

# Data Science Strategy

## for Injury and Violence Prevention



**Centers for Disease  
Control and Prevention**  
National Center for Injury  
Prevention and Control



# Data Science Strategy for Injury and Violence Prevention



## Contributors:

### ***Office of Strategy and Innovation***

Steven Sumner  
Christopher Jones  
John Thompson

### ***Office of Informatics***

Melvin Crum

### ***Office of Science***

Arlene Greenspan  
Kate Shaw  
Gwendolyn Cattledge

### ***Office of Communication***

Kelly Holton  
Jennifer Middlebrooks

### ***Office of Policy and Partnerships***

Elizabeth Solhtalab

### ***Division of Injury Prevention***

Judy Qualters  
Michael Ballesteros  
Amy Wolkin  
Royal Law  
Alexander Crosby  
Erin Sauber-Schatz  
Brandon Nesbit

### ***Division of Violence Prevention***

James Mercy  
Thomas Simon  
Kathleen McDavid Harrison

### ***Division of Overdose Prevention***

Grant Baldwin  
Erin Parker  
Puja Seth  
Lara DePadilla

**August 2020**

National Center for Injury Prevention and Control  
Centers for Disease Control and Prevention  
Atlanta, Georgia

## **Suggested Citation:**

Centers for Disease Control and Prevention. *Data Science Strategy for Injury and Violence Prevention*. Atlanta, GA: National Center for Injury Prevention and Control, Centers for Disease Control and Prevention, 2020.



# Background

## Rising Morbidity and Mortality Due to Injuries and Violence

Injuries and violence are the leading causes of death in the United States for children, adolescents, and adults ages 18 to 44 years and rank in the top 10 causes of death for persons 45 years or older.<sup>1</sup> In recent years, rates of deaths due to many forms of injury and violence—drug overdose, suicide, homicide, road traffic crashes, and falls—have increased, leading to recent declines in life expectancy in the United States.<sup>2</sup> Beyond rising mortality, injuries and violence contribute to substantial morbidity as well as social and economic costs each year.<sup>3,4</sup>

Preventing injury and violence is a public health imperative given the significant impact on individuals, families, and communities across the United States. However, primary challenges to rapidly addressing these public health problems include limitations of both public health data as well as prevention and response capabilities. Lack of timely information, inability to identify emerging health threats, limited capacity to target services, increasingly prevalent health misinformation, declining participation in and lack of representativeness of traditional data systems, and fragmentation of electronic health records and clinical data systems are examples of the challenges facing contemporary public health efforts.

**Preventing injury and violence is a public health imperative given the significant impact on individuals, families, and communities across the United States.**

## Data Science and Public Health

A growing body of research now indicates that application of novel data and data science tools, methods, and techniques can help address critical public health needs, including injury and violence prevention and related issues such as social determinants of health and health equity. Academic research has focused, for example, on the use of novel data sources such as internet search queries to assess disease-related trends in real-time, natural language processing to study electronic health records and other systems with unstructured text, machine learning to improve prevention programming, network analysis to better understand mortality risk, online surveys to improve data timeliness and response rates, and interactive data visualization to improve communication and dissemination of scientific findings.<sup>5-9</sup>

Although data science is an emerging field, academic, industry, and governmental organizations<sup>10</sup> have typically defined it by two consistent features: 1) a multidisciplinary approach that blends methodological techniques from computer science, statistics, and various subject matter domains and 2) a focus on large, complex, or otherwise novel data sources.

For the purposes of public health and injury and violence prevention, the National Center for Injury Prevention and Control (Injury Center) defines population-health data science as a multidisciplinary approach combining traditional epidemiologic methods and contemporary computer science techniques, with a particular focus on large and complex data sources, to improve the measurement and prevention of injury and violence in communities.



## Federal and Local Governmental Response to Data Science

Adoption of data science strategies and a recognition of the growth of digital advances are increasing within health and public health agencies at the federal, state, and local governmental levels. The National Institutes of Health (NIH) recently released the NIH Strategic Plan for Data Science,<sup>10</sup> which outlines NIH goals in five key areas: 1) data infrastructure (e.g., data storage, security, cloud technologies); 2) modernization of data ecosystems (e.g., data accessibility, integration, sharing, linkage); 3) data management, analytics, and visualization; 4) workforce development (e.g., internal and external capacity building); and 5) policy development for data science activities. In addition, the U.S. Food and Drug Administration (FDA) recently released its Digital Health Innovation Action Plan focused on advancing guidance, policies, and tools for digital health technologies.<sup>11</sup>

Local health agencies have also made advancements in testing health innovations using novel data sources (e.g., online data) and new statistical approaches (e.g., machine learning) to improve health department prevention and response activities. Health innovations that use data science approaches have been explored in diverse areas such as preventing foodborne illness, childhood lead poisoning, and housing fires, among others.<sup>12-15</sup>

To date, a strategy that focuses on applying data science to advance injury and violence prevention has not been developed. In early 2019, the Injury Center’s Office of Strategy and Innovation convened an internal working group, recognizing the potential public health benefits from implementing an injury and violence prevention data science strategy. The work group was comprised of a diverse group of leaders and subject matter experts from across the center to develop the first data science strategy for injury and violence prevention.

This report reflects the results of these discussions and outlines specific goals for data science activities at the Injury Center. Progress in each of these critical areas will help solve immediate needs that exist in the field of injury and violence prevention and contribute to lessening the burden of injury and violence.





# Strategy for Leveraging Data Science to Advance Injury and Violence Prevention

The Injury Center has identified eight goals to leverage data science for injury and violence prevention to more rapidly and effectively reduce the burden on public health. The specific goals are to:

- 1) Expand the availability and utility of more timely data for injury and violence prevention.
- 2) Improve rapid identification of health threats and response to communities.
- 3) Increase access to accurate health information and prevent misinformation.
- 4) Enhance the usefulness of current data systems by improving data linkage.
- 5) Share information in compelling, useful, and accessible ways.
- 6) Advance ethical practices for data science for injury and violence prevention.
- 7) Increase efficiency of analytic and scientific processes for injury and violence prevention.
- 8) Evaluate promising state and local data science efforts for injury prevention and expand the capacity of state and local health partners in data science methodologies.

The remainder of the strategy provides a discussion of each goal and specific activities the Injury Center will pursue to realize the eight goals of the strategy. In addition, this report includes a discussion of three factors that are critical for successful implementation:

- Strengthening the data science workforce
- Expanding data science partnerships
- Advancing information technology infrastructure

## **Goal 1** Expand the availability and utility of more timely data for injury and violence prevention

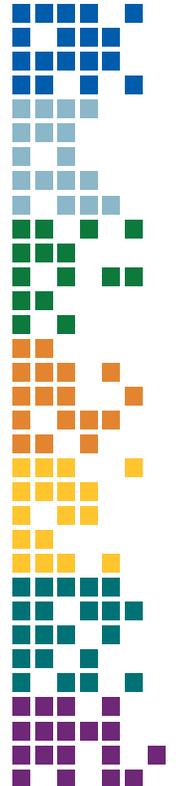
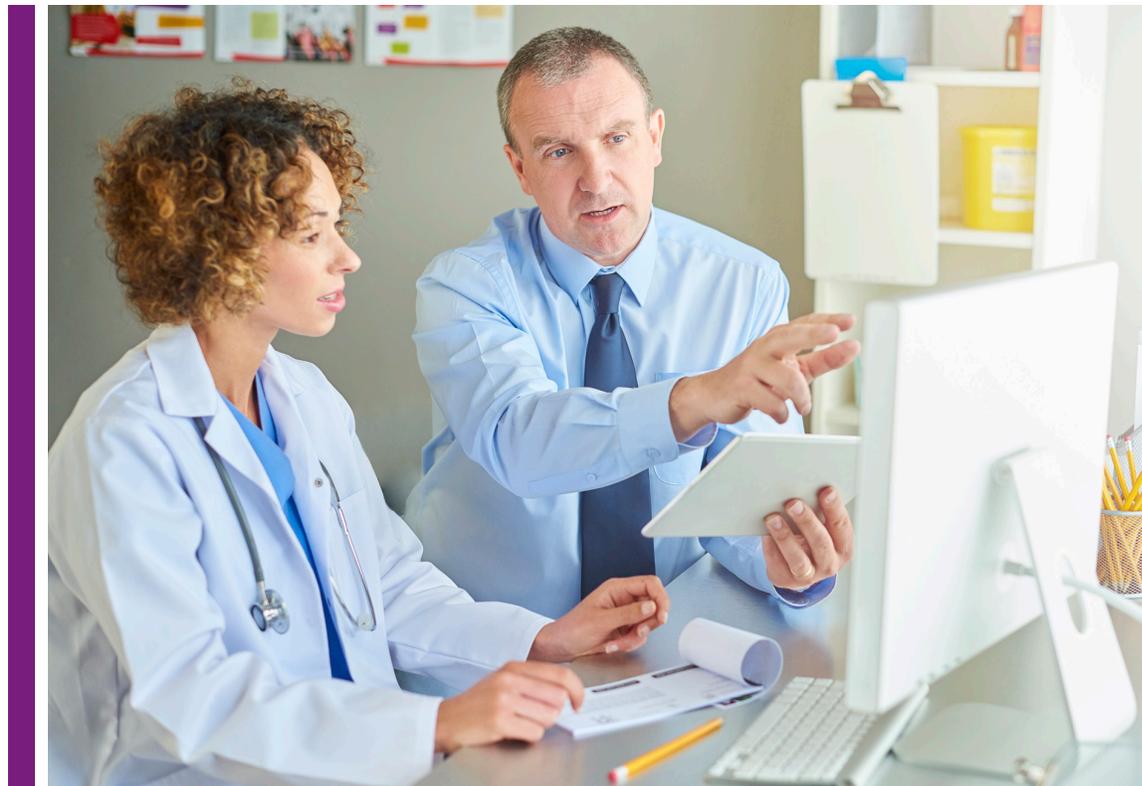
Review of current epidemiologic data for injury and violence prevention demonstrates moderate to strong data quality, but limitations in timeliness of information and the degree to which data measure a health state or behavior that is “upstream” of morbidity or mortality. Strategies to address these limitations to improve information at the national, state, and local level are needed.

Timely data is increasingly important for prevention activities to allow for more real-time feedback about whether current prevention policies, programs, practices, and funding are reducing injury and violence rates. Many current publicly available data for injury and violence outcomes are delayed by one or more years. Furthermore, the degree to which injury and violence data indicators are upstream of morbidity and mortality is also of growing interest for improving prevention efforts, especially those focused on primary prevention. Although fatalities and hospitalizations are important and objective measures of disease, such events often represent permanent or extreme health events for individuals and do not capture the full spectrum of the antecedents and impact of injuries and violence. Ideally, public health prevention, monitoring, and deployment of interventions occur upstream of or before such health consequences.

Recent Injury Center initiatives in public health data system development have supported advances in more current and specific data focused on upstream risk factors. For example, expansion of syndromic surveillance using near real-time text data from emergency department visits has resulted in more timely measures of trends in suspected drug overdoses and suicide.<sup>16,17</sup> Use of administrative data from prescription medication sales has helped provide a more timely picture of dispensed opioid analgesic prescriptions, medications to treat opioid use disorder, and overdose reversal medications.<sup>18</sup> And recent funding to states has accelerated reporting of mortality data, such as through the State Unintentional Drug Overdose Reporting System (SUDORS).

To achieve Goal 1, the Injury Center will:

- 1) Accelerate data collection, processing, and reporting from established public health data systems, such as mortality data systems;
- 2) Evaluate and invest in new data sources (such as administrative, business, or Web-based data), as well as new data collection modalities (such as online surveys);
- 3) Develop and evaluate forecasting and other advanced methodologies combining information from multiple data sources via machine learning models to estimate injury and violence burden in real-time; and
- 4) Advance the use of electronic health records for public health surveillance, including expanding the number of syndrome definitions available for injury and violence monitoring via the National Syndromic Surveillance Program (NSSP) and testing new methods, techniques, and tools such as natural language processing and machine learning to more routinely and efficiently use electronic health record information to monitor injury and violence.



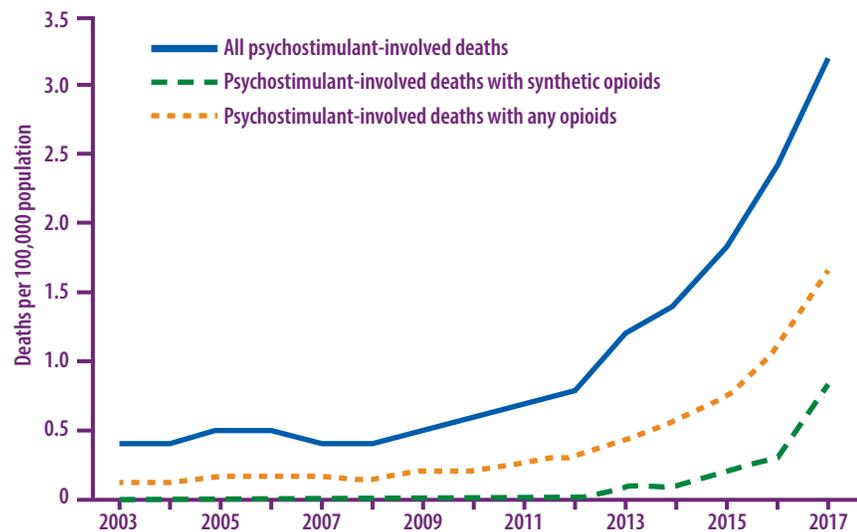
## Goal 2

### Improve rapid identification of health threats and response to communities

The Injury Center supports states and local communities through various means, including funding surveillance and prevention activities, deploying Epidemiologic Assistance (Epi-Aid) teams and Opioid Rapid Response Teams, and coordinating action on national health threats.<sup>19,20</sup> Effective action for these activities depends on rapidly identifying emerging health threats.

Identifying emerging and novel health threats remains a challenge for injury and violence prevention, as traditional injury surveillance and data systems are not generally designed to detect them. For example, the drug overdose epidemic, now the leading cause of injury-related mortality in the United States, has been characterized by a transition in the substances causing mortality over time. Overdose mortality was initially driven by prescription opioids, followed by heroin, and subsequently synthetic opioids such as illicitly manufactured fentanyl. Causes of overdose mortality in the United States continue to evolve as demonstrated by rising overdose deaths involving stimulants such as methamphetamine and other substances in recent years.<sup>21,22</sup>

**Causes of overdose mortality in the United States continue to evolve.**



These challenges are not unique to overdose deaths. Emerging threats such as youth suicide “games” and harmful online content challenge suicide prevention efforts,<sup>23</sup> and work in road traffic safety must keep pace with technological changes shaping road usage patterns.<sup>24</sup> Furthermore, deploying field response teams for any cause of injury depends on early identification of clusters of injury or violence.

Leveraging novel and timely data sources (e.g., online data sources) has shown promise in identifying emerging health threats earlier than more traditional data systems.<sup>23,25</sup> Although scientific methodologies to detect early signals in novel data sources has been more robustly developed for certain fields, such as pharmacologic or medical product safety monitoring, translating these approaches to injury and violence prevention is promising.<sup>26</sup>

To achieve Goal 2, the Injury Center will:

- 1) Advance the application of natural language processing methodologies to Injury Center text-based data systems and to novel streams of data;
- 2) Improve modeling approaches to detect space-time clusters of illness and injury for more rapid and accurate outbreak detection; and
- 3) Increase the timeliness of insights from Injury Center field investigations by expanding electronic data collection and automation of analyses and reporting.

## Goal 3

### Increase access to accurate health information and prevent misinformation

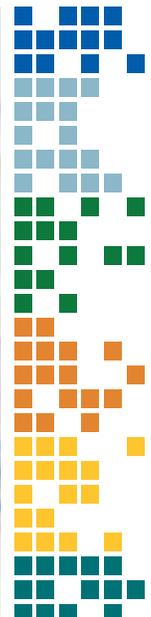
The importance of accurate and protective health information has recently gained visibility as a key component of public health efforts to combat the increasing proliferation of health misinformation, which is defined as “a health-related claim of fact that is currently false due to a lack of scientific evidence.”<sup>27</sup> The abundance of misinformation about vaccines is one example of how false information spreads.<sup>28</sup> Relatedly, content and information fostering stigma further prevent access to care and worsen health outcomes. Injury and violence prevention are similarly affected by these issues. For example, stigma and norms surrounding injury influence reporting and help-seeking behaviors for multiple forms of violence (e.g., bullying, sexual violence, intimate partner violence).<sup>29</sup> Treatment for mental health concerns as well as opioid use disorder, for example, is similarly influenced at the population level by social norms and health information about treatment options.<sup>30,31</sup>

Many health professionals and patients increasingly access health information online through novel channels and often in an environment that lacks clarity on which entities are trusted and scientifically valid. Indeed, social media has now supplanted print newspapers as a primary source of news.<sup>32</sup> While injury and violence prevention researchers have long recognized the importance of social norms and information on individual health decisions, these factors are increasingly essential given the changing dynamics of information sharing in online environments.

In addition to the need to identify specific types of health misinformation related to injuries and violence and the sources contributing to it, a need also exists to understand how this misinformation diffuses through formal and informal networks as well as effective strategies to counter misinformation.

To achieve Goal 3, the Injury Center will:

- 1) Quantify the prevalence and characteristics of both accurate information and misinformation related to injury and violence prevention;
- 2) Assess how health information, norms, and stigma diffuse through communities and networks; and
- 3) Leverage this information to expand awareness of evidence-based prevention policies, practices, and programs.

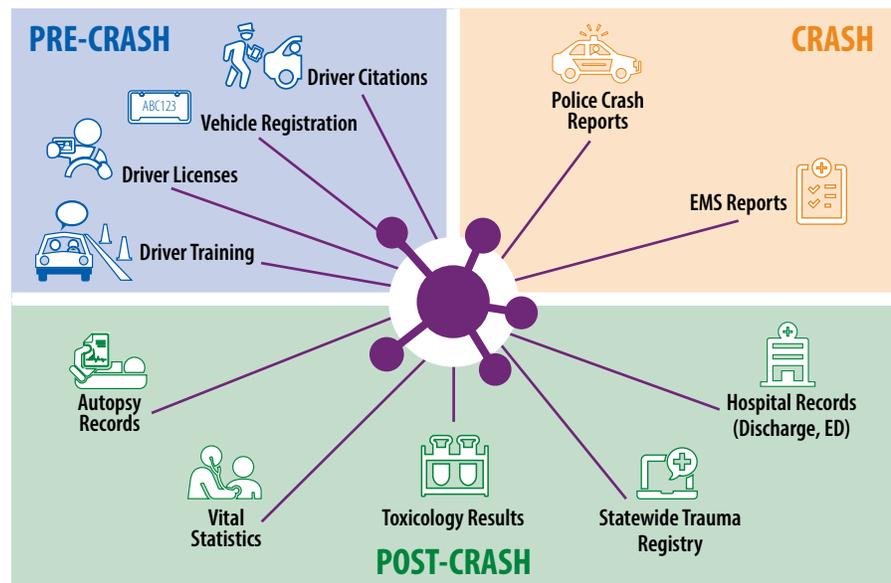


## Goal 4

### Enhance the usefulness of current data systems by improving data linkage

While many leading injury and violence prevention data systems collect robust information on the occurrence of injuries and violence, data on antecedent risk and protective factors and other contextual variables are less available. These limitations prevent the construction of a complete picture of the burden of injury and hinder the identification of prevention strategies. For example, in the absence of motor vehicle crash information linked to emergency department records, public health officials cannot identify the protective factors most strongly associated with health outcomes.<sup>33</sup>

**When motor vehicle crash information is linked to emergency department records, public health officials can identify the protective factors most strongly associated with health outcomes.**



While some Injury Center data systems, such as the National Violent Death Reporting System (NVDRS) and SUDORS, perform robust data linkage activities, many potential risk and protective factors for injury and violence have not been adequately examined because of an absence of linked data on them. This includes information on novel risk and protective factors as well as certain types of injury and violence outcomes. Addressing these data silos by linking multiple data sources can unlock a new understanding about risk and protective factors, help identify opportunities for prevention, and improve understanding of the effectiveness of intervention strategies. Additionally, data linkage can be performed at units of analysis beyond the individual level, such as by geography, further expanding our understanding of risk and protective factors.

Data science work has particular relevance to data linkage efforts given the proliferation of administrative data that can potentially be linked, as well as the consideration of informatics approaches to enable data system interoperability and data sharing. One such example is the Cardiff Model, a novel violence prevention program that brings together data from both health and law enforcement to better guide community planning.<sup>34</sup>

Cities, counties, and states are increasingly linking data systems to improve health outcomes due to the growing maturation and availability of electronic data systems. These advances can be leveraged for injury and violence prevention efforts.

To achieve Goal 4, the Injury Center will:

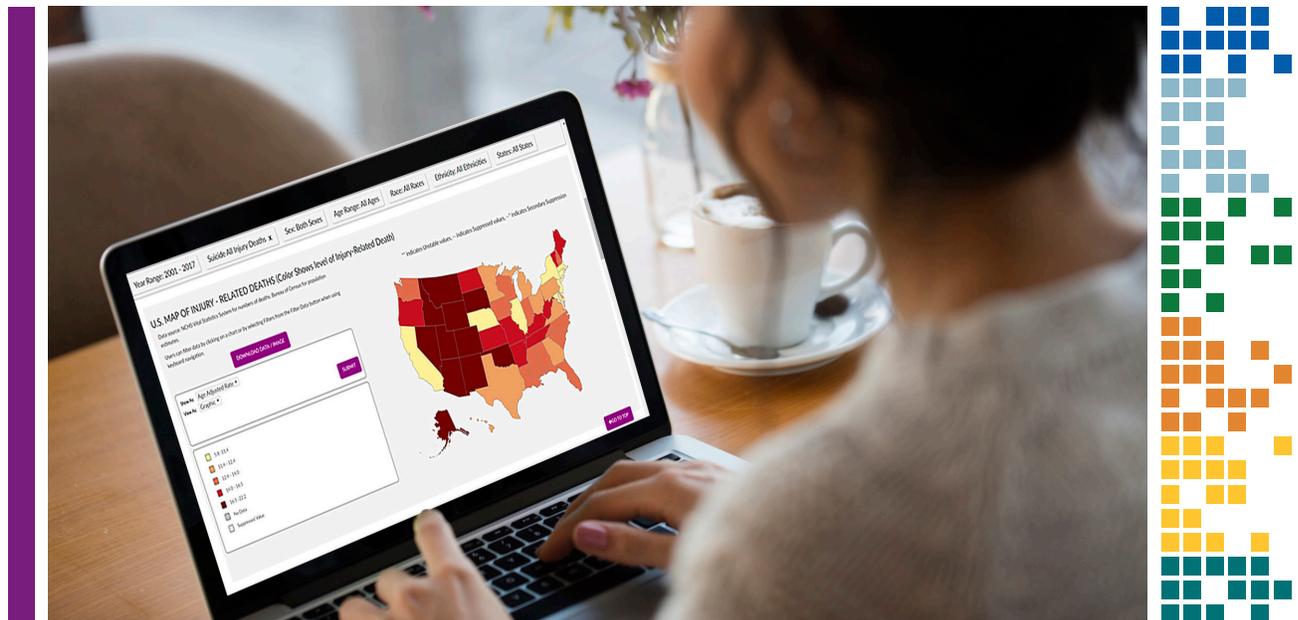
- 1) Advance the availability of linked data and data system interoperability by building on existing and scalable injury and violence data systems such as NVDRS and SUDORS;
- 2) Formally evaluate the benefits of linked data systems for injury and violence prevention in state and local settings; and
- 3) Improve awareness, understanding, and best practices related to the legal and ethical considerations for data linkage through collaboration with organizations active in this area.

## Goal 5

### Share information in compelling, useful, and accessible ways

In recent years, the Injury Center has expanded the use of both static and interactive data visualizations to communicate findings to public and scientific audiences. For example, the Center's Web-based Injury Statistics Query and Reporting System (WISQARS) now allows users to explore injury and violence data interactively via maps and charts.<sup>35</sup> Users can compare rates of injury deaths across states and export data for further analysis.

Data visualization is rapidly being recognized as a central component of shaping individual health behaviors and of simplifying complex data used to inform health policy and resource allocation. This growth is evidenced through multiple efforts, such as the Robert Wood Johnson Foundation–supported Visualizing Health project or the Global Burden of Disease program.<sup>36,37</sup> Within the field of injury and violence prevention, for example, the evidence-based Cardiff Model uses near real-time geospatial data visualization from multiple sources of the location of violence-related injuries occurring in public places to increase the timeliness and utility of prevention programs.<sup>38</sup>





CDC Injury Center data visualization efforts have focused largely on mortality data; however, additional work is needed in visualizing morbidity data; enabling exploration and overlay of data from multiple sources such as social determinants of health; and in conducting research to better understand what forms of visualization best advance policy, programs, and public understanding of injury data. Expanding data visualization activities will also increase access to CDC data for external users.

To achieve Goal 5, the Injury Center will:

- 1) Advance data visualization techniques for injury information; increase visualization of morbidity data; and study the effectiveness of data visualization approaches to achieve policy, programmatic, and practice objectives;
- 2) Develop more robust dashboards and interactive tools for a comprehensive understanding of injury and violence topics and for comparison of data across multiple sources;
- 3) Modernize communication strategies by adding interactive data visualizations and digital storytelling to the Injury Center website; and
- 4) Develop internal data visualization tools to allow Injury Center scientists to improve the efficiency of data exploration and analytic tasks and train staff to enhance visualizations used in scientific products.

## Goal 6

### Advance ethical practices for data science for injury and violence prevention

While new data sources and methods for advancing health have evolved rapidly, these developments have also raised important ethical questions for scientists and health practitioners. Important areas for consideration include transparency and interpretability of machine learning models, maintenance of user privacy in large-scale online data, data security and risk of identification with linked data, and bias or inaccurate information present in some administrative data sources.<sup>39,40</sup> These questions have not been definitively addressed by public health agencies to date. To address this gap, the Injury Center will identify best practices and promote the ethical use of data science methodologies for the field of injury and violence prevention.

To achieve Goal 6, the Injury Center will:

- 1) Conduct systematic reviews to comprehensively assess the current state of use of data science methods for injury and violence prevention research and practice;
- 2) Convene diverse stakeholders to identify and continually evaluate best practices for injury and violence prevention given the complexities of such work; and
- 3) Utilize Injury Center funding opportunities and partnerships to advance work benefitting vulnerable and at-risk populations.



## Goal 7 Increase efficiency of analytic and scientific processes for injury and violence prevention

Injury Center data systems for injury and violence prevention contain increasing amounts of unstructured data, including natural language, and rely upon human coding of records to identify cases and perform certain reporting functions, incurring significant time and human resources. These data systems include NVDRS, SUDORS, the Drug Overdose Surveillance and Epidemiology (DOSE) Program, the Emergency Department Surveillance of Nonfatal Suicide-Related Outcomes (ED-SNSRO) Program, the Firearm Injury Surveillance Through Emergency Rooms (FASTER) Program, the National Electronic Injury Surveillance System—All Injury Program (NEISS-AIP), and the School Associated Violent Death System (SAVD), among others. The Injury Center also routinely conducts and updates systematic reviews for various injury and violence topics; this entails manually coding thousands of scientific articles per year.

Opportunities exist to better use natural language in these datasets and systematic reviews to enable faster querying of records for case identification and automation of some manual coding tasks. For example, supervised machine learning models may be used to automatically classify text into predefined categories where enough labeled or annotated text data already exist, while unsupervised or semi-supervised machine learning approaches may help in data exploration where labeled categories for such data are not readily available.

To achieve Goal 7, the Injury Center will:

- 1) Build tools using natural language processing and machine learning to improve the efficiency of scientific tasks for these key data systems for timely reporting and to increase scientific output;
- 2) Facilitate the use of data science tools to streamline repetitive processes, such as article classification, updating injury-related living systematic reviews, deidentifying health information, and cleaning variables; and
- 3) Disseminate novel methods and tools to increase analytic efficiency externally to state and local public health partners, grantees, and researchers to support their work.

### Key CDC Data Systems for Injury and Violence Prevention



**NVDRS**  
National Violent Death Reporting System



**SUDORS**  
State Unintentional Drug Overdose Reporting System



**DOSE**  
Drug Overdose Surveillance and Epidemiology Program



**ED-SNSRO**  
Emergency Department Surveillance of Nonfatal Suicide-Related Outcomes Program



**FASTER**  
Firearm Injury Surveillance Through Emergency Rooms Program



**NEISS-AIP**  
National Electronic Injury Surveillance System—All Injury Program



**SAVD**  
School Associated Violent Death System

## Goal 8

### Evaluate promising state and local data science efforts for injury prevention and expand the capacity of state and local health partners in data science methodologies

Many CDC data systems and programs for injury and violence prevention began as local innovations. For example, NVDRS, now the leading national data source for information on violent deaths, began as a pilot project in a single state. Similarly, CDC epidemiologic field responses have led to early understanding of health threats that would subsequently increase across the nation, such as fentanyl-involved overdoses.<sup>41</sup> Opportunity exists to identify and expand innovations being pursued by state and local public health departments in injury and violence surveillance, prevention, and response programs. In addition, opportunities exist to more quickly bring to scale innovations that are evaluated and deemed effective.

To achieve Goal 8, the Injury Center will:

- 1) Identify and evaluate novel local public health efforts that employ data science tools, methods, and techniques;
- 2) Quantify the cost effectiveness of such innovations, determine feasibility for broader scaling, and help to disseminate effective practices nationally; and
- 3) Strengthen the capacity of state and local health partners to apply data science methodologies to their work, such as through technical assistance to health departments for data systems and health surveillance methodologies.





# Data Science Implementation Considerations

Successful implementation of data science goals in the Injury Center will depend on addressing three key areas: 1) strengthening the data science workforce; 2) expanding data science partnerships; and 3) advancing information technology infrastructure.

## Strengthening the Data Science Workforce

As part of its 2019 reorganization, the Injury Center established the Data Science Team in the Division of Injury Prevention to expand data science research and programmatic activities. This team will substantially increase the Injury Center's capacity to advance data science, including building and applying machine learning, natural language processing, and advanced analytic approaches across injury and violence prevention. The Data Science Team will serve as a resource for all injury and violence topics, provide consultation, and will help implement data science methodologies. The Injury Center, concomitant to the establishment of a data science team, is also elevating informatics activities with the creation of an Office of Informatics as part of the 2019 reorganization. The Office of Informatics will advance the information technology infrastructure for data science, as discussed below. The Office of Strategy and Innovation will coordinate, support, and strengthen initiatives across the center as data science activities grow across all areas of injury prevention.

The Injury Center will also work to increase capacity for applying data science techniques across all offices, divisions, branches, and teams in the center. This will be achieved through increasing awareness of data science approaches for injury and violence prevention through center-wide trainings and expert presentations, and through the translation of tools, methods, and techniques across injury and violence topics.

The Injury Center will also expand formal training opportunities by establishing a data science fellowship for students from computer science and related fields to gain public health experience. Additionally, the Injury Center will develop a data science-focused position for incoming Epidemic Intelligence Service Officers, a training program for physicians, scientists, and other health care professionals, to increase data science skills among this workforce. Injury Center trainees in other programs, such as Informatics Fellows and Prevention Effectiveness Fellows, will also receive training in data science. Lastly, the Injury Center will work to increase exchange opportunities with scientists from university, non-governmental, and private sector organizations to contribute to solutions for leading public health challenges.

## Expanding Data Science Partnerships

The Injury Center currently engages a broad range of injury and violence partners; however, expansion of data science activities as outlined in this strategy will entail expanded efforts to engage partners, including academic and non-governmental partners not typically engaged in injury and violence prevention. This includes university departments of computer science, non-governmental organizations focused on technology-based prevention efforts, and technology industry stakeholders. Broad partnerships in these areas will enable CDC to ensure implementation of state-of-the-art approaches for data science and to help advance best practices in ethical considerations for data science, which requires input from a diverse constituency.

Multiple pathways exist to strengthen data science partnerships and to build capacity to apply data science for injury and violence prevention.



The Injury Center's Injury Control Research Centers (ICRCs) and Youth Violence Prevention Centers (YVPCs) are academic consortiums currently conducting a wide range of epidemiologic studies to advance injury prevention; opportunities exist to increase application of data science methodologies as part of the ICRC and YVPC portfolio of activities. Similarly, CDC funds injury prevention activities at state and local health departments through the Core State Violence and Injury Prevention Program (Core SVIPP), as well as through funding opportunities specific to certain injury and violence topics such as opioid overdose or child abuse and neglect. Lastly, funding opportunities for external investigators, such as CDC extramural cooperative agreements, can be avenues to stimulate work by academic researchers on applying novel data science approaches for injury and violence prevention.

### **Advancing Information Technology Infrastructure**

The application of data science is contingent on having an information technology infrastructure that can support the receipt, storage, analysis, and visual depiction of large and disparate data sources. The Office of Informatics will lead multiple key activities related to the successful implementation of data science in the Injury Center. Specifically, the office will work to standardize and expand tools for the Injury Center, health departments, and other partners to share, clean, process, and analyze data in automated ways. The office will work to develop best practices for storage of increasing amounts of injury and violence related data and implement capabilities to allow more rapid querying of large volumes of data by expanding access to cloud-based environments. The Office also will play an important role in scaling local innovations that are found to be effective, such as data linkage pilot projects. Finally, the office will oversee information security standards and accessibility of products such as interactive data visualizations. These activities align with CDC's Public Health Data Modernization Initiative.

## **Summary: Data Science at NCIPC**

The development and use of data science is advancing rapidly in many fields. For maximum impact on injury and violence prevention, CDC's Injury Center must embrace and increase the application of these approaches for public health. This strategy outlines the goals and activities the Injury Center will pursue in order to fully realize the public health benefits that can be achieved through the application of data science. Accomplishing these goals will entail strengthening the internal workforce, expanding public health partnerships, advancing information technology infrastructure, and increasing investments in data science activities.

Research to date in the Injury Center has demonstrated the potential of novel data sources to enhance early detection of health threats,<sup>25</sup> the value of linked data sources in establishing effective prevention programs for injury,<sup>34</sup> the importance of understanding information shared via novel communication channels about health,<sup>42</sup> and the ability of data science approaches to enable rapid field responses to injury outbreaks. Nonetheless, myriad opportunities exist to further translate research and technology advances to public health.

In light of rising rates of death and disability in the United States from multiple forms of injury and violence, the expanded and strategic application of data science approaches to increase the usefulness of information, impact of programs, and scale of prevention efforts is imperative to population health. A commitment to pursuing and advancing innovation in public health will form a central part of Injury Center activities to prevent injury and violence.



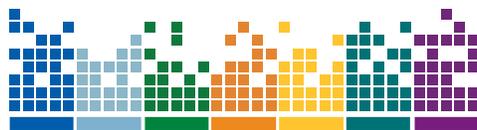
## References

1. Centers for Disease Control and Prevention, National Center for Injury Prevention and Control. Web-based Injury Statistics Query and Reporting System. Available at: <http://www.cdc.gov/injury/wisqars/index.html>.
2. Murphy SL, Xu J, Kochanek KD, Arias E. Mortality in the United States, 2017. NCHS Data Brief. No. 328. Nov 2018.
3. Florence C, Haegerich T, Simon T, Zhou C, Luo F. Estimated Lifetime Medical and Work-Loss Costs of Emergency Department-Treated Nonfatal Injuries--United States, 2013. *MMWR*. 2015;64(38):1078-1082.
4. Florence C, Simon T, Haegerich T, Luo F, Zhou C. Estimated Lifetime Medical and Work-Loss Costs of Fatal Injuries--United States, 2013. *MMWR*. 2015;64(38):1074-1077.
5. Brownstein JS, Freifeld CC, Madoff LC. Digital disease detection—harnessing the Web for public health surveillance. *N Engl J Med*. 2009;360(21):2153-2157.
6. McCoy TH, Jr., Castro VM, Roberson AM, Snapper LA, Perlis RH. Improving Prediction of Suicide and Accidental Death After Discharge From General Hospitals With Natural Language Processing. *JAMA Psychiatry*. 2016;73(10):1064-1071.
7. Kessler RC, Warner CH, Ivany C, et al. Predicting suicides after psychiatric hospitalization in US Army soldiers: the Army Study to Assess Risk and Resilience in Servicemembers (Army STARRS). *JAMA Psychiatry*. 2015;72(1):49-57.
8. Green B, Horel T, Papachristos AV. Modeling Contagion Through Social Networks to Explain and Predict Gunshot Violence in Chicago, 2006 to 2014. *JAMA Intern Med*. 2017;177(3):326-333.
9. Liu H, Cella D, Gershon R, et al. Representativeness of the Patient-Reported Outcomes Measurement Information System Internet panel. *J Clin Epidemiol*. 2010;63(11):1169-1178.
10. National Institutes of Health. NIH Strategic Plan for Data Science. 2018. Available at: [https://datascience.nih.gov/sites/default/files/NIH\\_Strategic\\_Plan\\_for\\_Data\\_Science\\_Final\\_508.pdf](https://datascience.nih.gov/sites/default/files/NIH_Strategic_Plan_for_Data_Science_Final_508.pdf).
11. Food and Drug Administration. Digital Health Innovation Action Plan. Available at: <https://www.fda.gov/media/106331/download>.
12. Harris JK, Mansour R, Choucair B, et al. Health department use of social media to identify foodborne illness - Chicago, Illinois, 2013-2014. *MMWR*. 2014;63(32):681-685.
13. Harris J, Hinyard L, Beatty K, et al. Evaluating the implementation of a Twitter-based foodborne illness reporting tool in the city of St. Louis Department of Health. *Int J Environ Res Public Health*. 2018;15(5):833.
14. Potash E, Brew J, Loewi A, et al. Predictive modeling for public health: Preventing childhood lead poisoning. Paper presented at: Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2015.
15. Madaio M, Haimson OL, Zhang W, et al. Identifying and prioritizing fire inspections: a case study of predicting fire risk in Atlanta. Bloomberg Data for Good Exchange, 2015.
16. Vivolo-Kantor AM, Seth P, Gladden RM, et al. Vital signs: trends in emergency department visits for suspected opioid overdoses—United States, July 2016–September 2017. *MMWR*. 2018;67(9):279.
17. Zwald ML, Holland KM, Annor FB, et al. Syndromic Surveillance of Suicidal Ideation and Self-Directed Violence - United States, January 2017-December 2018. *MMWR*. 2020;69(4):103-108.
18. Guy Jr GP, Haegerich TM, Evans ME, Losby JL, Young R, Jones CM. Vital signs: Pharmacy-based naloxone dispensing—United States, 2012–2018. *MMWR*. 2019;68(31):679.
19. Centers for Disease Control and Prevention, National Center for Injury Prevention and Control, Opioid Rapid Response Teams. Aug 27, 2019. Available at: <https://www.cdc.gov/opioids/opioid-rapid-response-teams.html>.
20. Perrine CG, Pickens CM, Boehmer TK, et al. Characteristics of a multistate outbreak of lung injury associated with e-cigarette use, or vaping—United States, 2019. *MMWR*. 2019;68(39):860.



21. Kariisa M, Scholl L, Wilson N, Seth P, Hoots B. Drug Overdose Deaths Involving Cocaine and Psychostimulants with Abuse Potential - United States, 2003-2017. *MMWR*. 2019;68(17):388-395.
22. Gladden RM, O'Donnell J, Mattson CL, Seth P. Changes in opioid-involved overdose deaths by opioid type and presence of benzodiazepines, cocaine, and methamphetamine—25 states, July–December 2017 to January–June 2018. *MMWR*. 2019;68(34):737.
23. Sumner SA, Galik S, Mathieu J, et al. Temporal and geographic patterns of social media posts about an emerging suicide game. *J Adolesc Health*. 2019.
24. Trivedi TK, Liu C, Antonio ALM, et al. Injuries associated with standing electric scooter use. *JAMA Network Open*. 2019;2(1):e187381-e187381.
25. Bowen DA, O'Donnell J, Sumner SA. Increases in Online Posts About Synthetic Opioids Preceding Increases in Synthetic Opioid Death Rates: a Retrospective Observational Study. *J Gen Intern Med*. 2019;34(12):2702-2704.
26. Sarker A, Ginn R, Nikfarjam A, et al. Utilizing social media data for pharmacovigilance: a review. *J Biomed Inform*. 2015;54:202-212.
27. Chou WS, Oh A, Klein WMP. Addressing Health-Related Misinformation on Social Media. *JAMA*. 2018.
28. Patel M, Lee AD, Redd SB, et al. Increase in Measles Cases - United States, January 1-April 26, 2019. *MMWR*. 2019;68(17):402-404.
29. Nguyen KH, Kress H, Atuchukwu V, et al. Disclosure of sexual violence among girls and young women aged 13 to 24 years: Results from the Violence against Children Surveys in Nigeria and Malawi. *Journal of Interpersonal Violence*. 2018.
30. Ward MC, Druss BG. Treatment considerations in severe mental illness: caring for the whole patient. *JAMA Psychiatr*. 2019;76(7):759-760.
31. Wakeman SE, Rich JD. Barriers to medications for addiction treatment: How stigma kills. *Subst Use & Misuse*. 2018;53(2):330-333.
32. Shearer E. Pew Research Center. Social media outpaces print newspapers in the U.S. as a news source. Available at: <https://www.pewresearch.org/fact-tank/2018/12/10/social-media-outpaces-print-newspapers-in-the-u-s-as-a-news-source/>. Published 2018.
33. Conderino S, Fung L, Sedlar S, Norton JM. Linkage of traffic crash and hospitalization records with limited identifiers for enhanced public health surveillance. *Accid Anal*. 2017;101:117-123.
34. Wu DT, Moore JC, Bowen DA, et al. Proportion of Violent Injuries Unreported to Law Enforcement. *JAMA Intern Med*. 2019;179(1):111-112.
35. Centers for Disease Control and Prevention. WISQARS - Explore Fatal Injury Data. Visualization Tool. Available at: <https://wisqars-viz.cdc.gov:8006/explore-data/home>.
36. Visualizing Health. University of Michigan. Available at: <http://www.vizhealth.org/>.
37. Global Burden of Disease. Institute for Health Metrics. Available at: <http://www.healthdata.org/gbd/data-visualizations>. Accessed.
38. Florence C, Shepherd J, Brennan I, Simon T. Effectiveness of anonymised information sharing and use in health service, police, and local government partnership for preventing violence related injury: experimental study and time series analysis. *Bmj*. 2011;342:d3313.
39. Rajkomar A, Hardt M, Howell MD, Corrado G, Chin MH. Ensuring fairness in machine learning to advance health equity. *Ann Intern Med*. 2018.
40. Moreno MA, Goni N, Moreno PS, Diekema D. Ethics of social media research: common concerns and practical considerations. *Cyberpsychol Behav Soc Netw*. 2013;16(9):708-713.
41. Mercado MC, Sumner SA, Spelke MB, Bohm MK, Sugerman DE, Stanley C. Increase in drug overdose deaths involving fentanyl—Rhode Island, January 2012–March 2014. *Pain Med*. 2017;19(3):511-523.
42. Sumner SA, Bowen DA, Bartholow B. Factors associated with increased dissemination of positive mental health messaging on social media. *Crisis*. 2019.





[www.cdc.gov/injury](http://www.cdc.gov/injury)