manner	Legal_Intervention
	Unintentional firearm–self-inflicted		UF_Selfinflicted
	Undetermined intent		UndeterIntent
	Homicide; terrorism homicide		Homicide
	Unintentional firearm–inflicted by other person; unintentional firearm–unknown who inflicted		Other
	Suicide or intentional self-harm; terrorism suicide		Suicide
Age Group	10–14, 15–19, 20–24, 25–29, 30–34, 35–44, 45–54, 55–64, 65–74, 75–84	AgeGroup	10–14, 15–19, 20–24, 25–29, 30–34, 35–44, 45–54, 55–64, 65–74, 75–84

Table 6

SUMMARY OF NVDRS VARIABLES USED IN GLM INVOLVING ETHNICITY

Variable	Description	Variable in R	Coding in R
State	Alaska, Arizona, Colorado, Connecticut, Georgia, Indiana, Iowa, Kansas, Kentucky, Maine, Maryland, Massachusetts, Michigan, Minnesota, New Hampshire, New Jersey, New Mexico, North Carolina, Ohio, Oklahoma, Oregon, Rhode Island, South Carolina, Utah, Vermont, Virginia, Wisconsin	State	Alaska, Arizona, Colorado, Connecticut, Georgia, Indiana, Iowa, Kansas, Kentucky, Maine, Maryland, Massachusetts, Michigan, Minnesota, New Hampshire, New Jersey, New Mexico, North Carolina, Ohio, Oklahoma, Oregon, Rhode Island, South Carolina, Utah, Vermont, Virginia, Wisconsin
Year	2016, 2017, 2018, 2019, 2020	Year	2016, 2017, 2018, 2019, 2020
Sex	Male	Sex	Male
	Female		Female
Race/Ethnicity	White, non-Hispanic	Ethnicity	E1
	Black or African American, non-Hispanic		E1
	Asian/Pacific Islander, non-Hispanic		E1
	American Indian/Alaska Native, non-Hispanic		E1
	Unknown race, non-Hispanic		E1
	Other/Unspecified, non-Hispanic		E1
	Two or more races, non-Hispanic		E1
	Hispanic		E2
Cause of Death	Legal intervention (by police or other authority)	Manner	Legal_Intervention
	Unintentional firearm–self-inflicted		UF_Selfinflicted
	Undetermined intent		UndeterIntent
	Homicide; terrorism homicide		Homicide
	Unintentional firearm–inflicted by other person; unintentional firearm-unknown who inflicted		Other
	Suicide or intentional self-harm; terrorism suicide		Suicide
Age Group	10–14, 15–19, 20–24, 25–29, 30–34, 35–44, 45–54, 55–64, 65–74, 75–84	AgeGroup	10–14, 15–19, 20–24, 25–29, 30–34, 35–44, 45–54, 55–64, 65–74, 75–84

Table 7 summarizes the data extracted from WONDER database with the corresponding coding in R. The one-to-one mapping in R is easily obtained for race. Four categories of race are considered, and they are ordered from the largest to smallest population size. Some age groups in WONDER are combined in R.

Table 8 shows the mapping between WONDER and R for the variable ethnicity. Two categories are used for ethnicity, and they are ordered based on the largest to smallest population size. The non-Hispanic/Latino group is used as the reference level.

The definition of age groups is based on the CDC report (Wilson et al. 2022). However, the age group 85+ has been removed because of its small size and uncertainty associated with the data collection for this age group.

Table 7

SUMMARY OF CDC WONDER STATE POPULATION BY RACE AND AGE GROUPS USED IN GLMS

Variable	Description	Variable in R	Coding in R
Race	White	Race	R1
	Black or African American		R2
	Asian or Pacific Islander		R3
	American Indian or Alaska Native		R4
Age Group	10–14 years, 15–19 years, 20–24 years, 25–29 years, 30–34 years, 35–39 years and 40–44 years, 45–49 years and 50–54 years, 55–59 years and 60–64 years, 65–69 years and 70–74 years, 75–79 years and 80–84 years	AgeGroup	10–14, 15–19, 20–24, 25–29, 30–34, 35–44, 45–54, 55–64, 65–74, 75–84

Table 8

SUMMARY OF CDC WONDER STATE POPULATION BY ETHNICITY AND AGE GROUPS USED IN GLMS

Variable	Description	Variable in R	Coding in R
Ethnicity	Not Hispanic or Latino	Ethnicity	E1
	Hispanic or Latino		E2
Age Group	10–14 years, 15–19 years, 20–24 years, 25–29 years, 30–34 years, 35–39 years and 40–44 years, 45–49 years and 50–54 years, 55–59 years and 60–64 years, 65–69 years and 70–74 years, 75–79 years and 80–84 years	AgeGroup	10–14, 15–19, 20–24, 25–29, 30–34, 35–44, 45–54, 55–64, 65–74, 75–84

4.1.3 ANALYSIS OF THE GLM MODELS

The following steps are undertaken in selecting the final most suitable GLM model among all models considered:

Step A: The chi-square goodness-of-fit test (GOF) is applied to all GLM models and those that passed these test are further considered.

Step B: Model diagnostics are examined for all models. The diagnostic tools include the half-normal plot of residuals and the density plot. Diagnostic plots are analyzed along with the findings in step A.

Step C: All models are additionally tested for overdispersion using the R function **check_overdispersion()** from the R package **performance** (0.10.3) (Lüdecke et al. 2021), and those models that did not pass the test are disregarded if possible.

Step D: The likelihood-ratio test and Akaike information criterion (see Appendix C) are used to compare the pairs of models in the selected subsets of model space.

Step E: The investigation and selection of the most suitable model are based on balancing consistency and the results across all the above steps.

Step F: The most suitable selected GLM model is used to explain the results and findings of this study.

All analyses were performed using R statistical software (version 4.1.2; R Core Team 2022). The parameters of the negative binominal GLM are estimated using R function **glm**() provided as part of the R package **stats** (4.2.2). Additional R packages used for data manipulation include **tidyverse** (1.3.2) and **tidyr** (1.2.1).

4.2. ANALYSIS OF THE GLM RESULTS

This section discusses proposed models that have been chosen based on several criteria outlined in Appendix C. Two models for each manner of death have been selected, and they are summarized in Table 9. The models that include "O" in the name have an additional factor indicating all other subcategories from manner of deaths, after suicide and homicide were considered. Model M8-S-R includes the interaction term between age group and race.

Table 9

MODELS CHOSEN FOR SUICIDE, HOMICIDE, AND OTHER

Model	Factors
M6-S-R	Race
M8-S-R	Age Group + Race + Age Group:Race
M5-S-E	Sex + Ethnicity
M2-H-R	State + Age Group + Sex + Race
M2-H-E	State + Age Group + Sex + Ethnicity
M2-O-R	State + Age Group + Sex + Race + Manner
M2-O-E	State + Age Group + Sex + Ethnicity + Manner

A summary of the results for each of these models as well as the interpretation of the results will be discussed in the following three subsections.

4.2.1 RESULTS FOR SUICIDE DEATHS

Two models for suicide are selected when race is considered: M6-S-R and M8-S-R. The summaries of the results for these models are shown in Table E-1 and Table E-3 in Appendix E, respectively. Both models passed the chi-square GOF test and the test for overdispersion. The half-normal plots in Figure G-1 and Figure G-2 show no outliers.

For the reader's convenience, the relative risks related to the baseline given the regression coefficients and *P* value for the significance of each factor in the M6-S-R model are reported in Table 10. The estimated coefficients are often interpreted on an exponential scale in terms of the relative risk for easy interpretation. The *P* value < 0.01 shows highly significant statistical results. All levels of the categorical variable race are significant determinants of suicide rates.

Table 10

RELATIVE RISKS AND COEFFICIENTS, *P* VALUES: SUICIDE-RACE, MODEL 6 (M6-S-R)

Variable	Relative Risk	Coefficient Estimate	*P* Value
Intercept	–	2.95	< 0.001
Race Black/African American (R2)	63%	−0.46	< 0.001
Race Asian/Asian American (R3)	76%	−0.27	< 0.001
Race Native American or Alaska Indigenous Resident (R4)	212%	0.75	< 0.001

Baseline for race is white (R1).

The reference level for suicide was category R1 (whites). This category was chosen because it is the largest racial group in the data, which is a common approach when none of the groups is representative of the whole dataset. The positive coefficient for Native Americans (R4) indicates higher suicide rates relative to whites. Native Americans historically experience higher rates of suicide because poorer health and socioeconomic outcomes as reported by the National Indian Council on Aging (2022). On the other hand, the negative coefficients for Blacks/African Americans (R2) and Asians/Asian Americans (R3) indicate lower suicide rates compared to the white category.

The authors interpret relative risk for model M6-S-R as follows:

- The highest risk of suicide is reported for Native Americans. A person of Native American descent is more than two times more likely (212%) to die by suicide than a white person, regardless of sex and age.
- The risk of suicide for Blacks/African American is, on average, about 63% that for whites regardless of sex and age.
- The risk of suicide for Asian or Asian American is, on average, about 76% that for whites regardless of sex and age.

A person of Native American descent is more than two times more likely to die by suicide than a white person.

The model M8-S-R is another model that is suitable for modeling the deaths due to suicide among all models considered. This model includes age group in addition to race as well as their interaction term. Table 11 shows the relative risk and the corresponding *P* value for the significance of each factor in the model. The results of this model show that the risk of suicide death increases by age group, from ages 10–14 to 45–54, but then slowly decreases toward older age groups. This study's results agree with the report published by Wilson et al. (2022), emphasizing that the rates of suicide are highest among adults aged 45–54.

The reference level for the age group is 10–14, which is picked automatically by the R software as the first level of an ordinal variable and is commonly used in the literature. However, it may not be the optimal group to use as a reference level, and the results can be transformed by multiplication. People in the age group 45–54 are almost eight times more likely to die by suicide compared to those aged 10–14. The risk of suicide varies by race and age. The risk of suicide across all age groups remains the highest for Native American across all age groups.

In the model containing the interaction terms, the interpretation is not straightforward. For example, to estimate the relative risk for a 45–54-year-old Asian/Asian American, one would need to multiply the relative risk associated with age 45–54, the relative risk associated with R3, and the relative risk associated with their interaction term, age 45–54, race R3. This calculation results in the relative risk of 374%. See Table E-3 for the full summary of coefficients in the model.

Table 11

RELATIVE RISKS AND *P* VALUES: SUICIDE-RACE, MODEL 8 (M8-S-R)

Variable	Relative Risk	*P* Value
Intercept	–	< 0.001
Age 15–19	406%	< 0.001
Age 20–24	617%	< 0.001
Age 25–29	662%	< 0.001
Age 30–34	682%	< 0.001
Age 35–44	710%	< 0.001
Age 45–54	777%	< 0.001
Age 55–64	717%	< 0.001
Age 65–74	587%	< 0.001
Age 75–84	710%	< 0.001
Race Black/African American (R2)	123%	0.072
Race Asian/Asian American (R3)	422%	< 0.001
Race Native American or Alaska Indigenous Resident (R4)	516%	< 0.001
Age 15–19: Race Black/African American (R2)	68%	0.005
Age 20–24: Race Black/African American (R2)	73%	0.024
Age 25–29: Race Black/African American (R2)	66%	0.002
Age 30–34: Race Black/African American (R2)	62%	0.001
Age 35–44: Race Black/African American (R2)	44%	< 0.001
Age 45–54: Race Black/African American (R2)	30%	< 0.001
Age 55–64: Race Black/African American (R2)	27%	< 0.001
Age 65–74: Race Black/African American (R2)	31%	< 0.001
Age 75–84: Race Black/African American (R2)	40%	< 0.001
Age 15–19: Race Asian/Asian American (R3)	29%	< 0.001
Age 20–24: Race Asian/Asian American (R3)	23%	< 0.001
Age 25–29: Race Asian/Asian American (R3)	18%	< 0.001
Age 30–34: Race Asian/Asian American (R3)	15%	< 0.001
Age 35–44: Race Asian/Asian American (R3)	10%	< 0.001
Age 45–54: Race Asian/Asian American (R3)	11%	< 0.001
Age 55–64: Race Asian/Asian American (R3)	13%	< 0.001
Age 65–74: Race Asian/Asian American (R3)	19%	< 0.001
Age 75–84: Race Asian/Asian American (R3)	28%	< 0.001
Age 15–19: Race Native American or Alaska Indigenous Resident (R4)	61%	0.027
Age 20–24: Race Native American or Alaska Indigenous Resident (R4)	50%	0.001
Age 25–29: Race Native American or Alaska Indigenous Resident (R4)	53%	0.003
Age 30–34: Race Native American or Alaska Indigenous Resident (R4)	49%	0.001
Age 35–44: Race Native American or Alaska Indigenous Resident (R4)	28%	< 0.001
Age 45–54: Race Native American or Alaska Indigenous Resident (R4)	21%	< 0.001
Age 55–64: Race Native American or Alaska Indigenous Resident (R4)	21%	< 0.001
Age 65–74: Race Native American or Alaska Indigenous Resident (R4)	29%	< 0.001
Age 75–84: Race Native American or Alaska Indigenous Resident (R4)	56%	0.089

Baseline for age is 10–14, for race is white (R1).

Model M5-S-E is the most suitable model selected for modeling suicide when the ethnicity and sex are considered as factors. See Table F-2 in Appendix F. The relative risk and the corresponding *P* value for the significance of each factor can be found in Table 12. The reference level for sex is female, and for ethnicity it is non-Hispanic/Latino. Both factors sex and ethnicity were found to be highly significant determinants of suicide deaths.

Males are more than three times more likely than females to die by suicide.

Figure 22 provides a graphical representation of the relative rates by ethnicity and sex, individually. Males are more than three times at higher risk of death from suicide than females considering all other factors being fixed. Han et al. (2016) reported an increase in suicide fatality rates among men as a result of the use of more lethal methods and higher intent to die. The authors' analysis finds that the risk of suicide among Hispanic/Latinos is about 72% that for the non-Hispanic/Latino population. A summary of the coefficients for this model is given in Table E-2.

Hispanic/Latinos are less likely than non-Hispanic/Latinos to die by suicide.

Table 12

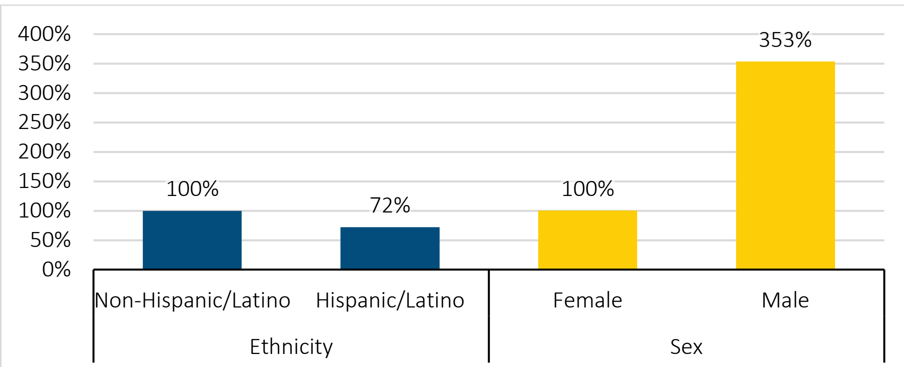
RELATIVE RISKS AND *P* VALUES: SUICIDE-ETHNICITY, MODEL 5 (M5-S-E)

Variable	Relative Risk	*P* Value
Intercept	–	< 0.001
Ethnicity Hispanic/Latino (E2)	72%	< 0.001
Sex Male	353%	< 0.001

Baseline for ethnicity is non-Hispanic/Latino (E1), for sex is female (F).

Figure 22

RELATIVE RISKS BY ETHNICITY AND SEX IN MODEL M5-S-E



4.2.2 RESULTS FOR HOMICIDE DEATHS

The M2-H-R is the most suitable model for modeling homicide rates when race is considered. This model includes state, sex, race, and age group as factors. Inclusion of states in the model allows comparing states in terms of the relative risk of violent deaths. All of the levels of factors are significant at the 0.05 significance level, except those for New Mexico and Vermont. Table 13 shows the relative risk and the corresponding *P* value for the significance of each factor in the model. See Table E-4 for a summary of the coefficients generated for this model.

The reference category for the state variable is Alaska. Alaska was picked automatically by the R software as the reference level because it is the first in the alphabet, but it is the worst for homicides. All estimate coefficients by state have a negative sign, indicating that the risk of dying because of homicide is lower in other states compared to that in Alaska. Thus, the relative risk for these states is below 1 relative to Alaska.

Coefficients for age groups, sex, and race are all significant, indicating that these variables and their categories are important factors of the violent manner of death due to homicide. Males have about a three times higher risk of dying because of homicide compared to women assuming that all other variables are fixed at a constant level.

Males are about 3 times more likely than females to die because of homicide.

Blacks/African Americans have nearly a 7.5 times higher risk of dying because of homicide relative to whites. Asians/Asian Americans and Native Americans have, respectively, a three and a four times higher risk of dying because of homicide relative to whites holding all other variables at constant level.

People of color are more likely than whites to die by homicide. For Black/African Americans, the risk is more than 7 times that of whites. For Native Americans, it is about 4 times, and for Asian/Asian Americans, it is about 3 times that of whites.

Homicide risk is high across all age groups relative to the reference level 10–14 years. Homicide risk increases from 4.5 for the age group 15–19 years to 6.6 for those 25–29. Based on that and including race as a factor, this analysis shows that the risk of homicide decreases starting at 4.7 for the 35–44 group up to a level of 2.0 for those in the age group 65–74. The age group 75–84 again experiences an increase of 2.8 in the relative risk. Somewhat similar trends in relative risk across age groups are found in the model that uses ethnicity as a factor. In this model the highest rate of homicide is reported for those 20–29. The young adults in the 20–29 group are more than seven times likely to die by homicide compared to those 10–14.

Similarly, the M2-H-E model is found to be the most suitable model for modeling mortality rates due to homicide when ethnicity is considered as one of the factors. Table 14 shows the relative risk and the corresponding *P* value for the significance of each factor in the model. See Table E-5 for a full summary of coefficients for this model.

All state coefficients are significant except those for Maryland, New Mexico, South Carolina, and Vermont. These coefficients have a negative sign, indicating a lower risk of dying because of homicide relative to Alaska when ethnicity is considered in the model along with sex and age groups. The Hispanic/Latino group has about a 1.2 higher risk of dying because of homicide compared to the non-Hispanic/Latino group that was used as a reference level. Males have about a 3.3 higher homicide risk compared to females holding all other variables constant. Finally, the risk of homicide varies by age groups. The age group 30–34 years old has a risk of 6.6, the highest risk among all age groups relative to the reference level 10–14 years. The lowest homicide risk is observed for the 65–74 group.

Hispanic/Latinos are 20% more likely than non-Hispanic/Latinos to die because of homicide.

Other studies on homicide reported similar findings. Haegerich et al. (2014) reported that the rate of homicide was nearly four times higher for men than women in 2010. Campbell (2002) emphasized that based on two large case-control studies of women in U.S. urgent care and emergency departments, risk factors for injury by an intimate partner were characteristics of the male perpetrator, including histories of arrest, substance abuse, poor education, unemployment, and expartner status.

The Morbidity and Mortality Weekly Report published by the CDC in 2021 (Petrosky et al. 2021) reported that intimate partner violence contributes to many homicides among American Indian and Alaska Indigenous Resident females based on NVDRS data for 2003–2018. Nearly half of the reported homicides were results of an argument for female victims, and 45% were due to intimate partner violence. The same report stated that the homicide rate was three times higher in Native American and Alaska Indigenous Resident males than females (12 versus 3.9), and the median age of Native American and Alaska Indigenous Resident victims was 32 years during the period 2003–2018.

Table 13

RELATIVE RISKS AND *P* VALUES: HOMICIDE-RACE, MODEL 2 (M2-H-R)

Variable	Relative Risk	*P* Value
Intercept	–	< 0.001
State Arizona	45%	< 0.001
State Colorado	41%	< 0.001
State Connecticut	27%	< 0.001
State Georgia	30%	< 0.001
State Indiana	49%	< 0.001
State Iowa	33%	< 0.001
State Kansas	50%	< 0.001
State Kentucky	54%	< 0.001
State Maine	43%	< 0.001
State Maryland	34%	< 0.001
State Massachusetts	20%	< 0.001
State Michigan	35%	< 0.001
State Minnesota	23%	< 0.001
State New Hampshire	41%	< 0.001
State New Jersey	21%	< 0.001
State New Mexico	90%	0.269
State North Carolina	33%	< 0.001
State Ohio	39%	< 0.001
State Oklahoma	51%	< 0.001
State Oregon	36%	< 0.001
State Rhode Island	37%	< 0.001
State South Carolina	41%	< 0.001
State Utah	30%	< 0.001
State Vermont	68%	0.083
State Virginia	26%	< 0.001
State Wisconsin	36%	< 0.001
Age 15–19	453%	< 0.001
Age 20–24	636%	< 0.001
Age 25–29	662%	< 0.001
Age 30–34	611%	< 0.001
Age 35–44	471%	< 0.001
Age 45–54	332%	< 0.001
Age 55–64	248%	< 0.001
Age 65–74	203%	< 0.001
Age 75–84	286%	< 0.001
Sex Male	297%	< 0.001
Race Black/African American (R2)	739%	< 0.001
Race Asian/Asian American (R3)	280%	< 0.001
Race Native American or Alaska Indigenous Resident (R4)	390%	< 0.001

Baseline for state is Alaska, for age is 10–14, for sex is female, and for race is white (R1).

Table 14

RELATIVE RISKS AND *P* VALUES: HOMICIDE-ETHNICITY, MODEL 2 (M2-H-E)

Variable	Relative Risk	*P* Value
Intercept	–	0.030
State Arizona	57%	< 0.001
State Colorado	46%	< 0.001
State Connecticut	39%	< 0.001
State Georgia	76%	0.001
State Indiana	73%	< 0.001
State Iowa	34%	< 0.001
State Kansas	64%	< 0.001
State Kentucky	75%	0.001
State Maine	37%	< 0.001
State Maryland	92%	0.320
State Massachusetts	28%	< 0.001
State Michigan	63%	< 0.001
State Minnesota	27%	< 0.001
State New Hampshire	38%	< 0.001
State New Jersey	38%	< 0.001
State New Mexico	103%	0.733
State North Carolina	70%	< 0.001
State Ohio	66%	< 0.001
State Oklahoma	79%	0.008
State Oregon	36%	< 0.001
State Rhode Island	52%	0.002
State South Carolina	97%	0.695
State Utah	28%	< 0.001
State Vermont	73%	0.272
State Virginia	53%	< 0.001
State Wisconsin	44%	< 0.001
Age 15-–19	521%	< 0.001
Age 20–24	754%	< 0.001
Age 25–29	754%	< 0.001
Age 30–34	655%	< 0.001
Age 35–44	481%	< 0.001
Age 45–54	332%	< 0.001
Age 55–64	232%	< 0.001
Age 65–74	155%	< 0.001
Age 75–84	175%	< 0.001
Sex Male	329%	< 0.001
Ethnicity Hispanic/Latino (E2)	121%	< 0.001

Baseline for state is Alaska, for age is 10–14, for sex is female (F), and for ethnicity is non-Hispanic/Latino (E1).

4.2.3 RESULTS FOR ALL "OTHER" MANNERS OF DEATH

This subsection analyzea the results for all "other" manners of violent deaths such as legal intervention by police or other authority (coded as LegaIntervention), unintentional firearm–self-inflicted (coded as UFSelfinficted) and unintentional firearm–inflicted by other person or unintentional firearm–unknown who inflicted (coded as Other). The legal intervention represents the reference level category for the manner of death variable.

The M2-O-R model is the most suitable model for modeling "other" manners of death when race is considered as one of the factors. Table 15 shows the relative risk and the corresponding *P* value for each factor in this model. A summary of the GLM output for this model is presented in Table E-6 in Appendix E.

Table 15

RELATIVE RISKS AND *P* VALUES: “OTHER”-RACE, MODEL 2 (M2-O-R)

Variable	Relative Risk	*P* Value
Intercept	–	< 0.001
State Arizona	38%	< 0.001
State Colorado	31%	< 0.001
State Connecticut	21%	< 0.001
State Georgia	16%	< 0.001
State Indiana	36%	< 0.001
State Iowa	35%	< 0.001
State Kansas	34%	< 0.001
State Kentucky	36%	< 0.001
State Maine	51%	< 0.001
State Maryland	266%	< 0.001
State Massachusetts	23%	< 0.001
State Michigan	59%	< 0.001
State Minnesota	26%	< 0.001
State New Hampshire	43%	< 0.001
State New Jersey	17%	< 0.001
State New Mexico	51%	< 0.001
State North Carolina	16%	< 0.001
State Ohio	23%	< 0.001
State Oklahoma	40%	< 0.001
State Oregon	33%	< 0.001
State Rhode Island	81%	0.186
State South Carolina	21%	< 0.001
State Utah	61%	< 0.001
State Vermont	105%	0.765
State Virginia	16%	< 0.001
State Wisconsin	23%	< 0.001
Age 15–19	126%	0.021
Age 20–24	162%	< 0.001
Age 25–29	197%	< 0.001
Age 30–34	232%	< 0.001
Age 35–44	170%	< 0.001
Age 45–54	172%	< 0.001
Age 55–64	151%	< 0.001
Age 65–74	105%	0.585
Age 75–84	123%	0.046
Sex Male	152%	< 0.001
Race Black/African American (R2)	239%	< 0.001
Race Asian/Asian American (R3)	165%	< 0.001
Race American Indian or Alaska Indigenous Resident (R4)	505%	< 0.001
Manner Other	77%	< 0.001
Manner UFSelfinflicted	66%	< 0.001
Manner UndeterIntent	199%	< 0.001

Baseline for state is Alaska, for age is 10–14, for sex is female (F), for race is white (R1), and for manner of death is legal intervention.

The results show that all categories for the manner of death are significant determinates of the number of death due to “other” manners. More specifically, the risk of dying because of undetermined intent is about two times higher than that for legal intervention. The risk of dying because of an unintentional firearm–inflicted by other person or an unintentional firearm–unknown who inflicted is about 77% of the risk for the legal intervention by police or other authority holding all other variables constant. The risk of dying because of unintentional firearm–self-inflicted is about 66% of the risk of dying because of legal intervention by police or other authority holding all other variables constant. In this model, males are one and a half times more likely to die because of these “other” manners of violent deaths compared to females.

For race, Native Americans have about a five times higher risk of dying because of these “other” manners of violent deaths compared to the whites. The relative risks for Blacks/African Americans and Asians/Asian Americans are 239% and 165%, respectively, compared to the risk for whites. The risk due to “other” manner of death also varies by age group, with the highest risk (232%) reported for the age group 30–34 and the lowest risk (105%) reported for the age group 65–74 relative to the reference level 10–14 years.

People of color are more likely than whites to die because of “other” manners of violent death—Native Americans are five times as likely, Blacks/African Americans are 2.4 times as likely, and Asian/Asian Americans are 1.65 times as likely as whites.

When ethnicity is considered as one of the factors, the most suitable model is found to be M2-O-E. Table 16 shows the relative risk and *P* value for each factor in the model. A full summary of results for this model can be found in Table E-7.

Although this model passed the chi-square GOF test and showed no overdispersion, a few categories of the factors are not significant at a 0.05 significance level. These include “other” manners of death and age groups 65–74 and 75–84 years old. The Hispanic/Latino group has about a 1.4 higher risk of dying because of “other” manner of violent deaths compared to the non-Hispanic/Latino group that was used as a reference level. Males have about 1.5 higher risk of dying because of “other” violent manners of death compared to the female population holding all other variables constant. Finally, the risk of “other” violent manner of death varies by age group. The age group 30–34 years old has a risk of 2.7, the highest risk among all age groups relative to the reference level 10–14 years. The lowest risk of dying due to “other” manner of violent deaths is observed for the age group 75–84 years old, but this risk is not statistically significant at a 0.05 significance level.

The literature related to the “other” manners of violent deaths is small. Wilson et al. (2022) reported 699 victims of legal intervention cause of death in the U.S. in 2019. Males represent the most victims in this group, particularly men aged 25–29 years. The authors found that the legal intervention death rate was the highest among Native American males followed by Black/African American males. A firearm was used in the majority of legal intervention deaths. Homicide is known to be the most frequent cause of legal intervention death. The three most common circumstances leading to legal intervention deaths were that the victim’s death was precipitated by another crime, the victim used a weapon in the incident, and the victim had a mental health or substance use problem (other than alcohol). A study by Barber et al. (2016) reported that for 2005–2012 the top three states with the most law enforcement homicides were George, Maryland, and North Carolina. Black/African American males are 21 times more likely to be killed by a police officer than white males. Krieger et al. (2015) examined trends from 1960 to 2010 for death by legal intervention by race and social class, and they found that high-income Blacks/African Americans are just as likely to be killed by police officers as low-income Blacks/African Americans.

Table 16

RELATIVE RISKS AND *P* VALUES: "OTHER"-ETHNICITY, MODEL 2 (M2-O-E)

Variable	Relative Risk	*P* Value
Intercept	–	0.001
State Arizona	36%	< 0.001
State Colorado	27%	< 0.001
State Connecticut	18%	< 0.001
State Georgia	16%	< 0.001
State Indiana	31%	< 0.001
State Iowa	29%	< 0.001
State Kansas	30%	< 0.001
State Kentucky	29%	< 0.001
State Maine	48%	< 0.001
State Maryland	294%	< 0.001
State Massachusetts	18%	< 0.001
State Michigan	59%	< 0.001
State Minnesota	18%	< 0.001
State New Hampshire	40%	< 0.001
State New Jersey	16%	< 0.001
State New Mexico	59%	< 0.001
State North Carolina	15%	< 0.001
State Ohio	19%	< 0.001
State Oklahoma	38%	< 0.001
State Oregon	28%	< 0.001
State Rhode Island	73%	0.102
State South Carolina	17%	< 0.001
State Utah	56%	< 0.001
State Vermont	99%	0.963
State Virginia	14%	< 0.001
State Wisconsin	17%	< 0.001
Age 15–19	135%	0.058
Age 20–24	184%	< 0.001
Age 25–29	229%	< 0.001
Age 30–34	275%	< 0.001
Age 35–44	205%	< 0.001
Age 45–54	201%	< 0.001
Age 55–64	177%	< 0.001
Age 65–74	122%	0.190
Age 75–84	126%	0.180
Sex Male	146%	< 0.001
Ethnicity Hispanic/Latino (E2)	142%	< 0.001
Manner Other	90%	0.234
Manner UFSelfinflicted	83%	0.048
Manner UndeterIntent	179%	< 0.001

Baseline for state is Alaska, for age is 10–14, for sex is female (F), for ethnicity is non-Hispanic/Latino (E1), and for manner of death is legal intervention.

4.3 SPOTLIGHT OF RELATIVE RISK BY RACE AND ETHNICITY ACROSS ALL THREE MANNERS OF DEATH

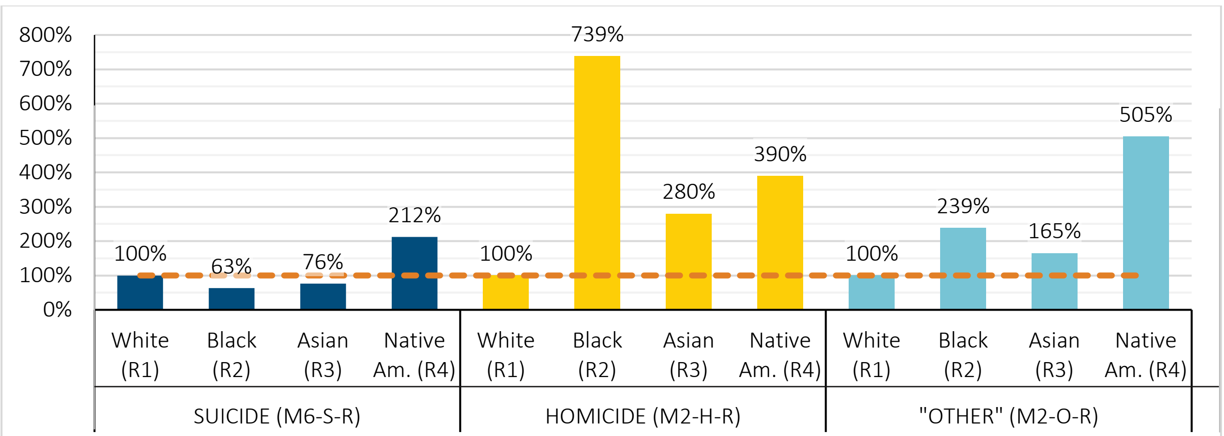
Race and ethnicity are important in the study of violent deaths in the U.S. because they are factors that have historically been linked to disparities in health outcomes, access to health care, and exposure to violence (Marzuk et al. 1995; Krieger et al. 2003; Petrosky et al. 2017; National Center for Health Statistics (U.S.) 2018). Understanding how race and ethnicity intersect with other factors such as age, sex, and geographical location can provide important insights into the underlying social determinants of violent deaths.

This subsection highlights the main findings on race and ethnicity across three GLM models: suicide, homicide, and "other" manners of death. The main results come from the models with a different set of factors, and so the interpretation of the relative risk for the factor of interest (race or ethnicity) will depend on whether other factors are considered in the model.

Figure 23 compares relative risks by race when other factors are fixed in the model. The highest risk of homicide is observed for Blacks/African Americans after controlling for age, sex, and state in the model. Native Americans have the highest risk of dying because of suicide when no other factors are considered in the model. Also, Native Americans have the highest rate of violent deaths due to "other" manners of death after controlling for age, sex, and state.

Figure 23

COMPARISON OF RELATIVE RISKS BY RACE WHEN OTHER FACTORS ARE FIXED IN THE MODELS

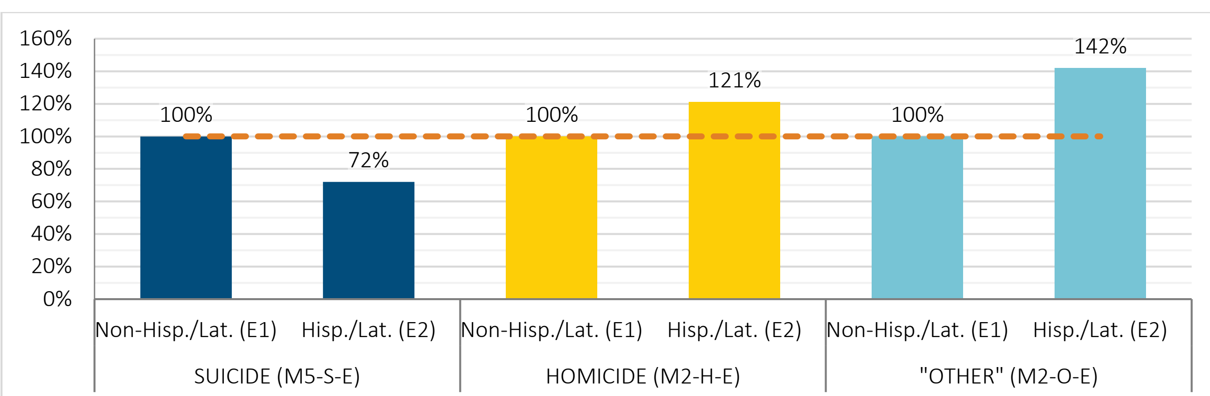


M6-S-R: no other fixed factors; M2-H-R: fixed factors are state, age, and sex; M2-O-R: fixed factors are state, age, sex, and manner.

Figure 24 compares relative risks by ethnicity when other factors are fixed in the model. Non-Hispanics/Latinos have a higher rate of suicide compared to their counterpart population after controlling for sex in the model. Hispanics/Latinos have the highest risk of death due to homicide compared to the non-Hispanic/Latino population when state, age, and sex are fixed in the model. Similarly, Hispanics/Latinos have the highest risk of death due to all other manners of death compared to the non-Hispanic/Latino population regardless of state, age, sex, and all categories of other manners of death.

Figure 24

COMPARISON OF RELATIVE RISKS BY ETHNICITY WHEN OTHER FACTORS ARE FIXED IN THE MODEL



M5-S-E: fixed factor is sex; M2-H-E: fixed factors are state, age, and sex; M2-O-E: fixed factors are state, age, sex, and manner.

Section 5: Study Limitations

The findings in this report are subject to several limitations. First, the NVDRS data are incorporated in the GLM modeling for 27 states for the period 2016–2022, representing about 43% of the total U.S. population. The biggest sample size over the past five years of data is represented by this selection of states because the NVDRS relies on state and local agencies to provide data; thus, the data quality and completeness vary across states and jurisdictions. The largest states based on the population are Texas, California, Florida, and New York, but they are not included in the NB-GLMs because of lack of the NVDRS data for this period. However, these top states often appear in news reports. The NB-GLM results would have more credibility if at least these four states were included in the modeling. Ideally, the available NVDRS data would include all states for a continuous period. However, the authors recognize that the availability and completeness of data depend on partnerships among the NVDRS and local health departments, vital statistics registrars' offices, coroners, medical examiners and law enforcement personnel.

Second, mortality rates are examined for a limited number of factors, including sex, age group, race, and ethnicity. However, violent deaths can arise from various causes, such as drug use, mental health issues, access to firearms, and social status. Although the data contain some of these factors, many of their values are missing. Expanding the scope of factors considered may affect the results, but it could also offer greater insights into the patterns of violent deaths in each state. For instance, including factors such as the poverty rate, state-level unemployment or other socioeconomic variables in this study's models may lead to significantly different results. Several previous studies linked mortality to social factors such as poverty or education. Barbieri (2020) developed the Socioeconomic Index Score by U.S. county and provided the county-level socioeconomic data related to the socioeconomic variables used in the scoring process as well as the county population during the period 1982–2018. Galea et al. (2011) found that the number of deaths attributable to social factors in the U.S. is comparable to the number attributed to pathophysiological and behavioral causes. One of the limitations of the NVDRS data is the lack of applicable socioeconomic variables associated with individual records. Thus, an opportunity for future studies is to consider other data sources (e.g., U.S. Bureau of Labor Statistics or U.S. Census Bureau) along with NVDRS data when modeling violent causes of death by race, ethnicity, and socioeconomic variables.

Third, some uncertainty surrounds the population data obtained from the WONDER database. The population data were aggregated by a preselected set of variables. Access to the raw data would facilitate understanding better the data structure.

Fourth, the small population for a given case when calculating the crude mortality rate could be problematic. For example, in 2020 the reported male-Asian/Asian American and Hispanic/Latino population for age 80–84 in Arkansas is only two people, but there is only one death due to suicide. Based on the NVDRS (2021) Data Analysis Guide for RAD Users_2021 document provided by the CDC, "For NVDRS, the report of counts and rates should be limited to instances where death counts are sufficiently large. Small numbers of events can vary considerably over time, resulting in unstable measurement, and could also pose concerns with respect to confidentiality and identifiable data."

Fifth, the authors have exclusively utilized negative binomial models after considering Poisson GLM models and quasi-Poisson GLM models because of their suitability for overdispersed count data, as well as advantages such as their flexibility in making fewer assumptions, the ease of interpreting their coefficients such as the impact of each factor on the response variable, and their ease of implementation. However, relying solely on negative binomial models may lead to model selection bias, where the choice of model is influenced by preconceived notions or biases. Furthermore, not comparing alternative models can hinder the identification of the best fit for the data. Thus, although the authors' models are useful for examining the relationship between the considered factors and the number of violent deaths, they should not be used to predict violent deaths.

Finally, the authors note several potential sources of modeling bias:

- Overdispersion: In many cases the variance may be much higher than the mean of the count variable, a phenomenon known as overdispersion. This can lead to biased parameter estimates and incorrect inference. This study adopts negative binomial models, which are more flexible than the Poisson regression models that require the variance of the count variable to equal the mean. The authors also utilize the goodness-of-fit test via the Akaike information criterion, the likelihood-ratio test, and diagnostic plots to select the most suitable model among all models considered and check the model assumptions to avoid overdispersion. Therefore, the potential modeling bias due to overdispersion is reduced.

- Temporal dependence: In modeling deaths one may find temporal dependence, where the number of deaths in previous years may influence the number of deaths in a given year. If this dependence is not accounted for, it can lead to biased estimates of the model parameters. However, in this study the variable year seems insignificant. Thus, the potential modeling bias due to temporal dependence is not a concern in this study.

- Spatial dependence: Deaths may also exhibit spatial dependence, where the number of deaths in neighboring areas may influence the number of deaths in one area. This spatial dependence can also lead to biased parameter estimates if not properly accounted for in the model. The most appropriate of the authors' models for homicide deaths and "other" manners of deaths include state as a factor, which accounts for the spatial dependence. We found that the models selected for suicide deaths did not include state as a factor, while other factors were in the models. This may be because the nature of the events is not highly associated with geographical locations..

- Measurement error: Measurement errors may exist in the count variable due to underreporting or misclassification of deaths, or in the population size, which can lead to biased parameter estimates. Unfortunately, the authors were not able to address measurement error in this study.

In summary, the authors strove to address potential sources of modeling bias regarding overdispersion, temporal, and spatial dependence to ensure accurate and reliable results.

Section 6: Discussion and Conclusion

This research study's statistical GLM models were built to model the mortality rates due to violent manners of death related to sex, race, ethnicity, and age in 27 U.S. states from 2016 to 2020. The study answered the authors' three research questions, summarized as follows:

- The mortality rates for all categories of violent deaths differ between underrepresented and white communities in the U.S.
- Sex, age, race, and ethnicity represent significant determinants of violent deaths in the U.S.
- When modeling "other" manners of deaths with race, age, sex, and state, all subcategories represent significant factors of "other" violent deaths. Categories of "other" manners of death include (1) legal intervention, (2) unintentional firearm–self-inflicted, (3) undetermined intent, and (4) other unintentional firearm deaths (inflicted by another person or unknown who inflicted). On the other hand, when ethnicity is considered along with age, sex, and state, unintentional firearm deaths are not significant factors of "other" deaths.

By employing the GLM models in this study, the authors can determine specific risk factors across all categories of race, ethnicity, sex, and age groups for each manner of death category. The authors can also compare states and determine which state has a significant number of deaths for a specific manner of violent deaths. For example, New Mexico and Vermont do not contribute significantly to homicide deaths when race is considered in the model. When ethnicity is included as a factor, Maryland, New Mexico, South Carolina, and Vermont do not contribute significantly to homicide deaths compared to all other states in the sample.

This study is essential to the literature on mortality studies of violent deaths in the U.S. The modeling approach that involves the GLM models is unique in this type of investigation using NVDRS data. Several limitations of this study have been listed, and so it is essential to carefully use the results when attempting to generalize these findings outside the sample of states considered in this study.

Additional research is required to understand the violent manners of death in the remaining states when complete longitudinal NVDRS data are available for all 50 states. Another research topic would include the impact of the COVID-19 pandemic on violent deaths. Several studies have already explored violent deaths with a specific focus on suicide (John et al. 2020) or homicide (Calderon-Anyosa and Kaufman 2021). However, these studies can be extended to investigate all manners of death, focusing on race, ethnicity, and other socioeconomic factors such as income, education, and occupation.

Section 7: Acknowledgments

The researchers express their deepest gratitude to those whose efforts and guidance enable successful completion of this project: the Project Oversight Group and others for their diligent work overseeing, reviewing and editing this report for accuracy and relevance.

Project Oversight Group members:

Mark Spong, FSA, CERA, MAAA, Chair

Dorothy Andrews, PhD, ASA, MAAA, CSPA

Kwabena Boamah-Acheampong, ASA, MAAA

Paul Boegel, ASA

Zara D'Antignac, ASA

Simone Davis, FSA, MAAA, CERA

Jean-Marc Fix, FSA, MAAA

Jevon Holder, ASA, MAAA

Ed Hui, FSA, MAAA, CFA

Jeffrey L. Johnson, FSA, MAAA

Uchechi Mitchell, Phd, MSPH

Ellyn M. Russo, MS

Tyler Yancey, ASA, CERA

At the Society of Actuaries:

Lisa A. Schilling, FSA, EA, FCA, MAAA, Senior Research Actuary

Appendix A: Specific Examples of Violent Manners of Death

Specific examples of violent manners of death are further described (Centers for Disease Control and Prevention, 2021) in this appendix.

(1) Suicide Death

Specific scenarios that should be classified as suicide:

1. A person engaged in a suicidal act, then changed his or her mind, but still died as a result of the act.
2. A person intending only to injure rather than kill him- or herself (e.g., someone shot him- or herself in the leg with intent to injure but severed the femoral artery and died).
3. Assisted suicide involving passive assistance to the decedent (e.g., supplying only means or information needed to complete the act).
4. Intentional, self-inflicted deaths committed while under the influence of a mind-altering substance taken voluntarily.
5. Intentional, self-inflicted deaths committed while mentally ill (e.g., acute psychotic episodes that may impair a person's judgment).
6. According to the CDC NVDRS manual, the following examples of deaths are classified as undetermined where suicide could be suspected but it could not be determined.

Examples of cases to classify as undetermined for the NVDRS:

1. Victim died of a drug overdose, and it could not be determined if it was unintentional (i.e., accidental) or a suicide.
2. Victim died from a fall off a cliff, and it could not be determined if it was unintentional or a suicide.
3. The victim was found in their home and had died of a head trauma sustained in a fall. Foul play was not ruled out. (Note: It may be unclear from narrative information what manners are being considered. In this case, there is the possibility of homicide and unintentional manners.)

Specific scenarios that should not be classified as suicide (the preferred category is shown in parentheses):

1. The physical consequences of chronic substance abuse, including alcohol or drugs (natural death).
2. Acute substance abuse including alcohol or drugs with less than a preponderance of evidence of using the substance(s) with intent to harm oneself (undetermined or unintentional injury death).
3. Death as a result of autoerotic behavior, for example, self-strangulation during sexual activity (unintentional injury death).

(2) Homicide Death

Specific scenarios that should be classified as homicide:

1. Deaths when the suspect intended to only injure rather than kill the victim.
2. Deaths resulting from heart attacks induced when someone uses force or power against the decedent.
3. A death resulting from a weapon that discharges unintentionally while being used to control or frighten the victim.
4. Deaths that result when a person kills an attacker in self-defense.
5. Deaths labeled "justifiable homicides" where the person committing the homicide was not a law enforcement officer.

6. Death that results from a variation of Russian roulette where one person aims a partially loaded gun at another person and pulls the trigger knowing that there is at least some chance that the gun would fire.
7. Death attributed to "child abuse" without an intent being specified.
8. Death of a child after birth that results from a direct injury due to violence sustained prior to birth.
9. Death that results from an intentional act of neglect or omission by one person against another.

Specific scenarios that should not be classified as homicide (the preferred category is shown in parentheses):

1. "Vehicular homicide" without a preponderance of evidence of intent to use force against another (unintentional injury).
2. Hunting accident with a gun (unintentional firearm injury).
3. Accidental deaths at shooting ranges (unintentional firearm injury).
4. A youth kills someone by playing with a gun he or she believes is unloaded (unintentional firearm injury).
5. Deaths that take place in combat in declared or undeclared wars (operation of war; not collected by the NVDRS).
6. Death of a child after birth that results indirectly from violence sustained by its mother before its birth, for instance, a death from prematurity following premature labor brought on by violence (coded as "condition originating in the perinatal period"; not collected by the NVDRS).
7. Accidental poisoning deaths due to illegal or prescription drug overdose, even when the person who provided those drugs was charged with homicide (unintentional deaths not involving firearms are outside the scope of the NVDRS; a death of this type might be within the scope of "undetermined manner of death," below, if it is impossible to determine whether the death was intentional or unintentional).

(3) Unintentional Firearm Injury Death

Specific scenarios that should be classified as unintentional firearm deaths:

1. Celebratory firing that was not intended to frighten, control or harm anyone.
2. A person shoots him- or herself when using a gun to frighten, control or harm another person.
3. A child less than the age of six shoots him- or herself or another person.
4. A soldier is shot during field exercises in peacetime.
5. A person mistakenly thinks a gun is unloaded and shoots him- or herself or another person while fooling around with it.
6. A child who dies after birth from an unintentional firearm injury that is sustained prior to birth, that is, in utero.

Specific scenarios that should not be classified as unintentional firearm deaths (the preferred NVDRS category is shown in parentheses):

1. A person unintentionally shoots someone while defending him- or herself against an aggressor (homicide).
2. A person unintentionally shoots another person while using a gun to commit a crime (homicide).
3. Firearm injuries caused by unintentionally striking a person with the firearm, for example, by dropping it on someone's head, rather than with a projectile fired from the firearm (potential homicide or unintentional).
4. Unintentional injuries from non-powder guns such as BB, pellet, and other compressed air or gas-powered guns (outside of system scope).

(5) Legal Intervention Death

The following scenarios fall within the definition of legal intervention deaths in the NVDRS:

1. Incidents in which the decedent was killed while fleeing from or being pursued by law enforcement, including some scenarios in which the victim was not directly injured by law enforcement officers. Examples include the following:

– Victim's death resulting from a car crash while being pursued by law enforcement in a high-speed chase.

– Victim's death resulting from attempting to escape law enforcement contact or arrest (e.g., victim runs away from officers, unintentionally falls off a bridge and dies).

– Death resulting from a victim being killed by another person unrelated to the event while being pursued by law enforcement (e.g., a motorist hits and kills a victim who was being pursued by law enforcement).

2. Incidents in which the decedent died as the result of force applied by law enforcement officers without clear lethal intent (e.g., restraint or use of a typically nonlethal weapon such as a taser).
3. "Justifiable" and "criminal" homicides meeting the above definition.
4. Bystanders who are inadvertently killed by law enforcement acting in the line of duty by mechanisms such as firearms, explosives, blunt objects (e.g., batons), sharp objects or personal weapons.

Appendix B: Data Summary Table

Table B-1

AVAILABILITY OF STATE DATA EACH YEAR

Year	States
2020	Alabama, Alaska, Arizona, Arkansas, California, Colorado, Connecticut, Delaware, District of Columbia, Georgia, Idaho, Illinois, Indiana, Iowa, Kansas, Kentucky, Louisiana, Maine, Maryland, Massachusetts, Michigan, Minnesota, Mississippi, Missouri, Montana, Nebraska, Nevada, New Hampshire, New Jersey, New Mexico, New York, North Carolina, North Dakota, Ohio, Oklahoma, Oregon, Pennsylvania, Puerto Rico, Rhode Island, South Carolina, South Dakota, Tennessee, Texas, Utah, Vermont, Virginia, Washington, West Virginia, Wisconsin, Wyoming
2019	Alabama, Alaska, Arizona, California, Colorado, Connecticut, Delaware, District of Columbia, Georgia, Hawaii, Illinois, Indiana, Iowa, Kansas, Kentucky, Louisiana, Maine, Maryland, Massachusetts, Michigan, Minnesota, Missouri, Montana, Nebraska, Nevada, New Hampshire, New Jersey, New Mexico, North Carolina, North Dakota, Ohio, Oklahoma, Oregon, Pennsylvania, Puerto Rico, Rhode Island, South Carolina, Utah, Vermont, Virginia, Washington, West Virginia, Wisconsin, Wyoming
2018	Alabama, Alaska, Arizona, California, Colorado, Connecticut, Delaware, District of Columbia, Georgia, Illinois, Indiana, Iowa, Kansas, Kentucky, Louisiana, Maine, Maryland, Massachusetts, Michigan, Minnesota, Missouri, Nebraska, Nevada, New Hampshire, New Jersey, New Mexico, New York, North Carolina, Ohio, Oklahoma, Oregon, Pennsylvania, Puerto Rico, Rhode Island, South Carolina, Utah, Vermont, Virginia, Washington, West Virginia, Wisconsin
2017	Alaska, Arizona, California, Colorado, Connecticut, Delaware, District of Columbia, Georgia, Illinois, Indiana, Iowa, Kansas, Kentucky, Maine, Maryland, Massachusetts, Michigan, Minnesota, Nevada, New Hampshire, New Jersey, New Mexico, New York, North Carolina, Ohio, Oklahoma, Oregon, Pennsylvania, Puerto Rico, Rhode Island, South Carolina, Utah, Vermont, Virginia, Washington, West Virginia, Wisconsin
2016	Alaska, Arizona, Colorado, Connecticut, Georgia, Hawaii, Illinois, Indiana, Iowa, Kansas, Kentucky, Maine, Maryland, Massachusetts, Michigan, Minnesota, New Hampshire, New Jersey, New Mexico, New York, North Carolina, Ohio, Oklahoma, Oregon, Pennsylvania, Rhode Island, South Carolina, Utah, Vermont, Virginia, Washington, Wisconsin
2015	Alaska, Arizona, Colorado, Connecticut, Georgia, Hawaii, Kansas, Kentucky, Maine, Maryland, Massachusetts, Michigan, Minnesota, New Hampshire, New Jersey, New Mexico, New York, North Carolina, Ohio, Oklahoma, Oregon, Rhode Island, South Carolina, Utah, Vermont, Virginia, Wisconsin
2014	Alaska, Colorado, Georgia, Kentucky, Maryland, Massachusetts, Michigan, New Jersey, New Mexico, North Carolina, Ohio, Oklahoma, Oregon, Rhode Island, South Carolina, Utah, Virginia, Wisconsin
2013	Alaska, Colorado, Georgia, Kentucky, Maryland, Massachusetts, New Jersey, New Mexico, North Carolina, Ohio, Oklahoma, Oregon, Rhode Island, South Carolina, Utah, Virginia, Wisconsin
2012	Alaska, Colorado, Georgia, Kentucky, Maryland, Massachusetts, New Jersey, New Mexico, North Carolina, Ohio, Oklahoma, Oregon, Rhode Island, South Carolina, Utah, Virginia, Wisconsin
2011	Alaska, Colorado, Georgia, Kentucky, Maryland, Massachusetts, New Jersey, New Mexico, North Carolina, Ohio, Oklahoma, Oregon, Rhode Island, South Carolina, Utah, Virginia, Wisconsin
2010	Alaska, Colorado, Georgia, Kentucky, Maryland, Massachusetts, New Jersey, New Mexico, North Carolina, Oklahoma, Oregon, Rhode Island, South Carolina, Utah, Virginia, Wisconsin
2009	Alaska, Colorado, Georgia, Kentucky, Maryland, Massachusetts, New Jersey, New Mexico, North Carolina, Oklahoma, Oregon, Rhode Island, South Carolina, Utah, Virginia, Wisconsin
2008	Alaska, Colorado, Georgia, Kentucky, Maryland, Massachusetts, New Jersey, New Mexico, North Carolina, Oklahoma, Oregon, Rhode Island, South Carolina, Utah, Virginia, Wisconsin
2007	Alaska, Colorado, Georgia, Kentucky, Maryland, Massachusetts, New Jersey, New Mexico, North Carolina, Oklahoma, Oregon, Rhode Island, South Carolina, Utah, Virginia, Wisconsin
2006	Alaska, Colorado, Georgia, Kentucky, Maryland, Massachusetts, New Jersey, New Mexico, North Carolina, Oklahoma, Oregon, Rhode Island, South Carolina, Utah, Virginia, Wisconsin
2005	Alaska, Colorado, Georgia, Kentucky, Maryland, Massachusetts, New Jersey, New Mexico, North Carolina, Oklahoma, Oregon, Rhode Island, South Carolina, Utah, Virginia, Wisconsin
2004	Alaska, Colorado, Georgia, Maryland, Massachusetts, New Jersey, North Carolina, Oklahoma, Oregon, Rhode Island, South Carolina, Virginia, Wisconsin
2003	Alaska, Maryland, Massachusetts, New Jersey, Oregon, South Carolina, Virginia

Appendix C: Methodology

Actuarial literature suggests various statistical procedures for estimating parameters in the claim count or frequency models (Ismail and Jemain 2007). The generalized linear models (GLMs) originally introduced by Nelder and Wedderburn (1972) are widely used in the actuarial field. The most popular GLMs for modeling count data are Poisson, quasi-Poisson, and negative binomial as well as their zero-inflated versions.[1] Haberman and Renshaw (1996) presented several examples of applications of the GLMs in a wide range of actuarial problems involving mortality, multiple-state models, lapses, premium rating, and reserving. This study explores GLMs for modeling count data due to violent manner of death with a specific focus on the negative binomial GLM, which allows for handling excessive overdispersion as observed in the NVDRS data.

C.1 BACKGROUND OF GLM

The GLMs are more flexible models than ordinary linear regression models because the distribution of the response variable is assumed to be one of the exponential family distributions such as Poisson, binomial, negative binomial or gamma. The distribution of count-based data is usually characterized by discrete non-negative integers and is often highly skewed to the right. The GLM is composed of three components: a random component, a systematic component, and a link function (Agresti 2018). The random component assumes the type of distribution of the response variable Y_i, for $i = 1, \dots, N$, in the modeling count-based observations. The systematic component defines the independent variables as a *p*-dimensional vector of factors denoted as $\boldsymbol{X}_{\boldsymbol{i}}' = \left(X_{1i}, \dots, X_{pi}\right)$. The linear combination of the product of the factors $\boldsymbol{X}_{\boldsymbol{i}}'$ and the vector of the model parameters ($\boldsymbol{\beta}$) represents the right-hand side of the model equation. In contrast, the left-hand side of the model equation includes the response variable. The link function links the random component to the systematic component by relating the mean of the response variable to the factors. In other words, the link function defines the random component by a function associated with the systematic component. The vector of the model coefficients ($\boldsymbol{\beta}$) is estimated using the maximum likelihood method. The R function **glm**(), as part of the R package stats (R Core Team 2022), is used to analyze of the GLMs for count data. One of the biggest issues relevant to the analysis of count data is related to overdispersion. Overdispersion occurs when $\mathrm{var}(Y_i)$ is greater than $E(Y_i)$**.** The negative binomial GLM (NB-GLM) allows for modeling the count data with high overdispersion. Although Poisson regression assumes that the response variable's mean and variance are equal, the negative binomial allows for modeling $(Y_i) = \emptyset E(Y_i)$, where $\emptyset > 1$ is a parameter that is estimated. Thus, the NB-GLM model allows for modeling the response that has significantly more variance than the mean, as observed in the NVDRS data.

It is important that the issue of overdispersion is not present in the selected NB-GLM model; that is, the observed variance in the data is not higher than the variance of a theoretical model. The R package **performance** (Lüdecke et al. 2021) provides the function **check_overdispersion()**, which uses the Pearson chi-squared statistics for checking overdispersion in the model. For more about modeling with overdispersion, the reader is referred to the book by Gelman and Hill (2006).

C.2 THE NEGATIVE BINOMIAL GLM

Let Y_i, $i = 1, \dots, N$ be the random variables for the number of death due to a violent event, and its realizations are denoted as $y_i = 0,1,2 \dots$. Let $\boldsymbol{X}_{\boldsymbol{i}}' = \left(X_{1i}, \dots, X_{pi}\right)$ be a *p*-dimensional vector of categorical factors with its realization $\boldsymbol{x}_{\boldsymbol{i}}' = \left(x_{1i}, \dots, x_{pi}\right)$, $i = 1, \dots, N$**.** When modeling mortality rates, the NB-GLM can be related to the linear model for the ratio response as follows:

[1] When the number of zeros is so large that the data do not readily fit standard negative binomial distributions.

$$\log\left(\frac{Y_i}{p_i}\right) = \beta_1 + \beta_2 x_{1i} + \beta_3 x_{2i} + \cdots \beta_p x_{i,p-1}$$

where p_i represents the population count associated with the number of deaths Y_i for the ith level of aggregation. For example, if the data are aggregated by sex (two levels), race (four levels), and age group (10 levels), there are 60 levels of aggregation; thus, $i = 1, \ldots, 60$. The model equation above can be rearranged to have $\log(Y_i)$ on the left side of the equation and $\log(p_i)$ on the right side. This case models the mortality rate while still maintaining the count response for the NB-GLM model, and the model itself is known as a rate model. The term $\log(p_i)$ on the right side of the equation is treated as a factor, and it is referred to as an offset with no parameters attached with its coefficient fixed at one.

The negative binomial distribution arises from a generalization of the Poisson distribution where the Poisson parameter λ is gamma distributed. One of the parametrizations of the negative binomial random variable can be expressed with the following equation:

$$P(Y = y) = \binom{y+k-1}{k-1} \frac{\alpha^y}{(1+\alpha)^{y+k}}, \qquad y = 0,1,2, \ldots.$$

Then $E(y) = \mu = k\alpha$ and $\mathrm{var}(Y) = k\alpha + k\alpha^2 = \mu + \frac{\mu^2}{k}$ where μ represents the mean of the response. Note that when $k \to \infty$, the second term in the variance goes to zero and the whole distribution approaches Poisson as a limit. This second term $\frac{\mu^2}{k}$ allows the negative binomial distribution to accommodate the situations where the variance is significantly higher than the mean (overdispersion) relative to the Poisson distribution. A convenient way to link the mean of the response to the linear combination of factors is by using the link function

$$\eta = x'\beta = \log\left(\frac{\alpha}{1+\alpha}\right) = \log\left(\frac{\mu}{\mu+k}\right)$$

where k denoted the dispersion parameter. The R package MASS introduced by Ripley et al. (2022) allows for estimation of k (labeled as theta) and standard error of k along with the coefficients of the negative binomial model with an offset variable. As a result of the maximum likelihood estimation (MLE) the **glm()** function also returns the value of the loglikelihood function at the MLEs. In addition, the output from the **glm()** function returns many important statistics that are used in the statistical inference.

C.3 GOODNESS-OF-FIT TESTS

The goodness-of-fit test allows assessing the model fit by comparing how well the observed data correspond to the expected (fitted or predicted) model values. The goodness-of-fit statistic tests the null hypothesis that the model M_0 fits versus the alternative hypothesis that the model M_A fits (or the model M_0 does not fit). In most cases, the observed data represent the fit of the most complex possible model called the saturated model and included under the alternative hypothesis. The test statistics used in this report to test the goodness of fit include the Pearson goodness-of-fit statistic, deviance statistic, and likelihood-ratio test statistic.

C.3.1 THE DEVIANCE STATISTIC

The deviance, also known as the likelihood-ratio statistic, for a fitted model plays the role of the residual sum of squares for the GLM model and is used for assessing the goodness-of-fit and for comparing models. The deviance is computed as follows:

$$D = 2\{l(\boldsymbol{b}_{\boldsymbol{max}}; \boldsymbol{y}) - l(\boldsymbol{b}; \boldsymbol{y})\}$$

where $\boldsymbol{b}_{\boldsymbol{max}}$ denotes the maximum likelihood estimator of $\boldsymbol{\beta}_{\boldsymbol{max}}$ in which $\boldsymbol{\beta}_{\boldsymbol{max}}$ denotes the parameter for the saturated model. The log-likelihood function for the saturated model evaluated at $\boldsymbol{b}_{\boldsymbol{max}}$ is $l(\boldsymbol{b}_{\boldsymbol{max}}; \boldsymbol{y})$. The likelihood function for the model of interest evaluated at $\boldsymbol{b}$ is $l(\boldsymbol{b}; \boldsymbol{y})$. Under appropriate regularity conditions, the deviance follows a χ^2 distribution with the degrees of freedom equal to the difference in the dimensions of the two models. The R function **glm**() computes the null deviance, which corresponds to the model without any factor (model with an intercept term only), and the residual deviance, which corresponds to the fitted model under consideration, the model with p factors. For more information about deviance, the reader is referred to the popular book by Dunn and Smyth (2018).

C.3.2 THE LIKELIHOOD-RATIO TEST

The likelihood-ratio (LR) test is used to compare two competing statistical models. The test statistic is expressed as the ratio of the two likelihoods, one found by maximizing over the entire parameter space and another after imposing some constraints. Often the LR test is expressed as the difference between two log likelihoods. Results of the LR test indicate which of the two competing models is preferred. Running the LR test in R requires generating output for both competing models before running the LR test. The LR test statistic follows the χ^2 distribution, under the null hypothesis.

C.3.3 THE GOODNESS-OF-FIT VIA AIC

Akaike (1969) introduced the information criterion now known as the Akaike information criterion (AIC). The AIC is used as a goodness of fit for any statistical model and is also used in the model selection when several models are considered. The AIC is defined as

$$AIC = -2l + 2p$$

where l denotes the loglikelihood evaluated under $\boldsymbol{b}$ and p denotes the number of parameters. The penalty term $2p$ penalizes the loglikelihood when the additional factors are added to the model. The smaller the AIC value, the better the model fit. The AIC is good only for comparing models.

C.4 DIAGNOSTIC PLOTS

Diagnostic plots are used to evaluate the model assumptions and investigate whether one has observations with a large influence on the model fit. These are subjective assessments about the fitted model under consideration. Many diagnostic plots have been suggested (e.g., Dunn and Smyth 2018). In this study the authors use two types of diagnostic plots:

1. The half-normal plot.
2. The plot of observed versus predicted values.

The half-normal plot provides a visual assessment of whether some observations in a set of positive observations are unusually extreme values (i.e., outliers). The plot displays the sorted absolute values of the residuals against the upper quantiles of the upper half of a standard normal distribution (Atkinson 1981).

One does not usually look for a strength line relationship because a positive normal distribution is not expected for quantities such as leverages or outliers. The outliers are the points that diverge substantially from the rest of the data and should be easily spotted on this plot.

The plot of observed values (y_i) versus precited values $\hat{y}_i$ represents one of the richest form of data visualization. The density of the observed and the predicted values should be about the same if the model fits well. Any differences between these densities would indicate a suboptimal model fit.

Appendix D: Model Names

Table D-1

SUMMARY OF GLM MODELS WITH THEIR CORRESPONDING FACTORS

Suicide/Race	Suicide/Ethnicity	Homicide/Race	Homicide/ Ethnicity	Other/Race	Other/Ethnicity
M1-S-R	M1-S-E	M1-H-R	M1-H-E	M1-O-R	M1-O-E
M2-S-R	M2-S-E	M2-H-R	M2-H-E	M2-O-R	M2-O-E
M3-S-R	M3-S-E	M3-H-R	M3-H-E	M3-O-R	M3-O-E
M4-S-R	M4-S-E	M4-H-R	M4-H-E	M4-O-R	M4-O-E
M5-S-R	M5-S-E	M5-H-R	M5-H-E	M5-O-R	M5-O-E
M6-S-R	M6-S-E	M6-H-R	M6-H-E	M6-O-R	M6-O-E
M7-S-R	M7-S-E	M7-H-R	M7-H-E	M7-O-R	M7-O-E
M8-S-R	M8-S-E	M8-H-R	M8-H-E	M8-O-R	M8-O-E
M9-S-R	M9-S-E	M9-H-R	M9-H-E	M9-O-R	M9-O-E

Appendix E: Coefficients of the Models Selected

Table E-1

COEFFICIENTS: SUICIDE-RACE, MODEL 6 (M6-S-R)

Variable	Relative Risk	Coefficients Estimate	Coefficients Standard Error	*Z* Value	*P* Value
Intercept	–	2.95	0.01	197.77	< 0.001
Race R2	63%	−0.46	0.03	−16.55	< 0.001
Race R3	76%	−0.27	0.03	−7.77	< 0.001
Race R4	212%	0.75	0.04	18.24	< 0.001

Baseline for race is white (R1).

Table E-2

COEFFICIENTS: SUICIDE-ETHNICITY, MODEL 5 (M5-S-E)

Variable	Relative Risk	Coefficients Estimate	Coefficients Standard Error	*Z* Value	*P* Value
Intercept	–	2.13	0.02	139.85	< 0.001
Ethnicity E2	72%	−0.33	0.02	−14.00	< 0.001
Sex Male	353%	1.26	0.02	63.94	< 0.001

Baseline for ethnicity is non-Hispanic/Latino (E1), for sex is female (F).

Table E-3

COEFFICIENTS: SUICIDE-RACE, MODEL 8 (M8-S-R)

Variable	Relative Risk	Coefficients Estimate	Coefficients Standard Error	*Z* Value	*P* Value
Intercept	–	1.15	0.05	21.23	< 0.001
Age 15–19	406%	1.40	0.07	20.15	< 0.001
Age 20–24	617%	1.82	0.07	26.68	< 0.001
Age 25–29	662%	1.89	0.07	27.65	< 0.001
Age 30–34	682%	1.92	0.07	28.21	< 0.001
Age 35–44	710%	1.96	0.07	29.06	< 0.001
Age 45–54	777%	2.05	0.07	30.54	< 0.001
Age 55–64	717%	1.97	0.07	29.34	< 0.001
Age 65–74	587%	1.77	0.07	26.05	< 0.001
Age 75–84	710%	1.96	0.07	28.51	< 0.001
Race R2	123%	0.21	0.11	1.80	0.072
Race R3	422%	1.44	0.19	7.40	< 0.001
Race R4	516%	1.64	0.19	8.67	< 0.001
Age 15–19: Race R2	68%	−0.39	0.14	−2.80	0.005
Age 20–24: Race R2	73%	−0.31	0.14	−2.26	0.024
Age 25–29: Race R2	66%	−0.42	0.14	−3.10	0.002
Age 30–34: Race R2	62%	−0.48	0.14	−3.45	0.001
Age 35–44: Race R2	44%	−0.82	0.13	−6.11	< 0.001
Age 45–54: Race R2	30%	−1.19	0.14	−8.69	< 0.001
Age 55–64: Race R2	27%	−1.31	0.14	−9.34	< 0.001
Age 65–74: Race R2	31%	−1.17	0.15	−7.75	< 0.001
Age 75–84: Race R2	40%	−0.91	0.17	−5.34	< 0.001
Age 15–19: Race R3	29%	−1.23	0.22	−5.57	< 0.001
Age 20–24: Race R3	23%	−1.45	0.21	−6.80	< 0.001
Age 25–29: Race R3	18%	−1.72	0.22	−7.99	< 0.001
Age 30–34: Race R3	15%	−1.88	0.22	−8.59	< 0.001
Age 35–44: Race R3	10%	−2.28	0.21	−10.67	< 0.001
Age 45–54: Race R3	11%	−2.17	0.21	−10.19	< 0.001
Age 55–64: Race R3	13%	−2.04	0.22	−9.29	< 0.001
Age 65–74: Race R3	19%	−1.68	0.23	−7.32	< 0.001
Age 75–84: Race R3	28%	−1.28	0.25	−5.20	< 0.001
Age 15–19: Race R4	61%	−0.49	0.22	−2.21	0.027
Age 20–24: Race R4	50%	−0.70	0.21	−3.25	0.001
Age 25–29: Race R4	53%	−0.64	0.21	−2.99	0.003
Age 30–34: Race R4	49%	−0.72	0.22	−3.32	0.001
Age 35–44: Race R4	28%	−1.26	0.21	−5.92	< 0.001
Age 45–54: Race R4	21%	−1.54	0.22	−6.97	< 0.001
Age 55–64: RaceR4	21%	−1.57	0.23	−6.73	< 0.001
Age 65–74: RaceR4	29%	−1.25	0.29	−4.32	< 0.001
Age 75–84: RaceR4	56%	−0.58	0.34	−1.70	0.089

Baseline for age is 10–14, for race is white (R1).

Table E-4

COEFFICIENTS: HOMICIDE-RACE, MODEL 2 (M2-H-R)

Variable	Relative Risk	Coefficients Estimate	Coefficients Standard Error	*Z* Value	*P* Value
Intercept	–	2.51	0.12	20.52	< 0.001
State Arizona	45%	−0.80	0.10	−8.28	< 0.001
State Colorado	41%	−0.90	0.10	−8.86	< 0.001
State Connecticut	27%	−1.30	0.11	−11.79	< 0.001
State Georgia	30%	−1.22	0.10	−12.45	< 0.001
State Indiana	49%	−0.72	0.10	−7.33	< 0.001
State Iowa	33%	−1.10	0.11	−9.73	< 0.001
State Kansas	50%	−0.69	0.11	−6.52	< 0.001
State Kentucky	54%	−0.61	0.10	−6.03	< 0.001
State Maine	43%	−0.85	0.17	−4.93	< 0.001
State Maryland	34%	−1.09	0.10	−10.90	< 0.001
State Massachusetts	20%	−1.62	0.11	−15.23	< 0.001
State Michigan	35%	−1.06	0.10	−10.74	< 0.001
State Minnesota	23%	−1.46	0.11	−13.65	< 0.001
State New Hampshire	41%	−0.89	0.18	−4.83	< 0.001
State New Jersey	21%	−1.56	0.10	−15.40	< 0.001
State New Mexico	90%	−0.11	0.10	−1.11	0.269
State North Carolina	33%	−1.10	0.10	−11.34	< 0.001
State Ohio	39%	−0.95	0.10	−9.71	< 0.001
State Oklahoma	51%	−0.67	0.10	−6.77	< 0.001
State Oregon	36%	−1.03	0.11	−9.37	< 0.001
State Rhode Island	37%	−1.00	0.17	−6.03	< 0.001
State South Carolina	41%	−0.88	0.10	−8.83	< 0.001
State Utah	30%	−1.21	0.12	−10.19	< 0.001
State Vermont	68%	−0.39	0.23	−1.73	0.083
State Virginia	26%	−1.33	0.10	−13.47	< 0.001
State Wisconsin	36%	−1.01	0.10	−9.87	< 0.001
Age 15–19	453%	1.51	0.08	17.92	< 0.001
Age 20–24	636%	1.85	0.08	22.24	< 0.001
Age 25–29	662%	1.89	0.08	22.64	< 0.001
Age 30–34	611%	1.81	0.08	21.65	< 0.001
Age 35–44	471%	1.55	0.08	18.64	< 0.001
Age 45–54	332%	1.20	0.08	14.40	< 0.001
Age 55–64	248%	0.91	0.08	10.73	< 0.001
Age 65–74	203%	0.71	0.09	8.09	< 0.001
Age 75–84	286%	1.05	0.10	10.84	< 0.001
Sex Male	297%	1.09	0.02	52.67	< 0.001
Race R2	739%	2.00	0.02	97.38	< 0.001
Race R3	280%	1.03	0.10	10.37	< 0.001
Race R4	390%	1.36	0.06	24.27	< 0.001

Baseline for state is Alaska, for age is 10–14, for sex is female, and for race is white (R1).

Table E-5

COEFFICIENTS: HOMICIDE-ETHNICITY, MODEL 2 (M2-H-E)

Variable	Relative Risk	Coefficients Estimate	Coefficients Standard Error	*Z* Value	*P* Value
Intercept	–	0.24	0.11	2.17	0.030
State Arizona	57%	−0.56	0.08	−6.72	< 0.001
State Colorado	46%	−0.77	0.09	−8.89	< 0.001
State Connecticut	39%	−0.93	0.10	−9.69	< 0.001
State Georgia	76%	−0.28	0.08	−3.43	0.001
State Indiana	73%	-−0.31	0.08	−3.69	< 0.001
State Iowa	34%	−1.08	0.10	−10.70	< 0.001
State Kansas	64%	−0.45	0.09	−4.96	< 0.001
State Kentucky	75%	−0.29	0.09	−3.33	0.001
State Maine	37%	−0.99	0.17	−5.91	< 0.001
State Maryland	92%	−0.08	0.08	−1.00	0.320
State Massachusetts	28%	−1.27	0.09	−13.92	< 0.001
State Michigan	63%	−0.46	0.08	−5.51	< 0.001
State Minnesota	27%	−1.30	0.09	−13.89	< 0.001
State New Hampshire	38%	−0.96	0.19	−5.06	< 0.001
State New Jersey	38%	−0.98	0.09	−11.50	< 0.001
State New Mexico	103%	0.03	0.09	0.34	0.733
State North Carolina	70%	−0.36	0.08	−4.37	< 0.001
State Ohio	66%	−0.41	0.08	−4.91	< 0.001
State Oklahoma	79%	−0.23	0.09	−2.64	0.008
State Oregon	36%	−1.02	0.09	−10.85	< 0.001
State Rhode Island	52%	−0.66	0.22	−3.08	0.002
State South Carolina	97%	−0.03	0.08	−0.39	0.695
State Utah	28%	−1.26	0.11	−11.92	< 0.001
State Vermont	73%	−0.31	0.28	−1.10	0.272
State Virginia	53%	−0.63	0.08	−7.37	< 0.001
State Wisconsin	44%	−0.83	0.09	−9.43	< 0.001
Age 15–19	521%	1.65	0.08	19.98	< 0.001
Age 20–24	754%	2.02	0.08	24.73	< 0.001
Age 25–29	754%	2.02	0.08	24.78	< 0.001
Age 30–34	655%	1.88	0.08	23.03	< 0.001
Age 35–44	481%	1.57	0.08	19.30	< 0.001
Age 45–54	332%	1.20	0.08	14.69	< 0.001
Age 55–64	232%	0.84	0.08	10.08	< 0.001
Age 65–74	155%	0.44	0.09	5.16	< 0.001
Age 75–84	175%	0.56	0.10	5.91	< 0.001
Sex Male	329%	1.19	0.02	61.09	< 0.001
Ethnicity E2	121%	0.19	0.03	7.26	< 0.001

Baseline for state is Alaska, for age is 10–14, for sex is female (F), and for ethnicity is non-Hispanic/Latino (E1).

Table E-6

COEFFICIENTS: “OTHER”-RACE, MODEL 2 (M2-O-R)

Variable	Relative Risk	Coefficients Estimate	Coefficients Standard Error	*Z* Value	*P* Value
Intercept	–	2.72	0.13	20.65	< 0.001
State Arizona	38%	−0.98	0.10	−9.72	< 0.001
State Colorado	31%	−1.17	0.11	−10.57	< 0.001
State Connecticut	21%	−1.57	0.16	−10.08	< 0.001
State Georgia	16%	−1.83	0.11	−17.14	< 0.001
State Indiana	36%	−1.01	0.11	−9.39	< 0.001
State Iowa	35%	−1.04	0.12	−8.41	< 0.001
State Kansas	34%	−1.08	0.13	−8.49	< 0.001
State Kentucky	36%	−1.01	0.11	−9.03	< 0.001
State Maine	51%	−0.67	0.15	−4.52	< 0.001
State Maryland	266%	0.98	0.10	9.83	< 0.001
State Massachusetts	23%	−1.49	0.12	−12.85	< 0.001
State Michigan	59%	−0.52	0.10	−5.10	< 0.001
State Minnesota	26%	−1.36	0.12	−11.45	< 0.001
State New Hampshire	43%	−0.84	0.17	−5.05	< 0.001
State New Jersey	17%	−1.78	0.12	−15.34	< 0.001
State New Mexico	51%	−0.68	0.12	−5.78	< 0.001
State North Carolina	16%	−1.82	0.11	−16.74	< 0.001
State Ohio	23%	−1.49	0.10	−14.38	< 0.001
State Oklahoma	40%	−0.91	0.11	−8.31	< 0.001
State Oregon	33%	−1.10	0.12	−9.48	< 0.001
State Rhode Island	81%	−0.21	0.16	−1.32	0.186
State South Carolina	21%	−1.57	0.12	−13.37	< 0.001
State Utah	61%	−0.49	0.12	−4.28	< 0.001
State Vermont	105%	0.05	0.16	0.30	0.765
State Virginia	16%	−1.85	0.11	−16.71	< 0.001
State Wisconsin	23%	−1.49	0.12	−12.78	< 0.001
Age 15–19	126%	0.23	0.10	2.31	0.021
Age 20–24	162%	0.48	0.09	5.18	< 0.001
Age 25–29	197%	0.68	0.09	7.38	< 0.001
Age 30–34	232%	0.84	0.09	9.03	< 0.001
Age 35–44	170%	0.53	0.09	5.89	< 0.001
Age 45–54	172%	0.54	0.09	5.95	< 0.001
Age 55–64	151%	0.41	0.09	4.44	< 0.001
Age 65–74	105%	0.05	0.10	0.55	0.585
Age 75–84	123%	0.21	0.11	1.99	0.046
Sex Male	152%	0.42	0.03	14.69	< 0.001
Race R2	239%	0.87	0.03	26.61	< 0.001
Race R3	165%	0.50	0.09	5.77	< 0.001
Race R4	505%	1.62	0.07	23.93	< 0.001
Manner Other	77%	−0.26	0.06	−4.33	< 0.001
Manner UFSelfinflicted	66%	-0.41	0.06	-6.73	< 0.001
Manner UndeterIntent	199%	0.69	0.03	20.85	< 0.001

Baseline for state is Alaska, for age is 10–14, for sex is female (F), for race is white (R1), and for manner of death is legal intervention.

Table E-7

COEFFICIENTS: “OTHER”-ETHNICITY, MODEL 2 (M2-O-E)

Variable	Relative Risk	Coefficients Estimate	Coefficients Standard Error	*Z* Value	*P* Value
Intercept	–	0.65	0.19	3.43	0.001
State Arizona	36%	−1.02	0.12	−8.57	< 0.001
State Colorado	27%	−1.32	0.13	−10.54	< 0.001
State Connecticut	18%	−1.74	0.23	−7.74	< 0.001
State Georgia	16%	−1.83	0.12	−14.96	< 0.001
State Indiana	31%	−1.17	0.12	−9.55	< 0.001
State Iowa	29%	−1.24	0.14	−8.71	< 0.001
State Kansas	30%	−1.20	0.16	−7.67	< 0.001
State Kentucky	29%	−1.24	0.13	−9.77	< 0.001
State Maine	48%	−0.74	0.19	−3.84	< 0.001
State Maryland	294%	1.08	0.12	9.34	< 0.001
State Massachusetts	18%	−1.70	0.13	−12.91	< 0.001
State Michigan	59%	−0.52	0.12	−4.47	< 0.001
State Minnesota	18%	−1.69	0.14	−12.41	< 0.001
State New Hampshire	40%	−0.92	0.21	−4.32	< 0.001
State New Jersey	16%	−1.83	0.13	−13.78	< 0.001
State New Mexico	59%	−0.53	0.15	−3.55	< 0.001
State North Carolina	15%	−1.91	0.13	−15.26	< 0.001
State Ohio	19%	−1.65	0.12	−13.88	< 0.001
State Oklahoma	38%	−0.97	0.13	−7.45	< 0.001
State Oregon	28%	−1.27	0.13	−9.57	< 0.001
State Rhode Island	73%	−0.32	0.19	−1.63	0.102
State South Carolina	17%	−1.75	0.14	−12.23	< 0.001
State Utah	56%	−0.58	0.13	−4.50	< 0.001
State Vermont	99%	−0.01	0.20	−0.05	0.963
State Virginia	14%	−1.94	0.13	−15.04	< 0.001
State Wisconsin	17%	−1.78	0.14	−13.13	< 0.001
Age 15–19	135%	0.30	0.16	1.90	0.058
Age 20–24	184%	0.61	0.15	3.98	< 0.001
Age 25–29	229%	0.83	0.15	5.50	< 0.001
Age 30–34	275%	1.01	0.15	6.64	< 0.001
Age 35–44	205%	0.72	0.15	4.79	< 0.001
Age 45–54	201%	0.70	0.15	4.66	< 0.001
Age 55–64	177%	0.57	0.15	3.76	< 0.001
Age 65–74	122%	0.20	0.16	1.31	0.190
Age 75–84	126%	0.23	0.17	1.34	0.180
Sex Male	146%	0.38	0.03	11.40	< 0.001
Ethnicity E2	142%	0.35	0.06	6.08	< 0.001
Manner Other	90%	−0.11	0.09	−1.19	0.234
Manner UFSelfinflicted	83%	−0.19	0.09	−1.98	0.048
Manner UndeterIntent	179%	0.58	0.04	15.07	< 0.001

Baseline for state is Alaska, for age is 10–14, for sex is female (F), for ethnicity is non-Hispanic/Latino (E1), and for manner of death is legal intervention.

Appendix F: Results of Models

Table F-1

SUMMARY OF SUICIDE GLM RESULTS WHEN RACE IS CONSIDERED

Variable	Relative Risk	Coefficients Estimate	Coefficients Standard Error	*Z* Value	*P* Value	Variable
M1-S-R	5,939	6,542.60	25.06	30,427.38	0.00	−30,343.4
M2-S-R	5,940	6,544.14	25.08	30,425.55	0.00	−30,343.6
M3-S-R	5,966	6,343.33	9.39	32,096.04	0.00	−32,066
M4-S-R	5,967	5,874.43	2.30	36,168.70	0.80	−36,140.7
M5-S-R	5,975	6,001.77	4.32	33,946.77	0.40	−33,934.8
M6-S-R	5,976	5,931.16	1.87	37,091.11	0.66	−37,081.1
M7-S-R	5,970	6,080.72	2.00	36,953.49	0.16	−36,931.5
M8-S-R	5,940	5,757.96	2.46	35,820.64	0.95	−35,738.6
M9-S-R	5,969	6,343.72	5.15	33,740.22	0.00	−33,716.2

Table F-2

SUMMARY OF SUICIDE GLM RESULTS WHEN ETHNICITY IS CONSIDERED

Variable	Relative Risk	Coefficients Estimate	Coefficients Standard Error	*Z* Value	*P* Value	Variable
M1-S-E	3,571	4,140.55	53.58	21,621.22	0	−21,541.22
M2-S-E	3,572	4,140.38	53.57	21,619.24	0	−21,541.24
M3-S-E	3,598	3,855.46	8.72	24,029.22	0	−24,003.22
M4-S-E	3,599	3,695.78	2.26	27,403.87	0.13	−27,379.87
M5-S-E	3,607	3,690.25	4.31	25,523.71	0.16	−25,515.71
M6-S-E	3,608	3,762.77	1.84	28,103.02	0.04	−28,097.02
M7-S-E	3,600	3,690.29	2.23	27,443.2	0.14	−27,421.2
M8-S-E	3,590	3,677.95	2.3	27,345.7	0.15	−27,303.7
M9-S-E	3,599	3,793.19	7.26	243,67.45	0.01	−24,343.45

Table F-3

SUMMARY OF HOMICIDE GLM RESULTS WHEN RACE IS CONSIDERED

Variable	Relative Risk	Coefficients Estimate	Coefficients Standard Error	*Z* Value	*P* Value	Variable
M1-H-R	3,309	3,127.71	6.50	18,317.63	0.99	−18,233.63
M2-H-R	3,310	3,127.02	6.42	18,336.00	0.99	−18,254.00
M3-H-R	3,336	3,284.98	4.19	19,236.98	0.73	−19,206.98
M4-H-R	3,337	3,325.80	2.19	20,779.56	0.55	−20,751.56
M5-H-R	3,345	3,319.52	2.58	20,342.49	0.62	−20,330.49
M6-H-R	3,346	3,423.25	1.66	21,616.19	0.17	−21,606.19
M7-H-R	3,340	3,826.08	0.87	24,025.49	0.00	−24,003.49
M8-H-R	3,311	3,305.00	2.33	20,651.63	0.53	−20,571.63
M9-H-R	3,339	3,668.11	1.20	22,826.52	0.00	−22,802.52

Table F-4

SUMMARY OF HOMICIDE GLM RESULTS WHEN ETHNICITY IS CONSIDERED

Variable	Relative Risk	Coefficients Estimate	Coefficients Standard Error	*Z* Value	*P* Value	Variable
M1-H-E	2,202	2,261.56	14.59	12,975.75	0.18	−12,895.75
M2-H-E	2,203	2,251.39	13.77	13,017.66	0.23	−12,939.66
M3-H-E	2,229	2,194.64	5.00	14,132.81	0.69	−14,106.81
M4-H-E	2,230	2,246.31	2.32	15,460.72	0.40	−15,436.72
M5-H-E	2,238	2,245.91	2.45	15,349.16	0.45	−15,341.16
M6-H-E	2,239	2,347.41	1.51	16,395.90	0.05	−16,389.90
M7-H-E	2,231	2,297.57	2.25	15,573.26	0.16	−15,551.26
M8-H-E	2,221	2,234.79	2.34	15,451.93	0.41	−15,409.93
M9-H-E	2,230	2,236.78	4.99	14,176.09	0.46	−14,152.09

Table F-5

SUMMARY OF “OTHER” MANNER GLM RESULTS WHEN RACE IS CONSIDERED

Variable	Relative Risk	Coefficients Estimate	Coefficients Standard Error	*Z* Value	*P* Value	Variable
M1-O-R	4,705	3,720.21	3.54	18,820.42	1.00	−18,730.42
M2-O-R	4,706	3,722.75	3.54	18,822.65	1.00	−18,734.65
M3-O-R	4,732	4,366.62	1.21	22,124.00	1.00	−22,088.00
M4-O-R	4,733	4,399.42	1.16	22,295.26	1.00	−22,261.26
M5-O-R	4,741	4,410.70	1.13	22,375.24	1.00	−22,357.24
M6-O-R	4,742	4,441.62	1.08	22,538.72	1.00	−22,522.72
M7-O-R	4,736	4,951.82	1.00	23,343.62	0.01	−23,315.62
M8-O-R	4,709	4,558.65	0.94	23,240.52	0.94	−23,158.52
M9-O-R	4,738	5,042.11	0.81	24,231.60	0.00	−24,207.60

Table F-6

SUMMARY OF “OTHER” MANNER GLM RESULTS WHEN ETHNICITY IS CONSIDERED

Variable	Relative Risk	Coefficients Estimate	Coefficients Standard Error	*Z* Value	*P* Value	Variable
M1-O-E	2,192	1,833.67	4.56	10,947.54	1.00	−10,861.54
M2-O-E	2,193	1,835.56	4.55	10,950.99	1.00	−10,866.99
M3-O-E	2,219	2,243.26	1.16	13,375.73	0.35	−13,343.73
M4-O-E	2,220	2,257.86	1.11	13,462.91	0.28	−13,432.91
M5-O-E	2,228	2,272.56	1.07	13,531.92	0.25	−13,517.92
M6-O-E	2,229	2,286.22	1.04	13,614.00	0.19	−13,602.00
M7-O-E	2,221	2,308.43	1.09	13,560.97	0.10	−13,532.97
M8-O-E	2,214	2,337.23	0.94	13,905.14	0.03	−13,863.14
M9-O-E	2,223	2,367.21	0.93	13,941.65	0.02	−13,917.65

Appendix G: Diagnosis Plots for Selected Models

Figure G-1

QUANTILE-QUANTILE PLOT: SUICIDE-RACE, MODEL 6

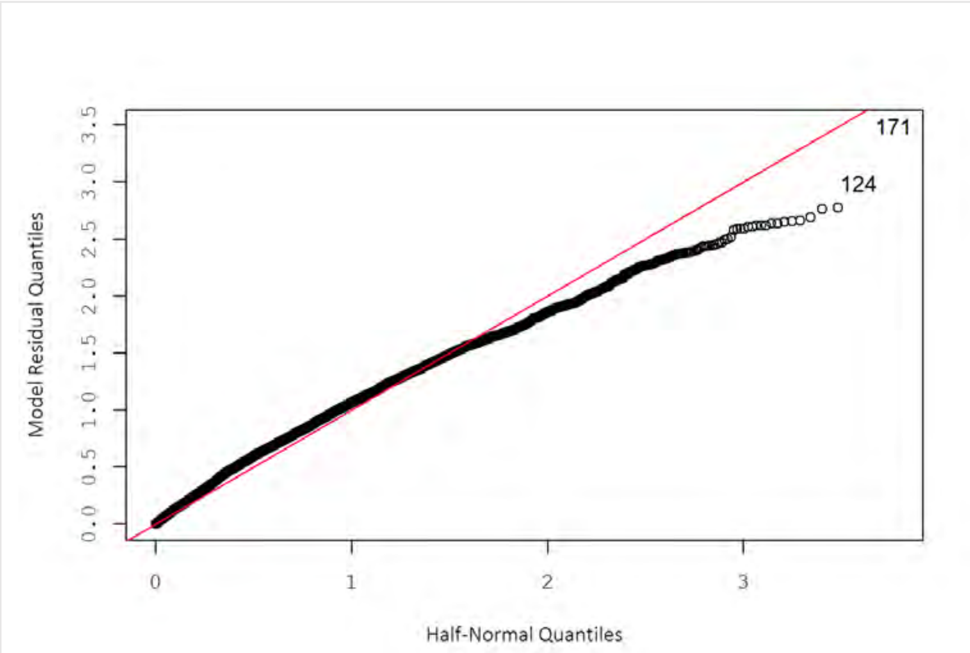
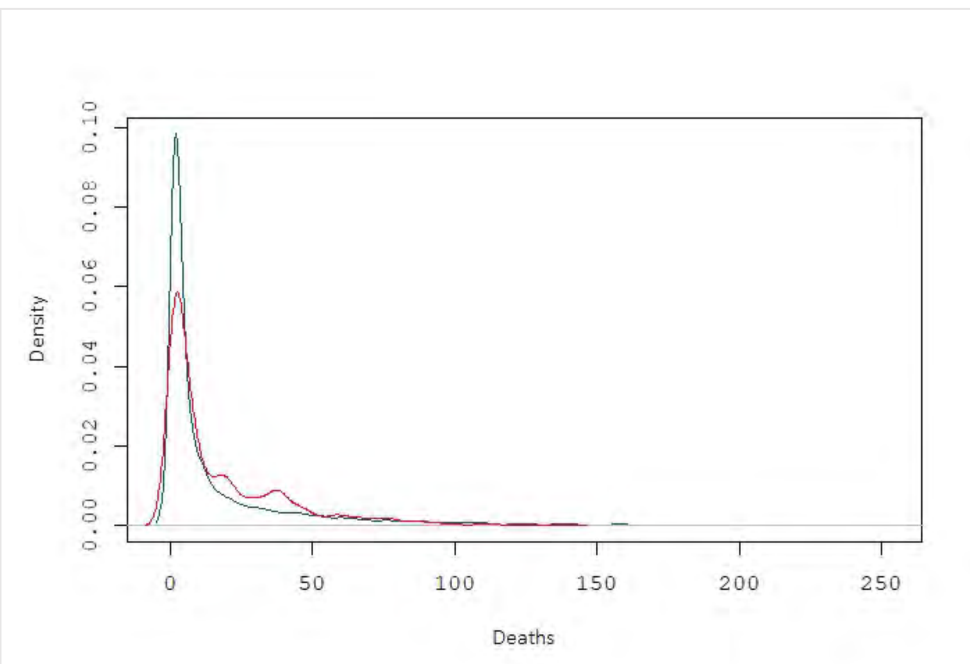


Figure G-2

PROBABILITY DENSITY PLOT OF DEATH: SUICIDE-RACE, MODEL 6



Blue line: empirical density from the data; red line: density from the fitted model.

Figure G-3

QUANTILE-QUANTILE PLOT: SUICIDE-RACE, MODEL 8

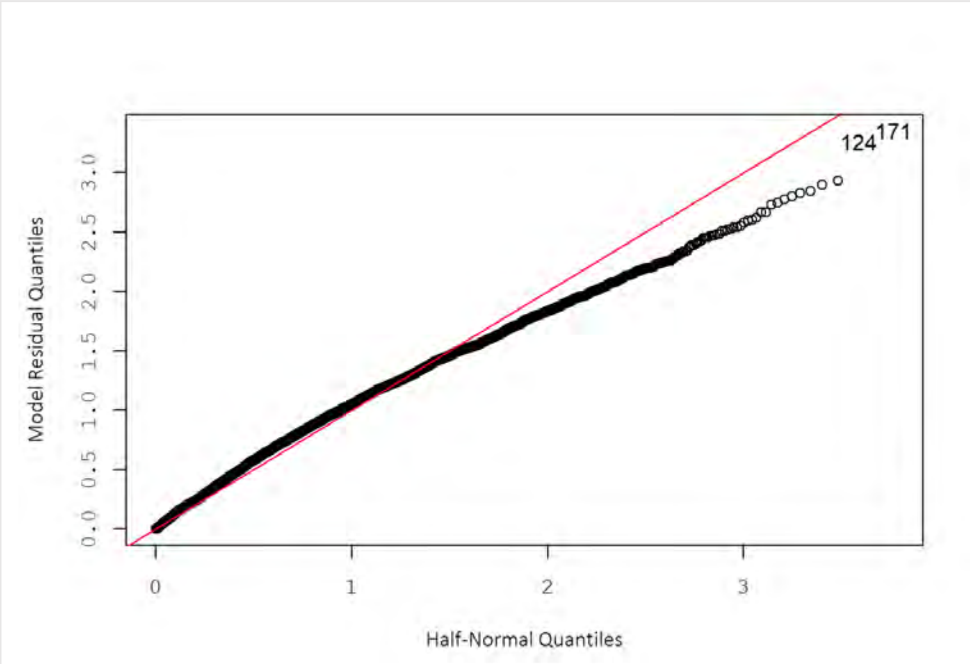
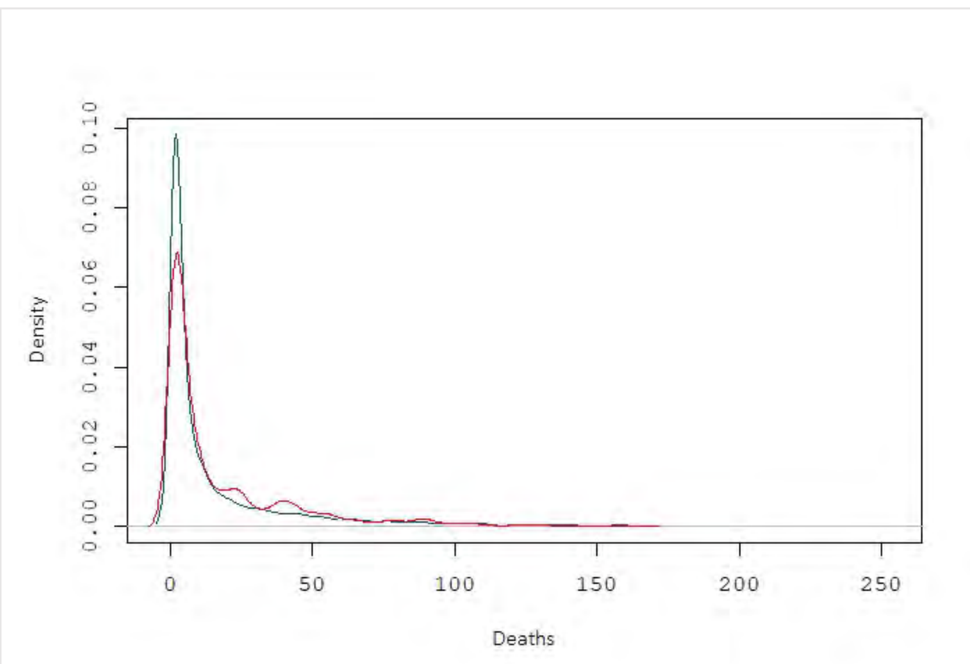


Figure G-4

PROBABILITY DENSITY PLOT OF DEATH: SUICIDE-RACE, MODEL 8



Blue line: empirical density from the data; red line: density from the fitted model.

Figure G-5

QUANTILE-QUANTILE PLOT: SUICIDE-ETHNICITY, MODEL 5

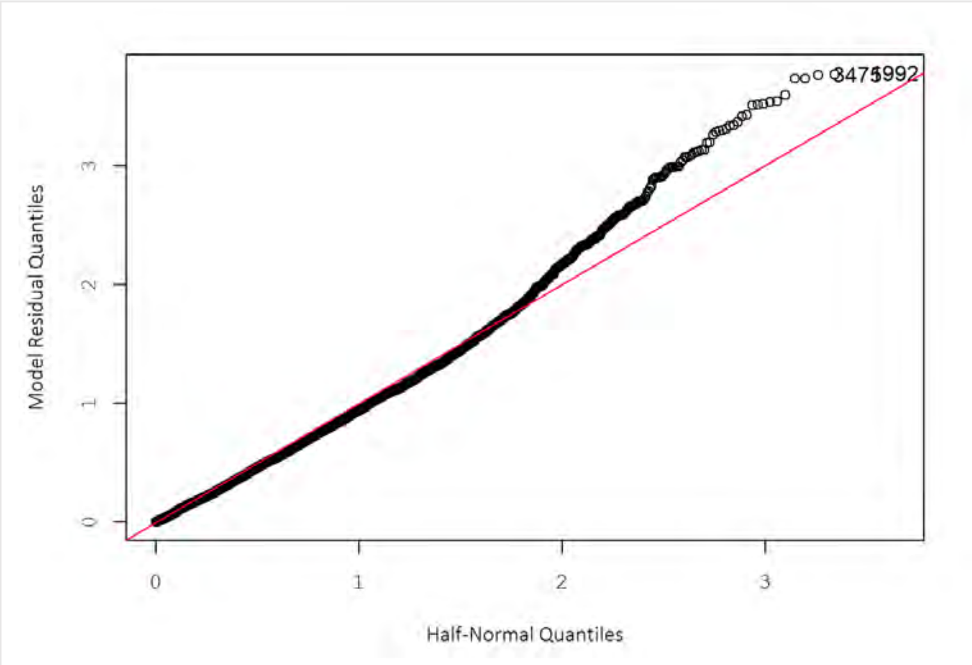
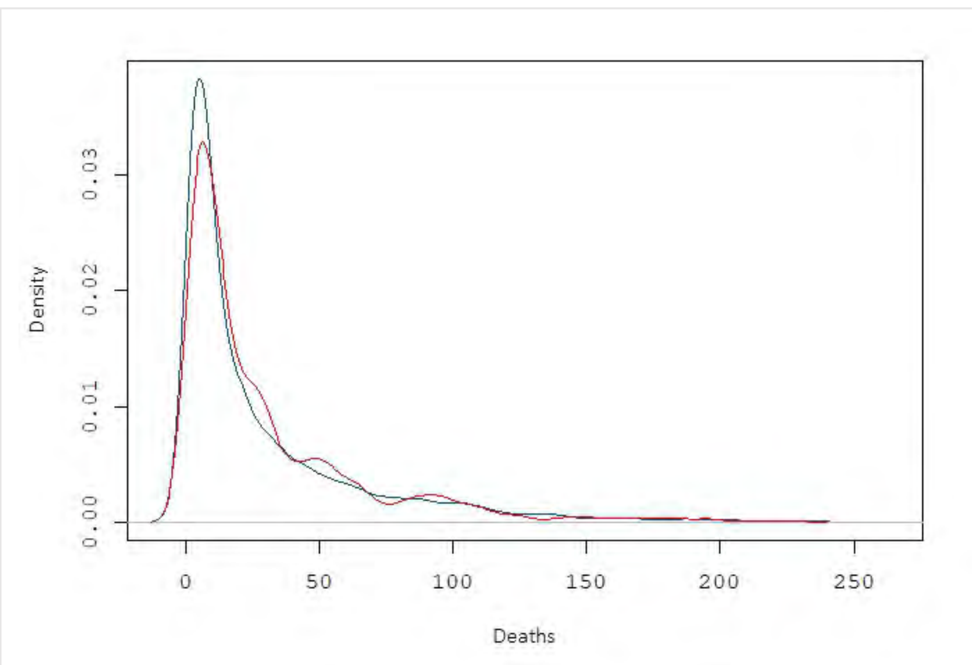


Figure G-6

PROBABILITY DENSITY PLOT OF DEATH: SUICIDE-ETHNICITY, MODEL 5



Blue line: empirical density from the data; red line: density from the fitted model.

Figure G-7

QUANTILE-QUANTILE PLOT: HOMICIDE-RACE, MODEL 2

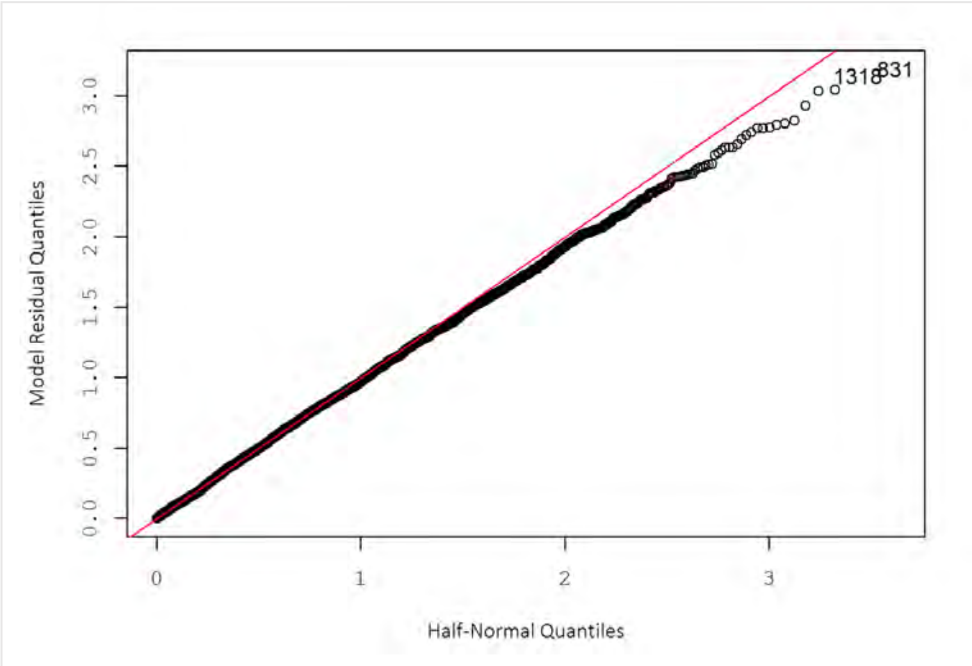
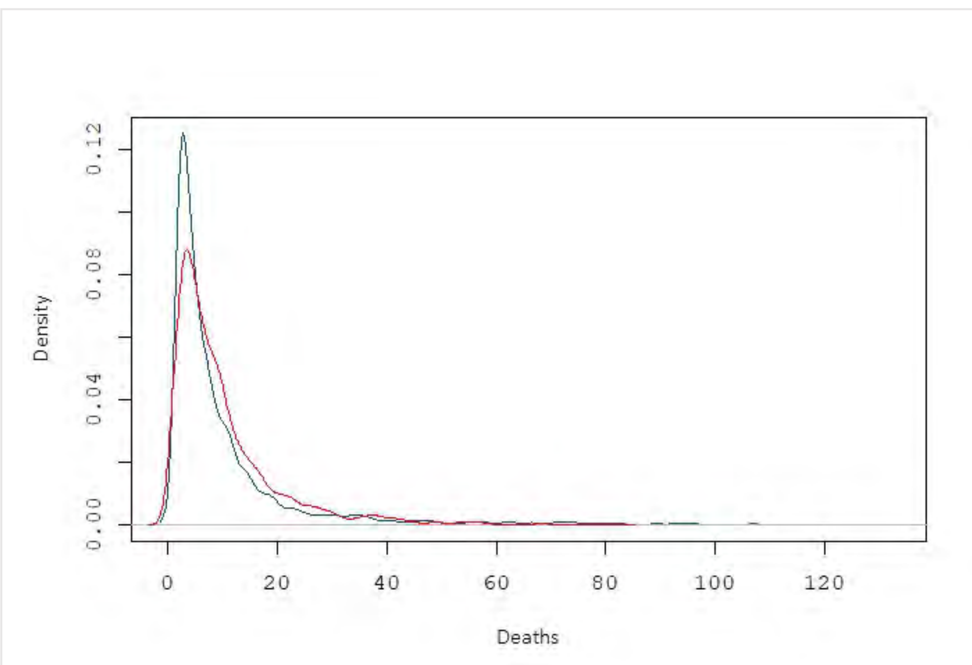


Figure G-8

PROBABILITY DENSITY PLOT OF DEATH: HOMICIDE-RACE, MODEL 2



Blue line: empirical density from the data; red line: density from the fitted model.

Figure G-9

QUANTILE-QUANTILE PLOT: HOMICIDE-ETHNICITY, MODEL 2

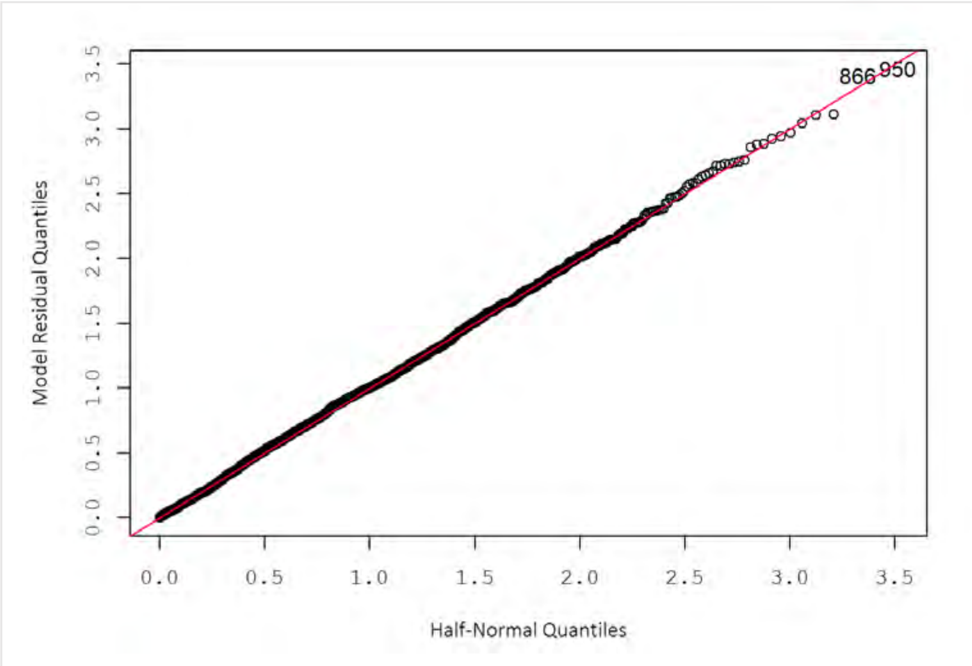
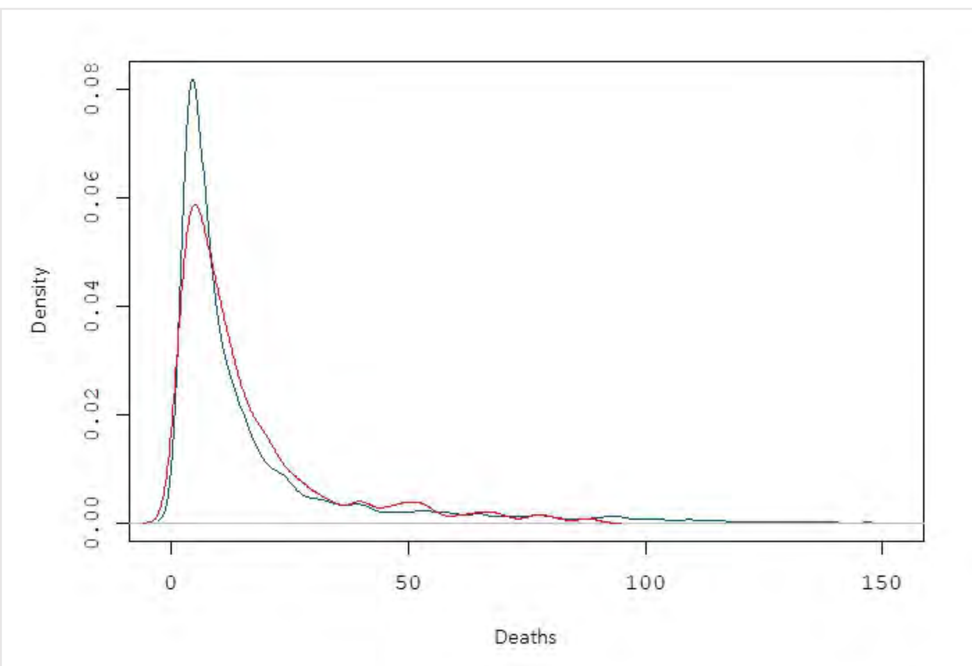


Figure G-10

PROBABILITY DENSITY PLOT OF DEATH: HOMICIDE-ETHNICITY, MODEL 2



Blue line: empirical density from the data; red line: density from the fitted model.

Figure G-11

QUANTILE-QUANTILE PLOT: "OTHER"-RACE, MODEL 2

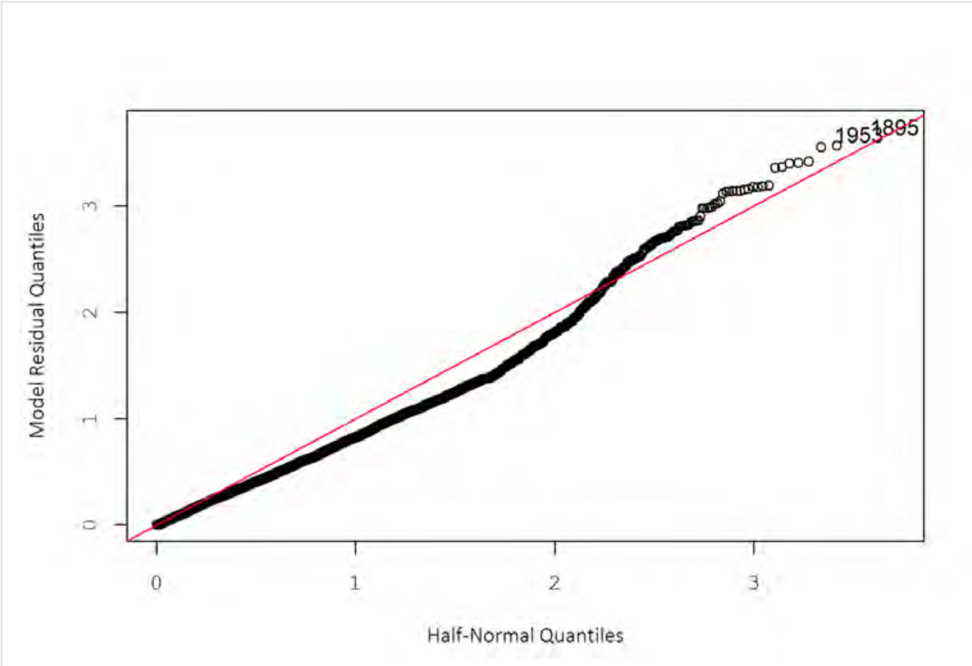
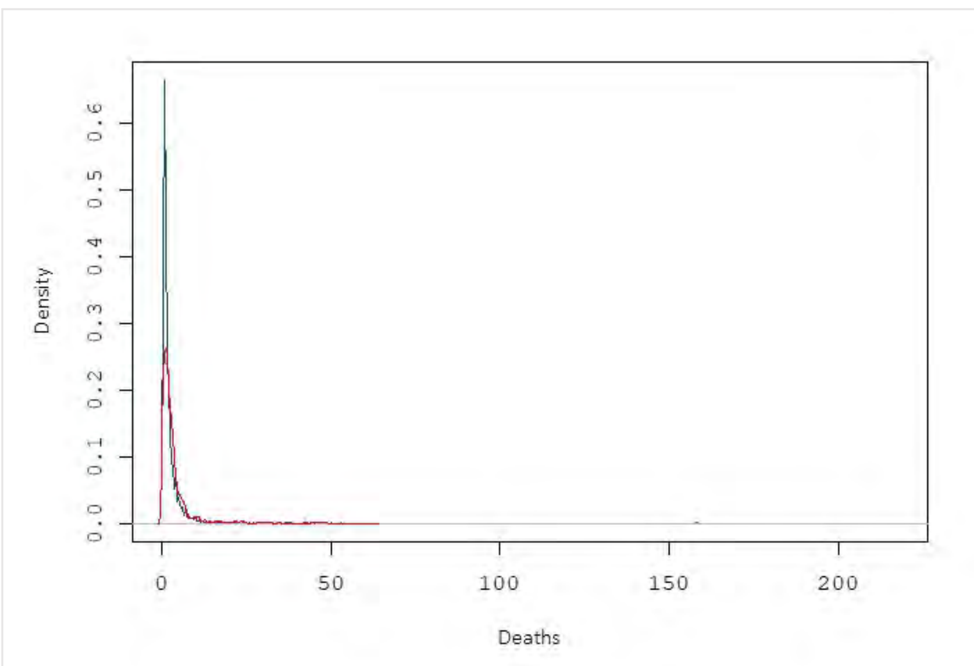


Figure G-12

PROBABILITY DENSITY PLOT OF DEATH: "OTHER"-RACE, MODEL 2



Blue line: empirical density from the data; red line: density from the fitted model.

Figure G-13

QUANTILE-QUANTILE PLOT: "OTHER"-ETHNICITY, MODEL 2

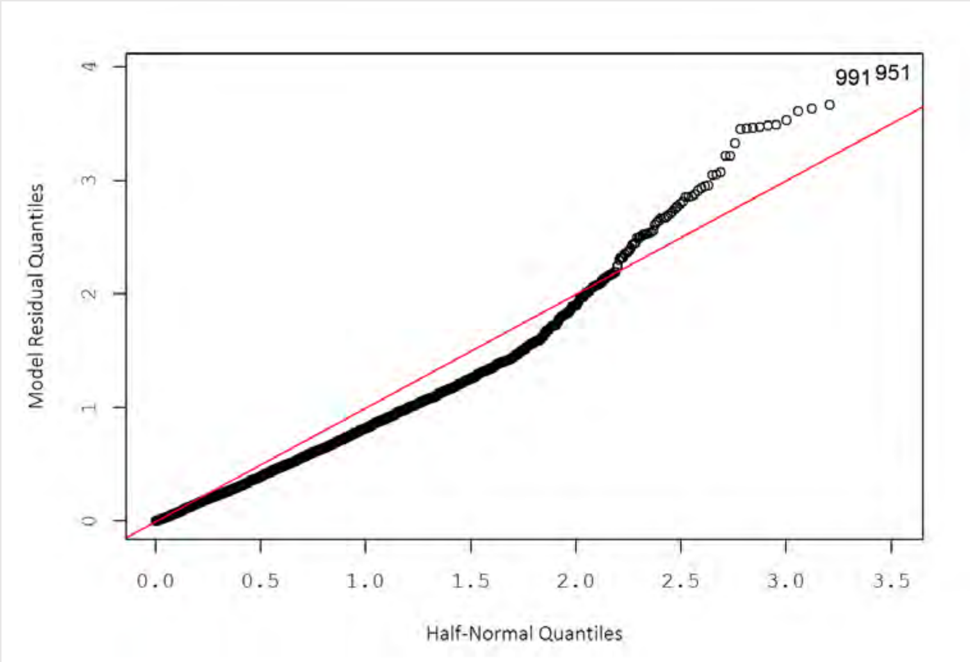
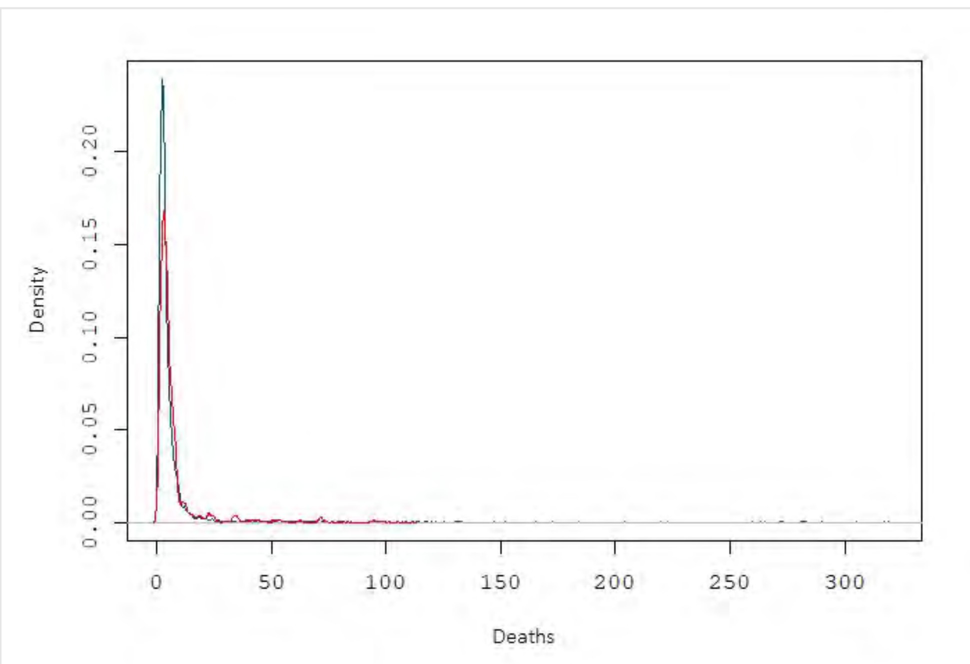


Figure G-14

PROBABILITY DENSITY PLOT OF DEATH: "OTHER"-ETHNICITY, MODEL 2



Blue line: empirical density from the data; red line: density from the fitted model.

References

Agresti, A. 2018. *An Introduction to Categorical Data Analysis.* Hoboken, NJ: John Wiley & Sons.

Akaike, H. 1969. Fitting Autoregressive Models for Prediction. *Annals of the Institute of Statistical Mathematics* 21, no. 1:243–247.

Atkinson, A. C. 1981. Two Graphical Displays for Outlying and Influential Observations in Regression. *Biometrika* 68, no. 1:13–20.

Barber, C., D. Azrael, A. Cohen, M. Miller, D. Thymes, D. E. Wang, and D. Hemenway. 2016. Homicides by Police: Comparing Counts from the National Violent Death Reporting System, Vital Statistics, and Supplementary Homicide Reports. *American Journal of Public Health* 106, no. 5:922–927.

Barber, C., D. Azrael, and D. Hemenway. 2013. A Truly National Violent Death Reporting System. *Injury Prevention* 19, no. 4:225–226.

Barbieri, M. 2020. Mortality by Socioeconomic Category in the United States. Schaumburg, IL: Society of Actuaries. Retrieved 4 29, 2023, from https://www.soa.org/resources/research-reports/2020/us-mort-rate-socioeconomic.

Boucher, J.-P., and M. Guillén. 2009. A Survey on Models for Panel Count Data with Applications to Insurance. *RACSAM—Revista de la Real Academia de Ciencias Exactas, Fisicas y Naturales. Serie A. Matematicas* 103, no. 2:277–294.

Calderon-Anyosa, R. J., and J. S. Kaufman. 2021. Impact of COVID-19 Lockdown Policy on Homicide, Suicide, and Motor Vehicle Deaths in Peru. *Preventive Medicine* 143:106331.

Campbell, J. C. 2002. Health Consequences of Intimate Partner Violence. *The Lancet* 359, no. 9314:1331–1336.

Centers for Disease Control and Prevention (CDC). 2021. National Violent Death Reporting System (NVDRS) Coding Manual. National Center for Injury Prevention and Control, Centers for Disease Control and Prevention (producer). Retrieved 4 29, from URL: www.cdc.gov/injury.

Centers for Disease Control and Prevention (CDC). 2022. Violence Prevention. Retrieved 4 29, 2023, from https://www.cdc.gov/violenceprevention/.

Dobson, A. J., and A. G. Barnett. 2018. *An Introduction to Generalized Linear Models.* Chapman and Hall/CRC.

Dunn, P. K., and G. K. Smyth. 2018. *Generalized Linear Models with Examples in R.* New York, NY: Springer.

Ehlman, D. C. 2022. Changes in Suicide Rates—United States, 2019 and 2020. *MMWR. Morbidity and Mortality Weekly Report* 71, no. 8:306–312.

Fridel, E. E., and J. A. Fox. 2019. Gender Differences in Patterns and Trends in US Homicide, 1976–2017. *Violence and Gender* 6, no. 1:27–36.

Galea, S., M. Tracy, K. J. Hoggatt, C. DiMaggio, and A. Karpati. 2011. Estimated Deaths Attributable to Social Factors in the United States. *American Journal of Public Health* 101, no. 8:1456–1465.

Gelman, A., and J. Hill. 2006. *Data Analysis Using Regression and Multilevel/Hierarchical Models.* New York, NY: Cambridge University Press.

Grinshteyn, E., and D. Hemenway. 2016. Violent Death Rates: The US Compared with Other High-Income OECD Countries, 2010. *American Journal of Medicine* 129, no. 3:266–273.

Haberman, S., and A. E. Renshaw. 1996. Generalized Linear Models and Actuarial Science. *Journal of the Royal Statistical Society: Series D (The Statistician)* 45, no. 4:407–436.

Haegerich, T. M., L. L. Dahlberg, T. R. Simon, G. T. Baldwin, D. A. Sleet, A. I. Greenspan, and L. C. Degutis. 2014. Prevention of Injury and Violence in the USA. *The Lancet* 384, no. 9937:64–74.

Han, B., P. S. Kott, A. Hughes, R. McKeon, C. Blanco, and W. M. Compton. 2016. Estimating the Rates of Deaths by Suicide among Adults Who Attempt Suicide in the United States. *Journal of Psychiatric Research* 77:125–133.

Ismail, N., and A. A. Jemain. 2007. Handling Overdispersion with Negative Binomial and Generalized Poisson Regression Models. *Casualty Actuarial Society Forum* 2007:103–158.

Jaffe, S. 2018. Gun Violence Research in the USA: The CDC's Impasse. *The Lancet* 391, no. 10139:2487–2488.

Joe, S., S. Canetto, and D. Romer. 2008. Advancing Prevention Research on the Role of Culture in Suicide Prevention. *Suicide and Life-Threatening Behavior* 38, no. 3:354–362.

John, A., J. Pirkis, D. Gunnell, L. Appleby, and J. Morrissey. 2020. Trends in Suicide during the COVID-19 Pandemic. BMJ, 371. Retrieved 4 29, 2023, from https://www.bmj.com/content/371/bmj.m4352

Kaufman, E. J., C. N. Morrison, C. C. Branas, and D. J. Wiebe. 2018. State Firearm Laws and Interstate Firearm Deaths from Homicide and Suicide in the United States: A Cross-sectional Analysis of Data by County. *JAMA Internal Medicine* 178, no. 5:692–700.

Kegler, S. R., T. R. Simon, M. L. Zwald, M. S. Chen, J. A. Mercy, C. M. Jones, M. C. Mercado-Crespo, J. M. Blair, D. M. Stone, P. G. Ottley, and J. Dills. 2022. Vital Signs: Changes in Firearm Homicide and Suicide Rates—United States, 2019–2020. *Morbidity and Mortality Weekely Report* 71, no. 19:656–663.

Krieger, N., J. T. Chen, P. D. Waterman, D. H. Rehkopf, and S. Subramanian. 2003. Race/Ethnicity, Gender, and Monitoring Socioeconomic Gradients in Health: A Comparison of Area-Based Socioeconomic Measures—The Public Health Disparities Geocoding Project. *American Journal of Public Health* 93, no. 10:1655–1671.

Krieger, N., M. V. Kiang, J. T. Chen, and P. D. Waterman. 2015. Waterman PD. Trends in US Deaths Due to Legal Intervention among Black and White Men, Age 15–34 Years, by County Income Level. *Harvard Public Health Review* 3:1–5.

Krug, E. G., J. A. Mercy, L. L. Dahlberg, and A. B. Zwi. 2002. The world report on violence and health. The lancet, 360(9339), pp.1083-1088.

Lüdecke, D., M. S. Ben-Shachar, I. Patil, P. Waggoner, and D. Makowski. 2021. {performance}: An {R} Package for Assessment, Comparison and Testing of Statistical Models. *Journal of Open Source Software* 6, no. 60:3139. doi:10.21105/joss.03139

Marzuk, P. M., K. Tardiff, A. C. Leon, C. S. Hirsch, M. Stajic, L. Portera, N. Hartwell, and M. I. Iqbal.1995. Fatal Injuries after Cocaine Use as a Leading Cause of Death among Young Adults in New York City. *New England Journal of Medicine* 332, no. 26:1753–1757.

Miller, M., D. Azrael, and D. Hemenway. 2002. Firearm Availability and Unintentional Firearm Deaths, Suicide, and Homicide among 5–14 Year Olds. *Journal of Trauma and Acute Care Surgery* 52, no. 2:267–275.

National Cancer Institute. 2023. NCI Dictionary of Cancer Terms. Retrieved 4 29, 2023, from https://www.cancer.gov/publications/dictionaries/cancer-terms/def/ethnicity.

National Center for Health Statistics (U.S.). 2018. Health, United States, 2017: With Special Feature on Mortality [Internet]. Hyattsville, MD: National Center for Health Statistics (U.S.). Retrieved 4 29, 2023, from https://pubmed.ncbi.nlm.nih.gov/30702833.

National Human Genome Research Institute. 2023. Race. Retrieved 4 29, 2023, from https://www.genome.gov/genetics-glossary/Race.

National Indian Council on Aging (NICOA). 2022. American Indian Suicide Rate Increases. Retrieved 4 29, 2023, from https://www.nicoa.org/national-american-indian-and-alaska-native-hope-for-life-day/.

National Indian Council on Aging (NICOA). 2023. CDC Study Shows Lower Life Expectancy for Natives. Retrieved 4 29, 2023, from https://www.nicoa.org/cdc-study-shows-lower-life-expectancy-for-natives/#:~:text=At%2071.8%20years%2C%20American%20Indian,81.9%20years%20for%20Hispanic%20individuals.

Nelder, J. A., and R. W. Wedderburn. 1972. Generalized Linear Models. *Journal of the Royal Statistical Society: Series A (General)* 135, no. 3:370–384.

NVDRS. 2021. National Violent Death Reporting System Data Analysis Guide. United States.

NVDRS. 2023. NVDRS Definition. Retrieved 4 29, 2023, from https://vdrs.us/resources/NVDRS_Definition.pdf.

Ogasa, N. 2022. Mass Shootings and Gun Violence in the United States Are Increasing. *Science News* 26. Retrieved 4 29, 2023, from https://www.sciencenews.org/article/gun-violence-mass-shootings-increase-united-states-data-uvalde-buffalo.

Petrosky, E., J. M. Blair, C. J. Betz, K. A. Fowler, S. P. Jack, and B. H. Lyons. 2017. Racial and Ethnic Differences in Homicides of Adult Women and the Role of Intimate Partner Violence—United States, 2003–2014. *Morbidity and Mortality Weekly Report* 66, no. 28:741–746.

Petrosky, E., L. M. Kollar, M. C. Kearns, S. G. Smith, C. J. Betz, K. A. Fowler, and D. E. Satter. 2021. Homicides of American Indians/Alaska Natives—National Violent Death Reporting System, United States, 2003–2018. *MMWR Surveillance Summaries* 70, no. 8:1–19.

R Core Team. 2022. R: A Language and Environment for Statistical Computing. Vienna: R Foundation. Retrieved 4 29, 2023, from https://www.R-project.org/.

RAND Corporation. 2022. Gun Policy in America. Retrieved 4 29, 2023, from https://www.rand.org/research/gun-policy.html.

Ripley, B., B. Venables, D. M. Bates, K. Hornik, A. Gebhardt, D. Firth, and M. B. Ripley. 2022. *Modern Applied Statistics with S.* New York, NY: Springer.

Siegel, M., C. S. Ross, and C. King III. 2013. The Relationship between Gun Ownership and Firearm Homicide Rates in the United States. *American Journal of Public Health* 103, no. 11:2098–2105.

Wilson, R. F., G. Liu, B. H. Lyons, E. Petrosky, D. D. Harrison, C. J. Betz, and J. M. Blair. 2022. Surveillance for Violent Deaths—National Violent Death Reporting System, 42 States, the District of Columbia, and Puerto Rico, 2019. *MMWR Surveillance Summaries* 71, no. 6:1–40.

About The Society of Actuaries Research Institute

Serving as the research arm of the Society of Actuaries (SOA), the SOA Research Institute provides objective, data-driven research bringing together tried and true practices and future-focused approaches to address societal challenges and your business needs. The Institute provides trusted knowledge, extensive experience and new technologies to help effectively identify, predict and manage risks.

Representing the thousands of actuaries who help conduct critical research, the SOA Research Institute provides clarity and solutions on risks and societal challenges. The Institute connects actuaries, academics, employers, the insurance industry, regulators, research partners, foundations and research institutions, sponsors and non-governmental organizations, building an effective network which provides support, knowledge and expertise regarding the management of risk to benefit the industry and the public.

Managed by experienced actuaries and research experts from a broad range of industries, the SOA Research Institute creates, funds, develops and distributes research to elevate actuaries as leaders in measuring and managing risk. These efforts include studies, essay collections, webcasts, research papers, survey reports, and original research on topics impacting society.

Harnessing its peer-reviewed research, leading-edge technologies, new data tools and innovative practices, the Institute seeks to understand the underlying manner of risk and the possible outcomes. The Institute develops objective research spanning a variety of topics with its strategic research programs: aging and retirement; actuarial innovation and technology; mortality and longevity; diversity, equity and inclusion; health care cost trends; and catastrophe and climate risk. The Institute has a large volume of topical research available, including an expanding collection of international and market-specific research, experience studies, models and timely research.

Society of Actuaries Research Institute
475 N. Martingale Road, Suite 600
Schaumburg, Illinois 60173
www.SOA.org