

## FINAL REPORT

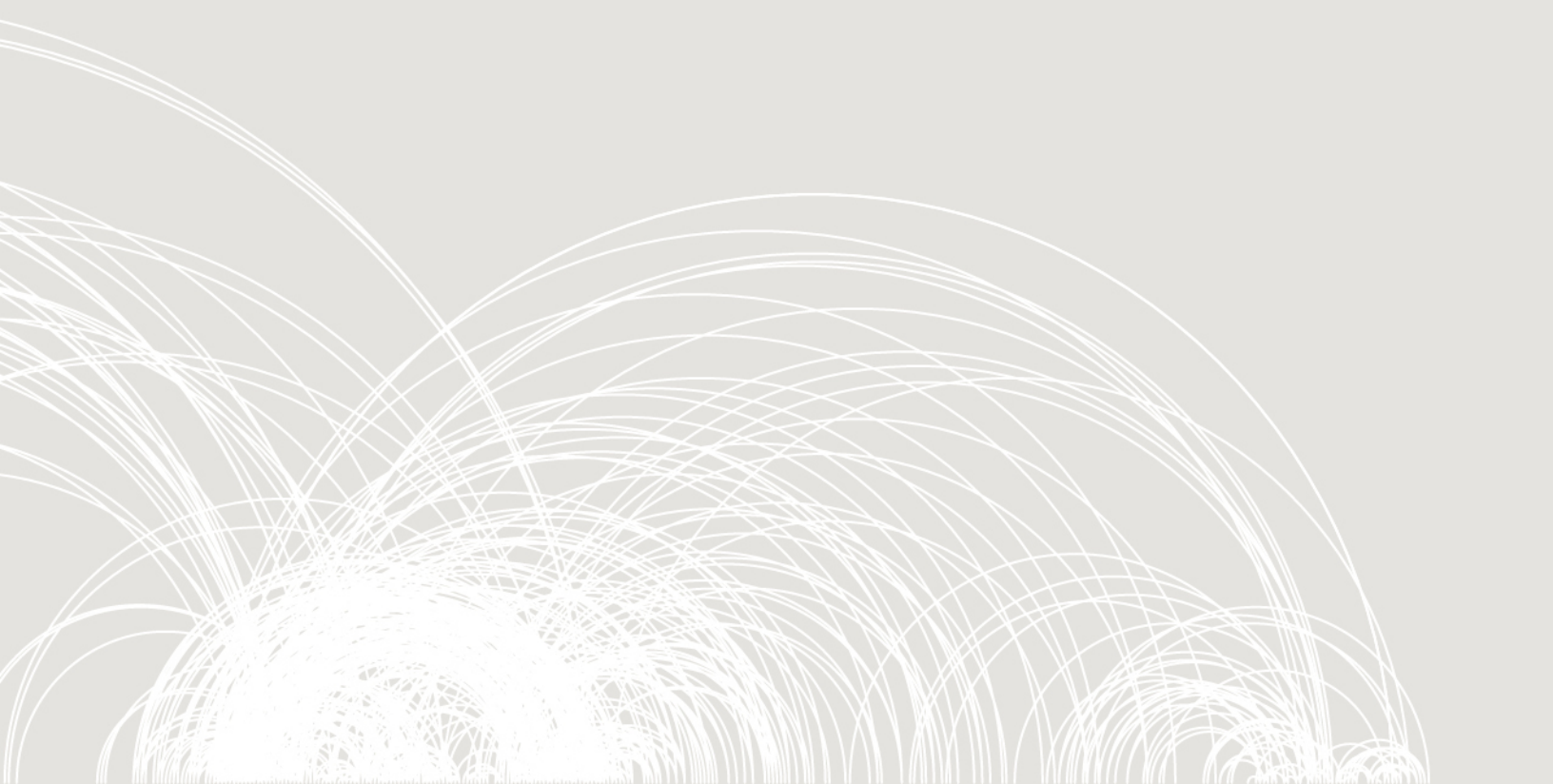
OCTOBER 2021

# Improving Data Infrastructure to Reduce Firearms Violence

**Editors:**

John K. Roman, PhD  
NORC at the University of Chicago

Philip Cook, PhD  
Duke University



# Acknowledgements

The report on Improving Data Infrastructure to Reduce Firearms Violence was written with funding from Arnold Ventures. NORC at the University of Chicago is grateful to the authors: Phil Cook, PhD; Catherine Barber, MPA; Nick Hart, PhD; Steve Marshall, PhD; Michael Mueller-Smith, PhD; Susan Parker, MPP, MS; Nancy Potok, PhD for their contributions to our knowledge on how to implement recommendations in the *Blueprint for a US Firearms Infrastructure*. NORC appreciates the support of Asheley Van Ness, our Arnold Ventures Project Officer, and Evan Mintz and Caitlyn Morrison, also of Arnold Ventures. NORC also thanks Julie Kubelka, Lisa Stein, Lynne Snyder, Mike Kalmar and Imelda Demus for editing and formatting this report.

This work is supported by Arnold Ventures, a philanthropy dedicated to tackling some of the most pressing problems in the United States.

# Table of Contents

Acknowledgements..... i

Executive Summary ..... 1

    Creating a Nonfatal Firearms Injury Database..... 2

    Increasing the Quality, Availability, and Usefulness of Firearms Data for  
    Research and Policy..... 4

    Practical Steps for Building State Capacity and Infrastructure to Use Data for  
    Evidence-Based Decision-Making ..... 6

    Conclusion..... 7

Chapter 1. Comprehensive Data on Gun Violence: Current Deficits,  
Needed Investments ..... 8

    Introduction ..... 8

    Potential Uses for Data on Firearms Injuries ..... 9

        Data Sources..... 9

        Hospital Data ..... 10

        Recommendations:..... 11

    Police Records of Criminal Incidents ..... 13

        Recommendations..... 15

    Summing Up: Comparing the Two Data Sources ..... 17

Chapter 2. Improving the Capacity of Hospital ED Data Systems to Track  
Nonfatal Firearm Injuries..... 18

    Introduction ..... 18

        Emergency Departments are an Important Data Source for Firearm Injury..... 19

        What Can and Cannot Be Expected from Hospital Surveillance ..... 20

    The NEDS and Statewide ED and Inpatient Databases ..... 24

        Background and History ..... 24

        Evaluating Sensitivity and Accuracy of NEDS for Firearm Injury Surveillance ... 26

        Conclusion and Recommendations ..... 31

    NEISS ..... 32

        Background and History ..... 32

        The Problem of Unstable Annual Estimates ..... 32

The CDC/CPSC Plan to Improve NEISS-AIP .....	34
Conclusion and Recommendations .....	36
Firearm Surveillance Through Emergency Rooms (FASTER).....	38
NSSP .....	38
Background and History .....	38
The FASTER Program .....	39
Conclusion and Recommendations .....	42
Appendix.....	43
Appendix 1 .....	44
Sources and Methods for Calculating Intent-specific CFRs for Firearm Injuries. 44	
Appendix 2 .....	46
Relevant ICD External Cause-of-Injury Codes for Firearm Injuries (Excluding War Operations).....	46
References .....	50
Chapter 3. Measuring Gun Violence Using Police Data .....	56
Introduction .....	56
Measuring Gun Violence in the U.S.....	57
Proportion of Gun Violence Reported to Police .....	60
Police and ED Overlap in Assault Gunshot Victim Reporting.....	60
Identifying Gun Violence in Police Data.....	60
National Police Data Systems for Measuring Gun Violence .....	62
UCR SRS .....	62
NIBRS .....	63
NIBRS Gunshot Injury Policy Recommendations Under Consideration.....	66
Transitioning to NIBRS .....	67
NIBRS Adoption Essential to Reliable Crime Data and Gun Injury Surveillance 68	
The Importance of State Uniform Crime Reporting Programs in the NIBRS Transition.....	71
The National Crime Statistics Exchange (NCS-X) and National Crime Estimates .....	72
The Urgent Need for Change in Federal Criminal Justice Data .....	73
National Crime Statistics Should Be Handled by a Well-Resourced Statistical Agency .....	74
The Importance of Accurate and Timely Police Data.....	74
NIBRS is a Flawed System That Should Be an Intermediary to a Better System .....	75

Conclusion and Recommendations .....	75
Priorities and Findings for Improving Police Data Surveillance of Gun Violence .....	76
Measuring Gun Violence in the U.S.....	76
The NIBRS Transition Jeopardizes Police Data Use.....	76
The Urgent Need for Change in Police Data .....	77
<b>Chapter 4. Studying Firearm Fatalities Using the NVDRS.....</b>	<b>81</b>
Synopsis .....	81
History and Development of NVDRS .....	82
Overview of NVDRS Methodology .....	82
State Variability in Death Investigation Systems.....	83
Strategies for Improving Death Investigation Systems.....	86
Published Research Using NVDRS Data.....	87
Methodologic Research .....	89
Recommendations for Increasing the Use of NVDRS by Researchers.....	90
Recommendation 1: Increase Support for Researchers Who Use NVDRS Data	90
Recommendation 2: Improve Timeliness by Providing Provisional Data	
Releases .....	91
Recommendation 3: Develop and Release Indicators of NVDRS Data Quality..	91
Recommendation 4: Fund NVDRS Methodologic Research .....	92
Recommendation 5: Strengthen the Death Investigation Systems That Underlie	
NVDRS.....	92
Summary .....	92
Literature Cited .....	94
<b>Chapter 5. Expanding Capacity and Capabilities to Monitor and</b>	
<b>Research Guns in the United States .....</b>	<b>101</b>
Introduction .....	101
Scaling Infrastructure With Machine Learning .....	103
Disciplining Data Construction to Avoid Bias.....	104
Benchmarking Strategies to Validate Data Quality.....	106
Diversified Access Mechanisms to Balance Research, Privacy, and Security .	107
Concluding Thoughts.....	108
References .....	109

**Chapter 6. Creating a Federal Gun Violence Interagency Working Group**..... 110

- Executive Summary ..... 110
- Introduction ..... 113
- Statutory and Other Mandates to Improve the Status Quo ..... 116
  - Foundations for Evidence-Based Policymaking Act ..... 116
  - The Paperwork Reduction Act of 1995 ..... 117
  - Information Quality Act of 2000 ..... 118
  - OMB Memorandum 14-03: Guidance for Providing and Using Administrative Data for Statistical Purposes ..... 119
  - Executive Order on Advancing Racial Equity and Support for Underserved Communities Through the Federal Government (EO 13985)..... 119
  - Federal Data Strategy ..... 120
  - Summary of Mandates ..... 121
- Implementation Options ..... 121
  - Establish a New Interagency Working Group Chaired by the Chief Statistician of the United States ..... 122
  - Establish an Interagency Subgroup Under the Umbrella of the Interagency Working Group on Equitable Data Established by Executive Order on Advancing Equity and Support for Underserved Communities (EO 13985) ..... 123
  - Establish an Interagency Federal Advisory Committee With Non-federal Members to Provide Advice Regarding Non-federal Datasets Relevant to Gun Violence..... 124
  - Establish an Interagency Pilot Project to Begin to Bring Data Together to Answer High-priority Questions About Gun Violence, With a Charge to Identify Needed Improvements in Content and Infrastructure, As Well As Barriers to Access. ... 125
- Role of State and Local Governments and Organizations ..... 127
- Current Federal and State Data Sharing Efforts ..... 130

**Chapter 7. Practical Steps for Building State Capacity and Infrastructure to Use Data for Evidence-Based Decision Making**..... 132

- Nancy Potok, PhD and Nick Hart, PhD | NAPx Consulting and Data Foundation 132
- Executive Summary ..... 132
- Introduction ..... 134
- What Can Be Learned from the Federal Experience ..... 136
  - Foundations for Evidence-Based Policymaking..... 136
  - The National Data Infrastructure and Evidence Ecosystem as Resource for States ..... 138

Federal-State Data Partnerships .....	140
Federal Data Modernization Efforts Relevant to States.....	143
Building State Capacity for Evidence Building .....	145
Data Leadership and Strategy .....	147
Legal and Regulatory Authority and Policies .....	149
Data Governance.....	151
Transparency of Processes and Uses.....	153
Accessibility for Data Use .....	154
Oversight and Accountability .....	155
Sustainability .....	155
A Roadmap for State Data Capacity and Infrastructure .....	156
Focus on Improving Collection, Access, and Use of Localized Firearms Data .....	160
Key Findings: .....	161
Conclusion .....	163
References and Resources .....	164

# Executive Summary

One of the great policy successes of the last decade is the increasing role of rigorous, objective, and transparent data and research in policymaking. Developing and implementing a data-driven government in which valid and reliable evidence informs solutions to our nation's most pressing health and safety challenges is more critical than ever as those challenges are ever more complex. Nowhere is that data foundation more needed than in the realm of firearms violence. Trustworthy data is a much-needed bridge to effective policymaking that can reduce the number of firearm accidents, suicides, homicides, and assaults. In an age of intense partisanship, shared facts are the cornerstone for building a shared purpose. The shared purpose of modernizing firearms data infrastructure is to improve public safety by reducing gun violence.

In the fall of 2020, Arnold Ventures, a philanthropy dedicated to maximizing opportunity and minimizing injustice, and NORC at University of Chicago, an objective nonpartisan research institution, released the *Blueprint for a US Firearms Infrastructure* (Roman, 2020)<sup>1</sup>. The *Blueprint* is the consensus report of an expert panel of distinguished academics, trailblazing practitioners, and government leaders. It describes 17 critical reforms required to modernize how data about firearms violence of all types (intentional, accidental, and self-inflicted) are collected, integrated and disseminated. This project, which is also supported by Arnold Ventures, takes the conceptual priorities described in the *Blueprint* and proposes specific new steps for implementation.

The first step in building a better firearms data infrastructure is to acknowledge where we currently stand. In *The State of Firearm Data in 2019* (Roman, 2019)<sup>2</sup>, the expert panel found that while there are a substantial number of data sources that collect data on firearms violence, existing datasets and data collections are limited, particularly around intentional injuries. There is some surveillance data, but health data on firearms injuries are kept separately from data on crimes, and there are few straightforward ways to link those data. Data that provide context for a shooting—where the event took place, and what the relationship was between victim and shooter—are not available alongside data on the nature of injuries. Valuable data collections have been discontinued, data are restricted by policy, important data are not collected, data are often difficult to access, and contemporary data are often not released in a timely fashion or not available outside of specialized settings. As a result, researchers face vast gaps in knowledge and are unable to leverage existing data to build the evidence base necessary to adequately answer key policy questions and inform firearms policymaking.

---

<sup>1</sup> Roman, John K. (2020). *A Blueprint for A U.S. Firearms Data Infrastructure*. Chicago: NORC at the University of Chicago.

<sup>2</sup> Roman, John K. (2020). *The State of Firearms Data in 2019*. Chicago: NORC at the University of Chicago.



In the *Blueprint*, the expert panel developed a set of recommendations organized around a reconceptualization of how data are collected and who collects data. The broad themes from the *Blueprint* are as follows:

- Almost all surveillance data in health and criminal justice is generated locally. It is a high-priority to provide information, technical assistance, implementation supports, and funding to state and local governments to improve their collections.
- Comprehensive monitoring of all federal data collections is needed to ensure that important data elements are being collected, data gaps are being addressed, and quality issues are quickly resolved.
- Timely dissemination of key data is important, including the development of guidelines to ensure consistency across collections and that resources are made available to speed reporting for collections with historical delays.
- Improvement is needed in strategic communication about the purpose and use of data to federal agencies, researchers and to the general public.

The current report builds on the *Blueprint* by developing implementation guidance for key recommendations. Where the *Blueprint* included actionable recommendations, such as naming discontinued surveys that should be resurrected, this report develops specific recommendations for implementation. The report is centered on three topics that were the highest priority for the expert panel but that required additional research before guidance could be disseminated. The research findings from that additional investigation are reported here, and recommendations to facilitate implementation are described. The three topic areas are as follows:

- The creation of a nonfatal firearms injury database
- Increasing the quality, availability, and usefulness of firearms data for research and policy
- Practical steps for building state capacity and infrastructure to use data for evidence-based decision-making

## Creating a Nonfatal Firearms Injury Database

The most glaring issue in building a U.S. firearms data infrastructure is the almost total absence of data on firearms-related injuries. In *Comprehensive Data on Gun Violence: Current Deficits, Needed Investments*, Philip Cook outlines the scope of the problem. Firearms injury data serves two purposes. Surveillance of firearms injuries would provide data on trends and patterns. It would also yield rich information about any underlying crime, which would better inform policy development, planning, and needs assessment. There are comprehensive sources of data on fatal shootings in public health (the National Violent Death Reporting System (NVDRS)) and in criminal justice (the Supplementary Homicide Report (SHR)) that provide trends and pattern data, as well contextual information for decision-making. However, there is no analog for nonfatal firearm injuries. In public health, there are three potential sources of data that draw

principally from emergency department data. There are limitations to each data source as the foundation for a nonfatal injury database and Cook details the improvements that would be needed—and the prospects for those improvements—for each. Cook also considers the challenge of developing a nonfatal database from police records, which provide rich data about the criminal incident but that lack pertinent information about whether an injury was from a shooting. Finally, Cook describes the challenges to national crime statistic data collection and reporting resulting from the ongoing transition by the Federal Bureau of Investigation (FBI) to a new way of collecting crime statistics and how that problem must be resolved before a nonfatal firearms database can be developed from police sources.

In, *Improving the Capacity of Hospital Emergency Department Data Systems to Track Nonfatal Firearm Injuries*, Catherine Barber examines how to build a firearm injury surveillance system from existing public health data. Barber notes that almost all shooting victims who are medically treated receive care in an emergency department, and that the coding system used for hospital billing already has the capacity to identify gunshot wounds. Three data systems could, through relatively modest tweaks, be used to greatly enhance monitoring, prevention and response to firearm injuries. The National Emergency Department Sample (NEDS) and the statewide emergency department databases from which it draws offer a rich source of firearm injury data. The challenge, however, is in the way firearm injuries are coded: currently far too many intentional injuries—mostly assaults—are coded as accidents. By contrast, the National Electronic Injury Surveillance System (NEISS) accurately records the cause of the injury but because of some problems with its sample design and small size, it yields imprecise estimates. The system could be substantially improved with additional funding for a new sample design and a larger number of reporting hospitals. Perhaps the most intriguing data source is the National Syndromic Surveillance Program (NSSP) which collects electronic health record information in near real-time on over 70% of emergency department visits nationally to track issues like disease outbreaks. The challenge with NSSP is that it is a new source of firearm injury data, with a pilot program (FASTER) having been launched this year in ten states. Barber concludes that investments should be made in all three systems, as each provides a slightly different perspective on firearms injury surveillance and that these improvements could likely be completed within three years.

In *Measuring Gun Violence Using Police Data*, Susan Parker describes the importance of police data as a unique source of information about gun violence. Police data measure the full scope of violence committed with a firearm, from threats to assaults to shootings -- even if no one is injured during a crime. Police record data on the location, circumstances, and perpetrators of gun violence, detail that is not tabulated in public health sources. However, the incumbent data system for national crime surveillance, the Uniform Crime Reporting (UCR) Program's Summary Reporting System (SRS) does not differentiate shootings from other criminal firearm use and gathers only monthly aggregate counts of crimes within a law enforcement agency's jurisdiction. The SRS replacement, the National Incident Based Reporting System (NIBRS), provides much richer incident-level data, such as the circumstances, relationship between victims and perpetrators, and other contextual information for each reported crime. While a handful of states

have for decades fully implemented incident-level NIBRS reporting, many states lag far behind. California and Illinois -- and by extension the Los Angeles Police Department and the Chicago Police Department -- do not report data to NIBRS in 2020 nor are they expected to in 2021. In contrast, more than 85% of law enforcement agencies annually report to the SRS. Despite this substantial reduction in the number of reporting agencies, NIBRS replaced the SRS on January 1, 2021 leading Parker to question whether NIBRS reporting is sufficient to generate reliable crime surveillance data. Parker notes that while several relatively small fixes, some already underway, would dramatically improve NIBRS measurement of gun violence, those issues are secondary to the larger problem of low NIBRS adoption. Parker offers several recommendations with the potential to address these substantial problems, which likely require a major review of NIBRS and federal police data collection systems.

## Increasing the Quality, Availability, and Usefulness of Firearms Data for Research and Policy

Firearms research has long been limited by a perceived prohibition on federal agencies to fund research related to the use of firearms. An amendment to the 1996 Omnibus spending bill (widely known as the Dickey Amendment) required that “none of the funds made available for injury prevention and control at the Centers for Disease Control and Prevention (CDC) may be used to advocate or promote gun control.” An amendment to the 2003 federal spending bill (widely known as the Tiahrt Amendment) similarly restricted the Bureau of Alcohol, Tobacco, Firearms and Explosives from sharing firearms trace data. While those policy restrictions have been lessened in recent years, the market for firearms research remains substantially constrained.

For the last three decades, these policy restrictions have severely limited the number of research projects about firearms and public safety. By the mid-2010s, there were only a handful of researchers dedicated to the study of firearms violence. Similarly, these restrictions limited federal, state, and local agencies’ experience sharing data on firearms and requesting research proposals to study firearms-related questions. In addition, while some of the prohibitions on research have been lifted, a cloud remains over this field of research. Further, the deep partisan division about the general role of guns in contemporary American society creates a culture of mistrust around even the most rigorous and transparent studies. Combined, these factors have created a lack of researcher and funder capability and capacity to study firearms.

Better research on the relationship between firearms ownership, storage, and use and suicide, assault, homicide, and accidental injury is critical to formulating a more coherent public policy that maximizes public safety. Addressing this constraint on the research market requires improvements related to both production and use of data. Demand may be considered as researcher interest in studying firearms-related questions—it can be increased in several ways, but perhaps most efficiently by increasing the quality and comprehensiveness of existing data, which effectively lowers the cost of conducting research. In the chapters that follow, two approaches are considered. One is the use case of the NVDRS, a valuable source of data to

understand the relationship between firearms and mortality that is a model for structuring other firearms-related data. The other is the Criminal Justice Administrative Records System (CJARS), which is a novel effort at the University of Michigan to integrate a wide swath of criminal justice data and, to integrate those data with wage and other data. On the supply side, the most productive approaches are to coordinate federal government activities related to firearms research and data collection, and one approach to doing so is considered below.

In *Studying Firearm Fatalities Using the NVDRS*, Steve Marshall describes the usefulness and limitations of an effective surveillance system that has recently been expanded to all 50 states. To develop the NVDRS, several obstacles apparent in less mature firearms data systems were overcome, including the need to standardize reporting to account for the variation between states in how law enforcement and medical examiners respond to and record information in violent death investigations. Efforts by the CDC to implement quality control procedures have generally been successful. As Marshall demonstrates, a large body of scholarship about the nature of violence has been produced using the NVDRS. Research facilitated by NVDRS data is increasing rapidly and is likely the single most important source of data currently available to understand the relationship between firearms and violence. Marshall makes five key recommendations to continue building the success of the NVDRS, including increased support for researchers, mechanisms for more timely data releases, release of quality metrics, and critically, and investment in improving the NVDRS and strengthening the underlying local and state systems that contribute data to the NVDRS. Those improvements align with the recommendations for improving the NVDRS in the Parker paper—indeed, implementing the NVDRS recommendations concurrent with basic investments in NIBRS would be the most efficient approach to improving that collection.

In *Expanding Capacity and Capabilities to Monitor and Research Guns in the United States*, Michael Mueller-Smith addresses the problem of increasing researcher demand for firearms data through a series of recommendations to address challenges in integrating data. Using CJARS as a model integrated research repository, Mueller-Smith recommends that designers of integrated systems take a strategic approach to data collection, building upon prior efforts to build momentum by tackling specific measurement goals rather than solving all policy questions simultaneously. Mueller-Smith describes four solutions for aggregation problems, as follows: 1) to use machine learning to scale data collected for operational rather than research purposes; 2) use strategies to integrate multiple sources into a single structure while avoiding duplication and pooled events; 3) creating an organizing framework for data with inconsistent definitions and data layouts and inadequate identifiers; and 4) diversifying means of accessing the data, to improve access for diverse audiences.

The last paper on increasing the quality, availability, and usefulness of firearms data for research and policy considers the strategy for developing firearms data, where the goal is to facilitate increased coordination among federal agencies. In particular, as Potok details in *Creating a Federal Gun Violence Interagency Working Group*, an expert panel made a number of recommendations to create a national strategy around firearms data and research. Key

elements include establishing clear and consistent priorities for firearms data and research; integrating public health, crime, and firearms data; and, reducing constraint on data sharing across agencies and the release of data to researchers and the public. Potok offers four recommendations to address these concerns, centered on the development of an interagency federal workgroup chaired by the Chief Statistician of the U.S. That workgroup would focus on data quality and coverage improvements and leverage a number of statutory mandates within the federal government, specifically the Evidence-Based Policymaking Act of 2018. Potok recommends that the workgroup coordinate with other federal workgroups, in particular the Equitable Data Working Group, and that a federal advisory committee be created as well. Further, Potok recommends that the centerpiece of the interagency’s mission be the creation of a pilot project that highlight the diversity of data collected across systems and the value of integrating those systems and reporting. The Potok paper aligns with the Mueller-Smith recommendations that targeted policy questions that create momentum are the most productive means to build a cross-sector firearms data system, rather than a one-sized solution to all policy problems.

## Practical Steps for Building State Capacity and Infrastructure to Use Data for Evidence-Based Decision-Making

The final paper in the collection addresses the foundational challenge for all firearms data improvements: How to build local and state capacity. While some data that can be used to inform key firearm research, such as the Healthcare Cost and Utilization Project (HCUP) billing data or electronic health records, are not recorded and reported by local governments, most critical data in this field originate at the local level. Each of the other papers in this collection describes a particular challenge about the variability in data quality. Whether data are generated by local law enforcement for use through the NIBRS or from local medical examiners for use through the NVDRS, consistent and accurate data collection and reporting at the local level—where data are input and collated—is critical to the success of any national firearms research database or surveillance system. In *Practical Steps for Building State Capacity and Infrastructure to Use Data for Evidence-Based Decision-Making*, Nancy Potok and Nick Hart create a roadmap for state and local governments to improve their data systems and structures. At the heart of their recommendations are the lessons learned from the creation of an evidence ecosystem in the federal government, built on recommendations from the Commission on Evidence-Based Policymaking. Those lessons include guidance on empowering data leaders, creating transparency, using local priorities to guide the development of the ecosystem, creating cross-agency strategies, and prioritizing transparency. Critical to this process are the federal-state partnerships and the insight that development of effective partnerships is bi-directional, with guidance from the states no less important than guidelines from the federal government.

## Conclusion

After reading these recommendations it would be natural to ask: What is the highest priority if we want to reduce gun violence? There are no easy answers to this question. In the short term, speeding release of federal data collections is likely the quickest means to spur researcher demand. Modest improvements in the three healthcare data collections is likely a much shorter route to the development of a national nonfatal firearms injury database. However, neither of these solutions alone solves the long-term problems of low quality and high variability in local and state data input and reporting. Solutions to that problem requires a much larger and longer-term investment. Similarly, a focus on health data for the nonfatal firearms database omits critical data that could only be captured from police data—critical information about assaults and robberies with a gun where no one is injured. And a shorter-term focus would put aside what is likely the biggest concern in this collection, namely, the giant step backwards in national crime statistics reporting, including firearms crime, that the United States will confront directly in 2022 when the insufficiencies of NIBRS become clear. The best recommendation then is from the papers that recommend focusing on one data problem at a time and to build momentum within a careful strategic framework.

# Chapter 1. Comprehensive Data on Gun Violence: Current Deficits, Needed Investments

Philip Cook, PhD | Duke University

## Introduction

The surge in gun violence starting in 2020 underscores the urgent need for a reliable system for documenting firearms violence. Prevention efforts require reliable and timely information. There are myriad questions that can be answered only with more and better data. How many people are shot, under what circumstances, and with what outcome—and are these rates trending up or down? How are shootings distributed across jurisdictions and among different demographic groups? What determines whether victims of gun assault and robbery are killed, wounded, or “merely” threatened? What do we know about the shooters in assault and robbery cases? Which programs and policies have been effective in reducing gun violence? Current data systems, and especially the NVDRS, provide detailed documentation for fatal shooting cases, including both suicides and homicides. But with respect to nonfatal shooting cases, including 80 percent of those shot in assaults and robberies, we remain largely in the dark.

This briefing identifies current deficits and promising avenues to develop a comprehensive firearms injury database. We focus on improving the quality and scope of existing data systems that use administrative records. There are two primary sources: hospital medical records and police crime records. These are documented and discussed in detail in the associated reports by Cathy Barber (see Chapter 2) and Susan Parker (see Chapter 3).

Any comprehensive data system for tracking and analyzing nonfatal firearms injuries will require data from both hospital medical records and police records. The current systems to compile hospital medical records are being refined and are on track to provide reliable national estimates within the next three years. Since hospital medical records are also compiled at the state-level, they could be used as a basis for state-level surveillance of firearms injuries. Police records have greater potential in some respects, but the existing UCR system is seriously deficient and the FBI has yet to specify a clear path forward.

## Potential Uses for Data on Firearms Injuries

A comprehensive data system on gun violence is needed for surveillance and for policy analysis.

- **Surveillance** provides reliable descriptive information on trends and patterns, to answer the “where,” “when,” “who,” “how,” and “why” questions. The ideal system would provide reliable statistics on a timely basis, aggregated to the local, state, and national level. In addition to informing the public, surveillance would be useful for planning and for needs assessment. A system limited to fatal shootings (i.e., suicide and homicide) does not tell the whole story, since nonfatal shootings exhibit much different patterns; for example, most fatal shootings are suicides (60 percent), but only five percent of nonfatal shootings are self-inflicted. In addition, most nonfatal shootings occur in the context of criminal assaults. The difference reflects the fact that almost all firearms suicide attempts result in death, while only about 20 percent of gunshot assaults result in death and leave many more survivors overall.
- **Policy analysis and planning offers** the data needed to understand the causes of gun violence—for policy design—and to evaluate the impacts of public and private actions intended to prevent or mitigate gun violence. For these purposes, it is important to have data on the shooters and the incidents that led to the shootings. A comprehensive data infrastructure should include measures of the underlying crimes of gun assault and robbery. The rationale follows from our understanding of gun violence prevention. In most gun assaults and robberies, the victim is threatened but not shot; however, a shooting is always a possible outcome. An intervention that reduces the overall volume of gun robbery and assault would likely reduce firearms injuries.

### Data Sources

The two primary sources of data on gun violence are hospital medical records and police crime records. In principle, hospital records are comprehensive, since about 90 percent of all firearms injuries are treated in a hospital emergency department (ED). The information in hospital records is routinely coded and compiled for billing purposes. The primary focus is on coding diagnoses and procedures, although injury cases coding is now required to include the external cause (e-code). Most gunshot wounds are identified as such in e-coding. However, accuracy in coding the *intent* of the shooting (assault, accident, suicide attempt) has proven a greater challenge.

Police records are comprehensive on shootings that would normally be investigated—cases in which the victim was shot by another person—and ordinarily provide more detail about circumstances and intent than medical records. Most nonfatal gunshot wounds occur in the context of criminal encounters. The remaining cases—accidental and self-inflicted shootings—are likely to be missing from police records.



A third potential source of information on assaults in which the victim is shot is the National Crime Victimization Survey (NCVS), conducted by the Bureau of Justice Statistics (BJS) since 1973. In practice the NCVS national estimates of gunshot assault are a fraction of the true total, due to intrinsic limitations of the sampling process. We do not know of any feasible revision to the NCVS that would solve this problem and do not view a household survey as a promising source.

## Hospital Data

Three existing systems have great potential to provide useful national surveillance systems for nonfatal gunshot injuries. Each has its own strengths and limitations.

### 1. HCUP NEDS

Most states with statewide hospital databases disseminate their own data locally and forward their data to the HCUP at the federal Agency for Healthcare Research and Quality. HCUP uses these data to construct the National Inpatient Sample (NIS) and the NEDS. NEDS is a stratified, single-stage cluster sample constructed by categorizing hospitals according to five strata: geographic region, urban/rural location, teaching status, ownership, and trauma-level designation. In 2019, the sample comprised 990 hospitals in 37 states that submitted data on nearly 36 million ED visits of all types, from which HCUP projected total visits of over 143 million for the nation. HCUP makes individual-level NEDS data available to researchers for a fee (as well as access to NIS and state-specific databases). In addition, HCUP disseminates aggregate state and national data via an online data-query interface (HCUP-Net).

**Strength.** NEDS is in place and can be used to generate national and regional estimates of the incidence of firearm injuries, together with patient demographics, seasonality, intent, nature of the wounds, procedures, payors, charges, length of stay, and outcome.

**Limitations.** For gunshot cases to be identified as such, medical records coders must include the appropriate coding for the mechanism of injury (for example, gunshot) rather than only the nature of the wound. This so-called e-coding has become common in recent decades and appears to be near universal in most states. Beyond identifying a gunshot as the mechanism of injury, e-codes indicate the intent and whether the gunshot was a suicide attempt, accident, or assault. In practice, the e-coding of intent is not accurate. A large proportion of firearm injuries (mostly assaults) are misclassified as accidents. This is a serious problem caused in part by limitations of the hospital coding software. If the coder does not specify that the injury is intentional (assault, self-inflicted, and so forth), the software in common use defaults to “accidental.” While NEDS is thought to provide a reasonably accurate estimate of the overall volume of nonfatal firearms injuries, it is not a reliable source concerning the breakdown of injuries by intent: NEDS-based estimates greatly underestimate injuries from gun assault and overestimate accidental injuries.

Other limitations include the multiyear lag in releasing data and associated estimates and the fact that NEDS cannot be used to generate estimates at the state and local level.

## Recommendations:

- NEDS is widely believed to provide reasonably accurate estimates of the national and regional volume of firearms injuries (albeit with long delay) and can be utilized as such. More research is needed on the sensitivity of e-coding to identify firearms injuries.
- NEDS estimates are not reliable in estimating the rates of firearms injury by specific intent. A study funded by Arnold Ventures will propose a change in coding guidance to the joint CDC National Center for Health Statistics (NCHS)/ Centers for Medicare & Medicaid Services (CMS) International Classification of Disease (ICD)-10 Coordination and Maintenance Committee in March 2022. Federal resources are needed to implement reforms and educate researchers on pitfalls of the existing data.
- NEDS is not structured to provide state-level estimates of firearms injuries, but hospital medical data are compiled in the HCUP State Emergency Department Databases (SEDD). Coupled with the HCUP State Inpatient Databases (SID), these data may be used for firearms injury surveillance and analysis in the 42 states that currently report these data. Public use is hampered by the fact that data are only available for public use at a substantial charge and are of variable quality. CDC could provide a useful service by purchasing these data, abstracting injury cases, and providing convenient public access to the resulting state-level files. The data limitations discussed above would remain.

## 2. NEISS Firearm Injury Surveillance System (FISS)

CDC collaborates with the U.S. Consumer Product Safety Commission to collect data on consumer product injuries through the NEISS. Two related systems collect data on firearm injuries through the NEISS-FISS, using a sample of 96 EDs intended to be nationally representative, and data on all injuries through the NEISS-All Injury Program (AIP), based on a two-thirds sub-sample.

**Strengths.** Working with a small sample of EDs enables CDC to employ and train expert coders. In particular, the classification of intent for firearm injuries is handled by a small number of coders at the CDC National Center for Injury Prevention and Control (NCIPC) and appears largely accurate. Data from AIP, which began in 2000, have been made available on Center for Disease Control and Prevention's (CDC's) user-friendly Web-based Injury Statistics Query and Reporting System (WISQARS)-Nonfatal Injury Data querying interface. Individual-level data from FISS are available to researchers from the Inter-University Consortium for Political and Social Research website.

In comparison with NEDS, NEISS estimates are more timely; annual estimates are available within 24 months.

**Limitations.** The sample of hospitals used by NEISS-AIP is too small to support stable national estimates, leading to suppressed cells on WISQARS-Nonfatal in recent years. Also, because firearm injuries are highly concentrated at a relatively small number of hospitals, even national estimates based on the full NEISS-FISS sample can shift abruptly from year to year as individual hospitals drop out of the sample and are replaced by hospitals with very different firearm caseloads.

**Recommendation.** The CDC is committed to reforming the NEISS-FISS program. Current CDC initiatives appear adequate to produce reliable national estimates of the volume of firearms injuries overall and by specific intent by 2023. These reforms should receive continuing support. The sample will be increased. When hospitals exit the sample, care will be taken to ensure that the replacement hospital from the same sampling tier has a roughly equivalent firearm injury caseload. The estimated coefficient of variation will be large (16.7 percent) but acceptable and far better than the current coefficient of variation of over 30 percent. More importantly, estimates will no longer be subject to large jumps and false trends due to the vagaries of hospital turnover in the sample.

### 3. NSSP

The purpose of NSSP is to “send early warning signals from EDs to public health” professionals in near real-time on threats such as infectious disease outbreaks, terrorism-related attacks, overdose spikes, etc. It is operated by the CDC’s Division of Health Informatics and Surveillance. Approximately 70 percent of hospital EDs now transfer data on all visits to a CDC-designed platform. Data elements include presenting complaints, triage notes, patient age and sex, and diagnosis codes and external cause-of-injury codes when available. NSSP defines specific syndromes (e.g., COVID-19, overdose) and uses natural language processing and artificial intelligence to identify ED visits. Currently the CDC NCIPC’s Firearm Injury Surveillance Through Emergency Rooms (FASTER) pilot in ten states will determine whether firearm injuries can be reliably detected and if so, whether they can be reliably classified as to intent.

**Strengths.** NSSP counts are available within 48 hours of a hospital visit and could be used to detect surges in gun violence, to help develop a timely response. Further, the system is in place, with 70 percent of EDs transferring data on a regular basis. The data are a census, rather than a sample, so that they could be used for state-level and small-area level surveillance.

**Limitations:** The FASTER pilot is only in its first year. It appears the system will be successful in identifying firearm injuries; however, accuracy at classifying by intent will likely be poor, at least in the short term. This is not necessarily a fatal flaw, since most ED-treated gunshot wounds are assaults, and spikes and dips in the number of gunshot wounds are driven by assaults. A second limitation is that the CDC supplies the platform for NSSP but does not have access below the national or regional aggregate level. Only hospitals and state/local health

departments can access the state and local aggregate and individual-level data except where they have given CDC explicit, active permission or where CDC has paid for the data.

**Recommendations:** Support the CMS current proposal to expand NISSP-participating EDs from 70 percent to nearly 100 percent. If the FASTER pilot proves successful, provide funding for the CDC NCIPC to access state and local data and to provide the public with convenient access to these data.

## Police Records of Criminal Incidents

Police records provide detailed information on violent crimes. Of the crime categories used by the FBI UCRs, gun robbery and gun aggravated assault include almost all cases in which one person is shot by another and survives. These crime categories also include the more frequent instances in which an individual is threatened with a gun but not shot.

Police data provide an alternative to hospital data for surveilling gunshot injuries or at least that subset of gunshot injuries resulting from one person shooting another, which are the bulk of all nonfatal shootings. Police agencies know of most such shootings. Reports from citizens (calling 911) are supplemented by reports from medical providers, who in almost all states are obligated to report to the police all gunshot cases that they treat. For their own purposes, some police departments have record systems that identify which violent crimes involved a gunshot injury, but no state or federal system currently distinguishes gunshot victims from other victims of gun crimes.

As of 2021, the FBI UCR—the primary source of national crime statistics—implemented a fundamental change. Since the early 1930s, the UCR has compiled and published summary reports from law enforcement agencies, with the reports comprising counts of certain types of crime. The list of crime types includes the following: murder, aggravated assault, and robbery, classified by type of weapon. The summary reports are widely used by the public and policymakers to track trends and patterns in crime. Summary reporting has been discontinued by the FBI, replaced by a requirement that agency reports use the NIBRS. NIBRS has been an option available since the 1980s but has not caught on. The expansion of NIBRS to all reporting entities would allow the UCR to continue reporting national counts of aggravated assaults and robberies known to the police. The FBI is considering a minor adjustment in coding instructions, to provide a comprehensive database on cases in which one person shot another. For such cases, NIBRS has greater detail than hospital records and provides information on the crimes committed in conjunction with the shooting and on any suspects.

Despite the change in direction by the FBI, and some effort by federal and state criminal justice agencies, a large share of police departments, including many of the largest departments, do not participate in NIBRS. (It should be noted that UCR use by law enforcement agencies has always been voluntary but was near universal when summary counts were all that was needed.)

It is unclear what is required to increase participation in NIBRS from the current level—55 percent of law enforcement—back to the level of participation enjoyed by the SRS of over 90 percent.

This interruption in comprehensive crime reporting is an urgent concern that transcends the problem of measuring gun violence. The UCRs have been essential to measuring levels, trends, and patterns of crime in the United States. The FBI's decision to discontinue the SRS in place for the last nine decades, and to replace it with an unenforced requirement for agencies to submit crime data in NIBRS, means that the nation has lost the leading measure of crime trends and patterns. The scope of this problem is far broader than the need to develop the infrastructure for gun violence prevention.

Until more agencies begin reporting in NIBRS, one alternative to generate valid national estimates of crime counts from law enforcement agencies is to create an estimation procedure based on a representative sample of agencies. BJS has pursued this possibility with the National Crime Statistics Exchange (NCS-X). A sample of 400 agencies, including the 72 largest agencies, is designed to produce reliable national estimates if all 400 agencies participate. Unfortunately, the grants program and offer of technical assistance have proven inadequate, and about half of the NCS-X sample is still not submitting NIBRS data.

An inexpensive tweak in NIBRS would make the data useful to surveilling and analyzing nonfatal shootings in those jurisdictions that do report. From the current NIBRS format, it is possible to identify assaults and robberies in which the perpetrator used a gun and to determine whether the victim was injured. What should be added is an item specifying whether the injury was a gunshot wound. Such a modification in NIBRS reporting was recommended by an FBI advisory board and will likely be adopted. With the modification in place, NIBRS data for participating jurisdictions could be used for analytic purposes.

**Strengths.** Police records of shooting cases should provide information that is not ordinarily included in hospital records—in particular, data on the shooter (when available) and the incident that led to the shooting. Such information is vital to understand the great majority of nonfatal shootings, those in which one person shoots another in criminal circumstances.

For local jurisdictions and entire states that are well represented by agencies that report crime through NIBRS, the resulting UCR data on gun crime can be used for surveillance purposes and to provide detailed data for policy analysis. These data will become much more useful for our purposes when and if NIBRS reporting is modified to identify gunshot wounds.

Police data also provide unique information on a larger set of gun crimes that have no counterpart in the hospital data—assaults and robberies in which the victim is not shot—that provide context for the shooting cases. These data are essential to investigating the epidemiology of gun violence.

Traditionally, the UCR data for each calendar year have been reported to the public within 10 or 11 months, yielding more timely results than NEDS or NEISS.

**Limitations.** UCR counts no longer serve as the basis for tracking national trends in violent crime. To date, the effort to generate a national estimate from a representative sample of agencies has failed due to low participation.

## Recommendations

The current version of NIBRS does not specify whether a violent crime victim was shot. We recommend revision of the NIBRS form so that firearms injuries can be identified and analyzed.

In the absence of a mandate, agencies must be persuaded to participate voluntarily in any crime reporting system. The low NIBRS participation rate may reflect a widespread judgment by law enforcement agencies that NIBRS lacks sufficient value to them or their communities to warrant making the switch from summary reporting. We recommend a national survey of agencies, both those now using NIBRS and those that are not, to ascertain what is needed to increase participation.

NIBRS is a notoriously complex and demanding system. A streamlined version of NIBRS may preserve the main advantages of incident-based reporting while offering a more attractive value proposition for law enforcement agencies. A NIBRS redesign should be on the table in planning next steps.

Our principal recommendation is that lead responsibility for generating national crime estimates be shifted from the FBI. Participation rates have dropped dramatically under the NIBRS requirement; for this reason, for the foreseeable future, the FBI's UCR counts will no longer serve as a valid indicator of national crime rates. Data currently available from state crime agencies and UCR (NIBRS) reports could be compiled and used to estimate state-level and national crime rates. However, the estimation process requires statistical sophistication. Requisite expertise may be available in a statistical agency such as BJS. BJS manages the NCS-X sample of 400 law enforcement agencies that was created to provide national estimates based on NIBRS data during the transition to full participation in NIBRS. (BJS is also responsible for the NCVS, which since 1973 has provided national estimates of some crime types based on a household survey.) The role of BJS could be expanded further to include issuing national compilations of crime data collected by state UCR agencies, many of which have continued summary reporting.

To accelerate compliance with the requirement to use NIBRS for UCR crime reporting, a combination of inducements is warranted. In addition to technical assistance and a grants program with a minimum of red tape, Congress may eventually require NIBRS reporting as a precondition for federal grants. One component of the grants program could be channeled

through the state agencies that current manage UCR. Many states are in close touch with individual agencies and may be able to use the funds to effect change.

In sum, we recommend the following:

1. Modify NIBRS so that the type of injury in violent crimes is identified and whether the victim was shot
2. Increase participation in NIBRS by law enforcement agencies
  - ▶ Solicit information from police agencies to identify barriers to NIBRS participation
  - ▶ Fund state UCR agencies to manage programs of grants and technical assistance for agencies willing to convert to NIBRS
  - ▶ Require NIBRS participation as a condition of eligibility for other federal grants programs
3. Generate national estimates of crime rates by statistical inference from the subset of agencies that do report to NIBRS
  - ▶ Transfer lead responsibility for crime reporting from the FBI to BJS or another statistical agency
  - ▶ Continue NCS-X sample of cities, with additional resources to encourage participation
  - ▶ Develop a statistical model for making inferences from the sample of agencies that report to NIBRS

## Summing Up: Comparing the Two Data Sources

Both hospital medical data and police crime data have certain advantages as the basis for tracking trends and patterns (surveillance) and for policy analysis.

Hospital medical data include all types of firearms injuries, not just assault. Perhaps most importantly, the federal data systems for compiling hospital medical records are in place. NEDS and NEISS are being revised and should be able to generate high-quality national estimates by 2023 with accurate coding of intent. The NSSP, if it proves successful in the pilot phase, could produce comprehensive counts of firearms injuries at national, state, and local levels with only a brief lag.

Police crime records include all types of crime committed with firearms, not just crimes in which the victim is shot. The records on assaults and robberies in which the victim is shot include more detail concerning the incident and the shooter than is available in medical records. Crime rates and incident-level data on individual jurisdictions are available to the public, for agencies that choose to participate in the FBI's UCRs; such reports are available with a lag of less than a year, far shorter than for NEISS and NEDS. However, participation rate in the UCR is currently only 55 percent, and there is no system currently in place for generating national or state-level estimates based on the sample of agencies that choose to report. A public investment in crime data is urgent.



# Chapter 2. Improving the Capacity of Hospital ED Data Systems to Track Nonfatal Firearm Injuries

Catherine Barber, MPA | Harvard University

## Introduction

Most firearm injuries lead to a trip to the ED, not a trip to the morgue. Yet today in the United States, we have no data system that accurately tracks how many nonfatal firearm injuries occur, where and to whom they occur, and whether they resulted from assault, accident, self-harm, or legal intervention. Currently, there are two long-standing national data systems for ED visits that come close to providing this information, and a third very new system shows promise.<sup>3</sup>

If the changes to three hospital data systems that we recommend in this report are implemented in the short term, we can reasonably expect that within three years the nation will have:

- Stable, annual estimates of ED-treated firearm injuries at the national level from a sample of about 1,000 emergency departments, and an annual census of ED-treated firearm injuries at the state and local level in nearly all states.
  - ▶ *FROM:* For national estimates: the NEDS from the HCUP. For state and local data: SEDDs and SIDs from state organizations or, in many states, from HCUP.
  - ▶ *PROBLEM TO SOLVE:* Currently, hospital medical records coding systems that supply data to NEDS, SIDs, and SEDDs misclassify a large proportion of intentional injuries as accidents.
  - ▶ *SOLUTION:* Recommend a new guideline to the national committee that governs how injuries are coded in hospital billing data, and work with relevant industry stakeholders (hospital information managers, coders, and the software companies whose products support coding) to ensure the new guideline is implemented.
- National estimates of firearm injuries that are accurately classified by intent type, based on a small sample of hospitals.
  - ▶ *FROM:* NEISS-FISS.

---

<sup>3</sup> A fourth, the National Hospital Ambulatory Medical Care Survey, exists but does not have sufficient case size to reliably estimate firearm injuries. Additionally, a private initiative, the National Trauma Data Bank, provides useful detail; however, because 30% of firearm injuries are not treated in trauma centers, we focus in this report on ED data systems from all acute care hospitals, regardless of trauma center designation.

- ▶ *PROBLEM TO SOLVE*: While coding of injury intent is good, sample limitations of the current public-facing data interface result in national estimates with such wide confidence intervals that they do not meet the CDC’s criteria for public release.
- ▶ *SOLUTION*: Support and maintain the sample design improvements currently underway.
- A near real-time census of ED-treated firearm injuries at the state and local level in ten states, with potential for nationwide implementation.
  - ▶ *FROM*: Firearm Surveillance Through Emergency Rooms (FASTER), based on NSSP.
  - ▶ *PROBLEM TO SOLVE*: While NSSP dates back nearly 20 years, FASTER only began in 2021 and must prove that it can rapidly and accurately identify firearm injuries from NSSP.
  - ▶ *SOLUTION*: If FASTER proves viable, expand to other states, and create a national, online data querying interface to facilitate access to aggregate state, local, and national data.

## Emergency Departments are an Important Data Source for Firearm Injury

An estimated 90 percent of people who sustain nonfatal gunshot wounds<sup>4</sup> (with the exception of minor graze wounds) are seen in the ED. Some evidence:

- 91 percent of jail inmates in five cities who were previously shot reported that they were treated in the hospital for their injuries.<sup>i</sup>
- 88 percent of people who received medical care for assaultive gunshot wounds were treated in a hospital, according to NCVS data.<sup>ii</sup> The others were treated at the scene or at a home; none were treated only in a doctor’s office or non-hospital facility.
- The total number of nonfatal gunshot wounds estimated by the NEDS appears to be in the right ballpark. How can we tell? By working backward from fatalities. That is, if we know the number of firearm-related homicides, accidents, and suicides, and we know that for firearm injuries roughly 1-in-5 assaults, 1-in-20 accidents, and 9-in-10 self-inflicted shootings result in death, we can estimate annual nonfatal injuries based on deaths.<sup>iii</sup>
- A current study<sup>iv</sup> is reviewing hospital charts for firearm injuries. Less than half a percent were cases in which a patient with a gunshot wound appeared for care days or weeks after being shot because they initially did not seek ED care for fear of being reported to police or some other reason.<sup>v</sup> Presumably, if a substantial proportion of shooting victims avoid hospital care, late presentations for pain or wound infection would be more common.

---

<sup>4</sup> Throughout we use the terms “gunshot wound” and “firearm injuries” to refer to injuries from a traditional projectile fired from a firearm. We exclude injuries caused by non-powder guns such as BB guns, air guns, and flare guns; non-projectile injuries like pistol whipping; or injuries from non-traditional projectiles like rubber bullets or bean bags.

## What Can and Cannot Be Expected from Hospital Surveillance

Surveillance<sup>5</sup> systems collect ongoing, accessible, comparable, and representative data and make it available (with confidentiality protections) to researchers, health and safety officials, and the public to help monitor and address public health problems. They differ from research studies, which are short term and cannot monitor trends over time. Research studies design instruments to collect exactly the information that they need. Many surveillance systems (including those described here) rely on existing administrative records, such as hospital billing data or medical charts, and the data they already contain. As such, users must recognize the strengths and limitations of these administrative systems. For example, hospitals may be an imperfect source of data on the specific firearm type that was used in an assault (look to police data for that), but an excellent source of information on medical severity or injuries that may not come to the attention of the police, like suicide attempts.

The nationally standardized format for hospital billing data uses a coding system that identifies injuries by their mechanism (e.g., firearm, sharp instrument) and intent (e.g., assault, accident, self-harm, legal intervention). These data are sent to statewide databases in nearly all states, and many of these states submit their statewide databases to the HCUP from which nationally representative datasets like NEDS are assembled. That this administrative data infrastructure already exists at the hospital, state, and national level is a tremendous strength. It will be an even greater strength when coding of intent improves and when the data are made more accessible.

Since the intent coding problem may prove difficult to solve, improving the NEISS system is also a prudent step, since the quality of intent coding in NEISS is already high and steps to expand and improve the sample of hospitals in NEISS are already underway. Both NEISS and NEDS (and the state databases from which NEDS draws) characterize *all* injuries by intent type, not just firearm injuries. This is important when evaluating whether changes in rates of assaults or self-harm are specific to firearms or apply to other methods as well. The purpose of preventing firearm injuries, after all, is to bring down the overall toll of violence and suicide.

Ideally, surveillance systems provide timely data. “Timely” is relative. For chronic disease, for example, annual reporting may be adequate, while for infectious disease, near real-time reporting is necessary. NEDS and NEISS report data out annually and have a one to three year reporting lag. CDC’s Len Paulozzi, former science officer at the NVDRS, has said that using sluggish systems like these for prevention efforts is like “trying to hit a tennis ball last seen two years ago.” The FASTER program, using continuous data uploads from the NSSP, could make near real-time firearm injury surveillance a reality. While FASTER may solve the need for timely

---

<sup>5</sup> The term “surveillance” is used differently in the public health field than in the public safety field. It is not about covertly gathering data on individuals for enforcement purposes. It’s about openly gathering information on populations to track the incidence and characteristics of health problems for prevention and treatment purposes.

data, the other two systems supply more detailed and—especially in the case of NEISS—quality-controlled data.

### Tracking Emergency Department-Treated Firearm Injuries: Three U.S. Data Systems

	NEDS	NEISS	FASTER/NSSP
<b>Title</b>	NEDS and SEDDs, SIDs	National Electronic Injury Surveillance System	Firearm Surveillance Through ERs, from NSSP
<b>Sponsor</b>	NEDS: HCUP, a project of the federal Agency for Healthcare Research & Quality SEDDs/SIDs: State health departments, hospital associations, or public/private consortia. Underlying data: individual hospitals' proprietary data.	Consumer Product Safety Commission. CDC's National Injury Prevention and Control Center collaborates with Consumer Product Safety Commission (CPSC) to gather data on all injuries, including more detailed data on firearm-related injuries.	FASTER (Firearm Surveillance Through ERs): CDC's National Center for Injury Prevention & Control NSSP: CDC's Division of Health Informatics and Surveillance Underlying data: individual hospitals' proprietary data.
<b>Scope of system</b>	All ED visits (medical, psychiatric, injury)	First time visits for injuries treated in the ED	All ED visits (medical, psych, injury)
<b>Data entered by</b>	Medical records coders review and code charts for hospital billing purposes	Abstractors under CPSC/CDC purview review charts and enter data; coding of intent overseen by a small number of coders at CDC based on reading brief narratives describing incident	Electronic health record data (presenting complaint, coded diagnoses, triage notes, and patient demographics) for all ED visits are automatically uploaded from participating hospitals' records to CDC BioSense platform
<b>Representativeness</b>	NEDS: National and four regions only SEDDs/SIDs: State and local	National only	Neither currently; potential for national, state, and local
<b>Census or sample</b>	NEDS: Sample (~1,000 hospitals) SEDDs/SIDs: Census	Sample (~100 hospitals)	Census among participating hospitals; NSSP currently neither census nor representative sample nationally
<b>Reporting lag</b>	2-3 Years	1-2 Years	Info: 1-2 Days Reports: Quarterly in FASTER

	NEDS	NEISS	FASTER/NSSP
<b>Title</b>	NEDS and SEDDs, SIDs	National Electronic Injury Surveillance System	Firearm Surveillance Through ERs, from NSSP
<b>Reasonably accurate estimates of overall counts?</b>	Yes	No Expected to be by 2022.	Unknown (likely Yes) Evaluation underway
<b>Reasonably accurate classification of intent (e.g., assault, accident)?</b>	No Far too many intentional firearm injuries classified as accidents	Yes	Unknown (likely No) Evaluation underway
<b>Details on shooting circumstances?</b>	No	A few	No
<b>Aggregate firearm data searchable by the public online?</b>	Yes Via HCUP-Net, but not user-friendly and only pre-2015	Yes NEISS-AIP available from the WISQARS-Nonfatal Injury Platform	Not yet
<b>Researchers can apply for individual-level data?</b>	Yes, but expensive NEDS via HCUP SEDDs via HCUP and states	Yes, free NEISS-FISS via Inter-University Consortium for Political and Social Research (ICPSR) website	Not yet
<b>STRENGTHS</b>	In 47 states today, SEDDs/SIDs provide a census of state and local ED-treated firearm injuries. Data from many of these states are centralized by HCUP, and HCUP creates stable national estimates drawn from close to 1,000 hospitals. It is likely this will eventually become a 50-state system.	Classification of intent in NEISS is reliable. Data are accessible on a convenient online data-query interface at the WISQARS-Nonfatal Injury Data Platform.	Huge strength is near real-time data. Currently data from hospitals seeing over 70% of ED visits nationally are uploaded to the NSSP platform. A current CMS initiative would boost that to close to 100%.
<b>FLAWS</b>	<i>Far</i> too many assaults are misclassified as accidents, seriously compromising an otherwise valuable data source. Also, data are difficult to access and expensive to buy.	Provides only national—not state or local—data. Small sample size and extreme geographic clustering of gunshot wounds leads to imprecise estimates. Annual changes in case estimates can be an artifact of individual hospitals entering or exiting the sample.	FASTER is very new; whether it can use NSSP to accurately detect and classify firearm injuries is currently unknown. Also, CDC has authority to access aggregate data from NSSP at national and regional level only, not more granular level (state, local, individual), with certain exceptions.

	NEDS	NEISS	FASTER/NSSP
<b>Title</b>	NEDS and SEDDs, SIDs	National Electronic Injury Surveillance System	Firearm Surveillance Through ERs, from NSSP
<b>RECOMMENDATIONS</b>	<p>A team investigating the intent classification problem will propose changes in coding guidance to the March 2022 meeting of the NCHS/CMS ICD-10 Coordination and Maintenance Committee (which governs ICD coding policy). If a new policy passes, use federal resources to:</p> <ul style="list-style-type: none"> <li>■ Publicize the change to relevant professional associations (e.g., American Academy of Professional Coders, the American Hospital Information Managers Association).</li> <li>■ Work with companies that design coding software, coding look-up tables, and electronic health records software to improve intent classification.</li> <li>■ Educate researchers and journal editors on the intent classification problem.</li> <li>■ Evaluate the proportion of hospital-treated firearm injuries that receive a) any firearm injury-related e-code, b) the appropriate intent code (assault, accident, etc.).</li> <li>■ Create a user-friendly data-query interface to access injury data at national level and at state/local level in states that contribute SEDDs/SIDs.</li> </ul>	<ul style="list-style-type: none"> <li>■ Support CSPC and CDC’s new sample plan and expansion of the NEISS-AIP to all 100 NEISS hospitals.</li> <li>■ Fund CDC to maintain the full sample after FY22.</li> <li>■ Publicize the existence of the more detailed, individual-level NEISS data available to researchers from the NEISS-FISS.</li> <li>■ For WISQARS-Nonfatal users, provide a simple explanation geared for the non-statistician on how to interpret wide confidence intervals in the data.</li> </ul>	<p>If FASTER program proves successful:</p> <ul style="list-style-type: none"> <li>■ Support near 100% participation of EDs in the NSSP.</li> <li>■ Fund CDC to expand FASTER to all states.</li> <li>■ Negotiate an agreement among CDC, the Council of State and Territorial Epidemiologists (CSTE), and NSSP Community of Practice Governing Board authorizing CDC or CSTE to 1) to make national data on firearm injuries available at the state and sub-state level on a public, online data querying interface with appropriate data confidentiality protections and 2) provide a mechanism for researchers to apply for access to individual-level data.</li> <li>■ Fund CDC or CSTE to create and maintain a data querying website and to provide dataset access and documentation.</li> </ul>

# The NEDS and Statewide ED and Inpatient Databases

---

*Sponsored by:* HCUP of the Agency for Healthcare Research and Quality and by the state organizations that assemble SEDDs and SIDs

---

## Background and History

### *ICD Codes and Uniform Hospital Discharge Data*

Hospital case mix databases, including the NEDS, SEDDs, and SIDs, are based on hospital billing data. After a patient leaves the hospital, medical records coders review the chart and use the ICDs<sup>vi</sup> coding system to summarize the patient's diagnoses and the procedures and services the patient received. For injury visits, diagnosis codes describe the type of injury (e.g., laceration, fracture) and the body part involved. Injury diagnoses are accompanied by an "external cause" code (e-code), which describes both the mechanism by which the injury was inflicted (e.g., sharp instrument, firearm) and the intent underlying the incident (assault, accident, intentionally self-inflicted, legal intervention, terrorism, war operations, and undetermined).

The ICD coding system is updated every several years. One set of codes (ICD) is used by vital statistics registries for deaths, and a more expanded set (ICD-Clinical Modification, or ICD-CM) is used by healthcare organizations to capture the larger universe of nonfatal conditions. For injury deaths, the underlying cause of death on the death certificate must be an e-code describing the injury's external cause. For healthcare data on injury encounters, it is the opposite: the primary diagnosis *cannot* be an e-code; e-codes are auxiliary, non-reimbursable codes.

### *Statewide Hospital Databases are a Census of Patient Visits*

In nearly every state, hospitals forward a de-identified version of their billing claims data to a statewide organization—usually the state health department but sometimes the state hospital association or a public/private consortium. These statewide databases date back to the 1970s when payors and health services planners began comparing hospitals' costs and quality indicators and called for comparability in hospital claims data. The first databases covered inpatient discharges only; over the decades, many states have added databases covering ED visits, observation stays, and ambulatory care surgeries. The databases became increasingly valued not only for services planning and cost comparison but for basic epidemiology and for measuring health outcomes.<sup>vii</sup>

Statewide databases represent a census, not a sample, of visits to non-federal, acute care hospitals. Their data include patient demographics (including zip code of residence in many states), transfer status, hospital identifier, mode of arrival, diagnoses and e-codes, procedures, physician type, charges, expected payor, length of stay, and disposition. Combining the SEDDs

(which cover only treat-and-release cases and transfers) and with SIDs (which cover inpatient cases) nets a complete ED dataset. Data can be used to calculate population-based incident rates in most states using patient's county of residence.

### *HCUP and National Hospital Databases*

Most states with statewide hospital databases both disseminate their own data locally and forward it to the HCUP at the federal Agency for Healthcare Research and Quality. HCUP uses these databases to construct the NIS and the NEDS.<sup>viii</sup> Currently for NEDS, 39 states and the District of Columbia submit data. For NIS, all jurisdictions except Alabama and Idaho submit, and HCUP anticipates that Alabama will soon. NEDS is a stratified, single-stage cluster sample of hospitals constructed by categorizing hospitals according to five strata: geographic region, urban/rural location, teaching status, ownership, and trauma-level designation. There are a total of roughly 4,000 acute care, non-Federal hospitals in the nation. In 2019, 990 hospitals in 37 states submitted data on nearly 36 million ED visits of all types, from which HCUP projected total visits of over 143 million for the nation. HCUP makes individual-level NEDS data available to researchers for a fee (as well as the NIS and state-specific databases). It also disseminates aggregate state and national data via an online data-query interface (HCUP-Net).

### *E-Codes & Injury Surveillance*

In the last quarter of the 20<sup>th</sup> century, after decades of decline in deaths from infectious disease, and with injuries the top killer of people under 40, state and federal public health agencies turned attention to preventing injuries, and the CDC established the NCIPC in 1992. Hospital databases were of limited utility in understanding the etiology of injuries at that time because e-codes were infrequently coded in claims data. Knowing, for example, the incidence of severe head injuries is of limited utility to injury prevention efforts without also knowing the extent to which those injuries were sustained in, say, unintentional motor vehicle crashes, firearm suicides, or blunt force homicides—information that e-codes provide. Therefore, a large focus of state and national injury prevention efforts in the 1990s and early 2000s was to boost hospitals' use of e-codes.<sup>ix</sup> Many states mandated e-coding by law or regulation. Over time, e-coding rates improved greatly. By 2011, an HCUP analysis found that the mean e-code rate for inpatient datasets submitted by 47 states was 92 percent, and for ED databases submitted by 30 states was 94 percent.<sup>x</sup> The two lowest were 71 percent (Indiana and Ohio).

Although e-codes are not reimbursable, they are now considered the professional standard for medical records coding. Software programs used by hospitals to improve coding speed and accuracy, such as 3-M's encoder software, produce an error message if coders enter an injury diagnosis without an accompanying e-code. Without the e-code, firearm injuries could not be identified in databases. E-codes also differentiate between handguns and various types of long guns, but the majority of gunshot wounds (65 percent in 2014)<sup>xi</sup> are coded to "unspecified" firearm type. (The full list of e-codes for firearm injury is in Appendix 2.) Auxiliary e-codes classify the type of place where the injury occurred (e.g., street, school) and the broad type of activity in which the patient was engaged at the time (e.g., walking/running, sports). These are



frequently not coded or coded to unknown and have thus received little attention in injury surveillance research.

## Evaluating Sensitivity and Accuracy of NEDS for Firearm Injury Surveillance

Given that the vast majority of nonfatal firearm injury victims are treated in the ED, and all EDs use a coding system that is capable of identifying them, a critical question is whether the system is in fact doing so accurately.

### *Sensitivity and Positive Predictive Value (PPV)*

A database's "sensitivity" is the proportion of all true cases that it correctly identifies as cases. "Positive predictive value" is the proportion of encounters coded as cases that truly are cases.<sup>xii</sup> True cases for our purposes are projectile injuries resulting from the discharge of a firearm. These are typically not difficult to identify in a chart because the patient's presenting complaint (e.g., "multiple GSWs," "shotgun wound") and clinical notes (e.g., "Through and through bullet wound to left shoulder") often explicitly mention the firearm's involvement. A few studies have evaluated the sensitivity or PPV of ICD-coded hospital data for firearm injury. A 2001 study in Washington state found that sensitivity for firearm injuries was 91.6 percent (61.1-98.7) and PPV was 93.8 (91.2 to 95.6).<sup>xiii</sup> Another early investigation in Oklahoma in 1998 found a sensitivity rate of 81 percent in inpatient data and very high PPV.<sup>xiv</sup> A study in Indianapolis found that while most nonfatal firearm injuries identified in police data were seen in the ED, a significant minority did not receive a firearm-related ICD code<sup>xv</sup> (as noted previously, Indiana has one of the two lowest e-code rates in the country). If we assume that overall sensitivity of NEDS today in capturing firearm injuries is roughly 90 percent, and that 90 percent or more of firearm injuries are treated in the ED, it follows that 81 percent or more of firearm injuries will be captured by national ED data.

### *The Big Problem: Misclassification of Intent*

One problem documented in an early study of hospital firearm e-coding and that continues today, is likely misclassification of intent in ICD-coded hospital data. As described below, at least three national data systems that record injury-related data with a focus on injury intent suggest that hospital case mix data (including NEDS and SEDDs/SIDs) classify far too few firearm injuries as assaults and far too many as accidents. The problem was first reported in 1998 when researchers in Massachusetts linked data from the statewide inpatient database to the state's Weapon Related Injury Surveillance System (WRISS<sup>6</sup>).<sup>xvi</sup> Classification of intent in WRISS was previously evaluated by state health department personnel and found to be largely

---

<sup>6</sup> WRISS is a system under which nurses, physicians, and/or clerks in the ED send reports to the police and state health department after treating a gunshot wound. The report form includes a checkbox for intent and space for a brief incident narrative. For quality control, health department personnel reviewed each report form to confirm the accuracy of the intent classification and, for a sample of cases, reviewed the original hospital chart as well.

accurate. According to WRISS, 5 percent of the firearm admissions were accidents, 81 percent were assaults, and the remainder were self-inflicted or undetermined. According to the linked e-coded inpatient data, however, 57 percent were accidents and 31 percent were assaults.

The distribution of firearm injuries by intent in NEDS also appears to over-report accidents (Table 1, next page). One way to test whether the NEDS distribution by intent is plausible, as well as to evaluate the accuracy of NEDS’ overall firearm injury estimate, is to estimate expected values for nonfatal cases based on deaths. We can do that by applying intent-specific Case Fatality Rates (CFRs) to firearm deaths. In injury surveillance, the CFR is the proportion of all injury incidents (fatal and nonfatal combined) in a given time period that are fatal incidents. If we know, for example, that for every one fatal firearm assault, there are roughly 4 nonfatal injuries from firearm assault, we can estimate nonfatal assault-related firearm injuries based on firearm homicides.

Table 1 compares NEDS estimates for 2016 with estimates based on applying CFRs to 2016 firearm deaths. (The CFRs used are drawn from data sources other than ICD-coded hospital data to avoid circularity. See Appendix 2 for methods and data sources.) Both CFR-based and NEDS-based estimates are remarkably similar for *overall* nonfatal firearm injuries (about 78,000). With 990 EDs contributing data, estimates in NEDS are reasonably stable (coefficient of variation for the firearms estimate is 8 percent). Where the two approaches differ dramatically is in distribution by intent, with half of the NEDS cases classified as accidents, compared with only 10 percent of the CFR-based cases.

**Table 1.** Firearm Deaths and Estimated ED-Treated Nonfatal Firearm Injuries, by Intent –U.S., 2016

	Deaths*	CFR <sup>^</sup>	Estimated Nonfatal Injuries			
			CFR-based		NEDS-based	
			n	col %	n	col %
Homicide/assault	13968	0.18	63,632	81%	33,356	43%
Suicide/self-inflicted	22938	0.85	4,048	5%	2,379	3%
Accident	495	0.06	7,755	10%	38,879	50%
Undetermined	300	0.15	1,700	2%	2,489	3%
Legal Intervention <sup>^^</sup>	957	0.41	1,377	2%	1,136	1%
<b>Total</b>	<b>38658</b>		<b>78,512</b>	<b>100%</b>	<b>78,240</b>	<b>100%</b>

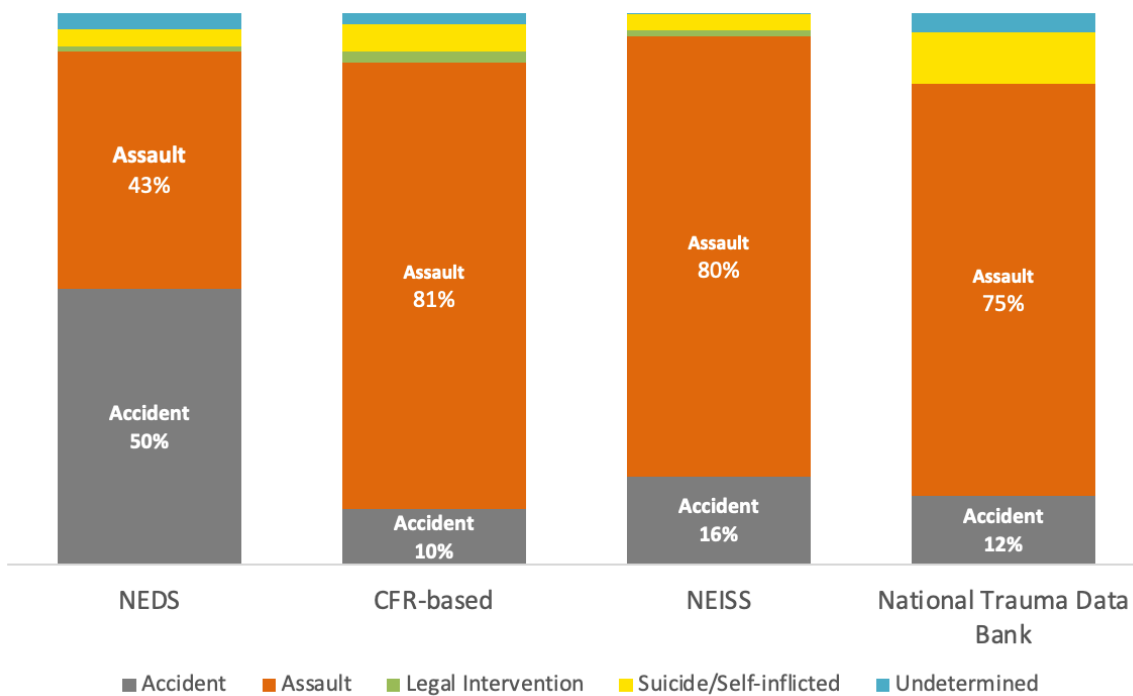
\* **Deaths** are from Vital Statistics with one exception. Because Vital Statistics miss nearly half of legal intervention deaths,<sup>xvii</sup> these deaths are from the *Washington Post*’s “Fatal Force” website, a validated data source,<sup>xviii</sup> and homicides are adjusted accordingly. <sup>^</sup>**CFR**, or the proportion of all injury incidents (fatal and nonfatal combined) that are fatal. See Appendix 2 for CFR data sources.

### Which Intent Distribution Is More Accurate: CFR-based or NEDS-based?

Figure 1 compares the distribution of nonfatal firearm injuries by intent from NEDS and CFR-based estimates with those from the NEISS (see section 2 on NEISS) and from trauma centers contributing to the National Trauma Data Bank.<sup>xix</sup>

The NEDS distribution stands out as the outlier, while the other three (like the earlier WRISS study) indicate that assaults make up three-quarters or more of nonfatal gunshot wounds. Preliminary data from a chart review study currently in progress<sup>xx</sup> also finds an intent distribution that aligns with the other data sources.

**Figure 1.** Distribution of intent classification for firearm injuries treated in the ED by data source. Data sources are for 2016 except from the National Trauma Data Bank, which is for 2010-2016.



### What Accounts for the Intent Classification Problem?

One clear smoking gun is explicit coding guidance. In the U.S., a joint committee of the NCHS and the Centers for Medicare and Medicaid Services oversee ICD coding policies. Protocols in 1994 stated that in the absence of specific documentation in the patient’s chart to the contrary, injuries should be coded as accidents. In 1996, the guideline changed and called for injuries with unclear information as to intent should be coded as “undetermined whether accidentally or purposely inflicted.”<sup>xxi</sup> The guidance changed again in October 2015, with the new ICD-10-CM release, returning to the pre-1996 position: “If the intent (accident, self-harm, assault) of the cause of an injury or other condition is unknown or unspecified, code the intent as accidental intent.”<sup>xxii</sup>

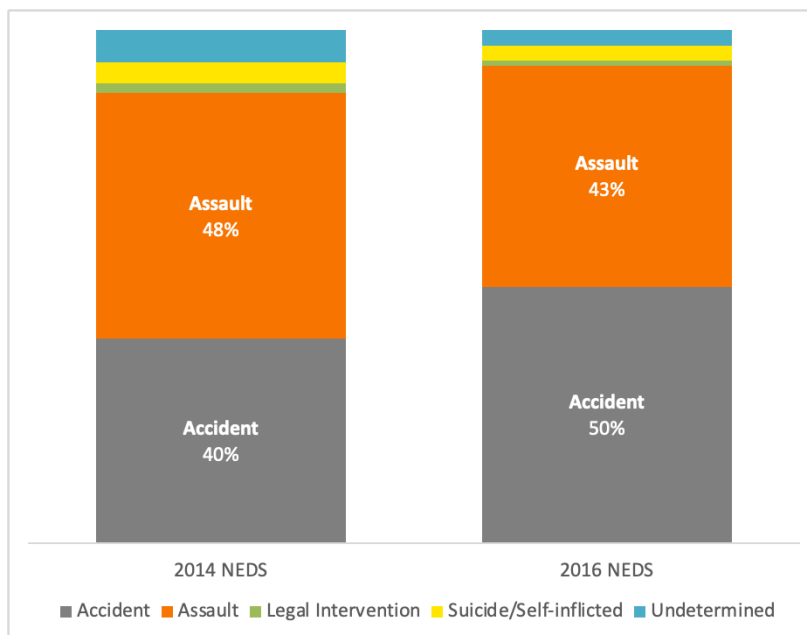
A reason for this guidance is to avoid overreliance on coding intent to unknown. Most injuries seen in the ED are accidental, but often neither the patient’s description of the incident nor the clinicians’ documentation explicitly state that. For example, in a medical chart that states, “Patient in a car crash 2 hours prior to arrival, complains of neck pain,” there is no indication whether the crash was accidental or intentional. However, accident is a good bet in the absence of evidence to the contrary. This is the case for many injury types (falls, motor vehicle crashes, burns, cuts) but not for gunshot wounds. A chart that says, “Patient sustained GSWs to left wrist and right shoulder; shot multiple times while he was walking down the street, did not recognize the shooter” again does not state intent, but in this case the chart is likely signaling an assault, not an accident.

*Coding Guidelines Alone Do Not Explain the Problem.*

ICD e-coding guidelines before 1996—like those after 2015—specified in effect that injuries should default to accident in the absence of evidence to the contrary. However, in the intervening years when that policy was not in effect, distribution by intent in NEDS (and other e-coded hospital data systems) still appeared to over-report firearm accidents, although not as dramatically. Figure 2 shows the distribution of nonfatal gunshot wounds by intent in the year before and the year after the coding guidance changed.

**Figure 2.** Distribution of nonfatal firearm injuries by intent and year, NEDS.

**NB:** In October 2015 new coding guidelines applied specifying that injuries with unknown intent be coded as accidents.



Potential sources of the problem are computer-assisted coding programs and look-up tables that steer users by default to e-codes in the unintentional range as well as coding training and

institutional custom. Whether, and the extent to which, each contribute to the problem is currently being investigated in a National Collaborative on Gun Violence Research (NCGVR) study.<sup>xxiii</sup>

### *Data Access*

HCUP provides free access to all of its data collections, including NEDS and many states' ED databases, via its online data querying website, HCUP-Net. The website is not well-suited to querying by external cause code and currently does not enable users to query by external cause beyond 2014. HCUP personnel state that they are working to resolve this. Individual-level datasets (both national and state-specific) are available to researchers for purchase and include excellent documentation. Costs, however, can be prohibitive at several hundred dollars per database per year, and not all states that have SEDDs and SIDs supply them to HCUP. States also make their own databases available to users, sometimes without cost.

## Conclusion and Recommendations

NEDS and the statewide ED and inpatient databases on which NEDS is based are enormously valuable, providing investigators with data on the incidence of firearm injury at the national level (NEDS) and at the state and local level (individual states' SEDDs/SIDs). However, the databases currently provide a distorted picture of the circumstances under which these injuries occur. Recent peer-reviewed publications using these data have seriously mischaracterized the firearm injury problem as a result.<sup>xxiv,xxv,xxvi</sup> This mischaracterization is an especially grave disservice to the communities that are most beset by firearm injury and could misdirect both funding priorities and outcome evaluations.

Steps can be taken over the next two years to solve the problem going forward:

The study team currently investigating source of the intent classification problem<sup>xxvii</sup> will propose a change in coding guidance to the March 2022 meeting of the joint NCHS/CMS ICD-10 Coordination and Maintenance Committee (which governs e-code coding practices used in U.S. hospital billing data). If a new policy passes, we recommend that federal resources over the next year be used to:

- Publicize the policy change to relevant professional associations and their members who are responsible for coding and managing hospital data (e.g., American Academy of Professional Coders, the American Hospital Information Managers Association)
- Work with software companies that design coding software, coding look-up tables, and electronic health records software to ensure that artifacts of the software are not contributing to inadvertent classification of firearm injuries as accidents.

In addition, we recommend that federal resources be used to:

- Educate researchers, journal editors, and other stakeholders on the intent classification problem in current and past NEDS data and other e-coded hospital databases.
- Conduct studies of ICD-coded hospital databases evaluating: the proportion of hospital-treated firearm injuries that receive a) any firearm injury-related e-code, b) the appropriate e-code with respect to intent (assault, accident, etc.).
- Support the creation and upkeep of a user-friendly data-query interface that gives users access to aggregate national, state, and sub-state injury data based on NEDS and individual states' SEDDs and SIDs.

# NEISS

---

*Sponsored by:* CPSC and CDC's NCIPC

---

## Background and History

The CPSC has operated the NEISS since 1971 to collect data on product-related injuries treated in the ED.<sup>xxviii</sup> NEISS is conducted in a stratified probability sample of U.S. hospitals that have over six beds and provide 24-hour emergency care. Over the years, the number of hospitals in the sample has hovered around 100 and is 90 today. The sample includes separate strata for very large, large, medium, and small hospitals, defined by the number of annual ED visits per hospital, and a separate stratum for children's hospitals. Data are collected by NEISS Coordinators at participating hospitals who review ED records for relevant injury visits and abstract data from patient charts.

Firearms are not regulated by the CPSC and as such were not included in the original NEISS scope of data collection. In 1993, CDC's NCIPC entered into an interagency agreement with CPSC to expand data collection to firearm injuries at all NEISS hospitals, establishing the NEISS-FISS.<sup>xxix</sup> Data collected as part of NEISS-FISS include patient demographics, region of body injured, type of incident location, relationship of victim to shooter, intent (assault-related, legal intervention, self-inflicted, unintentional, and unknown), whether a crime or argument or gang activity was known to be involved, and disposition. De-identified individual-level data are archived on the University of Michigan's ICPSR data repository website and available for download. Documentation on weighting processes from which to make national estimates are provided there.

Starting in July 2000, CDC's Injury Center collaborated with CPSC on a second expansion to NEISS, the NEISS-AIP. From a two-thirds sub-sample of NEISS hospitals, data were collected on *all* injuries and poisonings, regardless of whether associated with a consumer product or a firearm. NEISS Coordinators abstract the data and provide a very brief narrative summarizing the cause of the injury. For quality control, a small number of centralized master coders read these narratives (and FISS narratives from non-AIP sites) to confirm the accuracy of the NEISS Coordinator's coding of intent and mechanism of injury.

Aggregate data from the All-Injury Program are available on CDC's user-friendly WISQARS data querying interface. To date, firearm injury data on WISQARS-Nonfatal Injury Data platform have been from the AIP two-thirds sample, not the full NEISS/NEISS-FISS sample.

## The Problem of Unstable Annual Estimates

Given the small sample size for NEISS-AIP and wide variability in firearm injury caseloads at participating hospitals, annual estimates of firearm injuries have very wide confidence intervals

with Coefficients of Variation (CV) at 30 percent or higher in recent years (compared with 8 percent for NEDS). In the past, the WISQARS-Nonfatal website enabled users to see results on any data queries, including results with high CVs. Users could select the “advanced statistics” option which supplied Standard Error, CV, and confidence limits, to enable them to interpret the point estimates. Unstable estimates were asterisked, but visible to the user. This changed recently in response to a spate of media coverage highlighting the instability of the CDC’s nonfatal firearm estimates, and today the website will not show estimates for any injury type for which the CV is 30 percent or higher to prevent users without statistical training from misinterpreting the data (Figure 3).

**Figure 3.** Screen shot of WISQARS-Nonfatal output on estimated number of total nonfatal firearm injuries in the U.S., 2001-2019. Cells marked with an asterisk indicate Coefficient of Variation is 30 percent or higher.

Year	Estimated Number	Population	Crude Rate	Age-Adjusted Rate**	Number of Cases (Sample)	Standard Error	CV	Lower 95% Confidence Limit	Upper 95% Confidence Limit
2001	63,012	284,968,955	22.11	21.68	1,478	13,911	22.1%	35,747	90,278
2002	58,841	287,625,193	20.45	20.16	1,536	12,444	21.1%	34,451	83,231
2003	65,834	290,107,933	22.69	22.34	1,514	14,139	21.5%	38,122	93,547
2004	64,389	292,805,298	21.99	21.79	1,694	15,517	24.1%	33,976	94,802
2005	69,825	295,516,599	23.62	23.43	1,830	18,028	25.8%	34,490	105,160
2006	71,417	298,379,912	23.93	23.61	1,843	18,938	26.5%	34,299	108,535
2007	69,863	301,231,207	23.19	23.04	1,791	19,639	28.1%	31,370	108,355
2008	78,622	304,093,966	25.85	25.77	1,855	20,247	25.8%	38,938	118,306
2009	66,769	306,771,529	21.76	21.68	1,706	15,558	23.3%	36,275	97,262
2010	73,505	308,758,105	23.80	23.97	1,733	19,061	25.9%	36,146	110,864
2011	73,883	311,556,874	23.71	23.64	1,748	18,967	25.7%	36,708	111,058
2012	81,396	313,830,990	25.93	25.87	2,036	19,872	24.4%	42,446	120,345
2013	84,258	315,993,715	26.66	26.80	2,305	24,128	28.6%	36,968	131,548
2014	81,034	318,301,008	25.45	25.53	2,312	23,563	29.1%	34,850	127,219
2015	84,997	320,635,163	26.50	26.62	2,528	24,674	29.0%	36,636	133,357
2016	---	322,941,311	---	---	---	---	---	---	---
2017	---	324,985,539	---	---	---	---	---	---	---
2018	---	326,687,501	---	---	---	---	---	---	---
2019	---	328,239,523	---	---	---	---	---	---	---

### Controversy over CDC Firearm Injury Estimates

In 2017, a paper appearing in the *American Journal of Epidemiology*<sup>xxx</sup> reported a “hidden epidemic” of nonfatal, assault-related firearm injuries which increased during a time (2001-2013) when homicides had declined. An earlier commentary came to similar conclusions.<sup>xxxi</sup> Both used data from WISQARS-Nonfatal but ignored the unusually broad confidence intervals surrounding the estimates and failed to whether characteristics of the underlying sample from which the estimates were drawn might be drawing the apparent “epidemic.”



Cook and colleagues<sup>xxxii</sup> investigated the individual-level data in NEISS and rebutted the claims of an increase in nonfatal shootings.<sup>xxxiii</sup> Two features of the data led to the seeming increase. First, there was a steady decline over the study period in NEISS coders' use of the "undetermined" category. Second, although hospitals typically stay in the NEISS sample for multiple years, some do leave. When they are replaced, the new hospital is recruited from the same sampling category as the exiting hospital with respect to overall ED patient volume and region of the country but not with respect to gunshot wound caseload. Two replacement hospitals (out of 15 total replacements) accounted for most of the apparent increase in firearm assaults across the entire roughly 100-hospital sample during the period. When adjusting for both the downward trend in use of "undetermined" and the hospital replacement issue, they found no increase in nonfatal firearm assaults, but instead a small decrease, mirroring homicide trends. They made no claims that their adjustments better approximated the true absolute rates or trends in firearm injury. Rather, their analysis demonstrated that the apparent increase in assaults in the NEISS data were artifacts of reporting. The adjusted trends they reported were supported by studies from the same time period using hospital inpatient data,<sup>xxxiv</sup> trauma registry data,<sup>xxxv</sup> and ED data,<sup>xxxvi</sup> all of which showed no increase in nonfatal shootings.

The problems with the NEISS data drew a flurry of media attention critical of the CDC,<sup>xxxvii</sup> leading Senator Bob Menendez (D-NJ) and 11 senators to send a letter to the Department of Health and Human Services raising concerns about the WISQARS-Nonfatal data and about federal capacity to track nonfatal firearm injuries.

### *The Geographic Clustering Problem*

The problem at the core of using NEISS for firearm injury surveillance is that its sampling frame is designed to produce national estimates of ED injuries overall, not for any one specific injury type. For injuries as common as falls, which are seen in EDs at a rate of about 2,500 per 100,000 people in the U.S., and which are not expected to cluster geographically, the NEISS sample can be expected to deliver stable and relatively accurate point estimates each year. But at a rate of 27 per 100,000 people, firearm injuries are rare by comparison. And unlike most injury types, firearm injuries cluster dramatically not only in certain cities but in specific neighborhoods within those cities. Indeed, the clustering of nonfatal firearm injuries is so great that just 9 percent of the 953 hospitals in HCUP's NEDS sample in 2016 accounted for *two-thirds* of all firearm injury cases in unweighted data; 58 percent of the hospitals accounted for only 6 percent of cases.<sup>xxxviii</sup> Clustering like this poses a major challenge to sample design.

### *The CDC/CPSC Plan to Improve NEISS-AIP*

In 2019 the CPSC and CDC collaborated to improve the NEISS. CPSC awarded a contract to Westat, Inc. (CPSC contract 61320619F0134) to conduct an independent statistical analysis of the NEISS and NEISS-AIP samples and to recommend revisions to the sample frames to better meet injury data needs of both programs.<sup>xxxix</sup> Westat weighed the advantages and drawbacks to retaining the existing sample, expanding it, or drawing a new sample. The company

recommended that 1) CPSC redesign the NEISS sample, retaining current hospitals when possible for stability and cost savings, and 2) immediately expand the NEISS-AIP data collection from its current two-thirds sample to the full NEISS hospital sample.

These recommendations address several concerns. First, the NEISS data frame had not been re-examined since it was last updated in the 1990s and many hospitals have closed, merged, or migrated across sampling strata in the interim. Second, the Office of Management and Budget was concerned about data accuracy and accessibility of the NEISS-AIP estimates. CDC’s decision not to display injury estimates with CVs of 30 percent or larger on the website led to suppressed data cells for firearm injuries and many other injury types, causing concern by users and reporters. Expanding AIP to the full sample is estimated to result in CVs of 16.7 percent<sup>xl</sup> for firearm injuries and to improve reliability for several other injury types as well. At 16.7 percent, the confidence intervals will still be wide for firearm injuries but will be acceptable for release on WISQARS-Nonfatal. The new sample plan retains many of the existing NEISS sites but calls for recruiting 18 new sites.

**Figure. 4.** CPSC plan for new NEISS sample.

NEW NEISS SAMPLE						
Stratum	NEISS redesign	2021 NEISS: reporting (retained)	2021 NEISS: reporting (dropped)	2021 NEISS: replacements (retained)	2021 NEISS: replacements (dropped)	New
Small .....	43	30	0	8	3	5
Medium .....	26	14	1	1	0	11
Large .....	12	11	8	0	1	1
Very Large .....	11	9	0	2	0	0
Children's .....	8	7	1	0	0	1
<b>Total .....</b>	<b>100</b>	<b>71</b>	<b>10</b>	<b>11</b>	<b>4</b>	<b>18</b>

Source: Federal Register July 20, 2021;86(136):38315-8.

CPSC has filed the new plan on the Federal Register for public comment through September 21, 2021. The timeline for implementation is as follows:

- 2021 NEISS-AIP expands to the full current NEISS sample. This will likely lead to publishable firearm injury data for that data year and going forward.
- 2021 & 2022 CPSC and CDC recruit and train the new NEISS sample, adding 18 new sites for a total of 100. Data collection will continue in the old sample through 2023.
- 2023 Bridge year in which data are collected from both the old and new samples. This will accommodate time series studies that cross over the two NEISS samples and enable system designers to evaluate and adjust for impacts of the new sample on time trends.
- 2024 Official launch of the new sample only.

### *Ensuring Stability in Firearm Injury Caseload in Replacement Hospitals*

Under the new plan, when a hospital of a given size stratum exits (due to closure, new management, etc.), care will be taken going forward to ensure that the replacement hospital is not only from the same size stratum and region of the country but has a firearm injury caseload similar to the one it replaces. This will improve the stability of the estimates and help ensure that changes in incidence rates reflect actual changes in cases, not changes in hospital sample composition.

The direct costs of paying NEISS contractors (abstractors) and hospitals for NEISS data totaled \$4,448,000 in 2019.<sup>xii</sup> Expanding the AIP to the full NEISS sample is estimated at another \$650,000 in direct contractor and hospital costs. The CDC's NCIPC received funding in fiscal year 2021 to support the expansion of NEISS-AIP to the full NEISS sample but continued funding at the higher level to support the expansion has not been assured.<sup>xiii</sup>

### *Data Access*

NEISS data are highly accessible. The WISQARS-Nonfatal data querying website provides a user-friendly interface from which lay users can access aggregate NEISS-AIP data. Once the AIP data collection is expanded to the full NEISS sample, CDC expects that annual estimates of firearm injuries will no longer be subject to cell size suppressions due to wide confidence intervals. Individual-level data are available free to researchers from the ICPSR website, and datasets are well documented.

### *Conclusion and Recommendations*

With only 100 hospitals contributing data, national estimates from NEISS of the overall count of firearm injuries are far more vulnerable than estimates from NEDS (with its nearly 1,000 hospitals) to the three pronged-problem of extreme geographic clustering of shootings, low actual base rate, and sampling error. However, the intent classification problem in NEDS and other e-coded hospital databases is a long-standing one and may prove difficult to fix. Intent classification in NEISS, on the other hand, is more reliable and has centralized quality controls. While the NEISS-estimated *numbers* of firearm injuries may continue to have wide confidence intervals, the *distribution* of firearm injuries by intent appears to be in the right ballpark, given its general concordance with the distribution found in National Trauma Bank data, CFR-imputed estimates, the WRISS study, and preliminary data from the current NCGVR-funded study. In addition, both NEISS and NEDS support a myriad of public health information needs on topics as diverse as pedestrian injury, intentional self-harm, and drug overdose. Improving both systems—NEISS as well as the hospital billing systems that feed into NEDS, SEDDs, and SIDs—is warranted.

We recommend the following steps to improve estimates of firearm injuries in NEISS:

- Support CSPC and CDC's new sample plan and expansion of the NEISS-AIP to all 100 NEISS hospitals and that hospitals that leave the system over time are replaced by hospitals with similar firearm injury caseloads as the existing hospital.
- Allocate funding to CDC to maintain the full sample after FY22. (Funds are already allocated to cover current year costs of recruiting and training new sites and collecting all-injury data at full sample.)
- Publicize the availability of the individual-level NEISS-FISS data to researchers.
- For users of CDC's WISQARS-Nonfatal data querying website, provide a simple explanation geared for the non-statistician on how to interpret wide confidence intervals in the data.

# Firearm Surveillance Through Emergency Rooms (FASTER)

## NSSP

---

*Sponsored by:* CDC NCIPC (FASTER); CDC Division of Health Informatics and Surveillance (NSSP)

---

### Background and History

In response to the terrorist attacks of 2001, Congress directed the CDC to establish what is now called the NSSP to detect bioterrorism-related illnesses and other health issues.<sup>xliii</sup> The system has since come to be used for purposes as diverse as detecting health impacts of oil spills, tracking the incidence of opioid overdose, and detecting clusters of carbon monoxide poisonings resulting from power outages. The purpose of syndromic surveillance is to detect in near real-time emerging health problems for mobilization of a rapid response that can reduce attendant harms.<sup>xliv</sup> As such, these systems require active partnership and coordination across local, state, and national health agencies.

Currently participating hospitals account for 73 percent of ED visits nationally (Figure 3). Hospitals upload de-identified data from their electronic health information systems to state and local health departments or to data aggregators such as Health Information Exchanges.<sup>xlv</sup> These local health agencies in turn contribute the data to the federal BioSense platform, where it is received typically within 24 or 48 hours of the ED visit; data are updated with new uploads. Information fields from EDs include chief complaint, free text triage notes (when available), diagnosis codes (when available, either in ICD or SNOMED-CT<sup>xlvi</sup> codes<sup>7</sup>), patient characteristics, and location. EDs are not the only data source that NSSP uses, but we focus on them here given their relevance to firearm injury surveillance.

### *A Collaborative Model*

In the early years of NSSP's development, CDC attempted to have facilities report data directly to the CDC but had trouble recruiting hospitals other than federal hospitals. The software was difficult to use and added little value at the state or local level. Recruitment improved greatly when CDC redesigned the system on a more cooperative model in 2010, relying in large part on state and local health departments to recruit hospitals, involving local stakeholders in designing a more user-friendly system, and providing useful analytic tools.

---

<sup>7</sup> SNOMED-CT stands for "Systemized Nomenclature of Medicine—Clinical Terms." It is an international system available in multiple languages for standardizing and conceptually linking clinical terminology to assist in the exchange, processing, and communication of clinical information in electronic health records. It is governed by the International Health Terminology Standards Development Organisation (IHTSDO). See [www.ihtsdo.org/](http://www.ihtsdo.org/)

NSSP today relies on a “Community of Practice” model (coordinated by the CSTE via a cooperative agreement with the CDC) under which local, state, and national stakeholders work together to define an ever-expanding list of syndrome case definitions to track.<sup>xlvii</sup> A syndrome may be, for example, health events as diverse as anthrax poisoning, a suicide attempt, or an adverse reaction to a specific vaccine. The CoP develops and shares algorithms (using coded data and in some cases artificial intelligence techniques) to identify health encounters that meet those case definitions. They improve the algorithms by testing their sensitivity and positive predictive power. The analytic software program that pulls cases that meet the syndrome classifications from the BioSense platform is ESSENCE (Electronic Surveillance System for the Early Notification of Community-based Epidemics).<sup>xlviii</sup> ESSENCE provides users with tools to query and visualize the data. It also automatically detects and alerts users to potential adverse health events, employing a temporal alerting algorithm that uses a 30-day moving baseline to detect statistical anomalies.

## The FASTER Program

In September 2020, the CDC’s NCIPC awarded a total of \$2.23 million to ten state health departments (DC, FL, GA, NM, NC, OR, UT, VA, WA, WV) to test using NSSP as a data source for nonfatal firearm injury surveillance. The goal of Firearm Surveillance Through Emergency Rooms (FASTER) is to enable state and local health departments to rapidly track ED-treated firearm injuries in near real-time, to classify them by intent, and to use the surveillance data to help communities most affected by gun injuries respond. In the program’s first year, participants have developed syndrome case definitions to identify first visits to the ED for injuries caused by firearm projectile and are currently validating how these definitions perform.

While it is too early in the new program to say for sure, FASTER personnel are optimistic that about their ability to capture overall counts of ED-treated firearm injuries using NSSP. They are less optimistic about the system’s ability to classify firearm injuries by intent. This is both because there is limited free text data uploaded to BioSense that describes the circumstances under which the shooting occurred from which to classify intent, and because any ICD “external cause” codes that accompany a case often misclassify intentional gunshot wounds as accidents (as discussed in the NEDS section). The fact that FASTER may be successful only in identifying firearm injuries overall and not in classifying intent is not necessarily a fatal flaw. Roughly three-quarters or more of ED-treated, nonfatal gunshot wounds are assaults. Spikes and dips in firearm injury patients are driven by assaults. Detecting volatility in near real-time at the local level is where NSSP excels. Data from more traditional systems, like NEDS and NEISS, are 1-3 years behind.

A recent project similar to FASTER, which utilized NSSP for opioid overdose surveillance, concluded that while traditional ICD-coded discharge data offered many strengths, rapid syndromic data was able to detect changes quickly to alert public health and safety personnel to remediable events, like local overdose clusters tied to shipments of illicit drugs with higher toxicity than normal.<sup>xlix</sup>

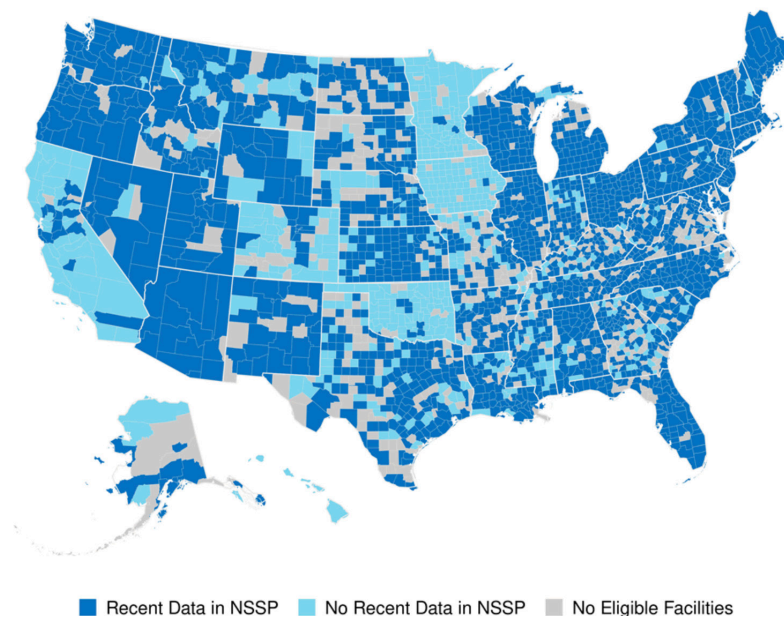
### *Federal Commitment*

The NCIPC is committed to the FASTER program and to continuing to support state and local health departments in monitoring firearm injury in the most timely and accurate way possible. The Injury Center has received additional federal funding for firearm injury research, which can help support FASTER. The CDC's Division of Health Informatics and Surveillance has a strong commitment to expanding and improving NSSP, listing it as the second of its current (2019-2021) top four priorities.<sup>i</sup>

### *National Coverage*

While the number of hospitals participating in NSSP is impressive, still one in four ED visits nationally is to a hospital that is not part of the system. Hospital participation is not randomly distributed. In six states (including California) very few hospitals participate (Figure 3). The Centers for Medicare and Medicaid Services in summer of 2021 passed a rule change (Federal Register, May 2021: 86 FR 25070) that requires acute care hospitals to participate in syndromic surveillance.<sup>ii</sup> The rule change (721 pages in length) covers an enormous range of topics, far beyond the syndromic surveillance requirement. The syndromic reporting expectation aligns with one of CMS's overarching goals, the Promoting Interoperability Program,<sup>iii</sup> which spurs hospitals and selected providers to adopt electronic health records (EHRs). CMS, as well as the U.S. Department of Health and Human Services (HHS) more broadly, views adoptions of EHRs that can communicate across various platforms as key to improving health care in many areas, such as electronic prescribing for better accuracy and tracking, health information exchange across providers, patient/provider communication, patient access to health records, clinical decision support, payment reform, public health reporting, tracking for adverse events, and outcomes research.<sup>iiii</sup> It appears likely that the impressive growth in the number of EDs participating in NSSP will likely continue to near 100 percent given the CMS rule change.

**Figure 3.** NSSP Participation, April 2020 – April 2021. Map identifies counties with at least one eligible facility that contributed at least one eligible record. (Map supplied by NSSP)



### Data Access

The FASTER program’s benchmark is for states to share aggregate state and local data on a quarterly basis with local prevention partners. No plans have been announced to date whether individual-level data would be made available to researchers, and data dissemination is up to each individual state. While CDC’s Division of Health Informatics and Surveillance developed the architecture for NSSP, CDC has access to its data only at the aggregate national and regional level. Access to more granular state, local, and individual-level data is controlled by the facilities and health departments that submit the data. Exceptions are when CDC provides funding for state and local health departments to share their data or when CDC has received explicit permission from local partners for access to the more granular data for specified syndromes. Tracking trends at the national and gross regional level is useful. But shootings, like many other public health problems (overdoses, infectious disease outbreaks) often cluster geographically. Rates of violence across different cities can move in opposite directions, driven by local conditions (e.g., a high visibility police shooting, shifts in drug markets, gang truces, etc.). Lack of national oversight on local trends hobbles the capacity to identify local drivers across states or to act quickly to allocate resources to areas that need it. Without timely local data, research and resource allocation will continue to be driven not by the problems of today but problems as they existed two years ago.

That said, a lesson from NSSP’s early experience seemed to be that a simple top-down approach, with CDC calling the shots, was unsuccessful. Hospital recruitment and trust fared better under a more collaborative model under which power and control was shared across



institutional, state, and federal levels. A solution that retains a shared power model while enabling national investigators access to more granular data appears warranted. There are already models for this where data that is controlled at the local and state-level (mortality data, hospital case mix data, NVDRS data) are conveyed to a national center (NCHS, HCUP, and NCIPC, respectively) and made available via online data querying interfaces at the national, state, and sometimes local levels (CDC WONDER, HCUP-Net, CDC WISQARS, respectively). In addition, each has a mechanism by which researchers can apply for access to the individual-level data.

## Conclusion and Recommendations

The FASTER pilot will likely establish that the NSSP is capable of efficiently delivering near real-time surveillance of firearm injuries, even if it is unable to successfully classify these events by intent type. The federal government is committed to NSSP, and the system's growth from currently capturing over 70 percent of ED visits to capturing nearly all seems likely in the next few years, given the recent CMS rule change<sup>liv</sup> requiring EDs to participate in syndromic surveillance. While NSSP offers an extraordinary timeliness advantage over NEISS and hospital case mix data, if it can supply data only at the aggregate national and HHS regional levels and depends only on a state-by-state approach to making state and local aggregates available, the system will be far less useful and informative than it could otherwise be.

If the FASTER program proves successful over the next year, we recommend the following steps to support firearm injury surveillance in NSSP should the FASTER pilot prove successful:

- Support near 100 percent participation of EDs in the NSSP, via the current CMS rule change and/or other—especially incentive-based—means.
- Negotiate an agreement among CDC, the CSTE, and the NSSP Community of Practice Governing Board authorizing CDC or CSTE to 1) to make national data on firearm injuries available at the state and sub-state level on a public, online data querying interface with appropriate data confidentiality protections and 2) provide a mechanism for researchers to apply for individual-level data, again with appropriate confidentiality protections.
- Provide funding to the CDC or CSTE to create and maintain the data querying website and to provide dataset access and documentation to researchers.
- Provide funding to the CDC to expand FASTER to all states.

# Appendix

## Table of Contents

Sources and methods for calculating intent-specific Case Fatality Rates for firearm injuries

List of firearm injury-related ICD-codes

References

# Appendix 1

## Sources and Methods for Calculating Intent-specific CFRs for Firearm Injuries

The body of this report presented the table below. Described here are 1) the rationale for estimating nonfatal injuries based on deaths and 2) the data sources used in the table for both the deaths and intent-specific CFRs.

**Table 1.** Firearm Deaths and Estimated ED-Treated Nonfatal Firearm Injuries, by Intent –U.S., 2016

	Deaths*	CFR <sup>^</sup>	Estimated Nonfatal Injuries			
			CFR-based		NEDS-based	
	n		n	col %	n	col %
Homicide/assault	13968	0.18	63,632	81%	33,356	43%
Suicide/self-inflicted	22938	0.85	4,048	5%	2,379	3%
Accident	495	0.06	7,755	10%	38,879	50%
Undetermined	300	0.15	1,700	2%	2,489	3%
Legal Intervention <sup>^^</sup>	957	0.41	1,377	2%	1,136	1%
<b>Total</b>	<b>38658</b>		<b>78,512</b>	<b>100%</b>	<b>78,240</b>	<b>100%</b>

\* **Deaths** are from Vital Statistics with one exception. Because Vital Statistics miss nearly half of legal intervention deaths,<sup>iv</sup> these deaths are from the *Washington Post*'s "Fatal Force" website, a validated data source,<sup>lv</sup> and homicides are adjusted accordingly. <sup>^</sup>**CFR**, or the proportion of all injury incidents (fatal and nonfatal combined) that are fatal. See Appendix 2 for CFR data sources.

### Estimating Nonfatal Injuries from Deaths

Roughly one out of every five shooting assault victim dies. That's a CFR<sup>8</sup> of about 20 percent. In a country with, say, 14,000 firearm homicides in a year, one can estimate that hospitals see about 70,000 nonfatal shooting assault victims ((14,000/.20) – 14,000). If in fact the actual CFR for gun assaults is as high as 25 percent or as low as 15 percent, the range of estimated nonfatal injuries will be 56,000 to 93,333. One's ability to project nonfatal injuries depends on having a good fix on the number of deaths and a good fix on the CFR, not just for assaults and homicides, but for each intent type.

<sup>8</sup> The term CFR is defined differently in injury surveillance than in traditional disease surveillance. In disease surveillance it refers to the proportion of people with a given condition who die from that condition over a specified time period such as one month or five years or ten years. In injury surveillance, it refers to the proportion of acute injury incidents in a given time period (usually a year) that are fatal incidents.

### Reliable Death Counts

Table 1 uses deaths from the National Vital Statistics System (NVSS) with one exception. Because close to half of legal intervention deaths are missed by Vital Statistics and are classified instead as homicides,<sup>lvii</sup> legal intervention deaths for 2016 come from the *Washington Post's* Fatal Force website, a database that has been validated against the NVDRS and other databases.<sup>lviii</sup> Homicides are adjusted accordingly. Firearm accidents are also frequently misclassified in Vital Statistics, according to a study using NVDRS data.<sup>lix</sup> Errors varied a great deal across states with some over- and others under-reporting accidents. However, summed across states, the numbers amounted basically to a wash, so Table 1 uses the total number from Vital Statistics.

### Reliable CFRs

To avoid circularity, the CFRs in Table 1 do not come from ICD-coded hospital data. Rather, they come from Massachusetts' WRISS and from police data. Massachusetts law requires hospitals to notify police when they treat a gunshot wound of any type, and these reports are the basis for WRISS. In the mid-1990s, state health department personnel conducted regular record reviews at all acute care hospitals in the state to determine the proportion of gunshot wounds that were successfully reported to the system and to test for reporting biases. WRISS published CFRs by intent type and demographic group using death certificates for deaths and hospital reports for nonfatal firearm injuries, adjusting for hospital under-reporting.<sup>lx</sup> Table 1 uses WRISS CFRs with two exceptions. The CFR for firearm accidents in the published WRISS table was 4.5%; however, WRISS noted that four deaths that were classified as homicides on the death certificate were accidents according to the ED and media reports (e.g., two boys playing with a gun). Reclassifying these as accidents bumped the CFR to 6.1 percent. The WRISS CFR for assaults was 17.6 percent, which is well in line with published reports using police data, including both current (17.9 percent)<sup>lxi</sup> and older (13-17 percent) data.<sup>lxii</sup>

WRISS did not use a separate category for legal intervention incidents, including them instead in the assault/interpersonal violence category. Table 1 therefore uses a CFR drawn from police data. In 2017, the news source VICE collected data on fatal and nonfatal shootings by police officers from the 47 largest urban police departments.<sup>lxiii</sup> We downloaded data from the VICE website and excluded departments where unknowns for fatal/nonfatal status were 25 percent or higher and departments that did not distinguish nonfatal injuries from subjects who were shot at but not hit. This left a total of 2,883 police shootings from which deaths, nonfatal injuries and misses could be identified. Among these, there was a 23 percent miss rate. Among hits, 41 percent were fatal, the CFR used in the table above.

While Table 1 projects very specific numbers, readers should understand that if the true value for the CFRs is somewhat higher or lower than those used here (which could well be the case), estimates will vary.

## Appendix 2

### Relevant ICD External Cause-of-Injury Codes for Firearm Injuries (Excluding War Operations)

#### *ICD-9-CM Codes (for Cases Discharged Before October 1, 2015)*

##### Assault, Terrorism

- E965.0 - Assault by handgun
- E965.1 - Assault by shotgun
- E965.2 - Assault by hunting rifle
- E965.3 - Assault by military firearms
- E965.4 - Assault by other and unspecified firearm
- E979.4 - Terrorism involving firearms

##### Self-inflicted

- E955.0 - Suicide and self-inflicted injury by handgun
- E955.1 - Suicide and self-inflicted injury by shotgun
- E955.2 - Suicide and self-inflicted injury by hunting rifle
- E955.3 - Suicide and self-inflicted injury by military firearms
- E955.4 - Suicide and self-inflicted injury by other and unspecified firearm
- E955.9 - Suicide and self-inflicted injury by firearms and explosives, unspecified

##### Unintentional

- E922.0 - Accident caused by handgun
- E922.1 - Accident caused by shotgun
- E922.2 - Accident caused by hunting rifle
- E922.3 - Accident caused by military firearms
- E922.8 - Accident caused by other specified firearm missile
- E922.9 - Accident caused by unspecified firearm missile

##### Legal intervention

- E970 - Injury due to legal intervention by firearms

## Undetermined

- E985.0 - Injury by handgun, undetermined whether accidentally or purposely inflicted
- E985.1 - Injury by shotgun, undetermined whether accidentally or purposely inflicted
- E985.2 - Injury by hunting rifle, undetermined whether accidentally or purposely inflicted
- E985.3 - Injury by military firearms, undetermined whether accidentally or purposely inflicted
- E985.4 - Injury by other and unspecified firearm, undetermined whether accidentally or purposely inflicted

## *B. ICD-10-CM Codes (for cases discharged on or after October 1, 2015)*

For the following, the 7<sup>th</sup> character refers to episode of care (A, D, S – initial, subsequent, and sequela). The use of “X” in the last or next-to-last character is used here to indicate any value, including missing. For surveillance of the incidence of firearm injury *events*, exclude cases where 7<sup>th</sup> character is D or S, and include those where it is A or missing. For surveillance of, say, overall burden of firearm injuries, use all.

### Assaults, Terrorism

- X93XXX Assault by handgun discharge
- X940XX Assault by shotgun
- X941XX Assault by hunting rifle
- X942XX Assault by machine gun
- X948XX Assault by other larger firearm discharge
- X949XX Assault by unspecified larger firearm discharge
- X958XX Assault by other firearm discharge
- X959XX Assault by unspecified firearm discharge
- Y384X1 Terrorism involving firearms, public safety official injured
- Y384X2 Terrorism involving firearms, civilian injured
- Y384X3 Terrorism involving firearms, terrorist injured

### Self-inflicted

- X72XXX Intentional self-harm by handgun discharge
- X730XX Intentional self-harm by shotgun discharge
- X731XX Intentional self-harm by hunting rifle discharge
- X732XX Intentional self-harm by machine gun discharge
- X738XX Intentional self-harm by other larger firearm discharge
- X739XX Intentional self-harm by unspecified larger firearm discharge
- X748XX Intentional self-harm by other firearm discharge
- X749XX Intentional self-harm by unspecified firearm discharge

## Unintentional

- W320XX Accidental handgun discharge
- W321XX Accidental handgun malfunction
- W3300X Accidental discharge of unspecified larger firearm
- W3301X Accidental discharge of shotgun
- W3302X Accidental discharge of hunting rifle
- W3303X Accidental discharge of machine gun
- W3309X Accidental discharge of other larger firearm
- W3310X Accidental malfunction of unspecified larger firearm
- W3311X Accidental malfunction of shotgun
- W3312X Accidental malfunction of hunting rifle
- W3313X Accidental malfunction of machine gun
- W3319X Accidental malfunction of other larger firearm
- W3400X Accidental discharge from unspecified firearms or gun
- W3409X Accidental discharge from other specified firearms
- W3410X Accidental malfunction from unspecified firearms or gun
- W3419X Accidental malfunction from other specified firearms

## Legal Intervention

- Y35001 Legal intervention by unspecified firearm discharge, law enfor. official injured
- Y35002 Legal intervention by unspecified firearm discharge, bystander injured
- Y35003 Legal intervention by unspecified firearm discharge, suspect injured
- Y35009 Legal intervention by unspecified firearm discharge, unspec'd person injured
- Y35011 Legal intervention by machine gun, law enforcement official injured
- Y35012 Legal intervention by machine gun, bystander injured
- Y35013 Legal intervention by machine gun, suspect injured
- Y35019 Legal intervention by machine gun, unspecified person injured
- Y35021 Legal intervention by handgun, law enforcement official injured
- Y35022 Legal intervention by handgun, bystander injured
- Y35023 Legal intervention by handgun, suspect injured
- Y35029 Legal intervention by handgun, unspecified person injured
- Y35031 Legal intervention by rifle pellet, law enforcement. official injured
- Y35032 Legal intervention by rifle pellet, bystander injured
- Y35033 Legal intervention by rifle pellet, suspect injured
- Y35039 Legal intervention by rifle pellet, unspecified person injured
- Y35091 Legal intervention by other firearm discharge, law enforcement official injured

Y35092 Legal intervention by other firearm discharge, bystander injured

Y35093 Legal intervention by other firearm discharge, suspect injured

Undetermined

Y22XXX Handgun discharge, undetermined intent

Y230XX Shotgun discharge, undetermined intent

Y231XX Hunting rifle discharge, undetermined intent

Y232XX Military firearm discharge, undetermined intent

Y233XX Machine gun discharge, undetermined intent

Y238XX Other larger firearm discharge, undetermined intent

Y239XX Unspecified larger firearm discharge, undetermined intent

Y248XX Other firearm discharge, undetermined intent

Y249XX Unspecified firearm discharge, undetermined intent



## References

---

- <sup>i</sup> May JP, Hemenway D, and Hall A. Do criminals go to the hospital when they are shot? *Inj Prev*. 2002;8(3):236-8. doi: 10.1136/ip.8.3.236.
- <sup>ii</sup> Unpublished data from the National Crime Victimization Survey, Concatenated Incident File, 1992-2015 (ICPSR\_36456/DS0003/36456-0003-Data.dta) available from <https://www.icpsr.umich.edu/web/NACJD/studies/36456/datadocumentation> Data are based on 93 respondents over the 24-year period who reported that they were victims of a gun attack and sustained a gunshot wound.
- <sup>iii</sup> See Appendix 2 for data and methods.
- <sup>iv</sup> National Collaborative on Gun Violence Research (NCGVR). “Our Grants: Evaluating and Improving the Accuracy of ICD-Coded Hospital Data Systems in Estimating the Incidence of Nonfatal Firearm Injuries by Intent Type” [webpage]. Available at: <https://www.ncgvr.org/grants/2020/hospital-data-systems.html>
- <sup>v</sup> Preliminary data from record reviews of ED visits to Mass General Brigham hospitals from 2001-2019 where at least one external-cause-of-injury code is in the firearm range (n=1,930).
- <sup>vi</sup> Centers for Disease Control and Prevention, National Center for Health Statistics. *International Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM)*. Accessed at: <http://www.cdc.gov/nchs/icd/icd10cm.htm> on Apr 12, 2020.
- <sup>vii</sup> Schoenman JA, Sutton JP, Kintala S, Love D, Maw R. The Value of Hospital Discharge Databases. NORC and the National Association of Health Data Organizations, 2005. Accessed 8/3/2021 from: [https://www.hcup-us.ahrq.gov/reports/final\\_report.pdf](https://www.hcup-us.ahrq.gov/reports/final_report.pdf)
- <sup>viii</sup> Personal communication, HCUP Technical Assistance Center, July 14, 2021.
- <sup>ix</sup> Annett JL, Fingerhut LA, Gallagher SS, et al. Strategies to improve external cause-of-injury coding in state-based hospital discharge and emergency department data systems: recommendations of the CDC Workgroup for Improvement of External Cause-of-Injury Coding. Centers for Disease Control and Prevention (CDC). *MMWR Recomm Rep*. 2008 Mar 28;57(RR-1):1-15.
- <sup>x</sup> Barrett M, Steiner C. Healthcare Cost and Utilization Project (HCUP) External Cause of Injury Code (E Code) Evaluation Report (Updated with 2011 HCUP Data). 2014. HCUP Methods Series Report # 2014-01 ONLINE. March 14, 2014. U.S. Agency for Healthcare Research and Quality. Accessed 7/30/2021: <http://www.hcup-us.ahrq.gov/reports/methods/methods.jsp>.

- 
- <sup>xi</sup> HCUPnet, Healthcare Cost and Utilization Project. Agency for Healthcare Research and Quality, Rockville, MD. <https://hcupnet.ahrq.gov/>. Weighted national estimates from HCUP National Emergency Department Sample (NEDS), 2014.
- <sup>xii</sup> Centers for Disease Control and Prevention. Guidelines for evaluating surveillance systems. *MMWR*. 1988;37(S-5):1-18.
- <sup>xiii</sup> LeMier M, Cummings P, and West T. Accuracy of external cause of injury codes reported in Washington State hospital discharge records. *Inj Prev*. 2001 Dec; 7(4): 334–338. doi: 10.1136/ip.7.4.334
- <sup>xiv</sup> Archer P, Mallonee S, Schmidt AC, Ikeda R. Oklahoma Firearm Injury Surveillance System. *Am J Prev Med*. 1998;15(3S):83-91.
- <sup>xv</sup> Magee LA, Ranney ML, Fortenberry JD, Rosenman M, Gharbi S, Wiehe SE. Identifying nonfatal firearm assault incidents through linking police data and clinical records: Cohort study in Indianapolis, Indiana, 2007-2016. *Prev Med*. 2021;149:106605. doi: 10.1016/j.ypmed.2021.106605.
- <sup>xvi</sup> Barber CW, Ozonoff VV, Schuster M, et al. Massachusetts Weapon-Related Injury Surveillance System. *Am J Prev Med*. 1998 Oct;15(3 Suppl):57-66. doi: 10.1016/s0749-3797(98)00065-8.
- <sup>xvii</sup> Barber C, Azrael D, Cohen A, Thymes D, Wang D, Hemenway D. Homicides by police: comparing counts from the National Violent Death Reporting System, Vital Statistics, and Supplementary Homicide Reports. *Am J Public Health*. 2016; 106(5):922-7.
- <sup>xviii</sup> Conner A, Azrael D, Lyons VH, Barber C, Miller M. Validating the National Violent Death Reporting System as a source of data on fatal shootings of civilians by law enforcement officers. *Am J Public Health*. 2019 Apr;109(4):578-584. doi: 10.2105/AJPH.2018.304904.
- <sup>xix</sup> Olufajo OA, Zeineddin A, Nonez H, et al. Trends in firearm injuries among children and teenagers in the United States. *J Surg Res*. 2020 Jan;245:529-536. doi: 10.1016/j.jss.2019.07.056. (The paper includes data on all age groups.)
- <sup>xx</sup> NCGVR, *op. cit.*
- <sup>xxi</sup> Barber *et al.*, *op. cit.*, 1998.
- <sup>xxii</sup> National Center for Health Statistics and Center for Medicaid and Medicare Services. ICD-10-CM Official Guidelines for Coding and Reporting. Accessed at: [https://www.cdc.gov/nchs/data/icd/10cmguidelines-FY2016\\_final.pdf](https://www.cdc.gov/nchs/data/icd/10cmguidelines-FY2016_final.pdf) on Apr 12, 2020.
- <sup>xxiii</sup> NCGVR, *op. cit.*

<sup>xxiv</sup> Kaufman EJ, Wiebe DJ, Xiong RA, Morrison CN, Seamon MJ, Delgado MK. Epidemiologic Trends in fatal and nonfatal firearm injuries in the US, 2009-2017. *JAMA Intern Med.* 2021 Feb 1;181(2):237-244. doi: 10.1001/jamainternmed.2020.6696. PMID: 33284327

<sup>xxv</sup> Avraham J, Frangos SG, DiMaggio CJ. The epidemiology of firearm injuries managed in US emergency departments. *Injury Epidemiology.* 2018;5:38 <https://doi.org/10.1186/s40621-018-0168-5>

<sup>xxvi</sup> Gani F, Sakran JV, Canner JK. Emergency department visits for firearm-related injuries in the United States, 2006-14. *Health Aff (Millwood).* 2017;36(10):1729-1738. doi: 10.1377/hlthaff.2017.0625.

<sup>xxvii</sup> NCGVR *op. cit.*

<sup>xxviii</sup> Consumer Product Safety Commission. Collection of Information; Proposed Extension of Approval; Comment

Request—Follow-Up Activities for Product-Related Injuries Including the National Electronic Injury Surveillance System (NEISS) [Docket No. CPSC–2009–0102]. *Federal Register.* July 20, 2021;86(136):38315-8.

<sup>xxix</sup> CDC. Surveillance for fatal and nonfatal firearm-related injuries --- United States, 1993—1998. *MMWR.* 2001;50(SS02):1-32.

<sup>xxx</sup> Kalesan B, Adhikarla C, Pressley JC, et al. The hidden epidemic of firearm injury: increasing firearm injury rates during 2001–2013. *Am J Epidemiol.* 2017;185(7):546–553

<sup>xxxi</sup> Jena AB, Sun EC, Prasad V. Does the declining lethality of gunshot injuries mask a rising epidemic of gun violence in the United States? *J Gen Intern Med.* 29(7):1065–9.

<sup>xxxii</sup> Cook PJ, Rivera-Aguirre AE, Cerdá M, Wintemute G. Constant lethality of gunshot injuries from firearm assault: United States, 2003-2012. *Am J Public Health.* 2017;107(8):1324-1328. doi: 10.2105/AJPH.2017.303837.

<sup>xxxiii</sup> Cook PJ, Rivera-Aguirre AE, Cerdá M, Wintemute G. RE: "The hidden epidemic of firearm injury: Increasing firearm injury rates during 2001-2013" (LTE). *Am J Epidemiol.* 2017;186(7):896. doi: 10.1093/aje/kwx279.

<sup>xxxiv</sup> Agarwal S. Trends and burden of firearm-related hospitalizations in the United States across 2001–2011. *Am J Med.* 2015;128(5):484–492.e1.

<sup>xxxv</sup> American College of Surgeons. National Trauma Data Bank 2015: Annual Report. Chicago, IL; 2015.

<sup>xxxvi</sup> Spitzer SA, Pear VA, McCort C, Wintemute GJ. Incidence, distribution, and lethality of firearm injuries in California from 2005 to 2015. *J Gen Intern Med*. 2014 Jul;29(7):1065-9. doi: 10.1007/s11606-014-2779-z.

<sup>xxxvii</sup> Campbell S, Nass D, Nguyen M. The CDC is publishing unreliable data on gun injuries. People are using it anyway. *FiveThirtyEight*, October 4, 2018. Accessed 8/4/2021 at: <https://fivethirtyeight.com/features/the-cdc-is-publishing-unreliable-data-on-gun-injuries-people-are-using-it-anyway/>

<sup>xxxviii</sup> Unpublished data from Healthcare Cost and Utilization Project, 2016 National Emergency Department Sample.

<sup>xxxix</sup> Office of Management and Budget, OMB.report. “National Electronic Injury Surveillance System (NEISS) and Follow-up Activities for Product Related Injuries. OMB: 3041-0029” [online document]. Accessed 8/10/2021 at: <https://omb.report/icr/202102-3041-001/doc/109385500>

<sup>xi</sup> Personal communication, Linda Dahlberg, Senior Advisor to the Director of the Division of Violence Prevention, CDC National Center for Injury Prevention and Control, July 12, 2021.

<sup>xii</sup> Consumer Product Safety Commission. Justification AIP Expansion National Electronic Injury Surveillance System (NEISS) and Follow-up Activities for Product Related Injuries 3041-0029 [online notice]. On Office of Management and Budget’s OMB.reports. Accessed 8/10/2021 at: <https://omb.report/icr/202102-3041-001/doc/109452300>

<sup>xiii</sup> Personal communication, Linda Dahlberg, Senior Advisor to the Director of the Division of Violence Prevention, CDC National Center for Injury Prevention and Control, August 30, 2021

<sup>xiiii</sup> Gould DW, Walker D, Yoon P. The evolution of BioSense: lessons learned and future directions. *Public Health Rep*. 2017 Jul-Aug; 132(1 Suppl): 7S–11S. doi: 10.1177/003335491770695

<sup>xliv</sup> CDC. Overview of syndromic surveillance: what is syndromic surveillance? *MMWR*. September 24, 2004 / 53(Suppl);5-11

<sup>xlv</sup> Personal communication, Loren Rodgers, PhD, Lead, National Syndromic Surveillance Program at Centers for Disease Control and Prevention, June 25, 2021.

<sup>xlvi</sup> de Lusignan S, Chan T, Jones S. Large complex terminologies: more coding choice, but harder to find data--reflections on introduction of SNOMED CT (Systematized Nomenclature of Medicine--Clinical Terms) as an NHS standard. *Inform Prim Care*. 2011;19(1):3-5. doi: 10.14236/jhi.v19i1.787.

<sup>xlvii</sup> The National Syndromic Surveillance Program Community of Practice (NSSP CoP) Charter. June 2020. Accessed August 10, 2021 at [https://nsspcommunityofpractice.org/wp-content/uploads/2020/06/NEW-NSSP-CoP-Charter-Final\\_signedfinal.pdf](https://nsspcommunityofpractice.org/wp-content/uploads/2020/06/NEW-NSSP-CoP-Charter-Final_signedfinal.pdf)

<sup>xlviii</sup> CDC Center for Surveillance, Epidemiology and Laboratory Services, Division of Health Informatics and Surveillance. *BioSense Quick Start Guide to Using ESSENCE*. December 2017. Accessed 9/9/2021 at <https://www.cdc.gov/nssp/biosense/docs/biosense-platform-quick-start-guide-for-essence.pdf>

<sup>xlix</sup> Vivolo-Kantor AM, Smith H, School L. Differences and similarities between emergency department syndromic surveillance and hospital discharge data for nonfatal drug overdose. *Ann Epi*. (In press, 2021.) <https://doi.org/10.1016/j.annepidem.2021.05.008>

<sup>l</sup> CDC Division of Health Informatics and Surveillance. “Division of Health Informatics and Surveillance Strategic Framework, 2019-2021” [online brochure]. Accessed 9/12/2021 at: <https://www.cdc.gov/csels/dhis/documents/DHIS-Strategic-Framework-Brochure-508.pdf>

<sup>li</sup> Newsroom, Centers for Medicare and Medicaid Services. “Fiscal Year (FY) 2022 Medicare Hospital Inpatient Prospective Payment System (IPPS) and Long Term Care Hospital (LTCH) Rates Proposed Rule (CMS-1752-P)” [online factsheet]. April 27, 2021. Accessed 9/11/2021 at: <https://www.cms.gov/newsroom/fact-sheets/fiscal-year-fy-2022-medicare-hospital-inpatient-prospective-payment-system-ipps-and-long-term-care>

<sup>lii</sup> Centers for Medicare and Medicaid Services. “Medicare and Medicaid Promoting Interoperability Program Basics” [webpage]. Accessed 9/12/2021 at <https://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms/Basics>

<sup>liii</sup> US Department of Health and Human Services Office of the National Coordinator of Health Information Technology. “Connecting Health and Care for the Nation: A 10-Year Vision to Achieve an Interoperable Health IT Infrastructure” [online “vision paper”]. Not dated. Accessed 9/12/2021 at: <https://www.healthit.gov/sites/default/files/ONC10yearInteroperabilityConceptPaper.pdf>

<sup>liv</sup> Proposed Rules, *Federal Register*. May 10, 2021:86(88):25070-25790.

<sup>lv</sup> Barber C, Azrael D, Cohen A, Thymes D, Wang D, Hemenway D. Homicides by police: comparing counts from the National Violent Death Reporting System, Vital Statistics, and Supplementary Homicide Reports. *Am J Public Health*. 2016; 106(5):922-7.

<sup>lvi</sup> Conner A, Azrael D, Lyons VH, Barber C, Miller M. Validating the National Violent Death Reporting System as a source of data on fatal shootings of civilians by law enforcement officers. *Am J Public Health*. 2019 Apr;109(4):578-584. doi: 10.2105/AJPH.2018.304904.

<sup>lvii</sup> Barber et al, 2016, *op cit*.

<sup>lviii</sup> Conner et al, 2019, *op cit*.

<sup>lix</sup> Barber C, Hemenway D. Too many or too few unintentional firearm deaths in official U.S. mortality data? *Accid Anal Prev*. 2011;43(3):724-31. doi: 10.1016/j.aap.2010.10.018.

<sup>ix</sup> Barber *et al.*, *op. cit.*, 1998.

<sup>ixi</sup> Braga A and Cook P. The association of firearm caliber with likelihood of death from gunshot injury in criminal assaults. *JAMA Netw Open*. 2018;1(3):e180833. doi: 10.1001/jamanetworkopen.2018.0833.

<sup>ixii</sup> Cook P. The case of the missing victims: gunshot woundings in the National Crime Survey. *J Quant Crim*. 1985;1(1): 91-102.

<sup>ixiii</sup> VICE News. Police Shootings: Get the Data. December 2017. Accessed 8/4/2021 at: [https://news.vice.com/en\\_us/article/a3jjpa/nonfatal-police-shootings-data](https://news.vice.com/en_us/article/a3jjpa/nonfatal-police-shootings-data)

# Chapter 3. Measuring Gun Violence Using Police Data

Susan Parker, MPP, MS | University of Michigan

## Introduction

Each year in the United States, over 110,000 people are the victims of gunshot injury and fatality. Most nonfatal victims are shot by another person in an interpersonal assault. Despite historically low firearm mortality rates over the last three decades, in 2020, more firearm homicides occurred than in any year since 1995, rising by 30 percent over the previous year's total.<sup>1</sup> Following the onset of a global pandemic that exacerbated social and economic inequalities, firearm violence in many cities is on track for record-breaking homicide rates in 2021 as well. In the first half of 2021, homicides rose an additional 16 percent.<sup>2</sup>

While reliable data systems exist for tracking homicides, no similarly valid and reliable data system exists for surveilling all forms of gun violence. Gun violence ranges from threats with a gun to violent assaults and shootings where a victim is wounded, sometimes fatally. Comprehensive information about nonfatal firearm injuries, including rates, location, victim characteristics and context, is unknown. Yet nonfatal firearm injuries occur more often than fatal firearm deaths. The gap in tracking gun violence has wide-ranging implications for gun violence prevention. Without nonfatal firearm injury data, it is not possible to determine if firearm lethality leads to additional firearm deaths and injuries. It is also impossible to assess if quality emergency medical response can prevent deaths. Lack of gun violence data harms national resource allocation, and law enforcement agencies on the front lines of responding to gun violence lack important measures of whether their efforts are effective.

Compounding the lack of insight into gun violence is the recent abrupt retirement of the nation's long-standing crime reporting system without strong adoption of its successor. The UCR SRS has for decades served as the measure of U.S. crime, in particular measuring violent crime committed with a firearm. While this system was not a perfect measure of crime incidence, for over 80 years policymakers, criminal justice leaders, and researchers relied on UCR SRS measures of violent crime committed with a firearm to measure gun crime and victimization. The new national standard for crime reporting, the NIBRS, addresses some UCR SRS inadequacies, but adoption among criminal justice agencies has been alarmingly low.<sup>3</sup> Despite limited adoption of the NIBRS, in 2015 the FBI opted to retire the legacy UCR SRS program as of January 1, 2021.<sup>4</sup> This decision is likely to hamstring the reliability and accuracy of crime trends in 2021 and in the years ahead. The unlikely best-case scenario is that one in four law enforcement agencies will be unable to report crime statistics to the new NIBRS system.<sup>5</sup> The loss of comprehensive local crime data, combined with imprecise national level estimates of

violent crime, is an urgent barrier to addressing violent crime, much less gun violence and victimization.

Gun violence surveillance should provide accurate counts of injuries resulting from gun violence, both fatal and nonfatal, as well as measure threats and assaults committed with a gun where no injury occurs. The purpose of this report is to consider recommendations to establish accurate and reliable national gun violence surveillance on the local and national level using robust police data systems. Critical to this objective is addressing the low adoption among law enforcement agencies transitioning to reporting incident-based crime data.

## Measuring Gun Violence in the U.S.

In the U.S., the majority of gun violence does not result in physical injury. It is measured in police records as assaults and robberies where the perpetrator is armed, and annually affects approximately 300,000 people.<sup>6</sup> Over 50,000 victims annually are shot by another person in a criminal assault.<sup>7</sup> Gunshot assault victims are most often young Black men aged 15-24. Their injuries are so severe that among those seeking emergency medical treatment, the majority result in inpatient hospitalization for additional care following the initial ED visit.<sup>7</sup>

The two principal data sources for gun violence surveillance purposes are ED records and police reports. Legal status, intent, and injury severity are important determinants of what kind of information is gathered about gun violence by which entity. Police investigation of interpersonal firearm violence determines whether a gun assault or other criminal violation occurred, and is recorded in police data. Police document the circumstances, victims, and other valuable details about gun violence in police data. While police document all gun violence known to them, other types of firearm injury intent are less comprehensive in police records. Because laws differ by state regarding the legal status of accidents involving guns, police do not document accidental shootings comparably in each state<sup>9</sup>. As such, unintentional and self-harm shootings are best captured in ED records. However, ED data do not always accurately differentiate between different types of intent (see Barber (2021) in this series for full detail), but such data are the most comprehensive source of all gunshot injury types. Finally, injury severity is also an important determinant of which data sources measure gun violence events. Gun assaults without injury requiring medical attention likely appear only in police data or may go undetected.

Law enforcement and hospital sources should have substantial overlap measuring nonfatal shootings, particularly for gunshot assault cases. The majority of states (45 states as of 2014) have enacted mandatory reporting laws which state that health providers treating gunshot wound injuries are required to report those injuries to law enforcement.<sup>8,9</sup> These injuries would mandate reporting regardless of whether the injury was the result of an assault, an accident, or

---

<sup>9</sup> For instance, states vary in their adoption of Child Access Prevention (CAP) laws which make negligent storage resulting in child firearm injury a criminal offense.



another cause. Among the five states that do not have a mandatory reporting law for all weapons injuries, three require reporting of all injuries that health providers treat stemming from “crimes” or “violently inflicted injuries.” In these laws, health providers were typically defined to include physicians, nurses, emergency medical technicians or other licensed medical professions. The expansive definition of health provider captures gun injuries seen not only in EDs but also in other, less commonly used health care settings, such as inpatient hospitals or urgent care clinics. Almost all states, therefore, mandate reporting of gunshot victims and coverage for assault is even more comprehensive. It is unknown, however, the extent to which health providers comply with reporting requirements.

Despite overlap in reporting gun violence, police reports and investigations capture unique, valuable information that is not tracked in ED settings. Police data record information about geographic location, circumstances, and perpetrators involved in gunshot injuries. ED data rely on coding standards set under the ICD diagnostic standards, which do not specify that the same detailed incident information is recorded. For instance, in a police investigation of a gunshot assault, police record whether the perpetrator is an intimate partner, a stranger, another relative, or even an acquaintance. The same gunshot assault treated in an ED would code a gunshot injury by assault, but ICD coding does not differentiate between an assaultive gunshot injury committed by a stranger or a close relative. ICD coding can record whether an incident may be associated with domestic or spousal abuse but in practice is often missing or underreported by the patient.<sup>10,11</sup> Police data is well-suited to capture information about assaults where no medical attention is required or where knowledge of the perpetrator is important for determining the nature of the assault.

Table 1 documents police and ED data sources for surveilling nonfatal gunshot wounds by intent, victim relation to perpetrator, and how data are aggregated. It is clear that an accurate, universal surveillance system would include both police and ED data. Police data captures important information about gun crime that ED data does not. ED data more comprehensively captures certain types of gun victimizations, including accidental and self-harm gun injury, as well as gunshot wounds that may not come to police attention.

**Table 1:** Comparison of Police and ED Data for Measuring Gun Violence

	Police Data	ED Data
<b>Intent type</b>		
Assault	Yes	Yes*
Accident	Not comprehensively	Yes*
Self-harm	Not comprehensively	Yes*
Legal Intervention	Yes	Yes*
<b>Gun violence sample</b>		
Violent crime committed with a firearm; no discharge; no wounds	Yes	No
Violent crime committed with a firearm; discharge; no wounds	Yes	No
Violent crime committed with a firearm; discharge; gunshot wounds	Yes	Yes
<b>Is data aggregated to federal database?</b>	Not comprehensively; only violent crime with gun	Yes; ED data is available at national and regional levels
<b>How timely is the release of data?</b>	9-month lag in national reporting; historically; quarterly updates	At least 18-month lag in national, regional reporting
<b>Strengths</b>	Police data include information on threats with a gun as well as shootings. Police record perpetrator information, relationships among victims and offenders, and linkages between crimes. Police have the ability through investigation to classify types of assaults.	ED data are likely to capture a large swath of gunshot victims. ED data include accidental and self-harm nonfatal firearm injuries for a full census of gun injuries beyond assaults.
<b>Flaws</b>	Not all gun violence is recorded by police (i.e., self-harm and accidental intent types). Not all crimes may be reported to police, though more serious violent crime is likely to be. Police data does not currently comprehensively measure gun discharges and gunshot wounds.	ED data may not accurately classify intent. It overestimates accidents, underestimates assaults and often fails to capture information about domestic violence.

\*ED coding of intent in the majority of sources such as the NEDS is not accurate for gunshot wounds. See Barber (2021) for full discussion.

## Proportion of Gun Violence Reported to Police

Research on the proportion of gun violence reported to police is limited. A commonly used comparison relies on survey data from the NCVS, comparing survey reports of violent crimes committed with a firearm to those reported to police agencies. While the most recent 2018 NCVS estimate of violent crimes committed with a firearm reported to police is 62 percent, this comparison does not measure actual gunshot victims, nor does it likely include individuals at high risk of gunshot injury.<sup>12</sup> Importantly, the NCVS sample is not well-suited to measuring gunshot victimization.<sup>13</sup> Another approach to estimating the incidence of gunshot reporting involves comparing gunshot detection systems to gunshots reported to police, which finds that 12 percent of gunshots detected are reported to police.<sup>14</sup> Reluctance to report crime to police is another barrier to comprehensive crime reporting that is exacerbated by poor police-community relations, police shootings, and other forms of misconduct.<sup>15,16</sup>

## Police and ED Overlap in Assault Gunshot Victim Reporting

In addition to uncertainty in the overall proportion of gunshot victimization captured in police records, overlap between police and ED records is not well-understood. In part, these studies can demonstrate a false reporting gap between police and EDs, driven by two reporting issues in ED records. First, firearm injury identification through medical coding does not always catch the gunshot wound. Second, the medical records personnel who classify hospital data frequently code intentional firearm injuries as accidental (see Barber 2021). Using ICD-coded ED data as a basis for comparing to police data will undercount assaultive gunshot wounds. Studies that use names and other identifiers to match the overlap between police and ED data still show gaps between the two sources. In an Atlanta study of gun injuries, 9 percent of all firearm injuries, possibly accidental gunshot injuries, could not be matched with a police report.<sup>17, 10</sup> Another study in Indianapolis carefully matches nonfatal gun assaults reported to police to ED records, finding substantial overlap between the two sources but that police records contain assaultive gunshot wounds that are unreported in ED records.<sup>18</sup> Importantly, medical provider compliance with mandatory reporting of firearm wounds or assaults is not known.

## Identifying Gun Violence in Police Data

While police records contain data useful for gun violence surveillance and standardize assaults or robberies with a firearm, reporting practices obscure incidence and victim counts. Since 1930, police agencies have reported crime data to the federal government through the FBI's UCR program in a format known as the SRS. The standards adopted over nine decades ago

---

<sup>10</sup> Georgia mandatory reporting laws in place at the time of the study require medical providers to report assaults, not all firearm injuries, to police. Not reporting accidental firearm injuries to police may have lowered gun injuries known to police compared to other jurisdictions with mandatory reporting of all gun wounds, but compliance with these laws for firearm injuries is not well understood.

continue to shape how police agencies record and report crimes. At the time, agencies opted to count incidents in discrete crime categories, requiring that one crime type per incident be reported. For incidents where multiple crimes take place, agencies may undercount reported crime. An additional problem is that no category for nonfatal gunshot assaults exists in the SRS. For criminal nonfatal shootings, the best definition has been to count them as part of two much larger categories of crime: gun robbery or aggravated assault with a gun. In this classification, it is not possible to know how many people are wounded by a gunshot or are shot at and not wounded.

Despite no explicit UCR SRS standards, police agencies already do track important information about nonfatal gunshot victims, though it is not aggregated to a standard reporting format. Because this information is outside the scope of the SRS, agencies differ in how this information is gathered, stored, accessed, and prioritized. In a recent study of nonfatal firearm injury in four cities, each jurisdiction had a different method of storing and accessing nonfatal firearm injury victimization.<sup>19</sup> For instance, the Baltimore Police Department created a new, separate crime category to track nonfatal firearm assault victims.<sup>20</sup> For other police agencies, gun violence and victim data is less accessible and may be stored in incident or investigation reports that require tallying nonfatal victims. Among the largest police agencies dealing with gun violence, nonfatal firearm injury data is for the most part accessible and recorded. Among the 40 largest police agencies in the U.S. that belong to the Major City Chiefs Association (MCCA), by 2018, over 70 percent reported nonfatal firearm events quarterly to MCCA summary data reports on violent crime.<sup>11</sup>

In addition to a lack of clarity about definitions, agencies face barriers related to access to resources for tracking nonfatal shooting victims. Solutions have included amending their Records Management Systems (RMS) to include an applicable nonfatal shooting category, tasking staff to tally crime incident reports, or automating reporting solutions. Resources and therefore access to nonfatal shooting data may vary by agency prioritization of nonfatal shootings or of incidence within a jurisdiction. However, police agencies are charged with investigating firearm assaults to determine if a crime has been committed and in the course of this central function, record detailed information about gun assaults. Access to data on nonfatal shootings should not be misconstrued as missing information on nonfatal firearm victims in police data.

---

<sup>11</sup> Source: <https://majorcitieschiefs.com/resources/> and calculations of proportion of agencies reporting nonfatal shootings.

## National Police Data Systems for Measuring Gun Violence

The FBI is the central coordinator of data on criminal offenses reported to police.<sup>21</sup> The FBI gathers annual agency level police data on criminal offenses, clearances, and arrests. Currently, this national police data system is undergoing its most significant transition in nine decades. In 2016, the FBI accepted a recommendation from its Criminal Justice Information Services (CJIS) Advisory Policy Board (APB) to transition fully to a new reporting platform, known as the NIBRS, starting in 2021. NIBRS gathers substantially more detailed and complex data about individual crime incidents, in contrast to the decades of aggregated incident counts.

For many state and law enforcement agencies, the transition to NIBRS is both complicated and voluntary, making adoption a challenging process. Some states and law enforcement agencies have managed the transition process and achieved close to full incident-based reporting. Others have faced adoption challenges or may not prioritize replacing existing reporting methods. In either case, it is likely impossible for a third of police agencies to continue reporting their data to the FBI in 2021. This change urgently jeopardizes national and local understanding of crime at a critical time when rising violent crime and victimization demand accurate and reliable data.

Further, difficult and slow adoption of NIBRS is a barrier to tracking gun violence. In June 2021, the FBI CJIS APB proposed two substantial changes to enable NIBRS to track nonfatal firearm victims. First, the APB proposed adding an injury category for gunshot wound victims. Second, the APB proposed a new data element to capture whether a firearm was discharged, accidentally or intentionally, in the commission of a crime. If these changes to NIBRS are approved by the FBI director, they will be added to NIBRS data collection scheme. However, if states and agencies continue to struggle to report their data to NIBRS, the important changes enabling better surveillance of gun violence will be unsuccessful.

### UCR SRS

Police data are compiled on the local, regional, state, and national levels through the FBI's UCR program. Since 1930, the UCR SRS has gathered data from law enforcement agencies to measure national, state, and local crime trends. The SRS gathers detailed data on crimes reported by law enforcement agencies, including violent crimes, rape, robbery, and aggravated assault as well as property crimes, in addition to police data on arrests and additional offenses. Originally conceived of as a paper-based system, the SRS by 2019 included over 18,000 law enforcement agencies nationwide who report their crime data and typically span upwards of 90 percent of the U.S. population.<sup>22,23</sup>

While the SRS is the most long-standing source of crime data, the limitations of the SRS spurred the transition to incident-based crime reporting. The SRS gathers data on a more limited set of offenses than is customary for most police agencies.<sup>24</sup> For instance, it does not gather offense information about assault by intimidation, human trafficking, or kidnapping. Additionally, the SRS tabulates the most serious offenses associated with each reported crime,

a convention known as the “hierarchy rule.”<sup>25</sup> The hierarchy rule ranks crimes in order of which is most severe and classifies a crime only according to the most severe offense. In effect, the SRS does not record all crimes that might have been committed in a single incident. In the event of a robbery where an aggravated assault also occurs, only the robbery would be recorded as an event. The SRS can therefore mask information about offenses and victimization.

The SRS does not gather information sufficient to conduct accurate surveillance of gun violence. Rather than recording information about crime victims, the SRS records the type of weapon used in the commission of a crime. Because relatively few gun robberies and aggravated assaults with a gun result in gunshot wound injuries, the SRS measure is far from measuring gunshot wound victimization. The SRS does measure two types of violent crimes committed with a firearm, robbery, and assault. According to the SRS, 37 percent of robberies are committed with a firearm and 26 percent of assaults are committed with a firearm annually, a figure that numbers on average approximately 310,000 annually.

## NIBRS

By the early 1980s, efforts were underway to create a national data collection system capable of recording detailed incident and arrest information. This system, known as the NIBRS, was intended to improve on the SRS. NIBRS records substantially more crime and offense types, gathering incident-level data on incident and arrest data on 52 offenses. It does not impose a hierarchy rule so each incident can be classified into ten criminal offenses per incident.

While NIBRS does not currently gather data on gunshot assault injury victims in a standardized manner, it does allow for more detail in recording information about crimes where a firearm is used than the prior SRS. NIBRS expands on the SRS to include additional crimes and allows for a firearm to be indicated in the commission of these crimes, including important offenses such as sexual assault and rape. However, NIBRS does not specify an injury type to record gunshot wound injury, so law enforcement agencies code gunshot wound injuries under a variety of injury types. Some agencies code “laceration” wounds, but others may specify injury types such as “other major injury.”

**Table 2:** Summary of Police Data Sources for Reporting Gun Violence

	UCR SRS	NIBRS
<b>Sponsor</b>	FBI	FBI
<b>Data gathered by?</b>	State UCR programs; individual law enforcement agencies	State UCR programs; individual law enforcement agencies
<b>Data format accepted by the FBI?</b>	No longer accepted by the FBI; however, many state and local agencies still use this format to track crime	Accepted by the FBI
<b>Agency representativeness</b>	Yes; upwards of 85 percent of law enforcement agencies report	No
<b>Population representativeness</b>	Approximately 90 percent	No
<b>State / Local / Regional representativeness</b>	Yes	No
<b>Details on circumstances?</b>	Limited	Yes
<b>Data available to track crime trends in 2021 and beyond?</b>	Unlikely. The FBI will generate the SRS tables in 2021 but they will be comprised of incomplete NIBRS-reported data	No
<b>Identifies gun assaults?</b>	Yes	Yes, but for a subset of agencies who report NIBRS data
<b>Identifies nonfatal gunshot wound victims?</b>	No	Not currently
<b>Strengths</b>	The SRS has been used since the 1930s to track crime trends and provides the most comprehensive repository of law enforcement agency data on robberies and aggravated assaults committed with a firearm.	NIBRS data is incident-based, providing detailed, linked information about crime. NIBRS does not classify crimes into a single event but allows for multiple classifications and accurate victim counts. Policy recommendations to change NIBRS to count gunshot wound victims are under consideration.
<b>Flaws</b>	The SRS uses the "Hierarchy Rule" to classify crime events to a single category, losing important detail about gun crime events. Additionally, incidents can result in multiple victims of gunshot wounds but in SRS constitute a single event. The SRS does not identify nonfatal firearm assault victims.	A large percentage of agencies have not adopted NIBRS, at least yet, and the FBI has not developed a method for estimating national or state crime rates from those agencies that do report. NIBRS currently does not identify gunshot wound victims.

In their current forms, neither the SRS nor NIBRS measure gunshot wounds reported to law enforcement. NIBRS allows a firearm to be associated with additional crimes such as kidnapping and simple assault. For NIBRS to improve on the SRS crime reporting, NIBRS should capture additional crimes committed with a firearm, for two reasons. First, because NIBRS captures additional crime types committed with a firearm, it expands firearm crime beyond aggravated assaults or robbery with a firearm. Further, NIBRS does not apply the hierarchy rule that the SRS employs, meaning it should not undercount aggravated assaults that occurred with a robbery.

Table 3 compares the estimates of crimes committed with a firearm between NIBRS and the SRS among six states whose population coverage in NIBRS is consistently at or above 98 percent. Agencies within each state that appear in both NIBRS and the SRS were then matched to compare crimes committed with a firearm between agencies reporting to both systems. Consistent with previous findings, NIBRS does not result in substantially more reported incidents. Robbery with a firearm is more comparable than aggravated assaults with a firearm between sources. The SRS rate of aggravated assaults committed with a firearm is higher than NIBRS.

**Table 3: NIBRS vs. SRS Measures of Crimes Committed with a Firearm (2019)**

	SRS		NIBRS	
	n	%	n	%
<b>Nationwide</b>				
Aggravated Assault with Firearm	201,617	26.0	-	-
Robbery with Firearm	89,768	34.8	-	-
Homicide with Firearm	10,258	73.6	-	-
<b>State-level Comparison</b>				
Aggravated Assault with Firearm				
Colorado	3676	31.8%	4338	25.3%
Kentucky	1714	34.0%	2118	29.2%
Michigan	8145	27.3%	9033	26.6%
South Carolina	7186	41.4%	8838	37.0%
Tennessee	10983	35.4%	14100	34.2%
Virginia	3229	31.5%	4356	27.5%
Robbery with Firearm				
Colorado	1315	38.6%	1337	38.0%
Kentucky	1008	46.7%	1037	45.0%
Michigan	2253	43.4%	2265	43.4%
South Carolina	1829	57.1%	1882	55.9%
Tennessee	3798	61.8%	3923	60.0%
Virginia	1771	50.3%	1807	49.2%

Sources: FBI Crime Data Explorer NIBRS 2019; Jacob Kaplan's UCR Data



## NIBRS Gunshot Injury Policy Recommendations Under Consideration

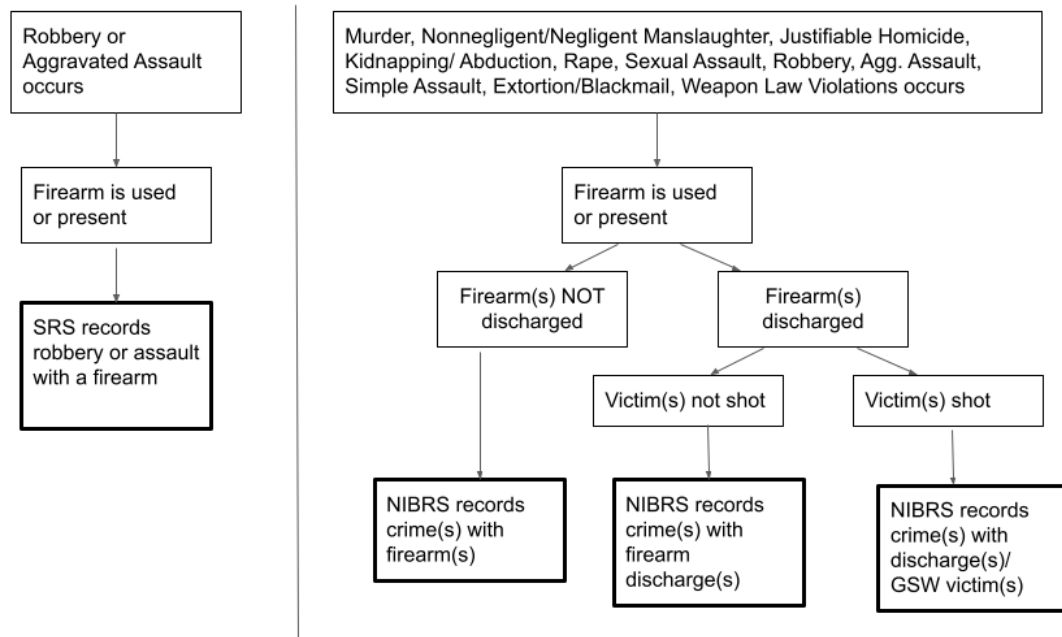
In June 2021, the FBI’s CJIS APB proposed several recommendations to improve federal crime data measuring nonfatal shootings. These recommendations are currently under FBI director consideration and address important NIBRS shortcomings in surveilling nonfatal firearm assault victims. A summary of the proposed changes to record gunshot wound incident information is in Table 4. These recommendations are a promising step toward using federal data sources to capture valuable police data about gun violence victims.

The first policy proposal is to add an injury category for “gunshot” wounds, clarifying the status of gunshot wound victims. This modification would allow for surveillance of gunshot victims reported to police. Second, the APB proposed the addition of a new data element within the NIBRS Offense Segment to track firearm discharges. Tracking firearm discharges allows for distinguishing between threatening with a firearm and actually firing the weapon. The proposed definition of firearm discharge is “the intentional, unintentional, accidental or negligent discharge of a firearm during the commission of a criminal incident.” If an agency records nonfatal gunshot events without a victim as a “shots fired” incident report according to agency protocol, the incident would be submitted for inclusion in NIBRS as a weapons law violation. Agency protocol may therefore make between-city comparisons difficult due to variation in agency reporting policy, but within agencies that do report “shots fired” events, it may be possible to distinguish between crimes when firearms are discharged and not. The differences in reporting firearms crimes in SRS and the NIBRS proposals are depicted in Figure 2, contrasting the event and data collection of recording firearm assaults and firearm assault victims.

**Table 4:** Summary of NIBRS Proposed Gunshot Wound Changes

Proposed Data Change	Detail	Element Type	Segment
Track gunshot wound injuries	Gunshot wounds inclusive of "incidents involving grazing or minor injuries received when the firearm was discharged" would be recorded as an injury code attributed to a crime incident.	Additional Category	Victim
Track firearm discharges	Track the "intentional, unintentional, accidental, or negligent discharge of a firearm during the commission of a criminal incident." Discharges may be recorded as "shots fired" to some agencies and will be recorded under "weapons law violations."	New Data Element	Offense

**Figure 2:** UCR SRS vs. proposed NIBRS firearm categories



While the proposal under consideration is an important step forward in tracking gun assault victims, two key problems remain. First, nonfatal shootings by police are not captured in NIBRS and therefore will not be tracked. The NIBRS perpetrator segment does not include an option for law enforcement to shoot in the line of duty. The resulting under-reporting may be substantial. Shootings by police claim approximately 1,000 lives each year, which comprise approximately 8 percent of homicides.<sup>26</sup> In 2015, the FBI created the National Use of Force Data Collection program, which should capture nonfatal police shootings from law enforcement agencies. Reporting is voluntary and agency reporting has been incredibly low. In the first quarter of 2021, only 89 agencies—less than 0.01 percent of law enforcement agencies—reported a use of force incident to the FBI.<sup>12</sup>

The second key barrier is the slow transition to NIBRS, documented in the next section.

## Transitioning to NIBRS

The NIBRS transition is essential to accurate national, state, and local crime statistics, but implementation progress is likely well-below the near-nationwide SRS reporting coverage. The slow transition compounded by a lack of transparency about agency reporting levels is likely to render 2021 crime data incomplete and unusable. It is likely to irreversibly harm knowledge about crime trends in the U.S., possibly for the next several years. Without detailed, accurate,

<sup>12</sup> Source: FBI Crime Data Explorer, available: <https://crime-data-explorer.fr.cloud.gov/pages/downloads>

and reliable information about violent crime and homicides, communities undergoing spikes in serious violent crime and homicide could be left without information crucial to problem identification and targeting of resources.

## NIBRS Adoption Essential to Reliable Crime Data and Gun Injury Surveillance

It is unclear how close to the robust nationwide reporting achieved under SRS that NIBRS will be for 2021, but much of the evidence is alarming. An April 2020 survey conducted by the FBI suggested that 75 percent of the nation’s law enforcement agencies were committed to transitioning to NIBRS by the January 1, 2021 deadline.<sup>5</sup> This estimate was taken prior to the start of a global pandemic and is unlikely, given the challenges to law enforcement agencies in 2020. In a statement, the BJS noted that “while the FBI’s assessments suggested the majority of agencies will transition by the 2021 deadline, some agencies have been delayed by COVID-19 and other social issues affecting resources and daily operations.”<sup>6</sup> The FBI has gathered “yearly evaluations” from agencies on their expected transition dates but states that “at this time, the FBI UCR Program cannot project how many agencies will report 2021 data or the population coverage that would be represented.”<sup>13</sup> Table 5 documents yearly reporting in NIBRS compared to the SRS. Slightly over half of the nation’s law enforcement agencies submitted data to NIBRS in 2020 representing less than two-thirds of the nation’s population.

**Table 5: NIBRS Reporting by Law Enforcement Agency, U.S. Population Coverage**

Year	NIBRS Agency Coverage		NIBRS Population Coverage		SRS Population Coverage	
	n	%	%		%	
2015	6823	40.6	31.6		96.4	
2016	7083	41.9	33.2		96.3	
2017	7146	43.6	35.7		91.3	
2018	7777	51.1	46.7		82.2	
2019	8536	57.0	53.2		81.8	
2020	9947*	57.7*	57.5**		unknown	

Source: FBI Crime Data Explorer. Denominator of law enforcement agencies and agency population coverage defined as agencies denoted as both participating and active and publishable.

\*Granular data on NIBRS participation for 2020 is unavailable at the time of publication but sourced from “UCR Participation 1960 - 2020”, available: [https://s3-us-gov-west-1.amazonaws.com/cg-d4b776d0-d898-4153-90c8-8336f86bdfec/ucr\\_participation\\_1960\\_2020.csv](https://s3-us-gov-west-1.amazonaws.com/cg-d4b776d0-d898-4153-90c8-8336f86bdfec/ucr_participation_1960_2020.csv)

\*\* Source: FBI response to inquiry about program coverage.

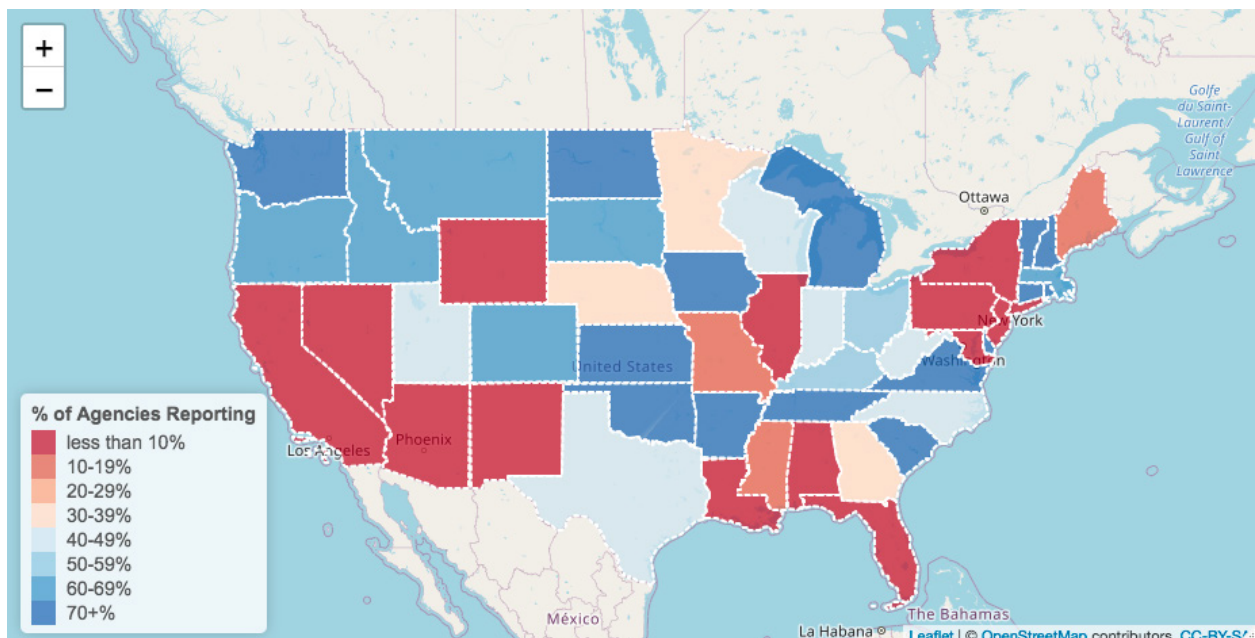
While some states and agencies are well within the timeline for NIBRS implementation, others will not meet the 2021 deadline for incident-based implementation. In some cases, the largest

<sup>13</sup> Correspondence from FBI National Press Office

states in the nation will not report their data to the FBI in 2021, including most New York agencies, and all of Florida, California, Alaska, and New Mexico. With little to no reporting from New York, Florida, and California alone, the FBI’s 2021 crime data will be reduced by almost two thousand agencies or almost a quarter (24.1 percent) of the U.S. population. The substantial loss in reporting is likely to render 2021 crime statistics produced by the FBI unusable at the national level and at the state- and local level. While the FBI will not accept data in the SRS format, many states including New York, Florida, Illinois, and California will continue to produce summary reporting statistics to inform crime trends in their respective states.<sup>14</sup>

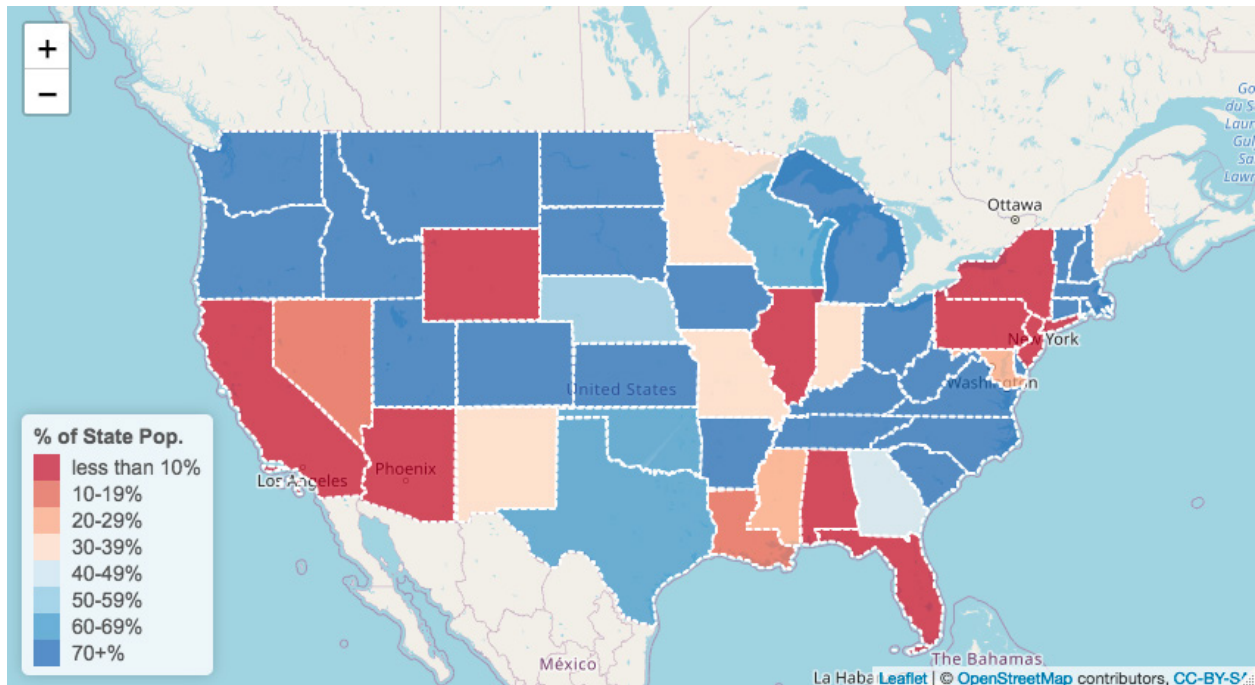
Figure 3 visualizes the percentage of agencies and state population that participating agencies reporting to NIBRS spanned in 2019. While FBI SRS data has been released for 2020, no update to NIBRS participation or data has been released at the agency level at the time of publication. For 2019, 13 states (in red) reported less than 10 percent of statewide law enforcement agency data in incident-based format. The median percentage of law enforcement agencies reporting to NIBRS is less than 50 percent, and the lowest quartile of states reporting incident-based data have less than 10% of agencies reporting to NIBRS. While median state population coverage is 77 percent, some of the largest states in the nation had less than ten percent population coverage, including California, New York, Florida, Pennsylvania, and Arizona.

**Figure 3:** 2019 Law Enforcement Agency NIBRS Reporting by State



<sup>14</sup> Sources: Freedom of Information Act Requests to Illinois and California State UCR Programs; Phone calls and correspondence with New York and Florida State UCR programs.

**Figure 4:** 2019 Population NIBRS Reporting Coverage by State



NIBRS participation levels underline alarmingly low participation particularly in some cities and counties. Table 6 reports 2019 NIBRS coverage by city and county in terms of agency and population. In cities with populations over 1,000,000, only 18 percent reported NIBRS data which constitutes two of the 11 largest police agencies in the nation. Small city and municipality coverage spanned less than half in terms of agency participation and population coverage.

**Table 6:** 2019 NIBRS Agency Participation

Population group description	Total agencies	NIBRS reporting agencies	Percent of NIBRS reporting agencies	Percent of Population Covered by NIBRS reporting agencies
Cities 1,000,000 or over	11	2	18.2%	13.4%
Cities from 500,000 thru 999,999	27	18	66.7%	70.5%
Cities from 250,000 thru 499,999	52	21	40.4%	41.3%
Cities from 100,000 thru 249,999	233	102	43.8%	43%
Cities from 50,000 thru 99,999	508	218	42.9%	43%
Cities from 25,000 thru 49,999	943	433	45.9%	46.3%
Cities from 10,000 thru 24,999	1996	922	46.2%	46.4%
Cities from 2,500 thru 9,999	4317	1812	42.0%	42.7%
Cities under 2,500	9231	2467	26.7%	33.9%

Metropolitan Statistical Area (MSA) counties 100,000 or over	181	78	43.1%	37.2%
MSA counties from 25,000 thru 99,999	484	262	54.1%	54.4%
MSA counties from 10,000 thru 24,999	295	163	55.3%	54.3%
MSA counties under 10,000	1799	433	24.1%	60.4%
Non-MSA counties 100,000 or over	3	1	33.3%	51.2%
Non-MSA counties from 25,000 thru 99,999	289	162	56.1%	55.5%
Non-MSA counties from 10,000 thru 24,999	649	384	59.2%	59.5%
Non-MSA counties under 10,000	2974	992	33.4%	63%
MSA State Police	236	24	10.2%	not applicable
Non-MSA State Police	168	42	25.0%	not applicable

NIBRS adoption likely means federal crime statistics at the state and national levels are unusable due to the volume of agencies that do not participate and the significant swaths of the U.S. population that are not represented. Because the FBI intends to use NIBRS as the basis of the annual SRS, no national crime statistics are likely to be available. As NIBRS will be the basis of state-level estimates, many states in the nation unable to report to NIBRS will not be counted. States with low levels of NIBRS reporting are likely to have uncertain crime estimates.

Unsuccessful NIBRS adoption will further harm the important objective of gun violence surveillance by criminal justice agencies. Slow adoption of NIBRS means no national, state, or local crime data on crimes committed with a firearm in 2021, even with imperfect summary reporting. In the future, should the proposals to amend the NIBRS gun injury data collection be adopted by the FBI director, a lack of nationwide NIBRS participation means that gun injuries are not surveilled in any systematic and reliable fashion on the part of law enforcement agencies. With many agencies already able to track nonfatal shootings, the slow adoption misses valuable years of data. Even with robust NIBRS adoption and FBI adoption of nonfatal gun victims as an injury type, it is likely to be 2-4 years before gunshot victim data would be gathered.

## The Importance of State Uniform Crime Reporting Programs in the NIBRS Transition

State UCR Programs play an essential role in reporting crime to the FBI and are deeply involved in facilitating the NIBRS transition. State UCR programs gather and submit crime data to the FBI on behalf of their state, streamlining FBI data collection tasks. Forty-three states mandate submitting crime data to the State UCR Program, and three states mandate cooperation.<sup>27</sup> Eighteen states mandate reporting to the national UCR program. The State UCR Programs provide technical assistance and support to local law enforcement agencies in reporting crime data meeting both state and federal guidelines for crime data reporting. State UCR programs

have been key to the NIBRS transition and are crucial for continued NIBRS adoption and implementation of future incident-based reporting changes such as tracking assault gunshot wound victims.

To support the NIBRS transition, state UCR Programs have performed a variety of functions to transition local law enforcement to reporting incident-based crime data. These functions have ranged from training, grantmaking, and technical support to consolidated bargaining with RMS to facilitate incident-based reporting. For most states, the NIBRS transition comes on top of state-level requirements for incident-based reporting. For instance, the Texas Incident-Based Reporting System gathers data on sexual assaults, drug seizures, and family violence.<sup>28</sup> The Michigan Incident Crime Reporting System gathers data on 104 criminal offenses, almost double the NIBRS list of offenses.<sup>29</sup> Many state UCR programs have adopted creative solutions to enable incident-based reporting in their respective states. To enable small agencies to adopt incident-based reporting, Wisconsin's UCR program adapted an existing police platform that agencies already use to record incident data.<sup>15</sup> Kansas' UCR program has approached prominent RMS vendors within the state to streamline incident-based reporting RMS changes among local agencies sharing the same vendors.<sup>16</sup>

## The National Crime Statistics Exchange (NCS-X) and National Crime Estimates

During the transition to NIBRS, the BJS and the FBI collaborated to support state and local law enforcement agencies transitioning to NIBRS in a variety of ways. A core focus has been the ongoing National Crime Exchange (NCS-X), an initiative to develop incident-based data by providing resources and technical support to state and local agencies to transition to NIBRS and ultimately provide incident-based data to the FBI.<sup>30</sup>

The NCS-X team assists agencies by coordinating information across federal and local agencies as well as producing resources to support the expansion of NIBRS reporting. Perhaps key among these resources have been technological guidance for agencies seeking to use federal funds to support the transition of their RMS to NIBRS compliance. Funding has supported local agencies in addition to providing funding and assistance to state UCR programs.

One of the most important aspects of the NCS-X program is the creation of a targeted sample of law enforcement agencies chosen to create population crime estimates from NIBRS data. The sample consists of 400 law enforcement agencies, which include the 72 largest law enforcement agencies in the U.S., as well as other agencies selected to produce nationally representative crime estimates. The NCS-X 400-agency sample was given additional resources

---

<sup>15</sup> Source: conversation with State UCR program director

<sup>16</sup> Source: conversation with State UCR program director

in the form of technical support, guidance, and grants assistance to achieve incident-based reporting by 2021. However, by August 2021, only 52 percent of the 400 agencies were certified to submit data to NIBRS.<sup>31</sup> Large agencies including Chicago, New York, and Los Angeles—the police agencies serving the largest metropolitan populations in the nation—are not able to submit to NIBRS. Other large police agencies unable to submit and in the NCS-X sample include Phoenix, Miami, Tucson, and St. Louis, among others. Table 7 below presents August 2021 NCS-X agency participation by NIBRS adoption status. In each population group, substantial proportions of agencies remain unable to submit incident-based data almost a year past the retirement of the SRS.

**Table 7: NCS-X Sample by NIBRS Reporting Status (as of August 2021)**

Population Group	Total Agencies	Agencies Reporting to NIBRS	
		n	%
All cities 250,000 or over	44	25	56.8%
Cities from 100,000 thru 249,999	36	16	44.4%
Cities from 50,000 thru 99,999	23	8	34.8%
Cities from 25,000 thru 49,999	38	22	57.9%
Cities from 10,000 thru 24,999	31	18	58.1%
Cities from 2,500 thru 9,999	57	31	54.4%
Cities under 2,500	63	32	50.8%
MSA Counties	77	42	54.5%
Non-MSA Counties	24	12	50.0%
State-level law enforcement	7	4	57.1%
<b>Total</b>	<b>400</b>	<b>210</b>	<b>52.5%</b>

## The Urgent Need for Change in Federal Criminal Justice Data

Starting in 2021, police data gathered by the FBI will almost certainly not be usable for basic crime trends and information. The FBI will not be able to report critical crime information such as how many murders occurred nationally or how many robberies will have taken place in New York City in 2021. Any change to gunshot injury reporting is rendered meaningless with incredibly low rates of adoption. This mismanagement of federal crime data collection underlines the urgent need for change in federal police data collection, governance, and roles.



## National Crime Statistics Should Be Handled by a Well-Resourced Statistical Agency

Federal statistical agencies are agencies within the federal government that disseminate statistical information for use by a variety of actors including governments, businesses, researchers, and the public. The federal government has 13 statistical agencies, including BJS. Federal statistical agencies “must provide objective, accurate, and timely information that is relevant to important public policy issues.”<sup>32</sup> In the 2018 National Academy of Sciences (NAS) Report “Modernizing Crime Statistics,” a primary conclusion was that federal crime data should be managed and gathered by an agency, or agencies, that act in the spirit of a federal statistical agency. It is important to note that the FBI is not a federal statistical agency, nor does its crime data collection adhere to a federal statistical agency’s fundamental role. Starting in 2021, the FBI will not be able to provide accurate or timely information relevant to measuring crime.

The 2018 NAS report concludes that, legally, the BJS “has the proper scope and range already built into its legislation” to coordinate and provide governance of crime data collections. Scholars have determined that “moving the UCR program to the BJS is a necessary but not sufficient condition to upgrade the nation’s crime monitoring capabilities.”<sup>33</sup> The largest drawback for the BJS is that it is a federal statistical agency that has never been funded in line with the duties it is tasked to perform.<sup>27</sup>

## The Importance of Accurate and Timely Police Data

Police data gathered by the FBI has historically contained problematic inaccuracies. In part, organizational culture may play a role in determining the responsibility the agency takes in providing accurate data. A former FBI director outlined the UCR role as: “we collect, we announce, we pass on; we do not analyze.”<sup>33</sup> Lack of emphasis on data quality and analysis is unlikely to resolve important problems in crime data quality, particularly in the context of new challenges posed by the incomplete NIBRS sample. The FBI’s county-level UCR SRS files “cannot be used with any degree of confidence” and require external development of algorithms for estimating county-level crime counts due to the erratic nature of county-level crime reporting that the FBI publishes.<sup>34</sup> , FBI procedures to correct for these missing data problems do not take into account basic, important information such as seasonality or demographic composition.<sup>35</sup> The FBI does implement some basic procedures to correct for problems with crime data reporting, focusing on “reasonableness” of data submissions and logical consistency.<sup>36</sup> However, these data quality checks often are only for outliers and fail to measure actual data quality. For instance, agencies report the ability to enter “NULL” into data fields to bypass cumbersome data quality checks.<sup>27</sup> Agencies using NIBRS report an “all or nothing” process whereby NIBRS incidents may be rejected for minor data entry inconsistencies. It is therefore not surprising that research finds that the FBI’s UCR SRS data do not match both what other national sources report or what states themselves report to the FBI.<sup>37,38</sup>

In addition to lack of data quality, supplying timely data is an additional barrier to informing crime policy. A 2008 National Academies of Science Workshop report considered the federal response to the 2005-2006 spikes in violent crime. Without timely knowledge of crime data, the Department of Justice dispatched “auditors” to examine local crime records across the nation due to lack of timely data. In comparison to other important federal data such as employment or housing, “the nation lacks timely information and comprehensive research on crime trends.”<sup>39</sup>

## NIBRS is a Flawed System That Should Be an Intermediary to a Better System

The NIBRS transition has been stagnant for many reasons. Primary among them is that NIBRS is a flawed data collection system whose usefulness is not always readily apparent to law enforcement agencies tasked with adopting it.

Another barrier to the transition to NIBRS is that switching to incident-based reporting by 2021 is essentially a mandate that came with little or no funding for many agencies. While NCS-X agencies received targeted funding, this funding came in the form of federal grants. Federal grants are onerous and generally mean larger agencies may be more able to apply for these dollars—and less likely to need them. In 2019, BJS made available \$21 million to support 29 agencies in adoption of NCS-X.<sup>40</sup> In contrast to the lack of funding for police agencies to switch crime reporting, over six years starting in 2010, hospitals received over \$30 billion in federal incentives to help acquire and use health information technology.<sup>41</sup>

Law enforcement agencies are willing and able to report crime data—high voluntary UCR SRS participation underlines agency ability and prioritization. The onerous and complicated NIBRS format may be only a medium-term solution to the problem of comprehensive, timely, and reliable national crime statistics. NIBRS data feature up to 99 victims and offenders each and have “intricately annoying victim-offender relationships between each possible pair.”<sup>27</sup> NIBRS data do not make for easy analysis. These core problems drive the 2018 NAP’s conclusion that NIBRS is an intermediate step to a modern crime incident reporting system that records crime attributes for more streamlined and simple data reporting.

## Conclusion and Recommendations

Law enforcement agencies across the nation are crucial to capturing data on gun violence. Police data have immense value for research and analysis of gun violence including assaults with a gun and gunshot assaults. For this reason, police data provide unique insight into gun violence, which can include invaluable data on nonfatal shooting victims, community violence, police shootings, and domestic violence.

Any comprehensive surveillance system of gun violence and victims should include both police data and medical data. Medical reports are valuable because they include all types of injuries from guns, such as self-inflicted gunshot wounds and accidental shooting victims. However,

police data may be more accurate and comprehensive in cases of assault. Further, police data are invaluable for measuring gun violence that may not leave a physical wound but that does lasting harm and reduces public safety, in turn deeply affecting individuals and communities.

## Priorities and Findings for Improving Police Data Surveillance of Gun Violence

### Measuring Gun Violence in the U.S.

- *Finding 1: Research is needed into the comprehensiveness of and overlap between police and ED data measures of gun violence.*
- *Finding 2: Little research explores mandatory reporting law compliance for gunshot wounds among police and emergency departments.*
- *Finding 3: NIBRS data—and in particular, proposed NIBRS improvements in measuring gun discharges and gunshot wounds—are promising improvements in better measurement of gun violence in police data.*
- *Finding 4: Even with proposed NIBRS changes, police shootings (or legal intervention) will not be comprehensively captured as they are included only in the FBI Use of Force database with low levels of agency reporting.*

### The NIBRS Transition Jeopardizes Police Data Use

- *Finding 5: Slow NIBRS adoption hampers timely and accurate police data measures of all crime incidents in the U.S.—not just of gun violence.*
- *Finding 6: State UCR programs are essential to timely, reliable, and accurate national police data collections.*
- *Finding 7: State UCR programs that are unable to report their data to the FBI in 2021 will continue to gather data statewide, which may provide a pathway to more comprehensive 2021 crime data.*

*Recommendation 1: Fund the BJS to partner with state UCR Programs to gather and produce accurate 2021 crime estimates.*

*Recommendation 2: Launch systematic inquiry into low NIBRS adoption to inform future fixes.*

## The Urgent Need for Change in Police Data

- *Finding 8: Investments in data quality, accuracy, and sophistication are important components of improved police data measures of crime and gun violence.*
- *Finding 9: Prior research emphasizes that police data improvements noted in Finding 1 may be best made by a federal statistical agency such as BJS.*
- *Finding 10: A primary barrier to a larger BJS role in national police data is chronic federal underfunding.*
- *Finding 11: NIBRS improves measurement of gun violence, but slow adoption of its complexity indicates a need for developing more usable incident-based future reporting systems.*

*Recommendation 3: Due to the increasing statistical complexity of generating crime estimates, the BJS should assume estimation oversight with appropriate levels of funding.*

*Recommendation 4: Fund state UCR Programs to scale up NIBRS participation and implement future improvements.*

*Recommendation 5: Start development of a simpler, streamlined incident reporting system.*

1. Richard Rosenfeld, Ernesto Lopez. *Pandemic, Social Unrest, and Crime in U.S. Cities: November 2020 Update*. Council on Criminal Justice; 2020.
2. Richard Rosenfeld, Ernesto Lopez. *Pandemic, Social Unrest, and Crime in U.S. Cities: July 2021 Update*. Council on Criminal Justice; 2021.
3. Emily J. Hanson. *The National Incident-Based Reporting System (NIBRS): Benefits and Issues*. Congressional Research Service; 2021.
4. FBI Letter on NIBRS Transition. Federal Bureau of Investigation. Accessed August 5, 2021. <https://www.fbi.gov/file-repository/ucr/fbi-letter-on-nibrs-transition-071018.pdf/view>
5. FBI Crime Data Could Go Away For One In Four Police Agencies. Newsy. Accessed August 4, 2021. <https://www.newsy.com/stories/fbi-crime-data-could-go-away-for-one-in-four-police-agencies/>
6. Crime in the U.S. 2018. FBI. Accessed October 11, 2021. <https://ucr.fbi.gov/crime-in-the-u.s/2018/crime-in-the-u.s.-2018>
7. Fowler KA, Dahlberg LL, Haileyesus T, Annet JL. Firearm injuries in the United States. *Prev Med*. 2015;79:5-14. doi:10.1016/j.ypmed.2015.06.002
8. Houry D, Sachs CJ, Feldhaus KM, Linden J. Violence-inflicted injuries: reporting laws in the fifty states. *Ann Emerg Med*. 2002;39(1):56-60. doi:10.1067/mem.2002.117759
9. *Mandatory Reporting of Non-Accidental Injury Statutes*. Victim Rights Law Center; 2014. <http://4e5ae7d17e.nxcli.net/wp-content/uploads/2021/01/Mandatory-Reporting-of-Non-Accidental-Injury-Statutes-by-State.pdf>
10. Calcia MA, Bedi S, Howard LM, Lempp H, Oram S. Healthcare experiences of perpetrators of domestic violence and abuse: a systematic review and meta-synthesis. *BMJ Open*. 2021;11(5):e043183. doi:10.1136/bmjopen-2020-043183
11. Olive P. Intimate partner violence and clinical coding: issues with the use of the International Classification of Disease (ICD-10) in England. *J Health Serv Res Policy*. 2018;23(4):212-221. doi:10.1177/1355819618781413
12. Criminal Victimization, 2018. Bureau of Justice Statistics. Accessed September 21, 2021. <https://bjs.ojp.gov/library/publications/criminal-victimization-2018>
13. Cook PJ. The case of the missing victims: Gunshot woundings in the National Crime Survey. *J Quant Criminol*. 1985;1(1):91-102. doi:10.1007/BF01065250
14. Carr J, Doleac JL. *The Geography, Incidence, and Underreporting of Gun Violence: New Evidence Using Shotspotter Data*. Social Science Research Network; 2016. Accessed November 5, 2018. <https://papers.ssrn.com/abstract=2770506>
15. Police Violence and Citizen Crime Reporting in the Black Community - Matthew Desmond, Andrew V. Papachristos, David S. Kirk, 2016. Accessed August 6, 2021. <https://journals.sagepub.com/doi/abs/10.1177/0003122416663494>
16. Kidd RF, Chayet EF. Why Do Victims Fail to Report? The Psychology of Criminal Victimization. *J Soc Issues*. 1984;40(1):39-50. doi:10.1111/j.1540-4560.1984.tb01081.x
17. Kellermann AL, Rivara FP, Lee RK, et al. Injuries Due to Firearms in Three Cities. *N Engl J Med*. 1996;335(19):1438-1444. doi:10.1056/NEJM199611073351906
18. Magee LA, Ranney ML, Fortenberry JD, Rosenman M, Gharbi S, Wiehe SE. Identifying nonfatal firearm assault incidents through linking police data and clinical records: Cohort study in Indianapolis, Indiana, 2007-2016. *Prev Med*. 2021;149:106605.

doi:10.1016/j.yjmed.2021.106605

19. Hipple NK, Huebner BM, Lentz TS, McGarrell EF, O'Brien M. The Case for Studying Criminal Nonfatal Shootings: Evidence from Four Midwest Cities. *Justice Eval J.* 2020;3(1):94-113. doi:10.1080/24751979.2019.1689152
20. Police, researchers: To address gun violence, U.S. needs to track the number of people shot - Baltimore Sun. Accessed August 6, 2021. <https://www.baltimoresun.com/news/investigations/bs-md-ci-shoot-to-kill-statistics-20161226-story.html>
21. National Academies of Sciences E. *Modernizing Crime Statistics: Report 1: Defining and Classifying Crime.*; 2016. Accessed January 9, 2017. <https://www.nap.edu/catalog/23492/modernizing-crime-statistics-report-1-defining-and-classifying-crime>
22. Crime in the U.S. 2019. FBI. Accessed October 10, 2021. <https://ucr.fbi.gov/crime-in-the-u.s/2019/crime-in-the-u.s.-2019>
23. Michael D. Maltz. *Analysis of Missingness in UCR Crime Data.* U.S. Department of Justice <https://www.ncjrs.gov/pdffiles1/nij/grants/215343.pdf>
24. *Benefits of NIBRS Participation.* U.S. Department of Justice - Federal Bureau of Investigation; 2017. [https://ucr.fbi.gov/nibrs/2017/resource-pages/benefits-of-nibrs-participation\\_gen-res.pdf](https://ucr.fbi.gov/nibrs/2017/resource-pages/benefits-of-nibrs-participation_gen-res.pdf)
25. SRS to NIBRS: The Path to Better UCR Data. Federal Bureau of Investigation. Accessed August 4, 2021. <https://www.fbi.gov/services/cjis/cjis-link/srs-to-nibrs-the-path-to-better-ucr-data>
26. Edwards F, Esposito MH, Lee H. Risk of Police-Involved Death by Race/Ethnicity and Place, United States, 2012–2018. *Am J Public Health.* 2018;108(9):1241-1248. doi:10.2105/AJPH.2018.304559
27. National Academies of Sciences, Engineering, and Medicine. *Modernizing Crime Statistics: Report 2: New Systems for Measuring Crime.* (Lauritsen JL, Cork DL, eds.). The National Academies Press; 2018. doi:10.17226/25035
28. NIBRS Technical Documentation | Department of Public Safety. Accessed August 5, 2021. <https://www.dps.texas.gov/section/crime-records-service/nibrs-technical-documentation>
29. *Counting Crime in Michigan.* Michigan State Police - Criminal Justice Information Center; 2018. [https://www.michigan.gov/documents/msp/Counting\\_Crimes\\_Feb\\_2019\\_Website\\_655796\\_7.pdf](https://www.michigan.gov/documents/msp/Counting_Crimes_Feb_2019_Website_655796_7.pdf)
30. National Crime Statistics Exchange. Bureau of Justice Statistics. Accessed August 5, 2021. <https://bjs.ojp.gov/programs/national-crime-statistics-exchange>
31. NCS-X Sample Agencies as of August 4, 2021. [https://bjs.ojp.gov/sites/g/files/xyckuh236/files/media/document/NCS-X\\_Sample\\_Agencies.pdf](https://bjs.ojp.gov/sites/g/files/xyckuh236/files/media/document/NCS-X_Sample_Agencies.pdf)
32. *Principles and Practices for a Federal Statistical Agency: Seventh Edition.* doi:10.17226/25885
33. Rosenfeld R. Transfer the Uniform Crime Reporting Program from the FBI to the Bureau of Justice Statistics. *Criminol Public Policy.* 2007;6(4):825-833. doi:10.1111/j.1745-9133.2007.00473.x

34. Maltz MD, Targonski J. A note on the use of county-level UCR data. *J Quant Criminol.* 2002;18.
35. Maltz M. Can We Trust the FBI's Crime Estimation Procedures? *ASC Criminol.* Accessed October 11, 2021. [https://www.academia.edu/39057721/Can\\_We\\_Trust\\_the\\_FBIs\\_Crime\\_Estimation\\_Procedures](https://www.academia.edu/39057721/Can_We_Trust_the_FBIs_Crime_Estimation_Procedures)
36. Data Quality Guidelines. FBI. Accessed October 11, 2021. [https://ucr.fbi.gov/data\\_quality\\_guidelines](https://ucr.fbi.gov/data_quality_guidelines)
37. Lauritsen JL, Rezey ML, Heimer K. When Choice of Data Matters: Analyses of U.S. Crime Trends, 1973–2012. *J Quant Criminol.* 2016;3(32):335-355. doi:10.1007/s10940-015-9277-2
38. Comer BP, Jorgensen C, Carter D. Reported Crime Frequencies: A Statistical Comparison of State Crime Reports and the UCR. *Am J Crim Justice.* Published online July 9, 2021. doi:10.1007/s12103-021-09623-y
39. National Research Council. *Understanding Crime Trends: Workshop Report.* (Goldberger AS, Rosenfeld R, eds.). The National Academies Press; 2008. doi:10.17226/12472
40. FY 2019 NCS-X Implementation Assistance Program: Support for Local Law Enforcement Agencies. Bureau of Justice Statistics. Accessed October 11, 2021. <https://bjs.ojp.gov/funding/opportunities/bjs-2019-15103>
41. Adler-Milstein J, Jha AK. HITECH Act Drove Large Gains in Hospital Electronic Health Record Adoption. *Health Aff (Millwood).* 2017;36(8):1416-1422. doi:10.1377/hlthaff.2016.1651

# Chapter 4. Studying Firearm Fatalities Using the NVDRS

Steve Marshall, PhD | University of North Carolina

## Synopsis

The NVDRS is the nation's primary source of public health information on fatalities resulting from violence, including firearm-related fatalities. The purpose of this chapter is to briefly review the methodology of NVDRS, examine trends in the research produced using the NVDRS, and recommend strategies for increasing the quantity and quality of NVDRS-related research.

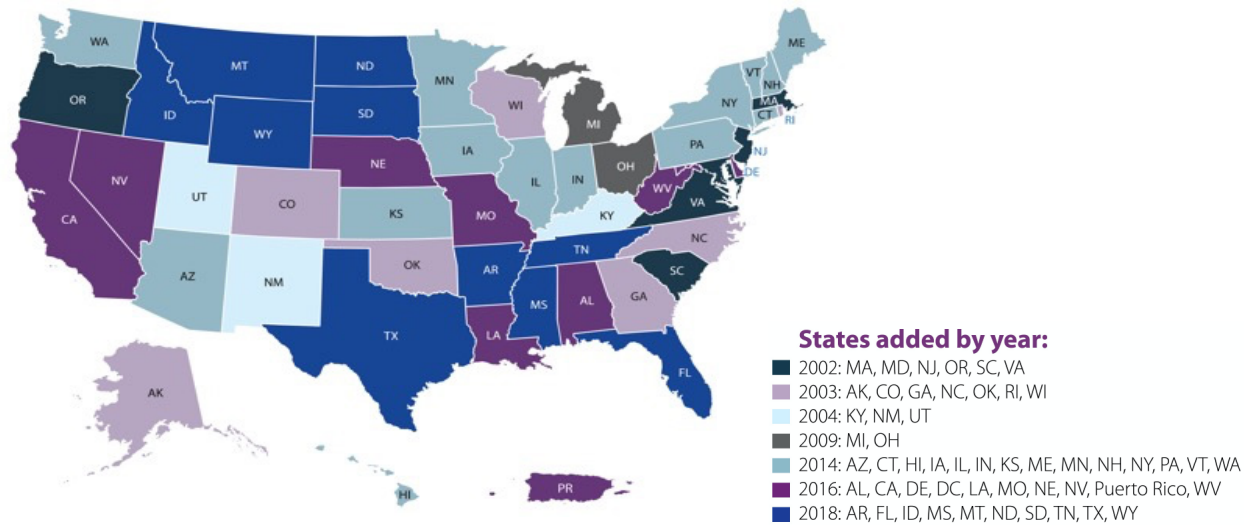
NVDRS is a mature surveillance system for deaths from violence. It provides researchers with a unified data resource that combines death investigation information from medical and legal systems into a single national database.<sup>1</sup> There has been a recent surge in the number of studies that use NVDRS data. This is a positive trend that should be fostered.

NVDRS is a secondary surveillance system that utilizes pre-existing records. For this reason, the timeliness and quality of NVDRS data depend on the rigor of local death investigation practices. Unfortunately, medical and legal death investigation practices vary considerably among states. Thus, NVDRS would benefit from standardized metrics that quantify state-to-state variation in data quality and timeliness. More fundamentally, the nation would benefit considerably from increased rigor and standardization of local practices and policies for medical and legal death investigators.<sup>2-6</sup>

This chapter concludes with recommendations for strategies to further strengthen NVDRS and motivate researchers to use this unique data resource. Specifically, we recommend further support for NVDRS to: 1) incentivize, support, and expand the use of NVDRS data by researchers; 2) develop and release quality metrics for the system; and 3) strengthen the death investigation infrastructure that underpins the NVDRS.



**Figure 1.** Growth of the NVDRS. Source: U.S. Centers for Disease Control and Prevention, NCIPC.



## History and Development of NVDRS

NVDRS is a key component in the nation’s public health data infrastructure. Over its two-decade history, it has evolved from an initial sixteen-state system (established between 2002 and 2004) to become a nationally system that is the primary source of firearm fatality data.<sup>1,7</sup> The system began in the late 1990s and early 2000s, when visionary researchers at the Harvard Injury Control Center, with support from a group of private foundations, commissioned a national group of experts to refine their concept of a national system for data on violent deaths.<sup>8</sup>

Similar to many other CDC systems, NVDRS is supported and implemented on a state-by-state basis.<sup>9</sup> NVDRS was launched in 2002 with funding for six states (**Figure 1**). A total of 18 states were added through expansions in 2003, 2004, and 2009, bringing the total to 24 states.<sup>8</sup> The horrific mass casualty event at Sandy Hook Elementary School in 2012 led to renewed calls to scale the system to cover the nation. By 2018, NVDRS support had been expanded to include all 50 states, plus Puerto Rico and the District of Columbia.<sup>10</sup>

## Overview of NVDRS Methodology

The NVDRS is a detailed and comprehensive source of data on firearm-related fatalities and other deaths from violence. The universe of deaths collected by NVDRS is comprised of all: a) homicides, b) suicides, c) deaths caused by law enforcement in the line of duty, and d) unintentional firearm deaths.<sup>7</sup> NVDRS complements the Bureau of Justice Statistic’s NCVS, which is limited to nonfatal events. Each NVDRS event report includes a large number (over 600) of relational data elements that provide context for each death. The details collected

include the weapons involved, the relationships between the people involved, and the circumstances of death such as recent conflicts, crises, or life stressors.<sup>7</sup>

To develop and implement prevention strategies for violence, it is critically important to understand the relationships between the people involved.<sup>1</sup> NVDRS uses a data architecture that captures the personal relationships in each casualty event and the weapons used. Multiple individuals who died in the same event, and the firearms and other weapons used in the event, are linked together in a single database. This innovation allows researchers to conduct analyses that illuminate the etiology of fatal firearm violence.<sup>10</sup> For example, events in which a man murders his intimate partner and then kills himself can be readily identified using NVDRS; it is unique in linking suicides and homicides that are part of the same incident.<sup>11-14</sup> No other national data source on fatal violence captures these important relationships between events, people, and weapons for fatal violence.<sup>1,6</sup>

A key aspect of NVDRS is that it is a secondary surveillance system. This means that it use pre-existing sources of data only; no primary data collection (such as contact with next of kin) is involved.<sup>9</sup> NVDRS uses data abstractors in each state to extract information from the pre-existing records generated by the death investigation systems in each state. Information includes data from vital statistics, law enforcement, and coroners and medical examiners. Data are combined by NVDRS data abstractors in a systematic manner to provide a more complete picture of each violent death than would be obtained from any source in isolation.<sup>7</sup>

One disadvantage to the secondary surveillance system approach is the drawn-out nature of some death investigations. Typically, the most recent NVDRS data available to researchers is at least 2 years old. The other main disadvantage is the variability inherent in death investigation systems among states. The impact of these issues, and strategies for addressing them, are discussed in detail below.

## State Variability in Death Investigation Systems

NVDRS combines pre-existing data sources and links them on a state-by-state basis, to a gap in our firearm violence data infrastructure<sup>9</sup> and provide a composite view of fatal violence across multiple medical and legal death investigation systems.<sup>7,15</sup> However, because it depends entirely on extant sources of information, NVDRS has brought into focus the limitations inherent in U.S. death investigation systems. These include the fragmented nature of our nation's medico-legal processes<sup>2</sup> and resultant heterogeneity in the quality of death investigation systems across the nation.<sup>4</sup>

Law enforcement response to an apparent death by violence varies considerably by local jurisdiction, the details of the event, and whether there is a perceived criminal intent. The priority of law enforcement is to establish the nature of the crime and, if relevant, apprehend perpetrators. Sharing data with public health professionals may vary by the perceived likelihood that a prosecution could be affected by such a collaboration. Variation by jurisdiction is

exacerbated by the fact that federal and state law enforcement typically adopt a weak stance to collaboration with other agencies and have done little to promote the integration of vital statistics and medical data into violence statistics.<sup>15</sup>

Medical response to violent death also varies considerably among states.<sup>5</sup> Prevailing medical death investigation practices in a specific locality depend on each state's requirements and current practices. They may also depend on the expertise and training of the certifying professional. In some rural areas, all death certifications may be done by only one or two providers. Most states require that local death investigations be conducted by medical professionals and have centralized state medical examiner systems that collate and review case reports. However, some states still permit elected non-medical professionals, known as coroners, to investigate and certify deaths. Irrespective of the medical background of the local examiner, investigation and certification of death is typically poorly compensated.<sup>3</sup>

To address these issues, CDC has implemented standard timelines for completing data acquisitions that are uniform across participating states. CDC has also implemented NVDRS quality control measures, including metrics to monitor data quality and blinded duplicate abstractions for a sample of cases in each state.<sup>7</sup> These are important and helpful advancements. However, national quality control procedures are not an antidote for the underlying problem of state-to-state variability in the quality of local death investigation systems.

As an example, a classic situation in medico-legal death investigation is the classification of a violent death as "Undetermined Intent," meaning that it is unclear if the death was a homicide or a suicide. Some deaths in this category result from areas of genuine uncertainty, such as a death in which a decedent appeared to deliberately provoke law enforcement agencies with apparent intent to be fatally wounded. Another ambiguity is a fatal single-vehicle crash that was preceded by repeated overt signs of suicide ideation and/or self-harm. Likewise, drug overdose deaths may reflect a complex etiology, sometimes including self-harm and crises, that can make determination of intent problematic.<sup>16-18</sup>

We acknowledge that there may be legitimate ambiguity in determination of intent. However, there is enormous state-to-state variability in the proportion of violent deaths that are classified by death investigators as belonging in the "Undetermined Intent" category. This variability is far too large to reflect regional variation in the incidence of truly ambiguous cases, such as those discussed in the preceding paragraph. Rather, it reveals the gross inconsistencies in death investigation practices among the states.

To illustrate how state-to-state variability in death investigation practices affects the NVDRS, we analyzed the proportion of "Undetermined Intent" deaths by state. We used 2018 NVDRS data, the most recent year NVDRS data are available. Our analysis used aggregated counts and was conducted using the NVDRS data publicly available on the CDC's WISQARS, an interactive user interface that provides customized reports of injury-related data. WISQARS enables the public to run custom table requests against NVDRS data. For more complex analyses, CDC

makes available a restricted access data (RAD) file, which can be obtained upon CDC receipt of a data request form.

CDC suppress small counts and totals based on small counts, to prevent deductive disclosure. For this reason, we defined a measure specific to this analysis, the “Undetermined Intent Ratio” (UIR), defined as 90 times the ratio of Undetermined Intent Deaths to the sum of the total homicides and suicides in each state. The sum of total homicides and suicides excludes unintentional firearm deaths, suicides preceded by a homicide, and legal intervention deaths, which are prone to CDC suppression for small counts. For most states, the UIR roughly approximates the proportion of violent deaths classified as undetermined intent.

**Table 1.** Analysis of Undetermined Intent Deaths by Jurisdiction, NVDRS 2018

UIR in the range 0.0 to 9.9	Alabama, Arizona, California, Colorado, Connecticut, Georgia, Illinois, Iowa, Kansas, Kentucky, Louisiana, Maine, Massachusetts, Minnesota, Missouri, Nevada, New Jersey, North Carolina, Ohio, Oklahoma, Oregon, Pennsylvania, South Carolina, Utah, Virginia, Washington, West Virginia, Wisconsin	n=28
UIR of 10.0 or greater; fewer than 100 Undetermined Intent Deaths	Alaska, Rhode Island, Vermont	n=3
UIR of 10.0 or greater; 100 or more Undetermined Intent Deaths	Indiana, Maryland, Michigan, New York	n=4
<i>UIR not computed due to small numbers (&lt;10 Undetermined Intent Deaths in 2018)</i>	<i>Delaware, District of Columbia, Nebraska, New Hampshire, New Mexico</i>	<i>n=5</i>
<i>Data not yet available via WISQARS (e.g. not funded to collect data in 2018)</i>	<i>Arkansas, Florida, Hawaii, Idaho, Mississippi, Montana, North Dakota, Puerto Rico, South Dakota, Tennessee, Texas, Wyoming</i>	<i>n=12</i>

*UIR is the ratio of Undetermined Intent Deaths to Total Homicides and Suicides Combined, multiplied by 90. Total homicides and suicides combined excludes unintentional firearm deaths, suicides preceded by a homicide, and legal intervention deaths.*

In 2018, the NVDRS data available in WISQARS classified 8.8 percent of violent deaths as Undetermined Intent, which is high, but not untenable, from a research standpoint. However, there is considerable state-to-state variability, and it is notable that most deaths classified as Undetermined Intent come from a handful of “outlier” states.

Most of the states reporting to NVDRS had an acceptable proportion of violent deaths classified as Undetermined Intent (**Table 1**) as evidenced by a UIR below 10 (n=28 states). However, seven states had a UIR above 10, and four of these (Indiana, Maryland, Michigan, and New York) were large states with more than 100 violent deaths in 2018 classified as Undetermined Intent. These four states accounted for 58 percent of all Undetermined Intent Deaths recorded by the entire NVDRS system in 2018. Maryland alone (less than 2 percent of the U.S. population) accounts for over one-third of all Undetermined Intent Deaths in NVDRS. This is a frustrating situation, given that Maryland was one of the first states funded by NVDRS, and concerns about the Maryland data were noted as early as the mid-2000s.<sup>5</sup>

This example demonstrates that there is substantial local variability in the death investigation practices that generate the information that NVDRS uses. Clearly, it is hard to generate prevention strategies in the absence of basic information on Intent.

At time of writing, a total of 11 states and Puerto Rico did not have retrievable 2018 data in the WISQARS system (**Table 1**). This includes the states that were funded for the first time in 2018 and includes the large states of Florida and Texas, which are notable for their lack of centralized resources.

## Strategies for Improving Death Investigation Systems

Over the past two decades, there have been repeated calls for progress towards strengthening the national standards for medico-legal death investigation, for example, in reports published by the National Research Council and the Institute of Medicine.<sup>3,4</sup> Two recent commentaries have called for improvements in national death investigation systems specifically in the context of NVDRS.<sup>6,15</sup> Kaplan et al.<sup>6</sup> advocate for improvements in death investigation systems so that all deaths from violence are “investigated, evaluated, and certified in the same way nationwide” with the goal of ensuring consistent NVDRS data.<sup>6</sup> They note that “NVDRS is limited by the assortment of medico-legal death investigation systems in the participating states,” and lament “extreme variations” due to the “idiosyncratic assumptions” of some death certifiers.<sup>6</sup> Mays and Cochran<sup>15</sup> draw attention to decades of underinvestment in local death systems and stress the potential for technological innovations that would improve timeliness of local and national data systems.<sup>15</sup> They point out that investments in technology, and other improvements in our death data systems, have the potential to yield “near real-time system of mortality surveillance,” thereby helping vital statistics fulfill “its public health function of being an early detector of emerging public health epidemics.”<sup>15</sup> For example, in an ideal world, a system such as NVDRS would be able to detect, in a very timely manner, upticks in specific types of fatal violence because of macro forces such as the COVID-19 pandemic. Recommendation 2 presents a strategy for addressing the issue of timeliness (see below).

To date, the resources that have been channeled into establishing and growing the NVDRS have been public health resources.<sup>6,10,19-22</sup> In order for NVDRS to reach its full potential as a comprehensive and timely source for national statistics on violent deaths, our national and local law enforcement, justice, and medical examiner systems must commit to NVDRS’s goals. This will involve a commitment to reforming local medico-legal death investigation procedures that are currently highly variable, fragmentary, and lack adherence to national standards.<sup>2-4</sup> Recommendation 5 presents specific strategies that would advance this goal (see below).

The development of NVDRS has influenced local death investigators to become more rigorous over time. For example, a detailed death investigation involving a firearm suicide typically would be of limited interest to law enforcement, beyond establishing that there was no perpetrator to apprehend. In some NVDRS states, such as North Carolina, the connection between public health and law enforcement created by NVDRS has meant that local law enforcement has

become more willing to collect and share data, once the importance of complete and timely reporting of violent deaths is fully understood. Ultimately, improvements in our medico-legal death investigation systems must occur at both the national and the local level.

## Published Research Using NVDRS Data

A small literature review was conducted to understand how NVDRS is being used by researchers and to examine trends in use of the data. We searched PubMed using the terms “National Violent Death Reporting System” or “NVDRS” and reviewed all identified articles that were published prior to December 31, 2020. Published letters to journals were excluded unless they pertained directly to a published study in that journal (e.g., Letters to the Editor following a publication). This search located 223 articles. While not intended to be comprehensive, this small search was considered appropriate for analysis of trends and to characterize the extant NVDRS-related literature.

We categorized the publications retrieved in our search based on the presence of one or more of the following three attributes:

- Publications that used NVDRS data to publish original research findings. This group comprises both studies that used data from a single state and studies that used CDC’s national RAD dataset, which CDC makes available to legitimate researchers (n=192).
- Methodologic research directly pertinent to NVDRS, such as evaluations of the system or details of statistical methods developed for use with NVDRS data (n=18).
- Commentaries, reviews, or editorials that described or publicized NVDRS, reviewed the methods and/or scope of NVDRS, or described and/or reviewed literature that used NVDRS data (n=17).

The three groups were exhaustive but non-exclusive (i.e., a study could be in more than one group). For example, a study that developed a new method, and applied it for the purposes of making scientific inference or descriptive analyses, would be counted in both group 1 and group 2 but would be counted only once in Figure 2.

Studies using NVDRS data have greatly deepened our scientific knowledge of violence. As noted above, NVDRS is unique among mortality databases because it captures data on the relationships between the people involved in the violence. Therefore, researchers have used the data to examine homicides resulting from intimate partner violence and situations in which homicides and suicides are linked.<sup>11,13,23-31</sup> NVDRS is a unique data resource for such analyses.

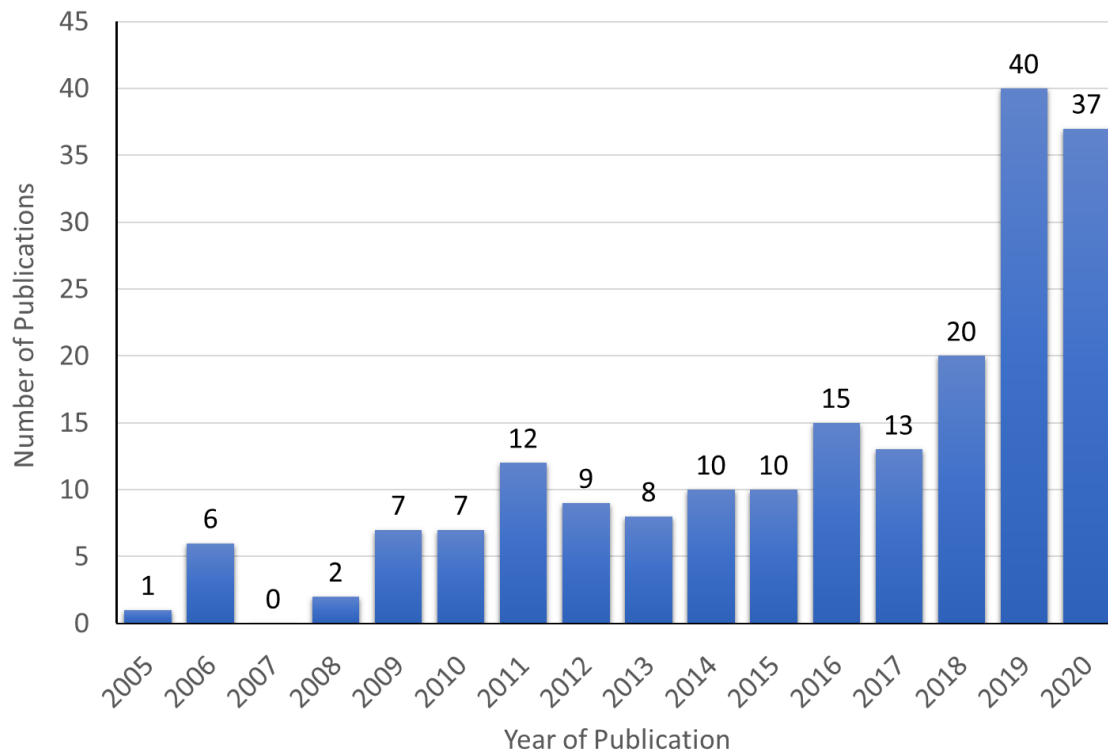
In addition, the large (and ever-growing) sample size of NVDRS has provided statistical power to examine violent deaths by subgroup, facilitating examination of violence among demographic subgroups such as children,<sup>23,32-36</sup> adolescents,<sup>11,24,36-38</sup> women,<sup>29,30,39-45</sup> older adults,<sup>46-53</sup> Hispanics,<sup>54-56</sup> LGBTQ+ groups,<sup>57-59</sup> and military/veterans.<sup>60-69</sup> Researchers have also used

NVDRS data to explore both the factors involved in homicides of civilians committed by law enforcement officers<sup>70-73</sup> and to examine homicides sustained by law enforcement officers.<sup>74</sup>

The majority of violent deaths in the U.S. involve firearms, and NVDRS captures details on all weapons involved. For this reason, an additional benefit of NVDRS has been the ability to characterize the toll of firearm violence in detail.<sup>13,33,38,62,75-82</sup> Of the articles that published original research findings using NVDRS data (n=192), a total of 83 (43 percent) mentioned firearms in the title, abstract, or keywords. Further, researchers have used NVDRS data to examine the impact of state-level firearm legislation on violent deaths.<sup>14,83</sup>

The number of research articles using NVDRS data has increased over time, with a steep surge in recent years (**Figure 2**). A preliminary review of 2021 publications indicates that the total number is likely to exceed that for 2020. The reasons for the recent surge in publications are not fully known but likely relate to the steady accrual of case numbers in the system over time and the increasing representativeness of the system as the number of states contributing data has grown. In addition, there is a growing understanding in the research community that some of the problems inherent in violence require population-based research using data sources such as NVDRS. There is increasing concern over the toll of violence in the U.S., so all these factors will likely continue to intensify over time. Now is an important time to implement strategies to support (and further increase) the use of NVDRS data by researchers. Recommendation 1 presents strategies for addressing this topic (see below).

**Figure 2.** Publications Using NVDRS Data



## Methodologic Research

As noted above, this review identified 18 papers discussing methodologic aspects of NVDRS and/or the data generated from the system. However, relative to other public health surveillance systems, there is little information on data quality in NVDRS. The literature (and extensive CDC documentation) lacks information on data quality measures. Specifically, there is little methodological research into statistical strategies to address known deficiencies and issues in the death investigation systems that underlie NVDRS.

Other large-scale federal surveillance efforts of similar scope have been able to invest in the development and application of advanced statistical methods. One point of comparison is the Fatal Analysis Reporting System (FARS), the national surveillance system for fatal motor vehicle crashes. FARS makes extensive use of the statistical technique of imputation to address known problems, such as incomplete alcohol and drug testing data from some states. The initial development of statistical methods to apply this technique to FARS alcohol data was supported by the federal agency that operates FARS (the National Highway Traffic Safety Administration).<sup>84,85</sup>

The NVDRS methodologic literature does not include contributions about using advanced statistical methods to address the timeliness and data quality issues that stem from the patchwork nature of death investigation systems in the U.S. As an example of the type of research that is needed, consider the problem of suicides from drug overdose that are misclassified under other causes, such as Undetermined Intent.<sup>5,17,18,86</sup> A recent methodologic paper<sup>87</sup> developed a statistical method to address this problem. Liu et al.<sup>87</sup> used three years of data from Utah to train a machine learning algorithm to recognize suicides from drug overdoses that were misclassified as Intent Undetermined deaths.<sup>87</sup> They estimated that drug overdose suicides in Utah are underreported by up to 33 percent, resulting in an overall undercount of all suicides in Utah by nine percent.<sup>87</sup>

Machine learning techniques, such as those used by Liu et al.<sup>87</sup> may be applied to NVDRS data in many ways. For example, they have potential to address the problem of outlier states (Indiana, Maryland, Michigan, New York) that are responsible for the majority of “Intent Undetermined” deaths nationally (**Table 1**). Another example is the potential to use machine learning techniques to facilitate release of timely provisional aggregate statistics on deaths from violence. This would allow NVDRS to generate sentinel advance reports to provide policymakers and researchers with timely provisional information on trends. This cost-efficient short term strategy would complement the long-term strategies of reforming local medical-legal processes across the nation. Recommendations 3 and 4 present specific strategies to advance methodologic research that would address data quality issues that result from variability in our nation’s death investigation systems (see below).



## Recommendations for Increasing the Use of NVDRS by Researchers

This section provides a set of recommendations for increasing the quantity, quality, and impact of scientific research that uses NVDRS data. These recommendations contain strategies to strengthen the rigor, sophistication, and timeliness of research using NVDRS data.

### Recommendation 1: Increase Support for Researchers Who Use NVDRS Data

It is important to ensure that there is a large and growing number of studies that use NVDRS data. To support an expanding pipeline of high-quality research products, we recommend establishing a national NVDRS research consortium dedicated to advancing the use of NVDRS data by the scientific community. The consortium would encourage new researchers to use NVDRS and assist existing NVDRS users in discovering new aspects of this unique data resource. In addition, the national consortium would be a forum to disseminate advanced statistical methods relevant to NVDRS data and innovative uses of NVDRS data. Some potential activities for this consortium are listed in **Table 2**.

Fortunately, CDC has created a solid foundation for expanded efforts in this area. The agency has invested considerable efforts to make NVDRS data resources accessible to researchers. (When the author contacted CDC regarding access to the data, he was provided with a comprehensive and detailed set of data documents in less than 15 minutes). CDC also deserves much credit for developing and presenting learning academy workshops at national forums, such as the American Public Health Association. In addition, some states have developed regional workshops to showcase their state's NVDRS to local researchers. There is considerable scope for scaling up these efforts so that they have greater impact and build a large and diverse community of NVDRS scholars.

**Table 2.** Potential Activities of Proposed NVDRS Research Consortium

- Maintain a list of published studies using NVDRS data
- Identify key research questions that are amenable to NVDRS data
- Maintain a list of databases that may be combined with NVDRS data to answer novel research questions
- Plan an annual NVDRS research conference to build linkages between researchers
- Provide technical support to emerging researchers, such as doctoral students and junior faculty, in using NVDRS data
- Provide a forum for discussing advanced analytical approaches to NVDRS data

## Recommendation 2: Improve Timeliness by Providing Provisional Data Releases

Violence is a dynamic and ever-evolving problem in the U.S. For example, the COVID-19 pandemic saw an escalation of firearm sales and an expansion of the demographic subgroups that typically purchase firearms. One concern is the modest pace at which detailed homicide and suicide data become available to researchers. NVDRS data files are typically finalized and made available to the researchers approximately two years after the violence occurred. The ability of researchers to monitor and respond to change is hampered by the time it takes to finalize NVDRS data. CDC has invested considerable efforts to ensure that complete data are acquired and finalized as rapidly as possible; the lag reflects the slow pace of local criminal justice and medico-legal death investigation infrastructure. CDC should provide more timely interim data releases of portions of the RAD file, as they become available. Interim releases could be limited to a subset of variables (i.e., the variables that are completed more rapidly by all states) or a subset of states (i.e., those geographic locations that generate high-quality data in the most rapid manner). If interim data releases were available, external researchers would be able to conduct data analyses limited to these variables or regions, noting that such work may not be nationally representative. Such a limitation is acceptable in exchange for more timely data, knowing that more complete datasets would become available in the future. As noted above, special statistical techniques, such as inverse probability weights, could be employed to address a lack of representativeness.

## Recommendation 3: Develop and Release Indicators of NVDRS Data Quality

Researchers who use NVDRS data receive very limited information about data quality. CDC has implemented *internal* quality control measures, including a quality control dashboard and re-abstractions to monitor data quality.<sup>7</sup> However, indicators of data quality are not included on the files released to external researchers. Markers of data quality, by state and time period, should be derived by CDC and made available to researchers who use the RAD files. Such measures could be based on metrics such as the proportion of violent and firearm deaths classified as Intention Undetermined in each state (moving average over a three-year period), and/or completeness of the circumstances variables for suicides and homicides in a state (moving average over a three-year period). It takes time to establish NVDRS within a state, so such measures should be reported only for states that have moved beyond the initial phase of system setup (e.g., states that have been funded for three or more years). Further, the scientific literature contains only one formal evaluation of a state NVDRS system.<sup>88</sup> There is a need to conduct and publish more evaluations of the system, either nationally or state-by-state, as peer-reviewed manuscripts or technical reports.

## Recommendation 4: Fund NVDRS Methodologic Research

Other large federal data collection systems, such as FARS (fatal motor vehicle crashes), have utilized statistical techniques such as multiple imputation to address data quality issues in their systems. Funding should be provided to researchers, via competitive mechanisms such as federal grants, to conduct methodological investigations into the use of advanced statistical techniques to address known limitations of the local death investigation systems that underlie NVDRS, such as data timeliness and variation in data quality. These techniques could include inverse probability weights, multiple imputation, and machine learning.

## Recommendation 5: Strengthen the Death Investigation Systems That Underlie NVDRS

As noted above, NVDRS depends solely on pre-existing administrative records, such as law enforcement reports, medical examiner reports, and death certificates. Thus, the timeliness and quality of NVDRS data depends heavily on the local criminal justice and medico-legal death investigation systems that generate the records and reports abstracted by NVDRS coders. The recent advent of a nationwide NVDRS system has highlighted the variations in quality due to the lack of standardized procedures for local criminal justice and medico-legal death investigation. These systems include medical examiners' and coroners' reports, law enforcement investigations, justice procedures, and vital statistics registration. Many of these systems take considerable time to assemble complete data. Simply put, our nation's death investigation systems are a patchwork quilt. The completeness, timeliness, and variability of local criminal justice and medico-legal death investigation systems should be fully documented through an independent review conducted by a national agency such as NAS. Such a review should detail the impact of these local variations in procedures on NVDRS and similar national systems. It should also provide an opportunity to make recommendations about improving the uniformity and rigor of criminal justice and medico-legal death investigations. Identification of the localities with the most rigorous procedures would inform future investments in future data infrastructure and suggest the most promising locations to build surveillance systems for nonfatal violence.

## Summary

NVDRS data is an invaluable source of data on deaths from violence. It is a unique resource that combines data from multiple sources into a comprehensive data resource that now includes all 50 states, plus the District of Columbia and Puerto Rico. The growth of NVDRS over the past two decades has led to an increase in number of published studies that use NVDRS data. There is ample scope for continued scaling up in the use of these data by researchers, and we provide strategies to facilitate that goal (Recommendation 1).

Because it relies on existing data sources, NVDRS depends on the timeliness and rigor of local death investigation practices, which vary markedly among states. To address data quality and timeliness, NVDRS should implement interim data releases of provisional statistics and data (Recommendation 2), the publication of standardized metrics to quantify state-to-state variation in data quality (Recommendation 3), and the development of statistical procedures to address the limitations inherent in our nation's death investigation systems (Recommendation 4). More fundamentally, both NVDRS and the nation would benefit from measures to strengthen the death investigation infrastructure and standardize it across jurisdictions (Recommendation 5).

## Literature Cited

1. Barber C, Azrael D, Hemenway D. A truly national National Violent Death Reporting System. *Inj Prev* 2013;19:225-6.
2. Ruiz L, Posey BM, Neully MA, Stohr MK, Hemmens C. Certifying Death in the United States. *J Forensic Sci.* 2018;63(4):1138-1145.
3. National Research Council. Strengthening Forensic Science in the United States: A Path Forward. Washington, DC: The National Academies Press; 2009.
4. Institute of Medicine. Medicolegal Death Investigation System: Workshop Summary. Washington, DC: The National Academies Press; 2003.
5. Breiding MJ, Wiersema B. Variability of undetermined manner of death classification in the US. *Inj Prev* 2006;12 Suppl 2:ii49-ii54.
6. Kaplan MS, Caetano R, Giesbrecht N, et al. The National Violent Death Reporting System: Use of the Restricted Access Database and Recommendations for the System's Improvement. *Am J Prev Med* 2017;53:130-3.
7. Blair JM, Fowler KA, Jack SP, Crosby AE. The National Violent Death Reporting System: overview and future directions. *Inj Prev* 2016;22 Suppl 1:i6-11.
8. Hemenway D, Barber CW, Gallagher SS, Azrael DR. Creating a National Violent Death Reporting System: a successful beginning. *Am J Prev Med* 2009;37:68-71.
9. Paulozzi LJ, Mercy J, Frazier L, Jr., Annett JL. CDC's National Violent Death Reporting System: background and methodology. *Inj Prev* 2004;10:47-52.
10. Nazarov O, Guan J, Chihuri S, Li G. Research utility of the National Violent Death Reporting System: a scoping review. *Inj Epidemiol* 2019;6:18.
11. Adhia A, Kernic MA, Hemenway D, Vavilala MS, Rivara FP. Intimate Partner Homicide of Adolescents. *JAMA Pediatr* 2019;173:571-7.
12. Logan J, Hill HA, Black ML, et al. Characteristics of perpetrators in homicide-followed-by-suicide incidents: National Violent Death Reporting System--17 US States, 2003-2005. *Am J Epidemiol* 2008;168:1056-64.
13. Schwab-Reese LM, Peek-Asa C. Factors contributing to homicide-suicide: differences between firearm and non-firearm deaths. *J Behav Med* 2019;42:681-90.
14. Sivaraman JJ, Ranapurwala SI, Moracco KE, Marshall SW. Association of State Firearm Legislation With Female Intimate Partner Homicide. *Am J Prev Med* 2019;56:125-33.
15. Mays VM, Cochran SD. Challenges and Opportunities for Modernizing the National Violent Death Reporting System. *Am J Public Health* 2019;109:192-4.
16. Rockett IR, Kapusta ND, Coben JH. Beyond suicide: action needed to improve self-injury mortality accounting. *JAMA Psychiatry.* 2014 Mar;71(3):231-2.

17. Rockett IAH, Caine ED, Connery HS, et al. Unrecognised self-injury mortality (SIM) trends among racial/ethnic minorities and women in the USA. *Inj Prev*. 2020 Oct;26(5):439-447.
18. Rockett IRH, Caine ED. Reconciling Suicides With "Accidental" Drug-Intoxication Deaths: A Behaviorally Based Definition of Self-Injury Mortality. *Am J Public Health*. 2020 Jan;110(1):73-74.
19. Butchart A. The National Violent Death Reporting System: a new gold standard for the surveillance of violence related deaths? *Inj Prev* 2006;12 Suppl 2:ii63-ii4.
20. Campbell R, Weis MA, Millet L, Powell V, Hull-Jilly D, Hackman H. From surveillance to action: early gains from the National Violent Death Reporting System. *Inj Prev* 2006;12 Suppl 2:ii6-ii9.
21. Crosby AE, Mercy JA, Houry D. The National Violent Death Reporting System: Past, Present, and Future. *Am J Prev Med* 2016;51:S169-s72.
22. Powell V, Barber CW, Hedegaard H, et al. Using NVDRS data for suicide prevention: promising practices in seven states. *Inj Prev* 2006;12 Suppl 2:ii28-ii32.
23. Adhia A, Austin SB, Fitzmaurice GM, Hemenway D. The Role of Intimate Partner Violence in Homicides of Children Aged 2-14 Years. *Am J Prev Med* 2019;56:38-46.
24. Adhia A, DeCou CR, Huppert T, Ayyagari R. Murder-Suicides Perpetrated by Adolescents: Findings from the National Violent Death Reporting System. *Suicide Life Threat Behav* 2020;50:534-44.
25. Bossarte RM, Simon TR, Barker L. Characteristics of homicide followed by suicide incidents in multiple states, 2003-04. *Inj Prev* 2006;12 Suppl 2:ii33-ii8.
26. Fraga Rizo C, Mennicke A, Van Deinse T. Characteristics and Factors Associated With Intimate Partner Violence-Related Homicide Post-Release From Jail or Prison. *J Interpers Violence* 2019;886260519888195.
27. Holland KM, Brown SV, Hall JE, Logan JE. Circumstances Preceding Homicide-Suicides Involving Child Victims: A Qualitative Analysis. *J Interpers Violence* 2018;33:379-401.
28. Logan JE, Ertl A, Bossarte R. Correlates of Intimate Partner Homicide among Male Suicide Decedents with Known Intimate Partner Problems. *Suicide Life Threat Behav* 2019;49:1693-706.
29. Palladino CL, Singh V, Campbell J, Flynn H, Gold KJ. Homicide and suicide during the perinatal period: findings from the National Violent Death Reporting System. *Obstet Gynecol* 2011;118:1056-63.
30. Reckdenwald A, Szalewski A, Yohros A. Place, Injury Patterns, and Female-Victim Intimate Partner Homicide. *Violence Against Women* 2019;25:654-76.
31. Schwab-Reese LM, Murfree L, Coppola EC, Liu PJ, Hunter AA. Homicide-suicide across the lifespan: a mixed methods examination of factors contributing to older adult perpetration. *Aging Ment Health* 2021 Sep;25(9):1750-1758.

32. Bennett MD, Jr., Hall J, Frazier L, Jr., Patel N, Barker L, Shaw K. Homicide of children aged 0-4 years, 2003-04: results from the National Violent Death Reporting System. *Inj Prev* 2006;12 Suppl 2:ii39-ii43.
33. Hemenway D, Solnick SJ. Children and unintentional firearm death. *Inj Epidemiol* 2015;2:26.
34. Hemenway D, Solnick SJ. The epidemiology of homicide perpetration by children. *Inj Epidemiol* 2017;4:5.
35. Klevens J, Leeb RT. Child maltreatment fatalities in children under 5: Findings from the National Violence Death Reporting System. *Child Abuse Negl* 2010;34:262-6.
36. Sheftall AH, Asti L, Horowitz LM, et al. Suicide in Elementary School-Aged Children and Early Adolescents. *Pediatrics* 2016;138(4):e20160436.
37. Bush AM. A multi-state examination of the victims of fatal adolescent intimate partner violence, 2011-2015. *J Inj Violence Res* 2020;12:73-83.
38. Johnson RM, Barber C, Azrael D, Clark DE, Hemenway D. Who are the owners of firearms used in adolescent suicides? *Suicide Life Threat Behav* 2010;40:609-11.
39. Carmichael H, Steward L, Velopulos CG. It doesn't just happen to "Other" people - An exploration of occupation and education level of women who die from intimate partner violence. *Am J Surg* 2019;218:744-8.
40. Choi NG, DiNitto DM, Sagna AO, Marti CN. Older women who died by suicide: suicide means, sociodemographic and psychiatric risk factors, and other precipitating circumstances. *Int Psychogeriatr* 2018;30:1531-40.
41. Huguet N, McFarland BH, Kaplan MS. A Comparison of Suicides and Undetermined Deaths by Poisoning among Women: An Analysis of the National Violent Death Reporting System. *Arch Suicide Res* 2015;19:190-201.
42. Ivey-Stephenson AZ, Blair JM, Crosby AE. Efforts and Opportunities to Understand Women's Mortality Due to Suicide and Homicide Using the National Violent Death Reporting System. *J Womens Health (Larchmt)* 2018;27:1073-81.
43. Ortega LA, Karch D. Precipitating circumstances of suicide among women of reproductive age in 16 U.S. States, 2003-2007. *J Womens Health (Larchmt)* 2010;19:5-7.
44. Petrosky E, Blair JM, Betz CJ, Fowler KA, Jack SPD, Lyons BH. Racial and Ethnic Differences in Homicides of Adult Women and the Role of Intimate Partner Violence - United States, 2003-2014. *MMWR Morb Mortal Wkly Rep* 2017;66:741-6.
45. Morrison PK, Pallatino C, Fusco RA, Kenkre T, Chang J, Krans EE. Pregnant Victims of Intimate Partner Homicide in the National Violent Death Reporting System Database, 2003-2014: A Descriptive Analysis. *J Interpers Violence* 2020;886260520943726.

46. Karch D, Nunn KC. Characteristics of elderly and other vulnerable adult victims of homicide by a caregiver: national violent death reporting system--17 U.S. states, 2003-2007. *J Interpers Violence* 2011;26:137-57.
47. Choi NG, DiNitto DM, Marti CN, Choi BY. Poisoning deaths among late-middle aged and older adults: comparison between suicides and deaths of undetermined intent. *Int Psychogeriatr* 2019;31:1159-69.
48. Choi NG, DiNitto DM, Marti CN, Kaplan MS. Older Suicide Decedents: Intent Disclosure, Mental and Physical Health, and Suicide Means. *Am J Prev Med* 2017;53:772-80.
49. Karch D. Sex differences in suicide incident characteristics and circumstances among older adults: surveillance data from the National Violent Death Reporting System--17 U.S. states, 2007-2009. *Int J Environ Res Public Health* 2011;8:3479-95.
50. Mezuk B, Ko TM, Kalesnikava VA, Jurgens D. Suicide Among Older Adults Living in or Transitioning to Residential Long-term Care, 2003 to 2015. *JAMA Netw Open* 2019;2:e195627.
51. Schmutte TJ, Wilkinson ST. Suicide in Older Adults With and Without Known Mental Illness: Results From the National Violent Death Reporting System, 2003-2016. *Am J Prev Med* 2020;58:584-90.
52. Sweeney HA, Fontanella CA, Steelesmith DL, Quinn C. Suicide in Older Adults in Ohio: Characteristics, Precipitants, and Mental Health Service Utilization. *Community Ment Health J* 2020;56:1549-56.
53. Wanta BT, Schlotthauer AE, Guse CE, Hargarten SW. The burden of suicide in Wisconsin's older adult population. *Wmj* 2009;108:87-93.
54. Betz ME, Krzyzaniak SM, Hedegaard H, Lowenstein SR. Completed suicides in Colorado: differences between Hispanics and non-Hispanic Whites. *Suicide Life Threat Behav* 2011;41:445-52.
55. Frankenfeld CL, Leslie TF. Descriptive Epidemiology of Homicides With Victim and Suspect Race or Hispanic Ethnicity Discordance in the United States: Analysis of National Violent Death Reporting System (NVDRS) 2005-2015. *J Interpers Violence* 2019;886260519861656.
56. Stansfield R, Mancik A, Parker KF, Delacruz M. County Variation in Intimate Partner Homicide: A Comparison of Hispanic and Non-Hispanic Victims. *J Interpers Violence* 2019;886260519861657.
57. Lyons BH, Walters ML, Jack SPD, Petrosky E, Blair JM, Ivey-Stephenson AZ. Suicides Among Lesbian and Gay Male Individuals: Findings From the National Violent Death Reporting System. *Am J Prev Med* 2019;56:512-21.
58. Ream GL. What's Unique About Lesbian, Gay, Bisexual, and Transgender (LGBT) Youth and Young Adult Suicides? Findings From the National Violent Death Reporting System. *J Adolesc Health* 2019;64:602-7.



59. Ream GL. An Investigation of the LGBTQ+ Youth Suicide Disparity Using National Violent Death Reporting System Narrative Data. *J Adolesc Health* 2020;66:470-7.
60. Basham C, Denneson LM, Millet L, Shen X, Duckart J, Dobscha SK. Characteristics and VA health care utilization of U.S. Veterans who completed suicide in Oregon between 2000 and 2005. *Suicide Life Threat Behav* 2011;41:287-96.
61. Horwitz AG, Smith DL, Held P, Zalta AK. Characteristics of Veteran and Civilian Suicide Decedents: A Sex-Stratified Analysis. *Am J Prev Med* 2019;56:e163-e8.
62. Kaplan MS, McFarland BH, Huguet N. Firearm suicide among veterans in the general population: findings from the national violent death reporting system. *J Trauma* 2009;67:503-7.
63. Kaplan MS, McFarland BH, Huguet N, Valenstein M. Suicide risk and precipitating circumstances among young, middle-aged, and older male veterans. *Am J Public Health* 2012;102 Suppl 1:S131-7.
64. Katz IR, McCarthy JF, Ignacio RV, Kemp J. Suicide among veterans in 16 states, 2005 to 2008: comparisons between utilizers and nonutilizers of Veterans Health Administration (VHA) services based on data from the National Death Index, the National Violent Death Reporting System, and VHA administrative records. *Am J Public Health* 2012;102 Suppl 1:S105-10.
65. Logan JE, Fowler KA, Patel NP, Holland KM. Suicide Among Military Personnel and Veterans Aged 18-35 Years by County-16 States. *Am J Prev Med* 2016;51:S197-s208.
66. McCarthy MS, Hoffmire C, Brenner LA, Nazem S. Sleep and timing of death by suicide among U.S. Veterans 2006-2015: Analysis of the American Time Use Survey and the National Violent Death Reporting System. *Sleep* 2019; 42(8):zsz094.
67. O'Donnell J, Logan J, Bossarte R. Ten-Year Trend and Correlates of Reported Posttraumatic Stress Disorder among Young Male Veteran Suicide Decedents-Results from the National Violent Death Reporting System, 16 U.S. States, 2005-2014. *Suicide Life Threat Behav* 2019;49:1473-87.
68. Patton CL, McNally MR, Fremouw WJ. Military Versus Civilian Murder-Suicide. *J Interpers Violence* 2017;32:2566-90.
69. Wood DS, Wood BM, Watson A, Sheffield D, Hauter H. Veteran Suicide Risk Factors: A National Sample of Nonveteran and Veteran Men Who Died by Suicide. *Health Soc Work* 2020;45:23-30.
70. Sivaraman JJ, Marshall SW, Ranapurwala SI. State firearm laws, race and law enforcement-related deaths in 16 US states: 2010-2016. *Inj Prev* 2020;26:569-72.
71. Wertz J, Azrael D, Berrigan J, et al. A Typology of Civilians Shot and Killed by US Police: a Latent Class Analysis of Firearm Legal Intervention Homicide in the 2014-2015 National Violent Death Reporting System. *J Urban Health* 2020;97:317-28.

72. Barber C, Azrael D, Cohen A, et al. Homicides by Police: Comparing Counts From the National Violent Death Reporting System, Vital Statistics, and Supplementary Homicide Reports. *Am J Public Health* 2016;106:922-7.
73. DeGue S, Fowler KA, Calkins C. Deaths Due to Use of Lethal Force by Law Enforcement: Findings From the National Violent Death Reporting System, 17 U.S. States, 2009-2012. *Am J Prev Med* 2016;51:S173-s87.
74. Blair JM, Fowler KA, Betz CJ, Baumgardner JL. Occupational Homicides of Law Enforcement Officers, 2003-2013: Data From the National Violent Death Reporting System. *Am J Prev Med* 2016;51:S188-s96.
75. Anestis MD. Prior suicide attempts are less common in suicide decedents who died by firearms relative to those who died by other means. *J Affect Disord* 2016;189:106-9.
76. Fowler KA, Dahlberg LL, Haileyesus T, Gutierrez C, Bacon S. Childhood Firearm Injuries in the United States. *Pediatrics* 2017;140(1):e20163486.
77. Hanlon TJ, Barber C, Azrael D, Miller M. Type of Firearm Used in Suicides: Findings From 13 States in the National Violent Death Reporting System, 2005-2015. *J Adolesc Health* 2019;65:366-70.
78. Hemenway D, Barber C, Miller M. Unintentional firearm deaths: a comparison of other-inflicted and self-inflicted shootings. *Accid Anal Prev* 2010;42:1184-8.
79. Hernandez-Meier JL, Akert B, Zheng C, Guse CE, Layde PM, Hargarten S. Status of legal firearm possession and violent deaths: methods and protocol for a retrospective case-control study. *Inj Prev* 2019;25:i49-i58.
80. Kalesan B, Sampson LA, Zuo Y, Galea S. Sex and age modify the relationship between life circumstances and use of a firearm in suicide deaths across 17 U.S. states. *J Affect Disord* 2018;236:105-11.
81. Kaplan MS, McFarland BH, Huguet N. Characteristics of adult male and female firearm suicide decedents: findings from the National Violent Death Reporting System. *Inj Prev* 2009;15:322-7.
82. Solnick SJ, Hemenway D. Unintentional firearm deaths in the United States 2005-2015. *Inj Epidemiol* 2019;6:42.
83. Gollub EL, Gardner M. Firearm legislation and firearm use in female intimate partner homicide using National Violent Death Reporting System data. *Prev Med* 2019;118:216-9.
84. Subramanian R. Transitioning to multiple imputation-A new method to estimate missing blood alcohol concentration (BAC) values in FARS 2002. NHTSA Technical Report HS-809 403. Washington, DC: 2002
85. Rubin DB, Schafer JL, Subramanian R. Multiple imputation of missing blood alcohol concentration (BAC) values in FARS 1998. NHTSA Technical Report HS-808 816. Washington, DC: 1999.

86. Rockett IRH, Caine ED, Connery HS, et al. Discerning suicide in drug intoxication deaths: Paucity and primacy of suicide notes and psychiatric history. *PLoS One* 2018;13:e0190200.
87. Liu D, Yu M, Duncan J, Fondario A, Kharrazi H, Nestadt PS. Discovering the Unclassified Suicide Cases Among Undetermined Drug Overdose Deaths Using Machine Learning Techniques. *Suicide Life Threat Behav* 2020;50:333-44.
88. Dailey NJ, Norwood T, Moore ZS, Fleischauer AT, Proescholdbell S. Evaluation of the North Carolina Violent Death Reporting System, 2009. *N C Med J* 2012;73:257-62.

# Chapter 5. Expanding Capacity and Capabilities to Monitor and Research Guns in the United States

Michael Mueller-Smith, PhD | University of Michigan

## Introduction

As information technology systems and survey operations have modernized over the last half-century, a range of novel data collection opportunities have emerged, advancing our ability to track and measure various forms of socioeconomic activity and outcomes critical to U.S. public policy. Numerous examples of cutting-edge data infrastructure have been founded and developed from administrative records not originally produced for research purposes but instead a byproduct of regular day-to-day operations of individuals, governments, nonprofits, and businesses. Examples can be found in a variety of policy domains: labor markets (Longitudinal Employer Household Dynamics program); education (National Student Clearinghouse); criminal justice (CJARS); and health care (HCUP). Such systems have successfully navigated a range of serious legal (e.g., HIPPA, FERPA) and privacy hurdles to provide major advances in data infrastructure, a prerequisite for building evidence to inform policy.

At the same time, survey data collection efforts are modernizing in response to changing conditions. Internet-based outreach efforts (including operations like Amazon MTurk) have lowered the cost of conducting surveys and expanded the possibility for new types of experimentation with subjects, with the caveat of having a nonrepresentative sample. Additionally, traditional representative surveys are gaining new life as expanded research possibilities have emerged for individual-level linkages with other survey and nonsurvey datasets, increasing the range of questions and research designs that can be explored with the data.

Leading efforts at data architecture integrate both survey and administrative data sources to flexibly approach data collection, capitalize on respective strengths, and minimize potential weaknesses. In the context of studying gun ownership, gun use, and its effects on the population, this is even more critical given the non-trivial legal and political barriers to progress. What is needed is a multifaceted strategy of interoperable, diversified, complementary collection efforts that together generate a sum greater than its parts.

**Highlighting a model data system:** The CJARS, founded in 2016, is an ongoing data collection effort and cutting-edge dissemination platform, designed to transform research and statistical reporting on the U.S. criminal justice system. It is the first nationally integrated research repository that follows individuals from arrest to charge to disposition to sanction.

Data come from all types of criminal justice agencies and from across the United States. At the University of Michigan, data are harmonized into a common schema that allows analysis across disparate jurisdictions. After secure transfer to the U.S. Census Bureau, CJARS data are anonymized and linked at the individual-level to confidential social, economic, and demographic survey and administrative records to produce novel empirical analysis of criminal justice caseloads.

The project's ultimate goal is to enable research and statistics that legislators and administrators can use to develop evidence-based criminal justice policy. Data from 23 states are currently held in the U.S.

Targeted efforts should tackle specific measurement goals without trying to solving all policy-relevant questions, providing a viable path forward that limits the opportunity for individual barriers to halt progress. Examples include:

1. Purchase or transaction data that may cover federally authorized gun dealers, including credit card transaction data
2. Health records (usage or claims) that indicate gunshot victimization
3. Mortality records that identify potential gun involvement
4. Weapons offenses in criminal justice records
5. Embedding gun-related questions (for instance, current and prior ownership) into ongoing nationally representative survey efforts: National Longitudinal Survey of Youth, General Social Survey, or the Children's Health Survey
6. Engaging modern survey platforms for social science research to assess willingness to pay

Central to the success of this approach is intentional planning to support linkage, such that any individual success builds broader momentum. Key linkage factors that should be considered include:

1. Individual personal identifiers or other personally identifying information variables
2. Unique gun identifiers
3. Local geographic information where applicable

Integrating multiple data sources produces timely, policy-relevant evidence: In response to the COVID-19 pandemic, Congress created the Paycheck Protection Program (PPP) to support small businesses. Original provisions from the Small Business Administration (SBA), however, made businesses ineligible for the PPP if an owner had a variety of recent contact with the justice system. Using secure data infrastructure in the Federal Statistical Research Data Center network, Finlay, Mueller-Smith, and Street (2020) investigated these restrictions using individual tax return data linked at the person-level with CJARS-covered criminal records. They found that as many as 3.2 percent of sole proprietorships may have been ineligible for PPP assistance due to current or prior criminal justice involvement. Black and Hispanic men with sole proprietorship income were significantly more likely to be PPP-ineligible than white men. Between 6.9 and 15.4 percent of former convicts rely on self-employment income, which is particularly pronounced for women.

In part due to this evidence, SBA later relaxed these provisions, expanding access to PPP support for over 1 million entrepreneurs with criminal histories.

## Scaling Infrastructure With Machine Learning

While administrative data have brought significant promise to several research and policy domains, they introduce serious challenges because the underlying information was produced for operational, not research, purposes. A common problem is the existence of free entry text fields or case notes that contain a wealth of information but typically require human review to extract relevant information for analysis purposes. When reviewing thousands if not millions of records, data harmonization becomes infeasible.

Some existing data collection efforts have invested significant time and resources in hand-coding raw data fields to formalize research schema. For example, the NEISS trains human coders to differentiate between “accidental,” “assault,” and “self-inflicted” injuries from free text fields, which is thought to substantially address overuse of the “accidental” classification. Other data resources are ripe for further investigation. For instance, national crime reporting systems like the Uniform Crime Reports or the NIBRS fail to differentiate between crimes involving gunshots versus gun threats, which some localities already differentiate and could be expanded through broader use of case notes for data processing and preparation. At scale, hand-coding records through human review for classification purposes becomes cost-prohibitive in many cases, limiting the available information from extant records for analysis purposes.

Recent years have seen an explosion of activity in machine learning and data science, which potentially provide a cost-effective path forward. Existing data series provide a rich source of *training data* from which machine learning models can be constructed. This would reduce the financial costs of operating those data collection efforts going forward, as the estimated models

can be used to algorithmically classify unambiguous records, focusing more resource intensive human review on ambiguous entries.

Machine learning also generates broader returns as complementary collection efforts may seek to analyze similar and overlapping field content at the state and local level. The production and broad distribution of such trained machine learning algorithms can empower local actors to leverage subnational data, which often entail lower access barriers compared to national datasets, and lower research costs overall. It also provides a common framework to develop estimates based on the same underlying definitional concepts, avoiding the problem of apples to oranges comparisons when different organizations use different classification criteria.

**Machine learning in practice to advance data infrastructure:** Currently, CJARS is built on over 2 billion lines of raw data, covering approximately 178 million unique criminal justice events, occurring in 23 states. Most agencies provide free entry text fields to describe the type of offense involved in a given criminal episode, resulting in over 4 million unique offense descriptions in the data. Choi, Kilmer, Mueller-Smith, and Taheri (2021) leverage a unique source of 386,906 classified offense descriptions produced by Measures for Justice (MFJ) to train a machine learning algorithm, known as the Text-based Offense Classification (TOC) tool.

TOC is now used by both CJARS and MFJ to support quality data processing at reduced costs. It also supports the research and analysis of a range of external organizations and will be launched in late 2021 as a public tool for all to use.

## Disciplining Data Construction to Avoid Bias

Administrative records hold significant promise for research and statistical purposes, but one must confront the fact that they are a byproduct of operational uses and not originally designed for analysis. This creates a number of unique challenges, including data and validation. In addition, researchers must have a clearly defined target observation unit for the intended data product and strategies to translate one or more source files into that structure.

To take a simple example: suppose a researcher were working with payment data and saw a series of regular monthly transactions with a licensed gun dealer. Should these be interpreted as distinct purchases indicating multiple gun transactions, or do the combined payments represent a single purchase being paid off in a monthly payment plan? What if there is a very large single purchase? Should this count as a single gun transaction or multiple? The answer depends on how one defines the intended unit of observation and the viability of defining a mapping from the source data to that target.

We can think of a second example in the health context: suppose a patient has multiple health events at a hospital over the course of several weeks that have been tagged as being

associated with gunshot wounds. Ultimately, the patient dies, which creates an additional death record that also notes a gunshot wound. Does this constellation of events represent a single shooting event leading to immediate care, follow-up care for complications, and ultimately death? Or, are the events distinct gunshot events for an individual in a crisis period?

While these examples might seem contrived, decisions on how to handle such situations will have fundamental implications for measurement. In the payment data example, gun prevalence rates could be dramatically over- or underestimated. Similarly, in the health example, the number of victimization events could be three times too high or one-third the true rate.

Such problems become even more complicated when combining multiple dataset sources, whether across non-mutually exclusive jurisdictional boundaries (e.g., a gunshot victim seeks treatment at both an urgent care facility and a hospital emergency room) or from the same provider over time (e.g., a health care event that was previously nonfatal is later reclassified as fatal).

Two features are critical to navigating these problems:

- Defining or developing unique identifiers that combine related events or observations from the same individual, household, or gun
- Implementing a strategy to disambiguate or deduplicate records down to the intended unit of observation for the target dataset (a process that can be disciplined by validation)

**Benchmarking CJARS against federal statistical series:** The United States lacks uniform rules across state and local jurisdictions on the privacy afforded to justice-involved. Likewise, there is substantial variation in the development of data access mechanisms for researchers. Lacking authority to compel data provision, CJARS relies on multiple strategies for opportunistic data acquisition, including data use agreements, public records requests, web scraping, bulk data downloads, and data donations. Data arrives in provider-specified formats and structures, which then have to be reconciled by staff at the University of Michigan. Due to the variation in data collection methods and the numerous creative solutions required to coherently process the data, there is a fundamental need to benchmark CJARS against other available data series to identify both the strengths and weaknesses of CJARS.

Papp and Mueller-Smith (2021) report the ability of CJARS to reproduce a range of statistical series published by the BJS, including the State Court Processing Statistics (SCPS), National Prisoners Statistics (NPS) Program, National Corrections Reporting Program, Annual Probation Survey, and Annual Parole Survey. Such comparisons have enabled CJARS to identify shortfalls in its data processing, and improve the quality and accuracy of the data product.



## Benchmarking Strategies to Validate Data Quality

A multifaceted approach to data collection provides flexibility and agility in response to changing legal and regulatory environments across jurisdictions and over time. For instance, combining information on transaction data, permit records, and self-identified ownership in survey responses could provide one of the most accurate measures of gun prevalence in the United States, collectively addressing the individual measurement and attrition biases of any individual source in isolation.

But, it also creates a number of serious challenges, including 1) numerous distinct native data layouts, especially if collecting information from state and local sources with locally defined data structures and formats; 2) inconsistent variable definitions, value codes, and free entry fields for categorical variables; 3) inadequate unique identifiers; and 4) potential duplicative coverage when receiving data from multiple providers with overlapping jurisdiction or when receiving multiple rounds of data over time from the same source.

Machine learning approaches previously described can help manage these types of data integration efforts operate at scale. Still, some tasks require tailored human engagement to understand the nature and content of a given data file.

A resulting data product is the consequence of a multitude of discretionary choices. Without an organizing framework to guide these decisions, the end result likely does not deliver on its promise.

In the context of a related data infrastructure effort, CJARS linkable person-level criminal justice records have been validated through replicating extant aggregate statistical series, including the SCPS, the NPS, and the Annual Probation and Parole surveys. Benchmarking aggregate information produced from CJARS microdata against accepted aggregate reporting programs achieves two goals. First, it provides a framework to guide data processing decisions. Second, it provides a benchmark against which to gauge data quality.

In the context of gun ownership, gun use, and its effect on the population, a number of plausible statistical series could be leveraged for benchmarking purposes to validate a new micro dataset, including:

1. General Social Survey (GSS)

- ▶ Times series variation in national prevalence of gun ownership among households since the 1970s

2. RAND State-Level Estimates of Household Firearm Ownership
  - ▶ State-level ownership estimates (1980-2016) built from integrating survey data sources (Behavioral Risk Factor Surveillance System [BRFSS], Gallup, GSS, and Pew Research Center) with administrative data from firearm-involved suicide rates, hunting licenses per capita, magazine subscriptions to *Guns & Ammo*, and the number of background checks from the National Instant Criminal Background Check System
3. NVSS
  - ▶ Gunshot fatalities over time and across geography
4. FBI's SHR, NVSS, NVDRS, Fatal Force database
  - ▶ Law enforcement use of force across geography and over time
5. NEISS
  - ▶ Nonfatal gunshot injuries

## Diversified Access Mechanisms to Balance Research, Privacy, and Security

A successful data platform should embrace multiple access mechanisms to serve multiple stakeholders.

At one end, consider the secure and confidential integrated microdata environment of the Federal Statistical Research Data Center network. This platform has been created to support research and statistical analysis of integrated data across multiple content domains and information owners. Examples could include:

1. Earnings trajectories before and after victimization events
2. Family and peer social spillovers in gun ownership rates
3. Victim/offender overlap between victimization and offender criminal justice records

While such a data environment provides the most significant promise for pushing the frontier of knowledge in this area, access and approval involve significant barriers, including potential financial costs, federal background checks, and physical limitations on where research can be performed.

Two complementary approaches have helped navigate the balance between privacy concerns and information availability. The first is a public data portal that curates aggregate statistics generated from linked microdata. The goal here is to remove any individual information, and with enhancements from differential privacy, protect the confidentiality of individuals covered in the data. Putting such information in the public domain increases transparency and data access, especially for efforts like evidence building to support data-driven policy, which often does not require individual-level information.

Second, for questions that aggregate statistics may not thoroughly answer, a synthetic data product that replicates the underlying variation in the confidential microdata but is artificially generated provides another avenue to lower access barriers without compromising privacy and confidentiality.

**Balancing availability and security in practice:** CJARS represents a significant (albeit growing) advancement in studying the U.S. criminal justice system. With support from the National Science Foundation, it is developing and deploying two modes of data access. First, secure access for qualified researchers on approved projects is supported through the Federal Statistical Research Data Center system. This access mechanism provides researchers with the ability to study confidential (anonymized) microdata that can be integrated with a range of other survey and administrative data held by the federal government.

In addition, CJARS is developing a synthetic data product composed of artificial records that preserve the underlying statistical information contained in the CJARS microdata. This will be publicly available via the University of Michigan's ICPSR. While the latter mode of access does not provide as many opportunities, its low-barrier approach will encourage broader adoption of CJARS in research and analysis.

## Concluding Thoughts

Evidence-based policymaking is a grounding principle of good governance. In the domain of guns and gun violence, there remains a striking gap between research capacity and societal impact. Lives and communities are transformed when shootings occur, yet the absence of critical data infrastructure required to study these topics renders political discourse unmoored and unproductive, lacking a range of empirical facts to guide debate.

The United States needs a new approach to address these shortcomings. Complementary survey and administrative data collection efforts should be engaged, with diverse strategies to ensure against single points of failure. Lessons from recent advances in data science and machine learning should be embraced to reduce production costs while enhancing data quality. Intentional capacity for interoperability will build and sustain momentum as these projects mature. Together, these efforts will promote improved evidence building capacity and support the adoption of policies that enhance the safety, productivity, and well-being of communities across the United States.

## References

Finlay, K., Mueller-Smith, M., Street, B. 2020. "Criminal Disqualifications in the Paycheck Protection Program." U.S. Census Bureau Work Paper Number ADEP-WP-2020-04.

Choi, J., Kilmer, D., Mueller-Smith, M., Taheri, S. 2021. "Hierarchical Approaches to Text-based Offense Classification." Working Paper.

Papp, J., Mueller-Smith, M. 2021. "Benchmarking the Criminal Justice Administrative Records System's Data Infrastructure." Working Paper. <https://cjars.isr.umich.edu/benchmarking-report-download/>.

# Chapter 6. Creating a Federal Gun Violence Interagency Working Group

Nancy Potok, PhD | NAPx Consulting

## Executive Summary

In October 2020, Arnold Ventures published “*A Blueprint for U.S. Firearms Data Infrastructure*,”<sup>1</sup> which contains the recommendations of an expert panel that had been convened by NORC at the University of Chicago. The panel examined how better use of public health and criminal justice data could help increase understanding of causes and prevention of gun violence. The panel also identified many barriers within the current research environment that hindered the ability to evaluate the effectiveness of preventive and intervention programs. Among the problems reported were that data on public health and gun violence are collected separately and siloed into narrow categories. Data on suicides, criminal use of firearms, crime victims, firearm acquisitions, attitudes towards firearms, and other related topics are collected by separate agencies and not easily shared—in fact, sometimes prohibited from being shared. In addition, the data being collected were often not sufficiently comprehensive, had variable quality, and were often missing important topics entirely. A key finding of the panel was that relevant data need to be looked at more holistically, to understand the many complex factors that can influence gun violence. Simply studying separate, individual datasets has been insufficient to effectively inform decision-makers.

This paper focuses on two of the panel’s recommendations: “*Increase federal data accessibility*” and “*Set up an interagency working group around data to create federal partnerships to address specific infrastructure gaps (that are not just reporting mechanisms).*” The recommendations are summarized below:

- **Increase federal data accessibility.** This recommendation calls on the federal government to prioritize data accessibility for qualified researchers with appropriate protections for confidentiality and use. Through the Interagency Council on Statistical Policy<sup>2</sup> chaired by the Chief Statistician of the United States in the Office of Management and Budget (OMB), the federal government should conduct regular reviews of the accessibility and usability of key firearms data and facilitate better access for researchers. Existing investments in data collection should be maximized by reducing barriers to use by researchers and assuring that data are timely, of high-quality, and being used appropriately.

---

<sup>1</sup> J.K. Roman, *A Blueprint for U.S. Firearms Data Infrastructure* (Bethesda, MD: NORC at the University of Chicago, 2020).

<sup>2</sup> Explain from Paperwork Reduction Act as amended by Evidence Act.

- **Set up an interagency working group around gun violence data.** This recommendation follows a model used by the federal government to tackle other cross-agency priority topics. The panel recommended that the interagency group be chaired by the Chief Statistician of the United States and include statistical officials, subject matter experts, and Chief Data Officers from agencies with relevant data.

The expert panel also created a conceptual framework through which various types of data were examined, including data that were available across the federal government and at the state and local level, and data that would result from evaluating demonstration programs focused on high risk groups or places. To put the framework in place, the panel developed several recommendations to help solve the problem of firearms data being difficult to access, collect, make publicly available, and integrate.

The recommendations, if implemented, would increase the number and timeliness of critical research questions that could be asked and answered to inform effective policymaking at all levels of government.

This paper examines implementation options for establishing an interagency working group on gun violence data. The charge of the group would be to develop and implement a strategy for 1) improving the completeness and quality of gun violence data, 2) bridging the data silos, and 3) increasing data accessibility for researchers inside and outside of government. The paper examines implementation steps that could provide the structure needed to achieve this goal, including:

- Establish a new interagency working group chaired by the Chief Statistician of the United States. (NORC expert panel recommendation).
- Coordinate the gun violence working group with the Interagency Working Group on Equitable Data established by Executive Order on Advancing Equity and Support for Underserved Communities (EO 13985) with an explicit link in its charter.
- Establish an interagency Federal Advisory Committee with non-federal members to provide advice regarding non-federal datasets relevant to gun violence.
- Establish an interagency pilot project to begin to bring data together to answer high-priority questions about gun violence, with a charge to identify needed improvements in content and infrastructure, as well as barriers to access. (NORC expert panel recommendation).

As highlighted through the discussion section, these four implementation options can be complementary activities rather than mutually exclusive. To take advantage of the synergy created by combining all these actions rather than approaching them as separate activities, this paper recommends beginning immediately rather than waiting for all pieces to be in place. Each element could be taken independently of the rest, although all of them together would be ideal.

**Recommended Immediate Actions for Implementation:** The Chief Statistician should establish an interagency technical working group on gun violence that coordinates closely with the Equitable Data Interagency Working Group established in EO 13985. The gun violence working group should begin by identifying a key research question related to an urgent problem and then design and conduct a pilot project to quickly demonstrate the value in linking datasets to answer additional high-priority questions. The pilot project would also be the launching point for identifying which agencies are responsible for data that are missing or need to be improved. One purpose of the pilot would be to begin collaborations that would include non-federal data, creating incentives to improve the quality of data reported by local and state entities to federal agencies.

The pilot project should also solicit advice from an outside federal advisory committee consisting of non-federal stakeholders, data owners, and data users but should not wait for such an advisory committee to be established before starting the pilot. The advisory committee should include representation from groups or communities greatly affected by gun violence, including suicide and crime, and gun owners, in addition to state and local entities such as police departments and public health offices that collect and provide data.

The Chief Statistician does not need additional legislation or authorization to establish an interagency working group. However, some preliminary actions should ideally be taken before the group is established:

1. The position of the Chief Statistician, vacant since January 2020, needs to be filled or a strong “acting” person needs to be in place.
  - ▶ The interagency working group needs high-level support from the Executive Office of the President, including OMB and the Domestic Policy Council (DPC), to increase its effectiveness and create leverage to direct agencies to participate and provide resources to the effort.
  - ▶ Additional resources to carry out a pilot project need to be identified and supported by agencies, OMB, congressional appropriators, and non-federal partners such as philanthropic organizations and state partners.
  - ▶ Transparency and oversight need to be part of the structure for the pilot project. If evidence-based policymaking is to progress, projects that use sensitive data must be accountable and uphold the public trust. Projects undertaken in the pilot must provide value to the public through a better understanding of gun violence and, ultimately, effective approaches to reducing violence. The value proposition for the first research project must be clear.

## Introduction

Gun violence is on the rise across the United States. Homicides increased 29% in 2020, but have been trending up since 2014.<sup>3</sup> The U.S. firearm homicide rate began climbing in 2015, leading to more than 14,000 deaths a year starting in 2017.<sup>4</sup> In 2020, 75% of homicides in the U.S. involved a gun.

Officials at all levels of government want the public to feel safe from crime and to enable communities to prosper without the disruptive and traumatic after-effects of violence and suicide. While there is a growing awareness that additional resources are needed to assist people in need of mental health interventions and treatment, resources are limited, and information is lacking on how best to invest public funds to tackle these complex problems. Decision-makers need timely, objective, and reliable data to understand the many interacting forces that lead to violence, injury, and death.

The expert panel convened by NORC at the behest of Arnold Ventures identified several barriers to accessing data that continue to limit the information that is available to make important decisions regarding public policies and programs. In its final report, the panel made several recommendations and noted:

*The key problem to be solved through these recommendations is that firearms data are often difficult to access, collections are narrow in scope, public release of data can lag by years, and few datasets and systems can be integrated. Firearms data often cannot be accessed because of policy restrictions (ATF [the Bureau of Alcohol, Tobacco, Firearms and Explosives] data and background check data), or firearms data can be accessed and do have valuable data but need a slight change in order to accurately identify firearm injury cases (such as the UCR [Uniform Crime Reporting], NIBRS [National Incident-Based Reporting System], and hospital data systems) or firearms data are simply not collected (e.g., state-level data on firearm ownership rates). As a result, the number of critical research questions that can be asked and answered in the service of more effective policymaking is severely constrained.<sup>5</sup>*

The panel was not the first to note these issues. As far back as 2005, the Committee on Law and Justice at the National Research Council of the National Academies convened the

---

<sup>3</sup> Jeff Asher, "Murder Rose by Almost 30% in 2020. It's Rising at a Slower Rate in 2021." New York Times, September 22, 2021. <https://www.nytimes.com/2021/09/22/upshot/murder-rise-2020.html> (Accessed October 15, 2021).

<sup>4</sup> Centers for Disease Control and Prevention, National Center for Health Statistics. Underlying Cause of Death 1999-2019 on CDC WONDER Online Database, released in 2020. Data are from the Multiple Cause of Death Files, 1999-2019, as compiled from data provided by the 57 vital statistics jurisdictions through the Vital Statistics Cooperative Program. Accessed at <http://wonder.cdc.gov/ucd-icd10.html>

<sup>5</sup> J.K. Roman, *A Blueprint for U.S. Firearms Data Infrastructure*, 2.



Committee to Improve Research Information and Data on Firearms.<sup>6</sup> The committee recommended that the federal government support a systematic program of data collection and research because the current data and research were lacking. In 2009, the Committee on National Statistics and the Committee on Law and Justice at the National Research Council convened a panel charged with reviewing the BJS programs to identify priorities for data collection.<sup>7</sup> That panel identified several gaps in the BJS portfolio and made several recommendations for improvements. Subsequent studies by the National Academies and RAND in 2013 and 2018 found that the earlier problems of insufficient, hard-to-access data persist. The NORC 2019 report built on these reports to develop a set of specific, actionable recommendations that could result in significant improvements in the quality and availability of data related to both the crime and public health aspects of gun violence.

Some of the biggest problems identified by the NORC expert panel were:

1. Federal data are not disseminated on a timely, regular schedule that can provide up-to-date information on where and what types of violence are occurring.
  - ▶ The data collected by federal agencies are highly restricted, not well documented, and hard-to-access.
  - ▶ Public health, crime, and firearms data are not looked at holistically by the various agencies that collect these data, resulting in an inability to formulate effective prevention strategies and monitor real-time outbreaks of violence, whether they consist of suicides or violent crimes committed with guns.
  - ▶ There are no official data on firearms ownership, which are crucial for understanding suicides better.
  - ▶ Current administrative data systems are inadequate for tracking nonfatal gunshot injuries, due to multiple federal agencies collecting different aspects of this information combined with inconsistent reporting from local entities providing input data.
  - ▶ Even marginal improvements could result in major quality increases in data collected on how firearms purchases are processed through federally licensed firearms dealers (FFL), how firearms are used in violent crimes, and how firearms contribute to unintentional injury and death, suicide, and homicide.
  - ▶ The federal government has not established clear and consistent priorities for state and local data collections and reporting, leading to inconsistent quality, incompleteness, and less utility of these data.

---

<sup>6</sup> C.F. Wellford and C.V. Petrie, eds., *Firearms and Violence, A Critical Review* (Washington, DC: The National Academies Press, 2005).

<sup>7</sup> R.M. Groves and D. Cork, eds., *Ensuring the Quality, Credibility, and Relevance of U.S. Justice Statistics* (Washington, DC: The National Academies Press, 2009.)

- ▶ Much of the data at the state and local levels are covered by multiple law enforcement and public health entities and jurisdictions. There is no consistent mechanism or infrastructure in place to facilitate cooperation across these multiple jurisdictions and federal agencies.
- ▶ States need assistance to improve their ability to collect and report higher-quality data. For example, the timeliness and quality of data entered by states into the National Instant Criminal Background Check System are insufficient and contains gaps.
- ▶ There is no focused national strategy for tackling these problems to gain systematic improvements.

## Statutory and Other Mandates to Improve the Status Quo

Although there is no national strategy to improve the collection, use, and dissemination of data related to gun violence, there are several new statutes, federal policies, and OMB guidance directed at improving federal use of data, supporting evidence-based policymaking, and increasing access to high-value data for researchers and the public. These new authorities and mandates can be used to address some of the traditional barriers identified by the NORC expert panel. They are explained below.

### Foundations for Evidence-Based Policymaking Act

The U.S. Commission on Evidence-Based Policymaking was created in 2016 through bipartisan legislation and charged with studying how government data could be used more effectively to inform public policy. The [final report](#) of the Commission included 22 unanimous recommendations, 11 of which were enacted into law in [PL 115-435](#), the Foundations of Evidence-Based Policymaking Act of 2018 (Evidence Act).

The Evidence Act enables federal agencies to better use and share data, making a distinction between non-sensitive or “open” data and the most sensitive data, such as data collected from individuals or businesses for statistical purposes that require appropriate privacy and confidentiality protections. Federal agencies and outside researchers are given expanded authority to link even sensitive data with appropriate protections to gain a more holistic picture of how well programs and policies are working and better understand the dynamics at play in the communities they serve. The act also established a governance structure for data management and stewardship. Key provisions of the Evidence Act include the following:

**Evaluation Plans and Learning Agendas.** Agencies are required to develop evaluations for their programs that are tied to learning agendas seeking to answer important questions in an agency’s strategic plan. To carry out the evaluations and advance the learning agenda, agencies need data and the ability to analyze those data to gain insights. In 2019 and 2021, the U.S. Office of Management and Budget issued guidance specifically targeted to helping agencies conduct high-quality evaluation studies (OMB Memoranda [M-19-23](#) and [M-21-27](#)).

**Open Data and Data Management.** Agencies are required to make data open and available to the public unless the data are otherwise determined to be sensitive or are prohibited by law from being shared openly, such as tax data. To encourage the use of their data, agencies are required to make available to the public a comprehensive data inventory, find ways to engage the public in making their data more useful, and include their data in a publicly available federal data catalogue, such as [data.gov](#), run by the General Services Administration.

**Strengthening the Confidential Information Protection and Statistical Efficiency Act (CIPSEA) and federal statistical activities.** CIPSEA was originally enacted in 2002 and created a special authority for principal federal statistical agencies to guarantee confidentiality to

respondents of federal statistical data collections. It standardized the confidentiality pledge federal statistical agencies use when collecting information for statistical purposes from the public; provided a uniform approach to protecting confidential information collected under the pledge; and required the application of sound scientific and statistical disclosure limitation techniques to minimize the risk of re-identification of respondents in statistical data products. The act named 13 principal statistical agencies (e.g., the Census Bureau, Bureau of Labor Statistics, BJS, NCHS), and the Chief Statistician was given the authority to grant the CIPSEA authorities to other agencies.

The 2018 amendments to CIPSEA in the Evidence Act created a presumption of accessibility for the statistical agencies: ***If a statistical agency requests data from another federal agency for statistical purposes, the agency must give those data to the statistical agency unless sharing is otherwise prohibited by law*** (such as sharing tax data or certain education data). The Evidence Act amendments also required the expansion of secure access to CIPSEA-protected data and that OMB issue regulations that would guide agencies in establishing tiers of sensitivity for their data so that appropriate access and protections would be put in place. In essence, these provisions build on the decades-long experience of the statistical agencies, making them trusted agents for creating new datasets from linked data that can provide valuable insights into decision and policymaking.

Importantly, non-statistical agencies are currently able to enter into agreements with each other to share data as well, even without the mandates of the Evidence Act. Several program agencies are engaging in data sharing activities that do not require the new authorities granted in the act. These will be discussed later.

To expand access to data for statistical purposes, the CIPSEA amendments also set requirements for OMB. The office had to establish criteria and issue guidance on how an agency might qualify for the CIPSEA designation; direct CIPSEA-designated agencies to expand secure access to qualified researchers to de-identified sensitive data; and set up one common application process for approving projects submitted by researchers who request access to sensitive data from a statistical agency.

## The Paperwork Reduction Act of 1995<sup>8</sup>

The [Paperwork Reduction Act](#) (PRA) requires coordination of federal information policy by OMB, with the intent of reducing the paperwork burden on the public from federal information collections. It has many little-known authorities. In particular, it created the position of Chief Statistician within OMB with statutory duties and responsibilities that include ensuring the integrity, objectivity, impartiality, utility, and confidentiality of information collected for statistical purposes. The Chief Statistician coordinates the 13 principal statistical agencies and 115 other

---

<sup>8</sup> 44 U.S.C. section 3501 et. Sec.

statistical offices across government; generates government-wide data collection standards (e.g., race and ethnicity, Metropolitan Statistical Areas, industrial and product classification systems used by the private sector); and develops methodological guidance and promotes innovation.

The PRA also created the Interagency Council on Statistical Policy (ICSP), whose membership consists of the designated statistical officials of the 24 largest federal agencies, including the heads of the 13 principal statistical agencies. This Council coordinates statistical activities across the government and has several working groups that tackle important methodological issues such as privacy protection and promoting innovation and implementation of the Evidence Act.

Under the authority of the PRA, the Chief Statistician has established many interagency working groups since 1995. Ongoing interagency working groups, with members appointed by the ICSP, make recommendations to the Chief Statistician on updating the [Standard Occupational Codes](#), [North American Industry Classification System \(NAICS\)](#), [North American Product Classification System](#), and the [Metropolitan Statistical Area](#) designations. There are also working groups that have been established to tackle particular issues such as changing how the [Poverty Rate](#) is calculated, how data on [race and ethnicity](#) should be classified and collected, and [consumer inflation measures](#). The membership of these interagency groups varies, depending on which agencies have programs and data related to the topic. For example, the membership of the working group looking at alternative ways to estimate inflation included agencies such as Housing and Urban Development, Social Security, and HHS, due to the significant impact inflation adjustments have on their programs.

## Information Quality Act of 2000<sup>9</sup>

The [Information Quality Act of 2000](#) requires OMB and other federal agencies to maximize the quality of information provided to the public. It required OMB to issue guidance that applied to the sharing and accessing of information disseminated by federal agencies. The OMB Government-wide Information Quality Guidelines, first issued in [2002](#) and updated in 2019 ([M-19-15](#)), require agencies to institute procedures to ensure the objectivity, utility, and integrity of information, including statistical information, provided to the public. The updated guidelines consider the growing use of administrative program records for creating evidence and conducting program evaluations by including a new section on re-use of existing agency program data. This section includes the following requirements:

*Update 2.3: Agencies should consider the potential for using existing data sources from both inside and outside the agency for statistical and research purposes, while protecting privacy and confidentiality.*

---

<sup>9</sup> Section 515 of the Consolidated Appropriations Act, 2001, Pub. L. 106-554.

*Update 2.4: When designing or improving data collection systems, Departments should actively solicit comment from their statistical, research, and evaluation agencies about potential downstream uses. Agencies should describe such uses in the Information Collection Request submitted to OMB for review under the PR. Implementation.*<sup>10</sup>

In addition, the updated guidelines encourage increased access to data while protecting privacy for sensitive data, for better transparency, reproducibility, and assessing the fitness of purpose for using the data.

## OMB Memorandum 14-03: Guidance for Providing and Using Administrative Data for Statistical Purposes

[M-14-03](#) was written with the goal “... to help both program and statistical agencies and components (including evaluation and analysis units) use administrative data more fully in a manner that respects privacy and protects confidentiality. Specifically, this guidance will help program agencies manage their administrative data with statistical purposes in mind.”<sup>11</sup> The memorandum calls for departmental and agency leadership to foster greater collaboration between program and statistical offices and encourages federal departments and agencies to promote the use of administrative data for statistical purposes. It specifically directs, “Heads of departments shall identify effective internal mechanisms to communicate the importance of identifying those administrative datasets with potential for statistical use. They shall establish an ongoing process for program and statistical agencies and components to collaboratively identify such datasets.”<sup>12</sup>

## Executive Order on Advancing Racial Equity and Support for Underserved Communities Through the Federal Government (EO 13985)

[Executive Order 13985](#) directs each federal agency to assess whether and to what extent its programs and policies perpetuate systemic barriers to opportunities and benefits for people of color and other underserved groups. The goal is for agencies to develop policies and programs that deliver resources and benefits equitably to all. Section 9 of EO 13985 establishes an Interagency Working Group on Equitable Data (Data Working Group). The Chief Statistician and the United States Chief Technology Officer are co-chairs of the Data Working Group and coordinate its work. The membership of the group includes representatives of OMB, the Council of Economic Advisors, Treasury, Commerce/Census Bureau, and other agencies as deemed appropriate by the co-chairs. The function of the Data Working Group is to identify and provide recommendations on “...inadequacies in existing federal data collection programs, policies, and

<sup>10</sup> OMB Memorandum M-19-15, 6. <https://www.whitehouse.gov/wp-content/uploads/2019/04/M-19-15.pdf>.

<sup>11</sup> OMB Memorandum M-14-06, 1. <https://www.whitehouse.gov/sites/whitehouse.gov/files/omb/memoranda/2014/m-14-06.pdf>.

<sup>12</sup> Ibid. p. 5.

*infrastructure across agencies, and strategies for addressing any deficiencies identified; and (ii) support agencies in implementing actions, consistent with applicable law and privacy interests, that expand and refine the data available to the federal government to measure equity and capture the diversity of the American people.”<sup>13</sup>*

## Federal Data Strategy

[The Federal Data Strategy](#) is intended to help agencies leverage their data as a strategic asset. OMB issued the mission statement, principles, and practices of the strategy as a memorandum to agencies, [M-19-18](#). The strategy is part of the President’s Management Agenda as a cross-agency priority goal and includes four components:

1. *Enterprise Data Governance* includes data management, standardizing metadata, creating inventories, safeguarding confidentiality and privacy, etc. The more expansive governance vision includes collaboration across agencies and agency program silos to bring multidisciplinary expertise together.
2. *Access, Use, and Augmentation* calls on agencies to make data available to the public more quickly and in more useful formats. In addition, agencies should increase access to sensitive, protected data while protecting privacy, confidentiality, and security, including the interests of the data providers. The strategy’s action plan calls for the creation of toolkits and methodologies to help agencies build their own competencies as well. Agencies are also expected to seek out new sources for building datasets, which could include commercially available data and data from state and local governments.
3. *Decision-Making and Accountability* addresses the need for policy- and decision-makers to increase their use of high-quality data and analyses to inform evidence-based decision-making and improved operations. In addition, increased government accountability and transparency should be achieved by providing accurate and timely spending information, performance metrics, and other administrative data. Agencies are expected to use the most rigorous methods possible. Using outside expertise is encouraged, and agencies need to facilitate the use of government data assets by external parties, such as academic researchers, businesses, and community groups.
4. *Commercialization, Innovation, and Public Use* requires agencies to reach out to partners outside of government to assess which data are most valuable and should be prioritized for public use. There are many examples of entrepreneurial companies that have taken public data to create new apps that benefit the public and found new economic engines, such as weather and geographic mapping companies. This part of the strategy seeks to accelerate that long-standing practice by releasing more data to the public.

---

<sup>13</sup> EO 13985 section 9 <https://www.whitehouse.gov/briefing-room/presidential-actions/2021/01/20/executive-order-advancing-racial-equity-and-support-for-underserved-communities-through-the-federal-government/>.

The current annual action plan for the Federal Data Strategy identifies actions that agencies must take to build their capacity, establish processes, and align their existing efforts to better leverage data. Of note, agencies must identify their data needs to answer priority agency questions, among other things. The strategy also includes quarterly reporting milestones for agencies and mechanisms for how progress will be reviewed in conjunction with agency budget requests.

## Summary of Mandates

The recommended actions to improve data related to firearms and gun violence fit well within the federal statutes, mandates, guidance, and data strategy. An interagency working group could get resources and recognition for a pilot project through the President's Management Agenda framework and individual agency funds. Several projects that advance the management agenda are funded through appropriations provided to OMB to distribute to agencies to carry out priority items on the President's Management Agenda. The Deputy Director for Management at OMB, who chairs the President's Management Council (made up of the Deputy Secretaries of the Cabinet agencies), is in charge of allocating these funds. If a pilot project were designed to make generalizable infrastructure or data sharing improvements that other agencies could use in advancing the data strategy, some funding for a pilot potentially could be obtained through these means to get a project started.

## Implementation Options

The expert panel recommended that an interagency working group be established by the Chief Statistician in OMB to tackle improving firearms data integration. The panel would include chief statistical officers from the 13 principal statistical agencies on the [Interagency Council on Statistical Policy](#) and other key agencies, such as the Centers for Disease Control and Prevention (CDC), ATF, and the FBI. The interagency workgroup would coordinate policy development and implementation for data collection and sharing, improve access to the data, and establish a pilot project with a small number of states and federal agencies. The pilot could also be used to create tools for data dissemination and a delivery system that supports the use of those products to optimize use. The expert panel recommended that the ICSP agencies consider funding and embedding the implementation supports through an intermediary organization that specializes in these issues.

The following discusses four possible approaches to implementing the expert panel recommendations.



## Establish a New Interagency Working Group Chaired by the Chief Statistician of the United States

As mentioned earlier, the Chief Statistician currently has authority under the PRA to establish interagency working groups as part of coordinating the federal statistical system. Indeed, such interagency groups already exist that focus on specific topics. However, the working group would need high-level support within the executive branch to maximize the gun violence interagency working group's effectiveness and have sufficient leverage to prioritize resources devoted to a pilot project, as well as implement data quality and coverage improvements. Ideally, relevant agency heads and OMB would support the goals of the working group and allocate funds in agency budgets to improve data through investments in infrastructure and data collection. In addition, support would be needed if funding legislation is required to incentivize efforts by states and local partners to improve and share their data for input into federal systems. Providing financial incentives and other value to states has been successful in other instances discussed in section VI.

The Chief Statistician, assisted by stakeholders, would need to garner support for the interagency workgroup. At OMB, it would be helpful for the Director of OMB or the Administrator of the Office of Information and Regulatory Affairs to send a letter to the relevant agency heads announcing the formation of the group and soliciting participation. The workgroup should also have a charter laying out its mission and authorities, including establishing a pilot project, developed by the Chief Statistician with input from the participating agencies.

The Chief Statistician also would need to regularly inform the ICSP of the group's progress and solicit input and advice from the council on the pilot project.

This approach would result in the working group functioning primarily at the technical level and giving recommendations to the Chief Statistician to be incorporated in ICSP activities and agency budgets. This would be advantageous from the standpoint of depoliticizing data on gun violence, but also runs the risk of not attracting sufficiently high-level support and resources from OMB and the agencies to advance the recommendations of the group and fund pilot project activities.

*Executive Order.* Another approach would follow the model of the Federal Interagency [Forum on Child and Family Statistics](#), which was chartered in 1997 by Section 6 of [Executive Order No. 13045](#). Twenty-three federal agencies participate in the forum, which collaborates to produce cross-agency statistics on children and families, including 41 indicators of well-being spanning seven domains. The forum was established by OMB, is run as one agency as designated in the executive order, and is guided by the Chief Statistician. However, this approach would require a new executive order, which could take a significant time and delay implementation. Forum activities do not include running an intergovernmental pilot project involving multiple agencies and states sharing data; however, a new executive order on gun violence could incorporate those activities for a forum.

## Establish an Interagency Subgroup Under the Umbrella of the Interagency Working Group on Equitable Data Established by Executive Order on Advancing Equity and Support for Underserved Communities (EO 13985)

As mentioned, section 9 of EO 13985 establishes an interagency working group on Equitable Data. The Chief Statistician and the Chief Technology Officer co-chair this working group. The membership of the Equitable Data Working Group is, for the most part, at the discretion of the co-chairs. Potentially, the Equitable Data Working Group could establish a small number of subgroups that focus on particular topics, with the first, for example, being data on gun violence. Gun violence does not affect all communities equally, and more data are needed to see whether federal programs and dollars are advancing equitable outcomes for the public. An in-depth look at gun violence data focused on improving data quality and data access would be in scope and appropriate for the Equitable Data Working Group.

However, becoming a subgroup of the Equitable Data Working Group has pros and cons. On the positive side, the relationship would give high-level support and visibility to the work. And because the Chief Statistician would be leading both groups, bringing them together could increase the ability of the Chief Statistician to coordinate efforts between the groups. The Chief Statistician would be responsible for keeping the effort nonpartisan.

Nevertheless, despite the best efforts of the Chief Statistician, a data subgroup on gun violence established through EO 13985 could take on a political aspect that may detract from needed bipartisan support for improving the data and the data infrastructure. Both firearms and advancing equity in underserved communities currently are polarizing issues, so it would be important for the interagency working group to conduct its work using a nonpartisan, objective approach to the data and the pilot project. The charter of the subgroup would need to clearly give the group a measure of independence to assure that it was perceived to be apolitical.

Another risk is that a different Administration could rescind the executive order and the Equitable Data Working Group would be disbanded. An interagency working group under the wing of the Chief Statistician and the ICSP would have a much better probability of surviving multiple changes in Administration.

Because the Chief Statistician would be leading both groups, a hybrid solution may accomplish both objectives of independence and high visibility. The Chief Statistician would convene the interagency working group on gun violence under the authority of the PRA, but that working group would stay in close contact with the Equitable Data Working Group and regularly report on its progress and findings to keep multiple components of the Executive Office of the President and interested agencies engaged and supportive.

## Establish an Interagency Federal Advisory Committee With Non-federal Members to Provide Advice Regarding Non-federal Datasets Relevant to Gun Violence

One way for the interagency working group to understand the value of sharing data from the perspective of states and localities would be to establish an advisory committee with representation from stakeholders, such as public health authorities, police departments, criminal justice nongovernmental organizations (NGOs), suicide prevention organizations, and researchers in these fields. A cross-disciplinary advisory committee could help identify approaches to improving data quality, access, and usage that would create value for the stakeholders and the federal agencies. Many federal agencies already have advisory committees, but they are not cross-agency in their focus. Just as the working group needs to have representation from multiple federal agencies, an advisory committee should also bring multiple disciplines together.

Additional legislative authority is not needed to set up a federal advisory committee. However, OMB itself is not set up to provide the necessary administrative support for such a committee. Two possible approaches would be 1) OMB sponsors the advisory committee but delegates the support to another agency, and 2) multiple agencies sponsor the advisory committee and one of the agencies provides the administrative support.

Examples of both of these models currently exist. For example, OMB established the [Advisory Committee on Data for Evidence Building](#), which is reviewing, analyzing, and making recommendations to the Director of OMB on how to promote the use of federal data for evidence building, with a focus on infrastructure and privacy. However, the day-to-day support of this advisory committee was delegated by the OMB Director to the Department of Commerce and is handled by its Bureau of Economic Analysis (BEA).

The Census Bureau, Bureau of Labor Statistics (BLS), and BEA share responsibility for the [Federal Economic Statistics Advisory Committee \(FESAC\)](#). The charge of the committee is to advise the three statistical agency heads on statistical methodology and other technical matters related to the collection, tabulation, and analysis of federal economic statistics. It is chartered by the Secretary of Commerce, but has participation from BLS, which is in the Department of Labor. FESAC is supported administratively by BEA.

Because OMB has a historical inclination not to sponsor federal advisory committees, the second model of having multiple agencies sponsor the advisory committee would likely be easier to implement. The HHS or Department of Justice leadership would need to be willing to take this on and assign one of their bureaus to provide the administrative support on behalf of the participating agencies. Most of the participating agencies have extensive experience managing federal advisory groups. Each agency could participate in the process of selecting from among the nominees to the committee and developing the agency charter. Establishing

this committee is likely to take at least 12-18 months, given the need to organize and then advertise for member nominations and give other public notice in the Federal Register.

### Establish an Interagency Pilot Project to Begin to Bring Data Together to Answer High-priority Questions About Gun Violence, With a Charge to Identify Needed Improvements in Content and Infrastructure, As Well As Barriers to Access.

Establishing a pilot project should be one of the first action items that the interagency working group takes up. It is through a pilot project that agencies will develop partnerships with states that provide value. Agencies will learn more about data sharing and access, and creating a blueprint for further collaborations. The pilot project should be relatively small initially to demonstrate value and bring in more interest and partners. The pilot project should tackle a high-priority area with major impact that can inform federal policy and programs and state operations of programs.

If start-up of the pilot waits to identify additional data needs until all the work is done, the interagency working group will have trouble sustaining support for its efforts. By contrast, if the pilot can demonstrate value, some of the barriers to data sharing and improving data will be much easier to tackle as the data owners will see a reason to invest resources in this effort.

A key question that the interagency working group will need to answer is how to pay for the pilot project. The pilot could be paid for through agency appropriations, states could contribute a share of funding, and philanthropic organizations could also contribute, particularly to sponsor sessions and workshops to organize the pilot. However, the cost and the sources of funding will depend on exactly what high-priority issue the pilot will address.

**Recommended Immediate Actions for Implementation:** The Chief Statistician should establish an interagency technical working group on gun violence that closely coordinates with the Equitable Data Interagency Working Group established in EO 13985. The gun violence working group should begin by identifying a key research question related to an urgent problem and then design and conduct a pilot project to quickly demonstrate the value in linking datasets to answer additional high-priority questions. The pilot would also be the launching point for identifying which agencies are responsible for data that are missing or need to be improved. One purpose of the pilot would be to begin collaborations that would include non-federal data, creating incentives to improve the quality of data reported by local and state entities to federal agencies.

The pilot project should also solicit advice from an outside federal advisory committee consisting of non-federal stakeholders, data owners, and data users but *should not wait for such an advisory committee to be established before start-up*. The advisory committee should include representation from groups or communities greatly affected by gun violence, including suicide and crime, as well as gun owners, in addition to states and local entities such as police departments that collect and provide data.

Some preliminary actions ideally should be taken before the group is established:

1. The position of the Chief Statistician, vacant since January 2020, needs to be filled or a strong “acting” person needs to be in place.
2. The interagency working group needs high-level support from the Executive Office of the President, including OMB and the DPC, to increase its effectiveness, and create leverage to direct agencies to participate and provide resources to the effort.
3. Additional resources to carry out a pilot project need to be identified and supported by agencies, OMB, congressional appropriators, and non-federal partners such as philanthropic organizations and state partners.
4. Transparency and oversight need to be part of the structure for the pilot project. If evidence-based policymaking is to progress, projects that use sensitive data must be accountable and uphold the public trust. Projects undertaken in the pilot must provide value to the public through a better understanding of gun violence and, ultimately, effective approaches to reducing violence. The value proposition for the first research project must be clear.

## Role of State and Local Governments and Organizations

State and local governments have important roles in the collection and use of data in the gun violence data ecosystem. Several important data collections originate with local programs that input data into a federally run data collection system. Examples examined by the expert panel include the FBI's Uniform Crime Reporting program that reports aggregate numbers on state-level totals of firearms-related crime; the NIBRS, which collects detailed data at the incident-level from police departments; and the CDC NVDRS, which compiles detailed, individual-level data on homicides and suicides.

However, in order for state and local organizations to devote resources to collaboration and improving data, the value proposition needs to be clear. Across other topic areas, state and local governments have found immediate value in sharing their data across silos. One example is the MidWest Collaborative. The MidWest Collaborative grew out of a need for states to understand the transitions of people from education to work across state lines. The collaborative began to organize in [September 2018](#), and piloted some training programs around specific high-value projects. The pilots were successful and the group began to monitor and support the health of the region's interconnected economies and societies. Initially the products and analyses were intended to assess workforce and education outcome measures, particularly student and worker in-flows and out-flows within their states and among the states in the collaborative, but this model could also be adapted to share and analyze data about gun violence and public health.

Some examples of the MidWest Collaborative's work include the Multi-State Postsecondary Report (MSPSR; [https://kystats.ky.gov/Reports/Tableau/2021\\_MSPSR](https://kystats.ky.gov/Reports/Tableau/2021_MSPSR)) and an unemployment dashboard created during the COVID-19 pandemic. The MSPSR is a state-driven dashboard produced by the Kentucky Center for Statistics (KYSTATS) using shared data to identify education through workforce flows in Indiana, Kentucky, Ohio, and Tennessee for Ohio and Kentucky postsecondary graduates. MSPSR allows the user to filter by the credential level, academic major, state of origin, and postsecondary institution to show employment and wages both in- and out-of-state for 1-, 3-, and 5-years out.

During the COVID-19 pandemic, the collaborative was able to quickly assemble dashboards to guide these decisions, helping states to quickly increase their capacity to interpret, analyze, and disseminate millions of Unemployment Insurance (UI) claims for evidence-based policy. Local workforce board administrators needed more information at lower levels of geography than could be provided through surveys, so they could quickly decide how and where to deploy scarce resources for remediation to best assist newly unemployed populations, particularly for traditionally underserved subpopulations.

The MidWest Collaborative's success demonstrates that the possibility of overcoming barriers and improving the utility of state and local data when there is a shared need for the information, an organizing force, and a ready platform and training for pilot projects to launch the efforts.

These general attributes can also be applied to gun violence data to create momentum to improve data quality and begin to combine data across silos in meaningful ways.

Another example of quick mobilization was in response to the COVID-19 pandemic. The American Rescue Plan Act of 2021 (Pub. L. 117-2) contains \$500 million in funding in section 2404 for the Director of the CDC *“to support public health data surveillance and analytics infrastructure modernization initiatives at the [CDC], and establish, expand, and maintain efforts to modernize the United States disease warning system to forecast and track hotspots for COVID–19, its variants, and emerging biological threats, including academic and workforce support for analytics and informatics infrastructure and data collection systems.”*<sup>14</sup> A similar approach to funding firearms data improvements could be enacted with sufficient congressional support, particularly if investments were geared toward implementing the recommendations for better data utilization and quality from a nonpartisan interagency task force. A bipartisan bill introduced in the House ([H.R. 8080](#), the Health STATISTICS Act of 2020) could also be used as a model for a gun violence data improvement. One key element that would be transferable to gun violence expanded the existing data linkage program at HHS for the *“...purpose of facilitating statistical public health research on trends and patterns across specifically defined, statistically relevant populations, with a particular focus on linking social determinants of health data, including with respect to—(1) food insecurity; (2) housing instability; (3) transportation access; (4) safety; (5) social connection and isolation; (6) financial resource strain; and (7) stress.”*<sup>15</sup> The demonstration project established in the bill was meant to assess the availability of datasets held by federal, state, local, and non-federal entities that would be useful to the research, and to use existing authorities and linkages of data, using the NCHS as the linking agent, as authorized by the Evidence Act. A similar multiagency approach to gun violence data could be housed at the BJS or the NCHS.

These examples illustrate the importance of a federal interagency working group to build state and local partnerships that also produce value for these partners because so much of the public health and crime data originate at the state and local levels. Without considering the value proposition for those organizations and recommending ways to provide resources, improving data quality and filling in the data gaps identified by the expert would be much more difficult. At a minimum, the interagency working group needs to incorporate state and local interests in data sharing among states, easing federal reporting requirements to make them more meaningful and less redundant, increasing capacity to conduct program evaluations, and improving their own program operations.

---

<sup>14</sup> <https://www.congress.gov/bill/117th-congress/house-bill/1319/text?q=%7B%22search%22%3A%5B%22CARES+Act%22%5D%7D&r=1&s=2> Section 2404.

<sup>15</sup> <https://www.congress.gov/bill/116th-congress/house-bill/8080/text?q=%7B%22search%22%3A%5B%22HR+8080%22%5D%7D&r=1&s=3> H.R. 8080, Section 3113.

Other important partners to bring in would be researchers from academia, particularly those with strong ties to state agencies. Academic partners can conduct research in privacy protection, work with states and federal agencies to link and analyze data, assist with quality measures development, advance and apply innovative computer science approaches such as machine learning and rich text analysis, and continue to conduct related social science, public health, and public policy research.



## Current Federal and State Data Sharing Efforts

Some existing successful state, federal, and academic data sharing partnerships are worth examining because their approaches could also apply to evidence building for gun violence. These are governance and data sharing models that a gun violence interagency working group could adapt for appropriate pilot projects. Notable examples include:

1. Statewide **Longitudinal Data Systems grants** to 41 states and the District of Columbia are administered by the National Center for Education Statistics and support the development of data to assess K–12 education outcomes and investments. These grants are used to create data systems at the state-level similar to those prohibited by the student unit record ban at the federal level. A similar approach could be considered for firearms data that have sharing restrictions at the federal level. These data help states, school districts, schools, and teachers make data-driven decisions and facilitate research on improving achievement and closing gaps. Mississippi LifeTracks is one state longitudinal data system that allows for the analysis of administrative data from multiple state agencies to assess education and workforce outcomes in the state. LifeTracks is funded through a combination of National Center for Education Statistics grants and annual state appropriations. Mississippi devotes a portion of its website to public accountability, listing approved projects and completed projects, and cites state-level statistics based on their results. This demonstration of the program's value and the useful information it provides has been suggested as the key to the system's sustainability.
2. **Census-Economic Research Service (ERS)–Food and Nutrition Service (FNS) Joint Project** is a long-term joint research project to acquire administrative data on the U.S. Department of Agriculture (USDA) food assistance programs from states and link them to Census Bureau surveys. The linked data provide insights on how program participation affects participants, who does not participate, and why. State Supplemental Nutrition Assistance Program (SNAP) and Women, Infants, and Children (WIC) agencies in participating states send their confidential microdata to Census in exchange for state-specific analyses and reports. ERS researchers participate in joint research projects. Census has been able to conduct linkage between Veterans Administration and Department of Defense data to gain specific insights into the decisions of veterans. The project successfully overcame several barriers to share records between two federal agencies and multiple states.
3. **Purchasing Patterns of Households Participating in the Women, Infants, and Children Program** was a 2019-2020 USDA-sponsored training program for agency employees to address questions about characteristics and buying habits of WIC and non-WIC households using commercial datasets. Commercial datasets may be available to help inform information about firearms.
4. **Temporary Assistance for Needy Families (TANF) Data Collaborative Pilot initiative** sponsored by the Administration for Children and Families (ACF) includes eight pilot sites supported for 30 months. It includes funding, intensive training, and technical assistance to

support state and local efforts and build strategic partnerships. The goal is to build TANF agency capacity to improve TANF program performance through applied data analytics. ACF also engages outside partners from academia and nonprofits. The data analytics training program has included TANF receipt data and Quarterly Census of Employment and Wages (QCEW) data for Illinois and Indiana and examines questions such as what characteristics increase an individual's risk of returning to TANF, what factors increase an individual's likelihood of not finding stable employment after leaving TANF, and what factors increase an individual's likelihood of not finding any employment after leaving TANF.

5. ***The UMETRICS Initiative: Universities Measuring the Impacts of Research on Innovation, Competitiveness, and Science*** effectively communicates the results of federally funded research to donors, policymakers, and other key stakeholders. The consortium of 31 universities was made possible by the federal STAR METRICS project and advances in the methods and tools to combine, mine, and analyze big data on federally funded research. UMETRICS examines the economic results generated by research in the form of 1) the benefits to and generated by students produced by universities, 2) spillovers to regional and national economies, and 3) the public value added to social well-being across the scientific spectrum, including innovations in health care, the environment, energy, and food system interventions, and improvements in policies from social science research. It provides individualized reports to each member of the consortium. Data come from the participating universities, the National Science Foundation (NSF) Survey of Earned Doctorates, the Census Bureau's LEHD and Non-Employer data, health care (Medicare), innovation (USPTO), finance (VentureXpert and CRISP, IPO databases); dissertation databases, industry announcements and information from curricula vitae, which are linked to each other. The resulting large-scale, structured, linked, updatable dataset permits new, high-quality, large-scale analyses of the scientific enterprise at a variety of levels. The STAR METRICS project was supported by the Census Bureau, NSF, NIA, and USDA. The Alfred P. Sloan and Ewing Marion Kauffman Foundation supported the establishment of the Institute for Research and Innovation at the University of Michigan, which manages the UMETRICS effort and continues to add new universities.
6. ***Local Employment Dynamics Partnership*** is a voluntary federal-state partnership started in 1999. States agree to share historical and ongoing administrative records of UI earnings data and QCEW data with the Census Bureau. The Census Bureau then produces a longitudinal data infrastructure from which new statistics about the dynamics of local employment and the locations of jobs and workers can be produced.

# Chapter 7. Practical Steps for Building State Capacity and Infrastructure to Use Data for Evidence-Based Decision Making

Nancy Potok, PhD and Nick Hart, PhD | NAPx Consulting and Data Foundation

## Executive Summary

This paper provides a roadmap based on specific examples taken from successful federal and state experiences that can help states and local governments build capacity and infrastructure to use data for evidence-based decision-making. Each state may be starting from a different level of experience and capability, and each has a different legal and regulatory landscape, but the approaches provided here are adaptable and scalable. The paper provides examples from recent federal efforts as well as successful federal-state partnerships and highlights from states that have put important foundational building blocks in place. Each section of the paper provides key takeaways or lessons learned. An essential element of all successful models is that they start with a value proposition. Experience has shown that there must be a clear payoff for the agencies providing data and the public they serve, for improvements in data infrastructure and analytical capabilities to be sustained. The examples presented below show a clear value, rather than creating processes without a focus on solving specific problems and knowledge gaps faced by agencies.

Public agencies face complex, multi-dimensional challenges in understanding how government resources can be deployed to achieve the best outcomes for their constituents in areas such as health, safety, education, jobs, social equity, and housing. Many of these issues are interconnected and require a holistic view that crosses traditional organizational boundaries. Similarly, understanding how to measure and improve the welfare of children may involve healthcare, housing, education, and crime data. Agencies need to demonstrate that tax dollars are, in fact, achieving stated program outcomes for the public.

Improved data science and computer science methods are enabling more timely, relevant data analysis. However, changing the way in which government agencies share and use data is not a simple undertaking. The U.S. Commission on Evidence-Based Policymaking noted in its report to Congress (2017) that sustained, concerted efforts are needed to strategically build the capacity and infrastructure to analyze data in meaningful ways. Goerge (2019) noted six key factors that lead to effective engagement with and among state and local governments to

access and provide data, including providing short term tangible benefits and a vision that connects authentically to the real-world experiences of government leaders.

Building capacity to better utilize data is a continuous learning process that has already been providing value. The federal government and many states have taken significant steps to build the needed capacity to produce information that can be combined from many sources across agencies to better serve the public. This paper discusses the important elements of a successful data strategy and implementation of increased capacity to analyze and utilize data.

These include the following:

- Data leadership and strategy
- Legal and regulatory authority and policies (framework)
- Data governance
- Transparency of processes and uses
- Accessibility for data use
- Oversight and accountability
- Sustainability

The discussion for each element includes specific examples that illustrate how each has been put into practice at the federal and state-level. The examples can be used to model similar efforts in other states and are designed to stimulate thinking and action to promote evidence building with data sharing across agencies, states, and other government jurisdictions. Also included is a detailed look at the topic of firearms violence. The paper concludes with a list of ten actions that states can take to maximize their opportunities for success. These include:

1. Establish and empower state and local data leadership roles
2. Identify local knowledge needs and data gaps based on priorities
3. Review legal authorities to share, protect, and use data
4. Provide feedback to federal partners about data needs and priorities
5. Establish a Cross-Agency Governance Process and Data Strategy
6. Execute on Data Governance by Integrating Local Data Where Feasible to Prepare for Linkage and Sharing
7. Institute transparent approaches to using data
8. Partner with existing entities to promote rapid progress and pilot projects to demonstrate value
9. Establish responsive engagement processes and oversight mechanisms
10. Plan for and implement robust privacy safeguards

**From “A Blueprint for a US Firearms Infrastructure”**

*“The development of a rigorous empirical research base to inform both citizens and policymakers requires a robust and sustainable data infrastructure. The most enduring data infrastructure is one that is comprehensive, flexible, and nonpartisan. Nowhere is that data foundation more needed than in the realm of firearms violence—reliable data are a critical bridge to effective policymaking that improves public safety by reducing the number of firearm accidents, suicides, homicides, and assaults.”*

*Roman JK. 2020. A Blueprint for a US Firearms Infrastructure. Bethesda, MD: NORC at the University of Chicago.*

## Introduction

States and localities face significant challenges in areas such as improving public health and safety, creating jobs and economic opportunity, tackling rising crime rates, addressing a lack of affordable housing, and measuring the long-term outcomes for students at public education institutions. The public depends on services delivered by a multitude of agencies at the federal, state, and local level. Complex, multifaceted, long-term challenges require approaches that are holistic and informed by information that crosses the traditional functional boundaries of agencies. Some state and local governments have taken steps to make better use of the data they collect from the public to improve service delivery and the policies behind them. These steps—often involving partners in academia, the private and non-profit sectors, and the federal government—have yielded positive results. This paper attempts to provide a roadmap of how state and local governments might learn from these successes: specifically, how to tackle systematically the seemingly overwhelming task of rationalizing and utilizing data to understand how to implement improvements and achieve better outcomes for the public.

Long-standing gaps in the national data infrastructure—including at the state and local levels—inhibit analytical capabilities that are long overdue and much-needed for contemporaneous policy debate, on topics ranging from the most efficient workforce development strategies to the most effective approaches to reducing firearm-related violence. An important illustration of this fragmentation and the gaps in data was identified by an expert panel commissioned by Arnold Ventures and convened by NORC at the University of Chicago. The expert panel issued three reports. One main finding in the [first report](#), on the state of firearms data, was that data on firearms violence—including health, public health, and crime data—are disordered and highly segmented, contributing to a lack of information that might otherwise inform public policy decisions.

At the federal level, many steps have been taken to focus attention on improving evidence-based policymaking and building capacity for data-informed decision-making. At the same time, several states have made great strides in improving access to and use of information not only to guide policy making but also to respond quickly to crisis situations, such as the COVID pandemic. Prioritizing better use of data provides an opportunity to modernize and update systems and analytical capabilities to reflect the public's expectations. Such a focus on data is an increasingly common practice among government officials.

Many states have learned hard lessons on what works and what doesn't as they have put in place capabilities to better use data. This paper attempts to consolidate and share examples of best practices that can save precious time for states and localities that want to make faster progress and avoid "reinventing the wheel." Each state is different. Laws, regulations, state government organization, local culture, and the maturity of existing capabilities may vary considerably among states. However, the information in this paper is general enough that it could be adapted in whole or in part to each state.

The paper provides an overview of the context for data modernization in 2021 and beyond, identifying successful national and state models that can be leveraged for further progress. It opens with background on federal developments or steps to advance the ability to use, share, analyze, and inform decision-making with data. Some steps can be replicated at the state-level, and some states have made significant progress with only a partial set of foundational steps in place. Next, this paper provides examples of successful state efforts and includes a brief overview of key characteristics for effective state and local capacity based on successful models. It discusses how those approaches can leverage federal investments and statutory authorities to promote data sharing and linkage. The paper then looks at possible examples of how states may use firearms violence-related data, with a roadmap for states to get started either building capacity or expanding existing capabilities. The paper concludes with key takeaways learned from the experience of existing intergovernmental efforts.

Finally, the paper emphasizes one key aspect of all successful models; that is, they start with a value proposition. Experience has shown that to marshal the resources needed to gain high-level attention, allocate sufficient funding and people, provide training, and change the culture in agencies where data silos have existed for decades, there must be a clear payoff for the participants, including the public. The examples presented here show a clear value for the agencies and the public they serve, above and beyond demonstrating compliance with mandates that may come from the federal government or the governor's office. In the examples, new processes focus on solving specific problems and knowledge gaps that agencies face.

## What Can Be Learned from the Federal Experience

The federal government has invested significant resources over the last decade to improve the ability of agencies to use data in policymaking. Individual agencies recognize that they can learn more from better use of their own program data and by combining their data with data from other sources such as statistical agencies and the private sector. However, scattered individual projects have yet to result in needed systemwide improvements. There are significant legal and cultural barriers to sharing data across agencies, as well as a lack of resources to create a data sharing infrastructure. Over the past six years, the situation began to change in important ways. The sections below describe some of the major initiatives that have propelled increased capability for evidence building at the federal level.

### Foundations for Evidence-Based Policymaking.

In 2016, the U.S. Congress passed a bipartisan bill to create a Commission on Evidence-Based Policymaking ([Pub. L. 114-140](#)). The Commission garnered bipartisan support because members of Congress, regardless of ideological viewpoint, realized that federal tax dollars were not being maximized, and there were few ways of knowing which programs were effective and why.

The bill established a 15-member Commission appointed by the President and Congressional leaders with consideration given to individuals with expertise in economics, statistics, program evaluation, data security, confidentiality, or database management. The Commission was required to submit a report and recommendations to Congress in 18 months, based on conducting a comprehensive study of data inventory, data infrastructure, database security, and statistical protocols related to federal policymaking and the agencies responsible for maintaining data. Several purposes for data were cited in the bill, such as to, “...*determine the optimal arrangement for which administrative data on federal programs and tax expenditures, survey data, and related statistical data series may be integrated and made available to facilitate program evaluation, continuous improvement, policy-relevant research, and cost-benefit analyses; make recommendations on how data infrastructure, database security, and statistical protocols should be modified to best fulfill those objectives; and make recommendations on how best to incorporate outcomes measurement, institutionalize randomized controlled trials, and rigorous impact analysis into program design.*”<sup>32</sup> In addition, the Commission was to consider whether a clearinghouse for federal program and survey data should be established.

The Evidence Commission was required to consult with several federal agencies about their areas of responsibility. As part of its mission, the Commission held seven public meetings and three public hearings around the country, to gather valuable insights from the public. In 2017, when the Commission issued its [final recommendations](#) to Congress and the President, it

---

<sup>32</sup> Evidence-Based Policymaking Commission Act of 2016, Pub. L. 114-140

provided a roadmap for improving data and analytical capabilities at the national level (CEP, 2017). The recommendations noted the need to bolster state and local government capabilities while supporting a broader ecosystem and infrastructure relevant for all levels of government. Many recommendations focused on leadership roles, data governance, data sharing, program evaluation, and planning processes to build evidence. Although the Evidence Act did not provide resources to implement the law’s mandates, agencies have prioritized their activities and worked within existing budgets and staffing levels.

### **Value to the Government of Evidence Building**

*Without evidence, the federal government is an ineffective fiduciary on behalf of the taxpayer. Unfortunately, in many instances, federal decision-makers do not have access to the data necessary to best inform decisions. In such instances, agencies are unable to show the benefits or impacts of the programs they administer and cannot determine what, if any, unintended consequences are created by programs, or whether programs can be improved.*

*Ensuring policymakers have access to high-quality administrative data is essential for evidence-based policymaking. Administrative data already exists and tested protocols are in place to facilitate merging of information and ensure confidentiality, yet agencies fail to share data with each other in ways that might improve program outcomes.*

*Further, Statutory restrictions often prevent agencies from sharing data with researchers who may be in a position to help the federal government identify needed solutions.*

*The first step in creating a culture of evidence-based policymaking is to determine what data is available and how to best get the data to policymakers. The Evidence-Based Policy Commission Act of 2015 will bring together leading researchers, program administrators, and experts to conduct a thorough study of existing infrastructure and statistical protocols. These*

*individuals will consider various methods of ensuring that policymakers have the access they need while balancing personal privacy and data integrity interests, and make recommendations*

*on how to best approach the issue of federal data access.”*

*[From House Report 114-211](#) Accompanying the Evidence-based Policymaking Commission Act of 2015*

Congress acted on the Evidence Commission’s report in an unusually short period of time. About one year after the report was issued, the bipartisan [Foundations for Evidence-Based Policymaking Act of 2018](#) (Evidence Act) was signed into law (Pub. L. 115-435). The law put



into statute 11 of the 22 Commission recommendations and mandated several structural improvements not covered by the Commission. These improvements included creating a Chief Data Officer for each agency and mandating that agencies create publicly available data inventories and establish data governance and management mechanisms. The law also required agencies to create evaluation plans to examine the effectiveness of their program outcomes, undertake an annual learning agenda to address big questions affecting public policy and program efficiency, and put in place an evaluation officer to oversee the work and build evaluation capacity.

The recommendations from the Evidence Commission and the legal requirements in the Evidence Act are relevant to state and local governments. Together, they may offer a model that could be adapted to build capacity to generate and use evidence at the state and local level across virtually every policy domain.

## The National Data Infrastructure and Evidence Ecosystem as Resource for States

Capacity for evidence building is multifaceted. It ensures organizations have available staff, resources, infrastructure, leadership, and processes to align the need for information with the ability to produce relevant, timely, and objective insights. However, simply allocating increased funding or staff to a data unit may not result in the type of capacity needed to support the cultural shifts, technical analysis, or legal and privacy considerations relevant for modern data analytics. The Evidence Commission outlined capacity in a governmental context at the national level along the dimensions of people, process, and the legal framework.

Within the U.S. federal government, extensive capacity already exists to support data sharing, linkage, and use for a range of analytical purposes. While this capacity is imperfect and brings with it ongoing internal challenges, operational components of this capacity include the federal statistical system; the performance, evaluation, and policy analysis infrastructure; and coordinated privacy and cybersecurity efforts, discussed below in turn.

**Federal Statistical System.** The decentralized statistical system in the U.S. is composed of 13 major agencies and dozens of smaller statistical units. Many of the agencies share a common data use and protection law—the CIPSEA—and others have their own unique legal authorities such as the Census Bureau (Title 13 U.S. Code) and National Center for Education Statistics (20 U.S. Code § 9573). Although it is decentralized, the system is coordinated by the Chief Statistician of the United States at the White House OMB, who maintains a community of practice for leaders in the system through the ICSP. Each agency and unit have its own resource constraints and priorities, but there have been some attempts to recognize the system more comprehensively in budgetary and staffing arrangements as part of the annual President’s Budget request. The statistical system also has multiple arrangements in place for bilateral and multilateral data sharing to capitalize on the capabilities and unique expertise in respective agencies and units. Some of the coordination of the system includes the ongoing development

of a common proposal application for researchers to apply for access to restricted data as well as the use of the [Federal Statistical Research Data Centers](#). The research data centers are managed by the U.S. Census Bureau and include over 30 locations where nongovernmental researchers can access highly restricted, de-identified data from multiple agencies to conduct approved statistical research projects. The infrastructure for accessing the data is part of the Census Bureau's internal information technology (IT) system.

**Performance, Evaluation, and Policy Analysis Infrastructure.** Federal agencies such as the Departments of Housing and Urban Development, Labor, and Education long ago established robust policy analysis offices and units to support the formulation of regulatory and administrative actions, as well as program evaluation activities. In the mid-1990s, as required by the Government Performance and Results Act of 1993 (GPRA), federal agencies bolstered capabilities to measure program performance and outputs. Building on performance measurement capabilities, the Evidence Act requires large federal agencies to maintain the program evaluation function to study program outputs and outcomes. Taken together, the performance, evaluation, and policy analysis infrastructure provide a venue for using administrative records, partnerships with the federal statistical system, and direct coordination with state and local governments to produce actionable insights and recommendations for policy actions.

**Privacy and Cybersecurity Coordination.** Federal agencies are required to maintain key leadership roles for privacy protections, such as the Senior Agency Official for Privacy and the Chief Information Security Officer. Increased attention in recent years to the need to bolster the cybersecurity infrastructure to safeguard systems and data has led to a proliferation of resources and staff. Cybersecurity is only one component of strong privacy safeguards, which also include confidentiality protections. Confidentiality is a subset of privacy and involves managing the risks of deidentification, deploying disclosure avoidance, and maintaining pledges of confidentiality. OMB also plays a central role through its privacy office, the Chief Statistician of the United States, and the Chief Information Officer, namely, to collaboratively and comprehensively deploy government-wide privacy safeguards that ensure information is both protected from potentially harmful uses while encouraging beneficial, authorized uses.

All federal activities are notionally coordinated at a policy level by the White House OMB, which is statutorily charged with coordinating data collection and management, IT systems, performance and evaluation, information policy, and privacy policy (NAPA, 2020). Technical details and standards may then be established by individual agencies, voluntary interagency councils, or voluntarily adopted based on industry or consensus approaches. Such collective approaches leverage the capacity within program offices and units, where administrative activities may supplement protections and capabilities to link or use data.

States will want to establish certain aspects of this capacity to address their unique contexts and needs, discussed in the roadmap below.

## Federal-State Data Partnerships

Federal agencies can serve as partners and resources in aligning and bolstering state and local capabilities nationally. Indeed, there are multiple examples where state partners recognize the critical role the federal statistical and administrative data capabilities play in supporting state capacity (Abazajian and Kleykamp, 2020). Below, examples are described of both long-standing and newer initiatives designed to provide value to all partners.

### *The NCS-X Initiative*

The NCS-X Initiative is a partnership between the U.S. BJS and the FBI. The goal of NCS-X is to transition all law enforcement agencies from reporting data into the older **Uniform Crime Reporting System (UCRS)**, which has been providing aggregate monthly crime statistics since 1930, to reporting into the newer **NIBRS**. **NIBRS** provides more information than the UCRS, such as the ability to provide circumstances and context for crimes like location, time of day, and whether the incident was cleared. Participation is voluntary for both systems, which contain data from cities, academic institutions, and state, county, tribal, and federal law enforcement agencies. The FBI and BJS provide incentives for governments and law enforcement agencies to participate. NCS-X leverages the existing infrastructure of NIBRS and is expected to increase the ability to monitor, respond to, and prevent crime by supporting production of timely, detailed, and accurate national measures of crime incidents. The NCS-X Implementation Team includes representatives from multiple organizations that provide technical assistance to NCS-X sample agencies and state UCR Programs, ranging from technical assessments and consultation to providing marketing assistance and facilitating agency interactions across states. One example of the types of assistance provided is the [NCS-X Playbook](#) for law enforcement agencies.

### *State Justice Statistics Program*

The State Justice Statistics Program administered by BJS through the Office of Justice Programs is a grant program that supports the establishment and operation of Statistical Analysis Centers (SACs) in the states and territories. The SACs collect, analyze, and report statistics on crime and justice to federal, state, and local government and share state-level information nationally.

### *The National Directory of New Hires*

The National Directory of New Hires (NDNH) is a system maintained by the Department of HHS as a repository of quarterly earnings, UI, and new hires records. States provide data to HHS under federal law and can access insights about noncustodial parents, to support implementation of child support enforcement programs. Shared use of the data provides actionable and programmatic uses of information and supports research and evaluation activities that generate summary statistical insights. For example, the Department of Housing and Urban Development used NDNH data to analyze long-term outcomes in the Family Options

Study (HUD, 2016); the statistical uses complemented an administratively collected data asset, while providing valuable insights to inform federal and state policy (Fletcher, 2019).

### *The State Longitudinal Data Systems (SLDS)*

The SLDS were established in 2005. SLDSs have demonstrated value to policymakers and educators in improving the ability to generate real-time, actionable information about student performance as well as capabilities to study long-term educational outcomes. Every state has an SLDS and can link basic educational data, and states are investing additional resources to connect education, workforce, and human services data to reduce barriers to service delivery (Garg et al., 2021). Data linked for administrative purposes produce tremendous benefits for research and evaluation activities, including the ability to study important outcomes and identify which educational strategies are most effective for groups of students. For example, the systems provide resources to state administrators and local practitioners for ensuring they can address major educational and workforce gaps, including learning disruptions from the global coronavirus pandemic. These systems are operated at the state-level; however, the major federal investment in development and operation places SLDS at the nexus of federal-state infrastructure. States receive substantial federal resources to develop education data capacity that generates benefits for both the state and federal governments.

### *The Vital Records System*

The Vital Records System is operated and maintained by states, with substantial federal investment from the HHS NCHS and the Social Security Administration (SSA). Vital statistics are collected to provide insights about morbidity and death, to help government agencies prevent improper payments and monitor health issues that may be leading indicators. The Vital Statistics Cooperative Program—the oldest federal-state data sharing partnership—is a partnership between NCHS and 57 state and local jurisdictions; the Program includes detailed microdata about individual characteristics and causes of death (Rothwell, 2017). States provide data, and in exchange, NCHS provides state funding under contract, and training and technical assistance to standardize data quality. SSA receives a subset of the data from states, including the record of death and date of death for individuals. States are the data provider to the federal government, and the benefits include resources to fund and operate the vital records system as well as electronically coded records based on provided information.

### *The Health Information Technology for Economic and Clinical Health Act (HITECH)*

The HITECH, enacted as part of the American Recovery and Reinvestment Act of 2009, authorizes funds to aid in the adoption and use of EHRs and linkages through Health Information Exchanges (HIE). The Centers for Medicare & Medicaid Services (CMS) and the HHS Office of the National Coordinator for Health Information Technology (ONC) have grant funds for states to close the gaps in interoperability, infrastructure, and other activities. Federal Medical Assistance Percentages funds at the 90/10 matching level support HIE activities such as EHR adoption, linking laboratory or other data sources for Medicaid eligible through HIE, and

supporting hardware and software EHR/HIE linkages at the provider site. CMS can provide funding for state administrative activities related to development of core HIE services (e.g., designing and developing a provider directory, privacy, and security applications, data warehouses), public health infrastructure, electronic Clinical Quality Measurement (eCQM) infrastructure, and provider on-boarding. CMS administration and enforcement of Health Information Portability and Accountability Act Administrative Simplification regulations also promotes interoperable data exchange by means of standards and operating rules. In addition, ONC has been advancing the Standards and Interoperability Framework, the State HIE Cooperative Agreement Program, the Direct Project, the Nationwide Health Information Network Exchange, and the ONC Health Information Technology Certification Program. States may use the 90/10 matching level HITECH administrative funding to update existing HIE infrastructure to align with ONC interoperability and security guidelines and to meet requirements for exchanging data with federal agencies.

### *QCEW*

The QCEW is conducted by the BLS primarily from state UI programs, supplemented by two BLS surveys. It is a quarterly count of employment and wages reported by employers covering more than 95 percent of U.S. jobs available at the county, MSA, state, and national level, by [detailed industry](#). BLS pays the states for these data and works closely with state workforce agencies to review and enhance the QCEW data before they are released. This product has been produced by BLS for decades. The **Local Employment Dynamics Partnership**, a voluntary federal-state partnership started in 1999, built on these efforts. States have agreed to share historical and ongoing administrative records of UI earnings data and QCEW data with the Census Bureau. The Census Bureau then produces a longitudinal data infrastructure from which new statistics can be produced about the dynamics of local employment and the locations of jobs and workers.

### *Census-ERS–FNS Joint Project*

The Census-ERS–FNS Joint Project is a long-term joint research project to acquire administrative data on U.S. Department of Agriculture (USDA) food assistance programs—the SNAP and the Special Supplemental Nutrition Program for WIC—from states and link the data to Census Bureau surveys. The linked data provide insights on how program participation affects participants and who is eligible for assistance but does not participate and why. State SNAP and WIC agencies in participating states send their confidential microdata to the Census Bureau in exchange for state-specific analyses and reports. ERS researchers participate in joint research projects. Census has been able to link Veterans Administration and DOD data to gain insights into veterans decision-making. The project overcame several barriers to share records between two federal agencies and multiple states.

### *TANF Data Collaborative Pilot Initiative*

The TANF Data Collaborative Pilot initiative sponsored by the HHS ACF is designed to build TANF state agency capacity to improve program performance through applied data analytics. The initiative includes eight pilot sites supported for 30 months with funding, intensive training, and technical assistance to support state and local efforts and build strategic partnerships. The selected states were California, Colorado, Michigan, Minnesota, New Jersey, New York, Utah, and Virginia. ACF's outside partners include the MDRC Center for Data Insights, Actionable Intelligence for Social Policy at the University of Pennsylvania, the Applied Data Analytics program of the Coleridge Initiative, and Chapin Hall at the University of Chicago. The Fall 2019 data analytics training program included TANF receipt data and QCEW data for both Illinois and Indiana. The program examined questions such as what characteristics increase an individual's risk of returning to TANF, what factors increase an individual's likelihood of not finding stable employment after leaving TANF, and what factors increase an individual's likelihood of not finding any employment after leaving TANF.

The examples above are just a few of many federal systems, or federally supported programs, operated in partnership with or by states. Each highlights the critical role that intergovernmental cooperation has in the national data infrastructure and evidence ecosystem. In addition, the examples present different models of how data are collected, managed, and used to support insights relevant for administrative decision-making, policymaking, and societal benefit.

### *Federal Data Modernization Efforts Relevant to States*

In recent years, several federal agencies pursued data modernization efforts that were intended to be cross-cutting, with resources and capacity that benefited state and local governments.

#### *Medicaid, Nutrition, and Human Services Interoperability Modernization*

Historically, human services, health, and nutrition assistance programs in most state and local governments had distinct administrative systems that could affect coordination when individuals applied for benefits; in addition, the lack of coordination affected the ability to analyze data across programs. In 2011, HHS and USDA announced a [major shift in policy](#) to encourage states to modernize systems and to facilitate interoperability across federally funded health, human services, and nutrition programs. The policy at the time was a waiver to federal rules requiring a certain match on technology and data investments, meaning that states could leverage the federal government paying for 90 percent of upgrades outside of Medicaid programs. Federal waiver of the rules for certain matches as an exception to federal government-wide guidance (i.e., OMB Circular A-87) gave states an incentive to upgrade systems for more efficient eligibility and enrollment systems expected to realized benefits for program administration and research activities. In practice, the waivers led states to better integrate Medicaid, SNAP, TANF, and other human services systems.

Waiver authority was available for states over a period of more than five years, meaning the funding was largely predictable and sustainable during that period. The waiver supports successful efforts to avoid duplication across traditional program boundaries and to improve access to benefits. Examples of [successful activities](#) include the use of electronic data matching to verify eligibility, conversion of paper-based systems to electronic records, the use of data management and analytics capabilities, and funding for state data hubs. Waiver of the federal rules allowed for blending of funding across the traditional program boundaries to support major systems upgrades and modernization, which in turn resulted in operational efficiencies and demonstrated benefits for the programs. Coordinated guidance across agencies and traditional program silos can create incentives for redesign of state systems.

### *Public Health Interoperability Modernization*

In 2020, the Coronavirus Aid, Relief, and Economic Security Act appropriated \$500 million for [public health data modernization](#) through the U.S. Centers for Disease Control and Prevention (CDC). The investments were intended to support federal, state, and local governments in strengthening the data collection, reporting, and analytical infrastructure for health surveillance and monitoring of diseases across the country. The focus has been designed around coordinating systems, modernizing the infrastructure to enable more rapid data collection, and building better partnerships and analytical capabilities. One major theme in the modernization effort is encouraging interoperability of systems for a seamless exchange of data. Coordination across systems should enable CDC to provide real-time information to the public, effective evaluation of health outcomes, and information for decision-makers in setting public health policy. Modernization also includes efforts to improve public health data standards, such as a strategy to reduce reporting burden from hospitals and states and more automated reporting for systems that involve lab results. In less than a year, the CDC is realizing progress, with new datasets published as open data, pilot projects to build pandemic-ready infrastructure to support states, and reduced time to apply data for contact tracing and other analytics.

The CDC data modernization efforts directly benefit the NCHS, specifically around vital records. Prior to modernization the vital records system was disconnected, error-prone, and sometimes duplicative. Under the CDC's modernization investments, the implementation of greater consistency in processes and data standards is facilitating bi-directional data sharing where both the states and federal government benefit from high-quality data. The vital statistics system and electronic case reporting receive 15 percent of total funding, with plans to allocate across all reporting jurisdictions to increase interoperability in 2021 and 2022. Similarly, the application of data science tools at NCHS is accelerating CDC's ability to produce relevant and timely health statistics. Other planned activities for NCHS include establishing a virtual data enclave to facilitate simpler, secure access to restricted data, for researchers to partner with NCHS in rapidly addressing major public health questions.

### *Data Modernization in the American Rescue Plan*

In March 2021, President Joe Biden signed the American Rescue Plan (ARP), a \$350 billion proposal the Administration presented to Congress to support economic recovery in the wake of the global pandemic. Throughout the Administration's engagement in developing and advocating for the proposal, there were strong signals about the need for state and local governments to apply data and technology to effectively implement ARP investments. In May 2021, the U.S. Department of the Treasury issued guidance to state, local, and tribal governments requiring certain data reporting about programmatic activities, specifically linked to certain requirements and expectations in the Evidence Act. For example, the guidance and corresponding federal regulation explicitly authorizes states to allocate funding received under ARP "to make improvements to data or technology infrastructure and data analytics, as well as program evaluations" (U.S. Department of Treasury, 2021). It is too early to say how states in general are using this funding flexibility, in coming months this will become clearer because the federal government requires a reporting category for data and evidence investments.

Some agencies also have announced targeted resources to support data infrastructure in states, in addition to the general ARP guidance from the Treasury Department. For example, the Economic Development Administration's Good Jobs Challenge includes \$500 million to support data and capacity that leads to scaling evidence-based interventions and for conducting evaluations (Economic Development Administration, 2021). In addition, given the guidance on implementation from the Treasury Department, additional investment opportunities for states will likely be forthcoming.

## Building State Capacity for Evidence Building

At the same time that federal initiatives have been underway in recent years, many states have built robust data capabilities, both within the state and in collaboration with other like-minded states. Both state and local activities have focused on enhancing capacity for evidence building and informing operational decision-making in real-time across many topical areas. Some efforts have increased transparency for the public about how government operates. Many states have realized that fragmented and siloed operational or programmatic structures can inhibit data sharing, linkage, and use. Even without legal barriers to sharing, organizational culture, scarce resources, and bureaucratic red tape can create frustrating roadblocks that can seem insurmountable. Yet some states have succeeded in overcoming these roadblocks. This section offers examples of pathways and foundational activities that can help states succeed.

One valuable resource is the annual [\*Invest in What Works State Standard of Excellence\*](#) published by Results for America, a national non-profit that works to promote data and evidence use. This voluntary self-assessment or framework reviews basic capacity characteristics across participating states, using a standard that includes explicit criteria about data governance and leadership, data policies, data infrastructure and data use (RFA, 2021). The framework



presented in the 2020 assessment identified 169 examples of promising systems and programs across 35 participating states (RFA, 2020).

We expand on the RFA framework to include other key aspects of capacity that are relevant at any level of government based on successful models, such as transparency, legal framework, and incentives for data sharing. Our approach is intended to reflect the collaborative nature of intergovernmental cooperation on data infrastructure, including the joint needs of states and federal agencies in using state-collected or managed data. Together, the key characteristics to weigh and consider in building, maintaining, and adapting state capacity are briefly described below.

### **Key Considerations for Increasing State Capacity to Build Effective Data Infrastructure**

- **Leadership and Strategy:** Have priority areas been identified by decision-makers to include in high-value data sharing and analytics projects? Have senior leaders expressed support for improving data infrastructure and have they identified, resourced, and empowered a responsible senior-level individual (e.g., chief data officer) for the state? Are the governance approach and planned action items articulated in a publicly accessible data strategy?
  - **Legal and Regulatory Authority and Policies:** Have experts reviewed appropriate state legal authorities to ensure sufficient capabilities exist in law to enable data collection, sharing, analysis, protection, and responsible use? Does the state have policies that outline expectations to use data for informing decisions, sharing data across organizations when appropriate, and safeguarding sensitive information?
  - **Governance:** Does the state have effective mechanisms to govern data practices and quality, including developing a data inventory and promoting accessibility of information? Are data sharing practices supported with common memorandum of understanding and with procedures generalizable to state agencies?
  - **Transparency of Processes and Uses:** Are mechanisms available to communicate openly with the public and stakeholders about data practices and uses, including the value and benefits of using data?
  - **Accessibility of Data:** Are procedures in place to enable applications and efficient approvals to qualified users, for access to government-managed data? Are open data assets and de-identified data available when possible, for public access?
  - **Oversight and Accountability:** Do external stakeholders routinely participate in processes that provide a diverse range of perspectives about data governance and use? Are legislative oversight procedures clear? Have routine processes been established to audit or periodically review state data infrastructure and practices?
  - **Sustainability:** Are resources available to provide for sustained systems, procedures, and personnel with expertise to collect, manage, protect, and use data? Is there a mechanism for training employees for ongoing collaborative work with their data?
-

## Data Leadership and Strategy

Data leadership is key to advancing more effective use of data. While many states have created the position of a Chief Data Officer, leadership can come from multiple sources. Often leadership arises from agency heads who understand that they need better data to improve their programs and the programmatic outcomes for the public. These leaders can have substantial influence and garner support from their state legislature and governor's office when the value of having better information is made clear. Value has often been clearest when focused on topics, such as tackling unemployment or reducing crime.

Data leadership requires having a strategic vision. At the federal level, a key driver of change has been the [Federal Data Strategy](#), developed by White House OMB officials including the Deputy Director for Management, the Chief Statistician of the U.S. and the U.S. Chief Information Officer. The work of developing the strategy was accomplished by a team of representatives from agencies across government and with important input from the public, gathered through roundtables, listening sessions, and solicitations for public comment. The Federal Data Strategy provides a government-wide framework that includes principles, practices, and an annual action plan. As stated on its website, "*The mission of the Federal Data Strategy is to fully leverage the value of federal data for mission, service, and the public good by guiding the Federal Government in practicing ethical governance, conscious design, and a learning culture.*"

The Federal Data Strategy is a model for how top-down leadership can drive change across government. Strategy development was incorporated into the President's Management Agenda. This agenda is taken up by the President's Management Council, comprising the Deputy Secretaries of the Cabinet agencies. Quarterly progress reports for agenda items are published on a website and delivered to the Council, which is chaired by the Deputy Director for Management at OMB. The cross-agency effort has high-priority and visibility. Deadlines were established, and progress was reported weekly. Funding was made available for the effort. The strategy was developed in about 18 months, including the first year action plan.

The action plan itself was organized into three categories: 1) actions taken by all agencies; 2) actions taken by the community of practice; and 3) shared solutions actions undertaken by one agency on behalf of other agencies.

The excerpt from the Federal Data Strategy below shows the types of actions taken by every agency and highlights many of the data standards of excellence discussed here. For example, agencies are required to set up a data governance body, assess the maturity of their infrastructure, and publish data inventories. Additionally, a community of practice was required to establish a cross-agency Chief Data Officer Council and focus on improvements in key areas such as financial management and geospatial data. Shared solutions included in the action plan include developing a data ethics framework, creating a toolkit for agencies on best methods for protecting privacy and confidentiality of sensitive data, and establishing a central data policy

coordinating committee. Each of the items in the action plan is supported by either a statute, an executive order, or official OMB guidance.

Agency Actions			
	Foundations for Evidence-Based Policymaking Act of 2018 and Associated OMB Guidance	Executive Order on Maintaining American Leadership in Artificial Intelligence	Improving Implementation of the Information Quality Act (M-19-15)
1. Identify Data Needs to Answer Priority Agency Questions	✓		✓
2. Constitute a Diverse Data Governance Body	✓		✓
3. Assess Data and Related Infrastructure Maturity			✓
4. Identify Opportunities to Increase Staff Data Skills	✓		
5. Identify Priority Data Assets for Agency Open Data Plans	✓	✓	✓
6. Publish and Update Data Inventories	✓		

Federal Data Strategy, First Year Action Plan, [www.strategy.data.gov](http://www.strategy.data.gov)

For states, leadership from the top can help create momentum and motivation remove barriers to data sharing and collaboration. Having a governor and state legislature drive innovation and change can free up resources and help cross-agency coordination. An executive order from the governor or enacted legislation can be helpful in focusing attention on high-priority policy areas and garnering resources to pay for people, training, and infrastructure.

One model of this top-down approach is Ohio. In 2019, the Ohio governor consolidated state data systems under an [executive order](#), establishing the **InnovateOhio Platform**. Through senior-level leadership provided by the Lieutenant Governor’s office, the platform was established as a shared resource, to link datasets intentionally across programs. The executive order established certain characteristics, many identified below, such as an expectation for data sharing across agencies unless prohibited in law. The Lieutenant Governor takes an active role in implementing the infrastructure, including serving as its director and [supporting public messaging](#) about the value of the system, the efficiencies achieved, and the benefits constituents directly receive from better data sharing.

Another model involves leadership shared across states, driven by agency leaders and engaged partners. One of the most successful examples of multistate collaboration is the [Midwest Collaborative for Evidence-Based Policy Making: Transitions in Education and Workforce](#). [The Collaborative focuses on](#) development of core data analytics projects and competencies for a group of Midwestern state workforce and education agencies, working with academic partners, that began convening in 2018. Several workshops and training sessions for state agency employees were held to identify a roadmap and strategy that would work for multiple states. The need was clear: members of the public were crossing state lines to live, work, and go to school, and to understand effective workforce and education strategies, states needed to share data with each other. The states used a secure shared cloud platform supplied

by a non-profit partner, the Coleridge Initiative, which also provided training to state employees on how to work with their shared data. The platform and training allowed the collaborative work to move forward more quickly, without having to coordinate multistate infrastructure investments. Hands-on work with the data also inspired new ideas about how to improve data quality and interoperability. One important aspect of the Collaborative is that they have implemented tiered access to protect privacy and confidentiality of sensitive data. The Governors and their staff, as well as the workforce boards, did not want synthetic or artificial data, so Coleridge produced confidential summary tabs of data that are password protected. Data have been supplied by the Missouri Departments of Employment Security, Labor, Corrections, and Higher Education; the Indiana Departments of Workforce Development and Higher Education; the Ohio Longitudinal Data Archive; and the Illinois Department of Employment Security. Kentucky, Tennessee, and Michigan are planning future training and work sessions. Funders have included several philanthropic foundations. A few of the topics addressed through the workshops and training sessions have included the earnings and employment outcomes of credentials and degrees, remedial skills training, vocational training, or apprenticeship training; how well postsecondary student majors match to in-demand occupations; and earnings across demographic groups and what factors explain observed differences.

## Legal and Regulatory Authority and Policies

Laws and regulations provide formal mechanisms for agencies to promote data access and use. The federal Evidence Act, for example, includes the requirement for OMB to issue regulations to implement several provisions, including the expanded sharing of data across agencies for evidence building activities. Many states implement federal laws through cooperative federalism, in addition to state-specific authorities; for this reason, state request for improved clarity and consistency in interpreting federal laws and regulations are common. Federal authorities offer examples of processes that can be leveraged at the state-level.

Similarly, important elements were incorporated in the Ohio executive order referenced in section 3.1 above. The Ohio executive order cited a section of Ohio state code that had been revised to allow data sharing between agencies ([Section 125.32, Ohio Revised Code](#)). This change in the law enabled “...an enterprise data management and analytics program to gather, combine, and analyze data provided by one or more agencies to measure the outcome of state-funded programs, develop policies to promote the effective, efficient, and best use of state resources, and to identify, prevent, or eliminate the fraudulent use of state funds, state resources, or state programs. Participating state agencies may use data gathered under the program for these purposes.” State agencies were required to provide data for use under the program. Ohio’s approach is similar to that of the Evidence Act, which has a presumption of accessibility to data for statistical agencies.

### Presumption of Accessibility to Data Assets

*“ACCESSIBILITY OF DATA ASSETS.—The head of an agency shall, to the extent practicable, make any data asset maintained by the agency available, upon request, to any statistical agency or unit for purposes of developing evidence. “(b) LIMITATIONS.—Subsection (a) does not apply to any data asset that is subject to a statute that— “(1) prohibits the sharing or intended use of such asset in a manner as to leave no discretion on the issue; or “(2) if enacted after the date of the enactment of this section, specifically cites to this paragraph. “(c) REGULATIONS.—The Director shall prescribe regulations for agencies to carry out this section. Such regulations shall— “(1) require the timely provision of data assets under subsection (a); “(2) provide a list of statutes that exempt agencies from the requirement under subsection (a) pursuant to subsection (b)(1); “(3) establish clear and consistent standards, to the extent possible, for complying with section 552a of title 5 (commonly known as the ‘Privacy Act of 1974’) and any other applicable law requiring the protection and confidentiality of individually identifiable information; and “(4) require a transparent process for statistical agencies and units to request data assets from agencies and for agencies to respond to such requests.”*

*Foundations for Evidence-Based Policymaking Act (Pub.L 115-435) Title III, Part D, Section 3581*

Ohio passed legislation enabling data sharing, reinforced by the governor through creation of a specific platform through which agencies must share their data. The federal Evidence Act mandated all agencies to share data only with the principal statistical agencies, but other agencies with programmatic data are not prohibited from sharing data with each other, to conduct statistical activities and conduct evidence building and evaluations. Federal data sharing remains decentralized, but the Evidence Act establishes an advisory committee to determine whether a centralized national data service should be established to combine sensitive data for statistical analytical purposes (including program evaluation).

Similarly, the governor of Indiana issued an [Executive Order](#) that mandated data sharing and established the governor’s Management and Performance Hub. The hub is discussed in more detail in section 3.4 below. Both the Ohio and Indiana initiatives, driven by Executive Order and statute, have been highly successful.

One important element of success for the Midwest Collaborative has been the use of common agreements and a clear legal framework. Having these in place is critical for sharing data across agencies and states. Further, bringing in other important state officials such as a Chief Privacy Officer or Chief IT Security Official can move forward a statewide effort that may involve multiple regulations around different types of data.

The Evidence Act mandates that OMB promulgate regulations and guidance around the data sharing enabled by the law. These regulations include establishing standards for agencies to: 1) categorize the sensitivity of their data and assign appropriate levels of accessibility; 2) determine whether a less sensitive data set could be produced that would increase accessibility; 3) conduct risk assessments for a data asset prior to release; and 4) be transparent about their processes and make the information easy to understand and publicly available. Although these requirements would apply to statistical agencies, they are generalizable to any agencies that make potentially sensitive data available to the public or to researchers.

## Data Governance

Data governance may be defined as “the specification of decision rights and an accountability framework to ensure the appropriate behavior in the valuation, creation, consumption and control of data and analytics” (Gartner, 2021). Put another way, data governance defines the roles and responsibilities of all participants involved in the collection, processing, use, and dissemination of an entity’s data and includes establishing policies, procedures, metrics, and accountability mechanisms.

The federal Evidence Act creates the position of Chief Data Officer (CDO) for each agency and specifies that this is to be a nonpolitical appointment. The CDO is required to be selected on the basis of qualifications such as “...*demonstrated training and experience in data management, governance (including creation, application, and maintenance of data standards), collection, analysis, protection, use, and dissemination, including with respect to any statistical and related techniques to protect and de-identify confidential data.*”<sup>33</sup>

Establishing and implementing data governance policies can be challenging in a government environment, where ownership of the data can be spread among multiple agencies. Only about 50 percent of states have appointed CDOs or their equivalent. Several states (e.g., Arkansas, Connecticut, and Texas) have adopted the idea of Agency Data Officers to support governance across departments. By establishing a CDO in each agency and assigning responsibility to that person for data governance, the states align with the Evidence Act in identifying data governance as important to using data to achieve better outcomes. The Evidence Act establishes a CDO Council composed of all the agency CDOs. The function of the Council is to “...(1) *establish Governmentwide best practices for the use, protection, dissemination, and generation of data;* (2) *promote and encourage data sharing agreements between agencies;*(3) *identify ways in which agencies can improve upon the production of evidence for use in policymaking;* (4) *consult with the public and engage with private users of Government data and other stakeholders on how to improve access to data assets of the Federal Government;* and (5) *identify and evaluate new technology solutions for improving the collection and use of data.*”<sup>34</sup>

---

<sup>33</sup> Pub.L 115-435 Title II Section 3520

<sup>34</sup> Ibid Section 3520A

Likewise, at the state-level, governance boards or steering committees are needed, to allow for sharing resources and knowledge in addition to data.

At the federal level, some of the most significant barriers to data sharing have yet to be addressed, but some states have moved forward. For example, putting in place a common agreement or Memorandum of Agreement for data sharing can save years of effort compared with the process of each agency negotiating unique agreements each time data are to be exchanged or shared. In addition, having a common application process for requesting permission to access sensitive data for specific projects can streamline the approval process considerably. There is no common federal Memorandum of Agreement mandated for agency use. Every agreement is negotiated separately, and the process of securing an agreement can be even more complicated if multiple agencies are involved. In 2014, OMB offered a complete model agreement for a standard Memorandum of Agreement for data sharing in guidance to agencies ([M-14-06](#)). However, adoption was not mandatory, and agency legal departments continued to craft unique agreements around each project.

The Commission on Evidence-Based Policymaking identified another challenge, namely, that, “Cumbersome and idiosyncratic data access procedures create confusion, impose unnecessary costs, and are a barrier to evidence building, without always providing significant privacy benefits.”<sup>35</sup> Many federal agencies provide direct access to researchers and have developed unique application and approval processes to grant that access. Such an approach becomes complicated and time consuming if researchers want to access data from multiple agencies. The Evidence Act required OMB to set up a single portal for applications for data from statistical agencies; however, this provision did not address the need to streamline the business process for approvals once an application is received.

States have established data governance structures to address some of these challenges. For example, North Carolina has used centralized data governance to increase productivity and effectiveness through the [North Carolina Government Data Analytics Center \(GDAC\)](#), which is a central organizational structure for comprehensive data management. As part of the state Department of IT, GDAC services are closely aligned with the Chief Information Officer (CIO) for systems modifications and analytical needs. GDAC’s approach facilitates collaboration with state agencies to define priorities, obtain access to relevant data assets, and provide for integration of analytical capabilities. For example, the [Criminal Justice Law Enforcement Data Services](#) support offered by GDAC facilitates an integrated criminal justice data system. GDAC also collaborates across agencies for operational activities related to compliance and fraud, and coordinates the state integrated data system for education and workforce data. Finally, GDAC’s authorization in [state law](#) led to the creation of a data governance program that includes the use

---

<sup>35</sup> U.S. Commission on Evidence-Based Policymaking (CEP). *The Promise of Evidence-Based Policymaking: Final Report of the Commission on Evidence-Based Policymaking*. Washington, D.C.: GPO, 2017. p35

of a standardized Memorandum of Agreement for sharing data across governmental jurisdictions, with strong privacy protections.

Similarly, the [Kentucky Center for Statistics](#) (KYSTATS) links data from across 15 state agencies to conduct analysis on education and workforce programs. KYSTATS is transparent about their capabilities, which include a [common portal for requesting data access](#), data use restrictions, a common data sharing agreement, and a data dictionary. Each of these practices streamline the efforts from data requestors—often researchers—to learn what data are available for access and use and how to request access; KYSTATS reflects comprehensive data governance activities. Further, in 2019, Kentucky began streamlining its data sharing across agencies in the [Enterprise Data Management Policy](#), which is implemented in conjunction with the state [Enterprise Privacy Policy](#).

**Arizona's Department of Administration** also developed and implemented an enterprise [memorandum of understanding](#) for data sharing, now used by dozens of state agencies. The approach aligns with Arizona's statewide [policy for interoperability](#) on government data and a policy that outlines requirements and expectations for data governance in the state.

## Transparency of Processes and Uses

Transparency is key to establishing and maintaining public trust in how the government uses data. There are several aspects to transparency to incorporate into building capability and infrastructure. However, transparency starts with the assumption that the public has a right to know how information is collected, used, and shared and how society benefits from these activities.

Statistics Canada, the centralized national statistics office of Canada provides a model for data transparency. Statistics Canada maximizes the use of administrative program data in producing statistics that supplement its surveys and censuses. The agency has created a set of [principles](#) with strong legal, policy, and organizational safeguards. These principles include strong privacy protections and consider the effects on privacy related to redirecting program data collected on individuals to new purposes, particularly when combined with other data. In addition, Statistics Canada has pre-approved certain types of data linkages with low risk of privacy violations. Further, Statistics Canada provides a summary of all linkages on its website.

Many states do not have open data portals. However, the COVID pandemic has demonstrated the value of providing important data to the public and to researchers. One approach is demonstrated by [Mississippi LifeTracks](#), a state longitudinal data system that enables analysis of administrative data from multiple state agencies. One primary purpose of LifeTracks is to assess education and workforce outcomes. The system received substantial initial funding through federal government investments from the Department of Education and receives annual state appropriations. When establishing LifeTracks, Mississippi engaged with stakeholders to prioritize transparency in the design and uses of the infrastructure, for example, transparently



sharing information on its public website, to promote and increase accountability. The LifeTracks website includes reports and projects completed from the system's data, demonstrating value while sharing relevant information.

The Indiana [Management and Performance Hub](#) is another example of a useful portal that provides transparency on the workings of state government and access to data. It includes the following features: a data hub with more than 160 secure and de-identified datasets of actionable data covering subject matter areas within state government; a searchable data catalog with definitions to available fields from datasets on the Indiana Data Hub; access to a secure collaboration environment that expedites research and analysis by bringing research teams, code, and data together; and a statewide data proficiency program for state employees.

Similarly, Maryland enacted the [Open Data Act](#) (State Government, Section 10-1501 through 10-1504) and as a result, established an [Open Data Council](#) with 37 members, including Cabinet agencies, five county representatives, and five private citizens. The Council Chair is the Secretary of the Department of IT, with the Director of the Governor's Office of Performance Improvement serving as the Vice-Chair. Maryland's [Open Data Portal](#) offers access to several useful data sets about demographics and business/economic activities in the state, government data to provide transparency of government operations, the ability to provide feedback, and mapping information.

## Accessibility for Data Use

Accessibility encompasses many functions and activities. It may mean providing information to potential data users through data inventories, clear instructions on how to request data, and data management policies that emphasize data quality and interoperability. In addition, states may collaborate to eliminate traditional barriers to access. The examples from Mississippi, Indiana, and Maryland above offer lessons on accessibility. Below, we highlight other state examples to enhance access to data for both the public and state employees.

**COVID UI dashboards**, initially created by the state of Illinois and the same non-profit organization supporting the Midwest Collaborative, are now being used by nine states in the Collaborative. The dashboards combine data from multiple sources, including up-to-date UI claims (initial and continuing) with crosstabs summarizing the education levels, age, race and sex of unemployed workers by county and industry, how much of their earnings have been lost, and estimates about how quickly their benefits are being used. The crosstabs can be linked to state TANF and SNAP records to determine the impact on low-income families. A philanthropy has provided funding to expand the existing infrastructure to enable the states and researchers to access, use, and build tools to inform decision-making about the loss of jobs in the pandemic; these tools may then be used at the state and sub-state-level across the country. The project will develop a standardized approach to employment and benefit use histories, industry characteristics, and outcome measures. Models are being developed on the impact of different sector-specific or educational interventions on getting successful jobs. Models forecast the

impact of the pandemic on individual and family participation in state income transfer, nutrition, and social service programs.

These innovations would not have been possible without a focus on increasing accessibility to key data. The Evidence Act recognized the need to improve accessibility by including provisions to encourage data sharing within a strong privacy framework. One particular provision, for example, was based on the Evidence Commission's recommendation that the federal government should change the default to encourage data sharing in conjunction with privacy safeguards, to realize the strategic value of data collected by the government when possible.

## Oversight and Accountability

Approaches for oversight and accountability are expected to vary by state, reflecting political priorities and structures and aligning with expectations from the state population and relevant stakeholders. Multiple mechanisms exist for enabling and encouraging oversight activities that promote accountability for the public from government's use and management of data. In general, such processes should include explicit participation, advice, and feedback from a diverse range of stakeholders related to all aspects of the data pipeline, including data collection, data protection, and data use. When Congress passed the federal Evidence Act, it included an [Advisory Committee on Data for Evidence Building](#) as one construct to proactively support new strategies moving forward and to provide strategic advice and consultation from experts about implementing key data provisions in the Evidence Act. The federal government also has established processes for publishing public notices about data collection requests required by the PRA and permissible uses under the Privacy Act of 1974.

Many states have established advisory boards and committees to support feedback on data priorities. In Kentucky, the Education and Workforce Statistics Board provides oversight and direction for the state longitudinal data system with integrated education and workforce data. The Board was established in [state law](#) with an advisory board to provide direction and oversee compliance with expectations, use limitations, and privacy protections.

Other approaches for oversight could include periodic annual reports to the legislature and public or an independent evaluation or audit process.

## Sustainability

A secure and stable funding source for state data activities is typically necessary to ensure continuity and the capability to meet future demands on the infrastructure. Because data activities may span multiple years, predictability in funding is an important characteristic for sustained capacity. Planning for and building capacity as an ad hoc exercise can be costly and inefficient. Instead, states should consider how resources and funding are provided, to enable a sustained presence for key priorities. A 2018 review of state data capacity noted that state

officials cited staffing and expertise as the single greatest challenge for using data in the state (Pew, 2018). Over a quarter of the respondents also identified funding as a major limitation.

In the District of Columbia, the capacity to organize and use data was bolstered through the creation of the [Lab at DC](#). Funded initially through a philanthropic contribution from Arnold Ventures, the Lab was eventually supported primarily through a sustained investment from DC government, following successful completion of multiple projects that demonstrated the value of the endeavor. While the Lab did not address all issues or data capacity in the District of Columbia, it did support capacity-building efforts and achieved a sustainable funding model to enable staff and expertise to support a range of priority projects.

There are multiple approaches for developing sustained capacity and resources to implement state data activities. As discussed above, funding may be available from federal agencies and programs to support a broad range of initial infrastructure and development activities. In addition, states may choose to identify philanthropic contributions for one-time or targeted priorities and activities that build capacity. However, sustained investment will likely require direct state appropriations, which could be supplemented by user fees for access to restricted data, for example.

One approach for enabling sustained resources may involve setting aside a portion of program implementation funds to ensure data infrastructure is adequately resourced to support program implementation. The Evidence Commission recommended this approach at the federal level, and the funding mechanism is equally relevant for state data capacity (CEP, 2017; Fatherree and Hart, 2019). Another successful approach is a shared resource model, where infrastructure and data, as well as training, are shared across government.

Finally, resources should be used to support individual- and organizational-level training on an ongoing basis. Staff at all levels of an organization should be exposed to resources for data literacy and use, including senior leaders and managers. Staff working with data collection or curation should also receive frequent updates and training to ensure the latest privacy and cybersecurity protections and practice are deployed.

## A Roadmap for State Data Capacity and Infrastructure

Every state and local government has a role to play in supporting evidence-based policymaking in its own context, to improve services and operations for local populations. For a state government seeking to build increased capacity and infrastructure to use data across agencies and programs, there are foundational building blocks. These building blocks can support state officials in implementing evidence-informed decision-making, to understand outcomes and to better describe characteristics or the economy, people, and programs in a state. Building this capacity and infrastructure is challenging, can be resource intensive, and always requires clear leadership.

Building blocks for the success of state data initiatives include the key characteristics of capacity outlined in section 3 above. These characteristics can be developed in tandem, though some may need to occur before others, based on specific state attributes, needs, and constituencies. The roadmap presented below for state capacity suggests approaches that can work well, though states should expect to modify steps based on their own situations. Roadmap steps may not be linear or sequential; many may require repetition or may be mutually inclusive to be developed concurrently with other steps.

As roadmap steps are considered in a state, a foundation should be established and expanded over time, with a clear project or focus in mind. Establishing priorities and areas of immediate attention will help build success stories and momentum for further success. For this reason, planning from the outset to scale approaches may prove fruitful for many states.

The key characteristics and examples identified in this white paper identify several clear opportunities for states to engage in evidence-based policymaking while building additional capacity and leveraging both national and peer capabilities. Roadmap steps reflect considerations in the federal Evidence Act along with state-specific considerations, as follows:

1. **Establish and empower state and local data leadership roles.** Anyone can be a leader when it comes to improving data quality and access; however, recognized senior-level leaders within states and agencies signal the effort as a priority to colleagues and partners. Leaders are instrumental in securing resources, personnel, and space on the policy agenda for improvements. States can establish a recognized CDO, ideally recognized, resourced, and empowered by elected officials and civil servants alike. In virtually every successful data initiative, a clear leader emerged to champion implementation of the project and to align interests, resources, and expertise to drive change. A recognized leader such as a CDO can also be instrumental in fostering new partnerships, for example, with industry, nonprofits, and research institutions in a state. Designated data leaders can coordinate with other C-suite officials such as the CIO to support effective implementation of systems necessary for robust data analysis and data governance. Initial leadership can come from agency heads who have a clear, immediate need for better information.
2. **Identify the Local Knowledge Needs and Data Gaps Based on Priorities.** All states will have limitations in resources and must prioritize how to implement data initiatives for modernization and capacity-building. There are multiple planning processes relevant to address priority data issues. A basic consideration is how to best identify where informational needs exist and how to devise a plan to address these needs. One approach is to develop a learning agenda, or a strategic plan for research and evaluation that asks first about the questions to be addressed, then identifies existing assets and areas where new data collection or data sharing may be necessary (Newcomer et al. 2021). To the extent possible, aligning such efforts with strategic planning or other administrative priorities will encourage leadership and development of resources about the selected topics.
3. **Review Legal Authorities to Share, Protect, and Use Data.** Identifying and understanding existing legal authorities in states is critical to determining what, if any, legal modifications

may be necessary to share and protect data while planning for uses. Such a review should determine whether data sharing is considered the default and whether unintentional prohibitions or limits exist on data sharing or use. Once a data champion has been identified and priorities established, conducting a legal review may be assigned to agencies, the CDO, a research institution, or an independent commission of experts. Reviews of legal authorities should be shared with relevant executive and legislative actors. This will encourage resolution of provisions that may unintentionally inhibit data sharing and areas that may need stronger privacy protections. Further, sharing such reviews may encourage recognition of how existing authorities may be used to simultaneously support program implementation and bolster data infrastructure or capacity.

4. **Provide Feedback to Federal Partners about Data Needs and Priorities.** States operate numerous programs that are federally funded or mandated. For this reason, states should be encouraged to actively share insights and perspectives about data needs and gaps, especially where there may be a role for the federal government to support states with resources, technical assistance, training, or interstate data coordination with standards and analysis. For example, states could participate in the federal agency learning agenda process required by the federal Evidence Act. One aspect of learning agenda feedback may be requesting support from federal partners to coordinate where there are data gaps across states. Another area might be making requests and clarifying areas where blending funding across multiple federal funding sources could help states develop and operate increasingly integrated systems to improve operations and research. States may encounter areas where federal law, regulation, or guidance impedes state-level priorities, sometimes unintentionally. In addition, state should be encouraged to use the formal learning agenda process to recognize where improved guidance from federal agencies could support state efforts.
5. **Establish a cross-agency governance process and data strategy.** State-level data governance encounters issues unique to each individual state, yet the presence of a clear, consistent, and coordinated governance process provides coherence to data activities for civil servants and partners. States can ensure that the capability to coordinate among senior officials in key agencies occurs frequently. Some states may choose to set up committees of data officials across agencies while others may choose a more centralized approach within a statistics unit or administrative office. Governance processes designed in different ways may each be paired with development of a data strategy to gather stakeholder input and feedback. The Federal Data Strategy received extensive feedback, including from industry and researchers; state-level strategies would likely receive similar attention.
6. **Execute on Data Governance by Integrating Local Data Where Feasible to Prepare for Linkage and Sharing.** Government agencies and institutions should know what information they have access to and the relevant quality of that information. Establishing robust data governance practices is one approach to requiring an inventory of available data assets along with relevant characteristics as metadata. Having data governance processes in place can simplify implementation while ensuring data assets are approached with future potential data sharing in mind. Further, integration is increasingly possible across education, workforce, criminal justice, and public health systems to support improved insights about

outcomes and to encourage individuals eligible for government benefits and services to apply for and receive them.

7. **Institute Transparent Approaches to Using Data.** While some states rapidly developed approaches for various initiatives that prioritized transparency, this was not a consistent theme across many organizations. Moving forward, as more individuals in states are subject to federal law and practices, it will be critical to ensure that transparent approaches are available to demonstrate the value of data management and use. Transparency can be encouraged using websites and portals to share insights about available data assets and particular projects, including how data are used for a benefit. It can also be championed by governors and other senior officials who speak to how the approaches benefit individuals in each state. Sharing information with the public or stakeholders should be viewed as a requirement for advancing the use of data collection, management, and use.
8. **Partner with existing entities to promote rapid progress and pilot projects.** The need to develop quick wins and success stories cannot be overstated. States could partner with other states, existing research institutions, or nonprofits such as the Coleridge Initiative to promote rapid progress in modernizing data infrastructure at the state-level. Pilot projects are especially valuable because they can be relatively low cost, reach a range of audiences, and provide clear narrative about the value proposition for better using data in practice. Many federal programs include waiver authorities that can be leveraged to test new ideas and support data infrastructure, for example, through Medicaid and some human services programs funded in part by the federal government.
9. **Establish responsive engagement processes and oversight mechanisms.** Data activities should be reviewed periodically by an independent organization to ensure that intended goals and objectives are being achieved. States can support this approach by explicitly recognizing the need for responsive engagement from knowledgeable individuals. In addition, states can encourage the institution of certain oversight mechanisms in legislative bodies or with voluntary advisory committees. For example, states could opt to establish a new independent advisory committee charged with assessing progress in implementing data governance procedures and with addressing underlying concerns about state data.
10. **Plan for and implement robust privacy safeguards.** Advancing data uses and data sharing at the state-level can responsibly occur alongside data access by ensuring that privacy safeguards are robust. The Evidence Act suggests that privacy and data access improvements can be reinforcing and concurrent. States must plan for strong cybersecurity approaches and to increase training, cultural responsiveness, and awareness of deidentification techniques and other privacy safeguards.

Roadmap steps can be implemented at the same time. However, identifying leaders who in turn can secure sustained resources may be a relevant starting point in most states that are beginning to build more robust data infrastructure and capacity. Similarly, publishing a comprehensive data strategy or learning agenda may generate many ideas for how states can most effectively participate in such initiatives.

## Focus on Improving Collection, Access, and Use of Localized Firearms Data

The expert panel commissioned by Arnold Ventures and convened by NORC at the University of Chicago to study firearms data observed fragmentation and gaps in the data ecosystem at the state and national level. Such fragmentation can be reduced by incorporating the roadmap from the previous section, regarding issues specific to firearms data. Indeed, some state and local jurisdictions have begun to improve capacity and infrastructure related to criminal justice topics; these efforts could be expanded upon to improve availability and access to information about firearms.

Allegheny County, Pennsylvania, which includes Pittsburgh, incorporates data from a wide range of programs into a portal for easy access and use by both the public and decision-makers. The portal includes data from criminal cases and the local jail as part of a larger warehouse of information used to identify strategies for early interventions and prevention among individual receiving services or with active criminal cases. Even with the initial infrastructure in place, the portal has gaps in coverage with more than 100 other police departments within the county. Summary tabulations are available to the public as open data through the county's QuickCount website, which enables the ability to monitor trends and view some additional detail for analysis by program, gender, age, ethnicity, and race.

In Maryland, a consortium formed the Baltimore Neighborhood Indicators Alliance with input from nonprofits, foundations, community members and the city government. The consortium brings together key data about the community, with visualizations and open data to support city accountability on specific indicators. For example, the portal includes a range of information for the city available by neighborhood on crime rates, shootings, gun-related homicides, arrest rates, and other relevant indicators.

Each of the local jurisdictions described above has practices in place to support access, governance of data, and transparency, in line with portions of the roadmap above. However, each also has room for improvement in addressing the needs for data linkage and sharing to strengthen the insights available using existing data. They offer examples of a starting point for other state and local jurisdictions that aspire to establish enhanced capacity for evidence-based decision-making.

National infrastructure may provide supports, efficiencies, and an economy of scale in supporting state and local jurisdictions in developing such an infrastructure on many topics, including firearms data. This is especially the case when a jurisdiction recognizes a need to integrate public health information while deploying strong cybersecurity safeguards and confidentiality protections. Roman (2020) has identified benefits for considering integrated approaches to firearms data and designing a data infrastructure that could facilitate improved data sharing capabilities in a secure environment. Leveraging the federal and more local

capacity, including available resources and systems, could accelerate approaches for bolstering the national, state, and local data ecosystem more rapidly and effectively in coming years.

Many states collect substantial data from local health and public safety departments, relevant to addressing firearms-related violence. States need not wait for the federal government to lead the way on data linkages. As demonstrated by the proactive and successful efforts of state consortia examining workforce and education data, sharing data in a secure environment that protects privacy and confidentiality of sensitive data is an effective approach to tackling these complex issues and does not require federal involvement.

## Key Findings:

This report offers a set of key findings for states, based on the lessons learned and experiences described above.

- Bipartisan support is important. There is value in having better information to understand outcomes that individuals from different political or ideological perspectives can agree on, such as better service delivery, more effective use of taxpayer dollars, increased emergency preparedness, economic development and job creation, improved health and safety, and full employment.
- A short-lived commission created by a legislature (or governor) could make recommendations for legislation. The commission should have members with technical expertise, not be overtly political, and should take input from a variety of public interests as well as relevant state and local agencies. It is important for groups with specific interests—such as privacy advocates, academic researchers, communities whose data may be shared across agencies because they participate in government programs, and others—to share their interests and concerns as part of the process.
- If a commission is created by legislation, key sponsors of the legislation should be kept apprised of progress and be committed to advancing legislation that would enact at least some of the commission's recommendations right away.
- States need not replicate the precise bureaucracy and infrastructure that the federal government has created, to create useful, actionable information from their data. States may leverage existing infrastructure and capitalize on their resources, systems, and data assets and even include some of those maintained by federal agencies. Capacity is built through central coordinating and sharing mechanisms. Significant infrastructure elements must be strategically aligned to focus on priority work that extends beyond a single agency, to build sufficient capacity to address complex challenges.
- Establishing standards and methods that can be adopted by collaborating agencies will have a synergistic effect to increase capacity. Fewer resources will be needed in the long run for sharing data, analytical methods, and infrastructure improvements. Ideally, standards could be adopted by consensus.



- Cultural change is difficult. Agencies have traditionally been focused on their own missions. Impetus toward culture change must come from the top but also must provide value to each agency engaged in creating a more collaborative culture.
- Intergovernmental partnerships can accelerate innovation and progress through mutually beneficial exchange of data and resources. Successful partnerships leverage the creation of value for all partners.
- Successful partnerships require sustained commitment over long periods of time, with continuous improvements as the partnership matures.
- Partnerships do not need to wait for all pieces to be assembled to begin. Often federal partnerships begin with a vision shared by one or two states and build from there, as the value of the partnership becomes evident.
- Targeted federal investments can provide resources to support data modernization and linkage at the state and local level. This approach may require knowledge of available flexibilities and federal laws.  
States and localities can make rapid progress on modernizing systems, standardizing data, and creating a data infrastructure by ensuring that federal funding is invested in alignment with a data strategy that will result in significant improvements to data quality, sharing, and analytics.
- States should maximize federal investments by taking a cross-agency, cross-functional approach to making improvements that comply with federal requirements. Such an approach can help remove state and local barriers to increased interagency data sharing.
- Agreement on core characteristics is necessary to support comprehensive capacity for data use at the state-level. Many states have established aspects of strong capacity that can serve as models for states seeking to improve data infrastructure.
- Building and improving capacity can be nonlinear and iterative, meaning there is likely no single path to success. Approaches should be tailored to meet relevant state priorities and constituent expectations or needs.
- Successful efforts have focused on specific topics, such as public health, education, and workforce issues. Such efforts have placed agency employees into environments where they can have hands-on training and gain experience working collaboratively across agency lines to improve data quality, create immediate insights and demonstrate value that can lead to sustainable processes. This type of collaboration can be expanded to firearms violence-related topics.

## Conclusion

As states continue to pursue data infrastructure and capacity improvements, they should leverage federal investments and opportunities from existing state-federal partnerships to enhance existing capabilities. Identifying one or two priority topics can be useful in setting up structures and processes for institutional and cultural change, yet all of these activities described in this report must be considered scalable when designed. The capacity to use data is not built in the abstract; it should be specific to align with existing priorities, staffing capabilities, and resource availability.

If approaches are successful, states can facilitate data leadership for policy priorities in a way that solves real problems. They can ensure that data are accessible for operational and evaluative purposes. They can plan for processes that align evidence building and use. They can develop substantially enhanced and expanded privacy-protective data infrastructure. Most critically, states that have lagged in devising approaches and identifying resources to build further infrastructure cannot afford to wait longer to recognize the strategic value of using the data they have already collected.

In the twenty-first century, all Americans needs access to reliable, valid information for decision-making at every level of government. Aligning some approaches at the state-level with recent progress and momentum from the federal government is an example of how resources can be more efficiently and effectively aligned to deliver the best services and policies possible for the public.

## References and Resources

Abazajian, K. and Kleykamp, T. *Leveraging Data for Economic Recovery: A Roadmap for States*. Washington, D.C.: Georgetown University Beeck Center, 2020.

Doar, R. and Gibbs, L. *Unleashing the Power of Administrative Data: A Guide for Federal, State, and Local Policymakers*. Washington, D.C.: Results for America and American Enterprise Institute, 2017.

Economic Development Administration. (2021). Good Jobs Challenge. Accessed at: <https://eda.gov/arpa/good-jobs-challenge/>.

Fatherree, K. and Hart, N. *What is the Role of States in Evidence-Based Policymaking?* Washington, D.C.: Bipartisan Policy Center, 2017. Accessed at: <https://bipartisanpolicy.org/wp-content/uploads/2019/03/What-Is-the-Role-of-States-in-Evidence-Based-Policymaking.pdf>.

Fletcher, A. What Will it Take to End Family Homelessness? The Family Options Study. In Hart, N. and Yohannes, M. *Evidence Works: Cases Where Evidence Meaningfully Informed Policy*. Washington, D.C.: Bipartisan Policy Center. Accessed at: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3766880](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3766880).

*Foundations for Evidence-Based Policymaking Act of 2018* (Evidence Act). Public Law 115-435. January 2017. Accessed at: <https://www.congress.gov/115/plaws/publ435/PLAW-115publ435.pdf>.

Garg, K., Guidera, A. and Hart, N. *Improving Data Infrastructure to Meet Student and Learner Information Needs*. Washington, D.C.: Day One Project, 2021. Accessed at: <https://www.dayoneproject.org/post/improving-data-infrastructure-to-meet-student-and-learner-information-needs>.

Gartner <https://www.gartner.com/en/information-technology/glossary/data-governance>, 2021

Goerge, Robert, 2019. "[Engagement with State and Local Government](#)," [SocArXiv](#) v6tc2, Center for Open Science.

National Academy of Public Administration (NAPA). *Advancing Information Policy and Use at the Office of Management and Budget*. Report of the Select Task Force: Organization of Information Policy and Use at OMB. Washington, D.C.: NAPA, 2020. Accessed at: [https://napawash.org/uploads/OMB\\_Information\\_Policy\\_and\\_Use\\_Working\\_Group.pdf](https://napawash.org/uploads/OMB_Information_Policy_and_Use_Working_Group.pdf).

Newcomer, K. Olejniczak, K., & Hart, N. *Making Federal Agencies Evidence-Based: The Key Role of Learning Agendas*. Washington, DC: The IBM Center for The Business of Government, 2021. <http://www.businessofgovernment.org/report/making-federal-agencies-evidence-based-key-role-learning-agendas>.

Pew Charitable Trusts. *How States Use Data to Inform Decisions*. Pew Charitable Trusts, 2018. Accessed at: [https://www.pewtrusts.org/-/media/assets/2018/02/dasa\\_how\\_states\\_use\\_data\\_report\\_v5.pdf](https://www.pewtrusts.org/-/media/assets/2018/02/dasa_how_states_use_data_report_v5.pdf).

Results for America (RFA). 2020 Invest in What Works State Standard of Excellence. Washington, D.C.: Results for America, 2020. Accessed at: <https://2020state.results4america.org/>.

RFA. 2021 Invest in What Works State Standard of Excellence: Criteria and Sub-Criteria. Washington, D.C.: Results for America, 2021. Accessed at: <https://results4america.org/wp-content/uploads/2021/05/2021-Invest-in-What-Works-State-Standard-of-Excellence-Criteria-and-Subcriteria-Final.pdf>.

Roman JK. 2020. The State of Firearms Data in 2019. Bethesda, MD: NORC at the University of Chicago.

Roman JK. 2020. A Conceptual Framework for a Firearms Data Infrastructure. Bethesda, MD: NORC at the University of Chicago.

Roman JK. 2020. A Blueprint for a US Firearms Data infrastructure. Bethesda, MD: NORC at the University of Chicago. 2 Roman JK. 2020.

Rothwell, C. The National Vital Statistics System. Presentation to the U.S. Commission on Evidence-Based Policymaking on January 13, 2017.

U.S. Commission on Evidence-Based Policymaking (CEP). *The Promise of Evidence-Based Policymaking: Final Report of the Commission on Evidence-Based Policymaking*. Washington, D.C.: GPO, 2017.

U.S. Department of Housing and Urban Development (HUD). *Family Options Study: 3-Year Impacts of Housing and Services Interventions for Homeless Families*. Washington, D.C.: HUD Office of Policy Development and Research, 2016. Accessed at: <https://www.huduser.gov/portal/sites/default/files/pdf/family-options-study-full-report.pdf>.

U.S. Department of the Treasury. (2021). Compliance and Reporting Guidance: State and Local Fiscal Recovery Funds. Washington, D.C. Accessed at <https://home.treasury.gov/system/files/136/SLFRF-Compliance-and-Reporting-Guidance.pdf>;

Treasury. (2021). Coronavirus State and Local Fiscal Recovery Funds Interim Final Rule. 86 FR 93 pp 26786. Washington, D.C. Accessed at: <https://www.govinfo.gov/content/pkg/FR-2021-05-17/pdf/2021-10283.pdf>.