Gaps Report

Challenges of Counting All Violent Deaths Worldwide

GReVD
A Global Registry of Violent Deaths

MAY 2020

Gulzhan Asylbek kyzy
Caroline Delgado
Gary Milante

This project was funded with aid from the UK government; the views expressed do not necessarily reflect the UK government's official policies.

www.grevd.org

The authors are thankful for input from Dan Smith (SIPRI), Tom Wheeler (DFID), Rachel Locke (Center for International Cooperation), Jessica Baumgardner-Zuzik (Alliance for Peacebuilding), and contributions from members of the GReVD consortium, including participants of the Washington, DC and Geneva Workshops (February 2019):

The Armed Conflict Location and Events Database (ACLED)
The Brookings Institution
The Center for International Development and Conflict Management (CIDCM)
The Center for Peace and Security Studies (cPASS)
Cline Center for Advanced Social Research
The Community Robotics, Education and Technology Empowerment (CREATE) Lab
The Global Terrorism Database (GTD)
The Igarapé Institute
The Peace Research Institute Oslo (PRIO)
Small Arms Survey (SAS)
Uppsala Conflict Data Programme (UCDP)

All errors and omissions remain those of the authors.

GReVD is convened by the Stockholm International Peace Research Institute (SIPRI) and The Brookings Institution.

The GReVD consortium includes the Armed Conflict and Location Database (ACLED), The Center for International Development and Conflict Management (CIDCM), The Center for Peace and Security Studies (cPASS), Cline Center for Advanced Social Research, The Community Robotics, Education and Technology Empowerment (CREATE) Lab, The Global Terrorism Database (GTD), The Igarapé Institute, The Peace Research Institute Oslo (PRIO), the Small Arms Survey (SAS) and the Uppsala Conflict Data Programme (UCDP).

CONTENTS

PART 1 Introduction

The scourge of war and violence affects everyone. In recognition, the world collectively committed, through the United Nations (UN) Development Programme Sustainable Development Goal (SDG) 16.1, to "Significantly reduce all forms of violence and related death rates everywhere." Although adopting this target has been an important step taken by the global community, national and international reporting mechanisms of conflict deaths and homicides, using today's methods for reporting and monitoring on these indicators, tell us only part of the story of the real trends in violent deaths over the next decade.

Today's methods need to adapt to monitor this goal for several reasons. Conflict deaths are unlikely to be reported voluntarily; intentional homicides are not reported consistently in many countries; and, even if both of these were reported voluntarily and consistently, they would likely leave out violent deaths associated with hybrid wars, violence perpetrated by organized crime and even by states (e.g., drone strikes, political assassinations and extra-judicial killings, and increasingly, "grey zone" conflict deaths).[1]

Discrepancies exist in how conflict deaths are monitored and measured even among the experts. Although these datasets (produced mostly by members of the GReVD consortium) are very high quality and are internally comparable over time, they are not currently reconcilable across datasets or among types of sources (administrative data and media-based data may not be comparable, for example). Some violent deaths would be double-counted if deaths in databases were aggregated; deaths from drone strikes and many deaths from criminal action, including trafficking, are often not counted at all; deaths from terrorist events are counted differently by case (or not at all), depending on expert and political definitions. Last, new types of violence, new means of reporting, and new methods for coding, including machine-learning applications, require adaptive solutions over the next decade. This report details these challenges and methodological considerations.

As a result of consistency issues between sources and datasets, the current methodologies for monitoring violent deaths can consistently tell us trends within individual data sets, but we cannot say with precision how many people die each year around the world due to violence. The existing approaches to count violent deaths are admittedly imperfect and sometimes only proxy measures exist. This means there is no reliable global picture of violent deaths, which makes it harder to assess overall progress in reducing violence, the first target of SDG 16. The United Nations Secretary-General published the *Special Edition: Progress towards the Sustainable Development Goals* following the High-Level Political Forum (HLPF) in July 2019, stating, "Renewed efforts are essential to move towards the achievement of Sustainable Development Goal 16."[2] We must invest now in the reporting and methodology that will be necessary for monitoring progress toward the goal through 2030.

1 Rachel Kleinfeld and Robert Muggah, "The State of War," *Foreign Policy (*March 18, 2019), https://foreignpolicy.com/2019/03/18/the-state-of-war/. See also part 2 of this report on *Intentional Homicide* and *Conflict-Related Death.*

2 United Nations Economic and Social Council, *Special Edition: Progress towards the Sustainable Development Goals, Report of the Secretary-General*, High-level Segment: Ministerial Meeting of the High-Level Political Forum on Sustainable Development (May 8, 2019), https://undocs.org/E/2019/68.

SIPRI and the Brookings Institution have convened a global consortium of experts to solve this data problem, through an initiative called the Global Registry of Violent Deaths (GreVD). The consortium consists of the Armed Conflict Location and Event Data project (ACLED), Cline Center for Advanced Social Research, Carnegie Mellon University Community Robotics, Education and Technology Empowerment Lab (CMU-CREATE) Lab, the Global Terrorism Database (GTD), Igarapé Institute, Peace Research Institute Oslo (PRIO), Global Violent Deaths Database (GVD) at the Small Arms Survey (SAS), the Uppsala Conflict Data Programme (UCDP), the Centre for Peace and Security Studies (cPASS) at UCSD, and START at UMD (see About the GReVD Consortium at the end of this report). These institutions collectively represent the leading global expertise on monitoring violent deaths.

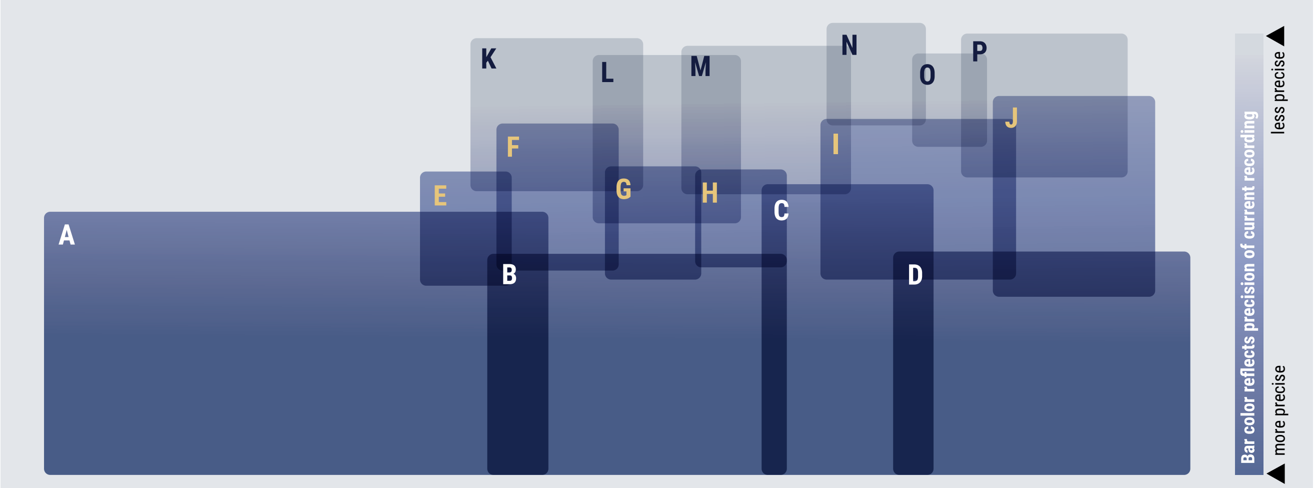
PART 2 The State of Violent Death Monitoring

WHAT IS A VIOLENT DEATH?

The types, patterns, and dynamics of violent deaths vary widely as do categorization by state, non-governmental, and international actors. Categorizations of violent deaths typically include deaths from armed group violence; the activities of organized crime; intentional homicides, including gender-based violence; assassinations; extra-judicial killings; drone strikes; terrorism; explosive remnants of war like landmines; protests that turn violent; and communal violence. Some "grey-zone" or "hybrid" conflict deaths might also be included in these categories, such as legal police killings and legitimate self-defense. The "edge" of what is a violent death may differ by culture, norm, or definition in national and international law. It is unclear whether deaths from executions (capital punishment), unsolved political disappearances, prison deaths, deaths in internment camps and detention centres, and migrant deaths in transit would or should be categorized as violent deaths (see Figure 1) by all stakeholders. Other commonly disputed categories include deaths from unintentional vehicular manslaughter, suicide, or indirectly as a result of war or other violence (indirect conflict deaths[3]). In some cases, deaths are already categorized in state recording of violent deaths.

FIGURE 1

What is a violent death?



A. Intentional Homicides
B. Unintentional Homicides
C. Killings in Legal Interventions
D. Direct Conflict

E. Assassinations
F. Extra-Judicial Killings*
G. Deaths during protests that turn violent
H. Communal violence that results in death
I. Deaths from drone strikes
J. Deaths due to explosive remnants of war (landmines and unexploded ordinance)

K. Police Killings*
L. Unsolved Disappearances
M. Deaths in Prison, Internment and Detention Centres*
N. Capital punishment*
O. Unexplained deaths by poison, chemical or biological weapons*
P. Deaths of migrants, refugees and the displaced in transit or camps

Sources: GVD: Items A—D; E—P are authors' estimates. Categories roughly to scale.
* Deaths counted by some countries as "killings in legal interventions."

3 Conflict-related deaths are often divided between **direct deaths**—deaths occurring as direct consequences of armed conflict, such as those caused by a weapon or an act of aggression—and **indirect deaths**—deaths caused by indirect consequences, such as the destruction of vital health and sanitation infrastructure or disruption of food supplies that may cause famine or disease outbreaks.

Deaths alone do not define violence.[4] In many cases, reporting deaths can be biased and thus, result in an inaccurate count of violence. Also, not all violence results in deaths. The first target of SDG 16 is reducing violence and violence-related deaths, not simply deaths. The consortium does not take the view that simply creating a registry of violent deaths will be sufficient for measuring global violence; rather GReVD is a necessary first step toward resolving multiple definitional and methodological issues that will lead to better research on violence and conflict.

The "edges" of what are considered a violent deaths matter for monitoring, because national actors and the international community require consistent and comparable data over time for evidence-based decision-making and measuring the impact of interventions aimed at reducing violence. Analysis on global trends based on current data requires greater scrutiny, particularly where methodologies remain disparate and definitions remain incomparable. The following sections review approaches to measuring violent deaths to locate the GReVD approach in the current context.

Understanding the UN Definition: SDG 16.1

The two main indicators for measuring progress on SDG 16.1 are 16.1.1—number of victims of intentional homicide per 100,000 population, by sex and age; and 16.1.2—conflict-related deaths per 100,000 population, by sex, age, and cause. A significant difference between the two indicators is that international bodies and participating member states clearly define intentional homicide, and more data are available for intentional homicides for monitoring purposes than for conflict-related deaths.[5] Therefore, indicator 16.1.1 is designated as a tier 1 indicator by the Inter-Agency and Expert Group on SDG Indicators (IAEG-SDGs), whereas indicator 16.1.2, which pertains to conflict-related deaths, is classified as a tier 2 indicator.[6]

Intentional Homicide

In defining intentional homicide, the IAEG-SDGs refers to the International Classification of Crime for Statistical Purposes (ICCS), which defines intentional homicide as the unlawful death inflicted upon a person with the intent to cause death or serious injury.[7] These sources are primarily administrative (see Channel 1: Administrative Sources in part 3 below). Data on intentional homicides are well summarized and described by the United Nations Office on Drugs and Crime (UNODC) in *The Global Study on Homicide 2019* (see Box 1).

Notwithstanding the exhaustive work by UNODC to collect this administrative data on intentional homicides, there remain multiple issues with comparing numbers among countries. As noted by ICCS, there remains "a lack of standardised concepts and internationally agreed statistical frameworks" to ensure consistency across countries. For example, the "same

4 See Fatalities, https://www.acleddata.com/wp-content/uploads/2017/12/Fatalities_FINAL.pdf.

5 Robert Muggah and Katherine Aguirre Tobón, "Citizen Security in Latin America: Facts and Figures," Strategic Paper 33 (April 2018), https://igarape.org.br/wp-content/uploads/2018/04/Citizen-Security-in-Latin-America-Facts-and-Figures.pdf.

6 See SDG Indicators, https://unstats.un.org/sdgs/metadata/. Tier 2 designation for 16.1.2 was updated in March 2019 and may not yet be reflected online. At the time of writing this report, the IAEG-SDG process was defining the methodology for reporting on SDG 16.1.2 through UN agencies, led by OHCHR.

7 UNODC, *International Classification of Crime for Statistial Purposes* (Vienna: UNODC, March 2015), https://www.unodc.org/documents/data-and-analysis/statistics/crime/ICCS/ICCS_English_2016_web.pdf.

act may be criminalized under different legal provisions in different countries" or "may be considered a criminal offence in one country but not in another."[8] Most countries omit accidental deaths (such as vehicular accidents) in their homicide statistics although in some countries, for example Switzerland, Turkey, and Ukraine, such deaths are included.[9] Other countries have included these numbers and then stopped, complicating time series analysis. Homicide data registered according to the Criminal Code of the Russian Federation, on the other hand, do not separate between committed and attempted homicides.[10] As a result, aggregate trends that combine numbers from multiple countries may be inconsistent.

8 Ibid.

9 Paul R. Smit, Rinke R. de Jong, and Catrien C.J.H. Bijleveld, "Homicide Data in Europe: Definitions, Sources, and Statistics," in *Handbook of European Homicide Research* (Springer, 2012).

10 See http://crimestat.ru/offenses_chart.

BOX 1: SUMMARY OF THE 2019 UNODC *GLOBAL STUDY ON HOMICIDE*

Based on the ICCS definition of intentional homicide, *The Global Study on Homicide 2019* reports intentional homicide statistics and estimates by geography and other variables. Where possible, it describes lethal violence disaggregated by situational context (homicide related to interpersonal conflict, homicide related to criminal activities, and homicide related to socio-political agendas); disaggregates homicides by gender (including a special focus on femicide), age; and analyzes the intentional homicide counts, rates, and trends at multiple levels—from global through national to subnational. The study includes analysis on the method of homicide, including firearms and sharp objects, and also reports on drivers of homicide at the "individual" (age and sex of a person) and "macro" (e.g., unemployment, inequality, lack of rule of law) levels. The UNODC Homicide Statistics 2019 dataset covers 202 countries and territories using this definition.

UNODC homicide estimates are based on preferred source at the country level as the basis for sub-regional, regional, and global estimates. Reliable data from criminal justice sources are typically given preference over public health data because public health data do not fully consider intentionality and legality. Intentionality and unlawfulness are central components of ICCS definition of lethal violence. In selected countries, mortality data produced by public health/civil registration authorities and collected by WHO have been used as preferred estimates for intentional homicide when these are based on actual counts of deaths by cause. Based on available national data, regional and global estimates were produced based on modeling techniques.

The UNODC study developed a data quality score for each country and territory in the database based on five quality components (comparability, completeness, timeliness, internal consistency, and external consistency) to assess the quality of published homicide data. The 2019 dataset also extends time coverage (1990–2017) compared to previous editions.

Nigeria is one example of the complexity of reporting and collecting homicide records. Large discrepancies exist in the figures reported: the Nigeria Police Force (2,712 homicides in 2012), Nigeria Watch (4,127 homicides in 2016), and the National Bureau of Statistics (average of 8,264 persons per year imprisoned from 2013 to 2016 for committing "murder"). These numbers are largely inconsistent with modeled estimates from WHO, which reports 17,059 in 2012. To assess the level of uncertainty, the UNODC together with the National Bureau of Statistics of Nigeria, launched a representative large-scale survey in April–May 2016 with a sample of 33,067 households. An adult household member was asked about any incidents of violent deaths in the household during 2013–2016. The results of the survey combined with a simple modeling exercise resulted in an estimated 64,000 incidents of intentional homicides annually for that period. The yielded result of this methodology reinforces the need for improved data collection in countries with unreliable or missing data.

Source: Authors' summary of UNODC (2019).

Conflict-Related Death

Defining what constitutes a conflict-related death is even more complex than that of intentional homicide. The general definition of armed conflict is the involvement of states and/or well-defined armed groups contesting states for control of territory or government. By any definition, the nature of warfare is changing; one important feature of contemporary armed conflicts is the increased blurring between state and non-state actors (NSAs), and the possible support of NSAs by states.[11] The evolving nature of armed NSAs can result in events that lead to violent deaths as a result of confrontation between organized armed groups or those that target a specific unarmed segment of the population, without the direct involvement of the state.[12] These events are inconsistently captured by current coding approaches. The sources for these data are primarily media and secondary sources (see Channel 2: Media Sources in part 3 below).

The motivation for violence may also be changing. New armed actors using violence may be driven by economic or material motives, rather than ideology or political objectives. Civilians may become the target of intentional or unintentional violence, both because of counter-insurgency tactics and as NSAs seek to control both legal and illegal economies. Armed violence can also change from one form to another over time, such as the transition of wartime violence into non-conflict crimewaves.[13] As a result, the distinctions between combatant and non-combatant, legitimate violence, and criminality break down, which in turn has implications for conflict and homicide data. Against such complexities, armed conflict is defined differently within and across disciplines. Some definitions are based on legal instruments, principally the preconditions that trigger applying the different Geneva Conventions governing international and non-international armed conflicts. Some databases use stated purpose or motivation of actors as a rule for inclusion, including ACLED,[14] GTD,[15] and UCDP. In some cases, they also rely on levels of conflict, such as UCDP and the Correlates of War (COW)[16] datasets.

Adding to the complexities described above, violent deaths—specifically, conflict-related violent deaths—are frequently distinguished between those caused by direct and indirect violence. Although in theory it may be easy to distinguish between direct and indirect violence, it is oftentimes difficult to do so in practice.[17] Deaths caused by direct violence commonly refer to those caused by a weapon or other acts of aggression, whereas deaths due to indirect violence might include deaths caused by conflict-induced famine; starvation as a political weapon; and other worsening of social, economic, and health conditions in the conflict-affected area.

11 Charles W. Mahoney, "Splinters and Schisms: Rebel Group Fragmentation and the Durability of Insurgencies," *Terrorism and Political Violence* (2017): 1–20, https://www.tandfonline.com/doi/abs/10.1080/09546553.2017.1374254.

12 Taylor B. Seybolt, Jay D. Aronson, and Baruch Fischhoff, "Introduction," in *Counting Civilian Casualties: An Introduction to Recording and Estimating Nonmilitary Deaths in Conflict*, ed. Taylor B. Seybolt, Jay D. Aronson, and Baruch Fischhoff (New York: Oxford University Press, 2013).

13 Ibid.

14 See http://www.acleddata.com/.

15 See Global Terrorism Database, https://www.start.umd.edu/gtd/.

16 See Correlates of War Project, http://www.correlatesofwar.org/.

17 Hana Salama, *Counting Casualties: Operationalising SDG Indicator 16.1.2 in Libya* (Switzerland: Small Arms Survey, 2018).

Health Approaches

Health specialists differentiate between natural and non-natural deaths. Natural deaths are deaths caused by natural causes like disease or illness, whereas non-natural deaths are deaths due to accidents, injuries, homicides, suicide, assaults,[18] falls, poisoning, or other means. Some authors have suggested that measures of unnatural deaths can be used to monitor a population's social, physical, and mental health.[19]

To understand the magnitude and pattern of unnatural deaths, forensic studies are usually conducted, using retrospective surveys, autopsies, and other health information systems. Although non-natural causes of deaths constitute a small fraction of all types of mortality, the trend has been changing worldwide over the past two decades.

It is important to highlight that the focus of public health organizations on *cause* of death differs from peace and conflict researchers' focus on the *context* of death.[20] For instance, the Global Burden of Disease (GBD)[21] study included in the Global Health Data Exchange (GHDx) focuses on causes of disease and risks, including cause of death.[22] Another example is WHO's Global Health Observatory (GHO) data, which provides health-related statistics, including for SDG targets, and uses causes of death to categorize data. The GHO also monitors such health-related priorities as mortality and burden of disease.[23]

Both GHDx and WHO differentiate causes of deaths into three main categories: (1) communicable, maternal, neonatal, and nutritional conditions (11.4 million deaths in 2016); (2) non-communicable diseases (stroke, cancer, and other causes; 40.5 million deaths in 2016); and (3) deaths due to injuries (4.9 million). Within this typology, injuries are further disaggregated by (3a) intentional injuries, defined as interpersonal violence (homicide, sexual assault, neglect and abandonment, and other maltreatment), suicide, and collective violence (war) (1.5 million deaths in 2016); and (3b) unintentional injuries (mostly road traffic injuries; 3.4 million deaths in 2016).[24]

Excess Mortality

The presence of violent deaths may also be indicated by the presence of excess mortality. Excess mortalities may be caused by armed conflict, when, for example, non-combatants may be displaced or become otherwise unable to access healthcare, food, and shelter due to destroyed infrastructure or weak governance.

18 Assaults include assault by smoke, fire, crashing, drugs, bodily force, etc. The International Classification of Disease also includes assault by unspecified means (homicide, manslaughter) but excludes injuries due to legal interventions and operations of war.

19 T.S. Mohan Kumar et al., "Profile of Unnatural Deaths in Manipal, Southern India 1994–2004," *Journal of Clinical Forensic Medicine* 13, no. 3 (2006); Prithvirajsinh Vaghela, "Profile of Unnatural Deaths in Bhuj (Gujarat): A Retrospective Study," *National Journal of Integrated Research in Medicine* 3, no. 2 (2012).

20 This distinction was expressed by UMD researchers during the preparation for the GReVD Workshop in Geneva.

21 In addition to cause of death, GBD includes data on years of life lost (YLLs), years lived with disability (YLDs), disability-adjusted life years (DALYs), life expectancy, probability of death, healthy life expectancy (HALE), and maternal mortality ratio (MMR). These data can also be used to assess impact of violence and conflict beyond death.

22 See Global Burden of Disease Study 2017 (GBD 2017) Data Resources, http://ghdx.healthdata.org/gbd-2017.

23 See About the GHO, https://www.who.int/gho/about/en/.

24 *Global Health Estimates 2016: Disease Burden by Cause, Age, Sex, by Country and by Region, 2000–2016* (Geneva: World Health Organization, 2018).

Excess mortalities estimations are based on expectations of mortality for a given population in a given situation. Excess mortalities are the difference between "expected deaths" from baseline mortality and "observed deaths" during a supposed causal event (which reflects a deviation from the norm observed prior to the event). Thus, the measure of excess mortality can indicate the severity of events.

Excess deaths are those that would not have occurred in the absence of conflict and include both direct and indirect deaths.[25] Typically, this is indicated by crude mortality rates (CMR), which are above and beyond deaths that would have occurred in the absence of war or other events. Of course, measuring this effect depends greatly on the availability and quality of population data before, during, and after armed conflict. Excess mortality may be inaccurate if the quality of CMR data (determined by health infrastructure and underlying social and economic conditions before conflict onset) are linked to the onset of conflict, which would most likely result in undercounting. Furthermore, it is difficult to distinguish between indirect and direct conflict deaths using excess mortality unless cause or context of a death are included in the data.

Excess mortality is difficult to estimate in countries with weak statistical capacity, where there are few reliable data on baseline mortality. Moreover, excess mortality rates may provide little practical information on perpetrators, specific types, and patterns of violence, which are required for attribution and intentionality. Population-based surveys estimate more violent deaths than do incident-reporting techniques, or rapid epidemiological surveys or demographic assessments.[26] Excess mortality might also be caused by disaster, epidemic, or other non-violent cause of death. Although excess mortality-based methods generate multiple types of mortalities that may be preventable, not all preventable deaths or those captured by excess mortality may be considered violent deaths.

TOWARD A CONSORTIUM APPROACH TO DEFINING VIOLENT DEATHS

In consideration of the above, defining violent deaths is a fundamental first step in monitoring SDG 16.1. Resolving these definitional issues is necessary and should include establishing a common research ontology, infrastructure, and methodology, which the consortium and the broader peacebuilding community can use. As a first step toward solving these definitional issues, the consortium, through its workshops, identified the two most salient axes for differentiating violent deaths from other deaths: *preventability* and *intentionality*.

Preventability: Although competition and conflict may be a natural part of the human condition, if violence is not necessary to resolve conflict, then violent deaths are preventable. Systems and tools are in place to prevent the outbreak of armed conflict, ranging from high-level preventive diplomacy and mediation, structural and operational prevention, to the strengthening of local capacities for peace. Arguably, preventing violence requires doing

25 Robert Muggah, *Counting Conflict Deaths: Options for SDG 16.1*, https://igarape.org.br/wp-content/uploads/2015/10/counting-conflict-deaths-muggah-2015.pdf.

26 Geneva Declaration Secretariat, *Global Burden of Armed Violence: Every Body Counts* (Geneva: Small Arms Survey, 2015).

more of these, and better.[27] Furthermore, the implicit assumption prevails that non-natural deaths are more preventable than deaths due to natural or medical causes.

Intentionality: Although explicitly included in indicator 16.1.1, intentionality is also relevant when considering deaths associated with conflict environments. All three data-coding partners in the consortium (ACLED, GTD, UCDP) rely upon purpose or motivation definitions for coding political violence. In addition, intentionality may be used to identify violent deaths that may not immediately register as violence (e.g., indirect deaths may be the outcome of intentional strategies to suppress populations, such as using starvation as a weapon of war).[28]

These challenges are not new; peace researchers have long struggled with these issues. GReVD's work will build upon the foundations created by others, particularly the consortium partners. The sections below review the current approaches of consortium partners ACLED, GTD, Igarapé Institute, and UCDP and how these are reconciled within the current GVD methodology. The consortium does not yet have solutions to these classification problems. The current consortium approach is to create a registry that allows for all these salient characteristics to be included, so that the consortium and users can define violent deaths.

Armed Conflict Location and Event Data (ACLED) Project: Version 8

ACLED collects information on armed organized violence, specifically political violence and demonstrations across Africa, South and Southeast Asia, and the Middle East with new coverage from 2019 for Europe, Latin America and the Caribbean, and Central and East Asia.

Political violence is defined as the use of force by a group with a political purpose or motivation.[29] ACLED records political violence through its constituent events; its intent is to produce a comprehensive overview of all forms of political disorder, expressed through violence and demonstrations, within and across states. A politically violent event is a single altercation where often force is used by one or more groups toward a political end, although some non-violent instances—including protests and strategic developments—are included in the dataset to capture the potential precursors or critical junctures of a violent conflict. ACLED recognizes a range of actors including state forces, rebels, militias, identity groups, demonstrators, civilians, and external and other forces. All actors have an official name, a political purpose, and use violence or protest for political means. For inclusions, organizations must be cohesive and are not assembled for single events, except for riots and protests. Furthermore, the events of organizations must be connected to each other as a means to achieve a larger political purpose. This necessary and sufficient definition of actors allows for monitoring campaigns and trajectories of movements.

---

27 United Nations/World Bank, *Pathways for Peace: Inclusive Approaches to Preventing Violent Conflict* (Washington, DC: World Bank, 2018).

28 This is not to suggest that other indirect—and largely unintentional—deaths are of less concern. However, the reason for focusing on intentional and preventable violent deaths is to enhance precision and avoid, for the purpose of monitoring SDG16.1, the inclusion of deaths that are predominantly the outcome of suboptimal governance, even if such governance is an outcome of armed conflict. Where the actual edges of violent deaths will fall will be defined by the consortium in an upcoming initiative on ontology.

29 For a detailed account on definitions and coding procedures, see ACLED codebook (https://www.acleddata.com/resources/general-guides/).

ACLED data are gathered and coded by researchers globally, who have knowledge of local languages and contexts. Coding is done weekly. To ensure the most accurate data are cleaned and are analysis-ready, ACLED cleans and checks inter-coder, intra-coder, and inter-code reliability every week before releasing real-time conflict data. Coding is reviewed for inter-coder reliability by running a series of coding scripts that checks for correct numbering, event types, locations, etc., ensuring that coding is systematic across researchers. Coding is also submitted to test intra-coder reliability, which includes checks on each researcher's coding specifically. Finally, inter-code reliability is tested to ensure that notes relating to each conflict match the conflict event itself, and whether the event should ultimately be included within the dataset.

Furthermore, for specific events that are becoming more common in different conflict environments, there are known problems with the details around the violent event. Limitations exist in both local and media reporting thoroughness: there are issues of bias, false and fictional information, and a lack of verification. This is why ACLED deems local partners to be vital for any collection, although it also underscores that sourcing data in these environments is not a technical issue, and does not have purely technical solutions.

Global Terrorism Database (GTD)

The GTD defines terrorism as "the threatened or actual use of illegal force and violence by a non-state actor to attain a political, economic, religious, or social goal through fear, coercion, or intimidation."[30] The current version of the GTD builds on a wider inclusion criterion than other terrorism databases (for example, including attacks on combatant targets) that gives freedom to users to choose more narrowly defined acts of terrorism according to the user needs. GTD excludes violence by state actors.

Because various types of violence occur on the periphery of definitions of terrorism, the GTD team frequently encounters reports of attacks for which there is conflicting or unclear information regarding whether the inclusion criteria are satisfied. The team marks these attacks as "doubt terrorism proper" and records an alternative designation, such as "other crime type" (e.g., hate crime, organized crime, or interpersonal crime), "lack of intentionality" (e.g., the perpetrator may suffer from mental illness or be under the influence of drugs or alcohol, or the event was an accident), "intra/intergroup violence," or "state actor." Likewise, for attacks that targeted combatant entities in the context of an insurgency, the GTD notes "insurgency" as the alternative designation. These alternative designations are simply notes to provide context for cases that fall on the margins of the definition and to allow users to exclude them if a particular analytical question calls for it. The GTD is not a comprehensive source of data on these other types of violence (see Purgatories in part 4 below).

Methods for collecting data can have an effect on the comprehensiveness of the resulting dataset. For example, the GTD team made improvements in data collection and coding in 2012, specifically, relating to (1) the population of sources that are used to compile the database; (2) the procedures that are used to identify potentially relevant information in those sources; and (3) the workflow and technologies that are used to identify and code the events that are included in the GTD.[31] Above all, it has expanded the coverage in terms

---

30 National Consortium for the Study of Terrorism and Responses to Terrorism (START), *Codebook: Inclusion Criteria and Variables*, July 2018, https://www.start.umd.edu/gtd/downloads/Codebook.pdf.

31 Michael Jensen, "Discussion Point: The Benefits and Drawbacks of Methodological Advancements in Data Collection and Coding: Insights from the Global Terrorism Database (GTD)," https://www.start.umd.edu/news/discussion-point-benefits-and-drawbacks-methodological-advancements-data-collection-and-coding.

of the number and types of news sources used to produce the database. Formerly, GTD targeted an already identified set of news sources to collect data. New methods from 2012 have expanded the source pool, improved filtering, and use new methods for identifying unique articles with higher relevance. Furthermore, categorization now allows for filtering by geography, attack type and weapons, targets and perpetrators, and casualties. These workflow improvements used advancement in tools and technology to handle expanded data processing.

The Igarapé Institute

Since 2015, the Igarapé Institute has hosted the Homicide Monitor, one of the world's largest datasets of publicly available data on homicides. The Homicide Monitor collects data exclusively on intentional homicide, as defined by government police departments, with a focus on Latin America.

The Igarapé Institute acquires detailed micro-level data for the most violent countries, states, and cities in the region. The underlying data are retrieved from public sources, including the UNODC Homicide Statistics 2019 database. To date, information is available for 220 countries and territories, 904 states, and 680 cities with populations of 250,000 or more.

The Homicide Monitor restricts its data collection to authoritative administrative sources, including official reports supplied by state authorities and direct contact with primary sources, such as criminal justice departments (i.e., police statistical offices and general attorney's offices) and public health sources (i.e., department or ministry of health services). In most cases, states discriminate between intentional homicide and other forms of violent death. In many cases, information is already prepared in statistical format. The Homicide Monitor database includes all datasets, while the one featured in the data visualization platform is the most complete.

For almost all cases, data in the Homicide Monitor are supplied by the criminal justice (and legal medicine) departments at either the national or municipal level. In some cases, the next best dataset is supplied by the department or ministry of public health, although such data typically feature a lag of between one and two years. If no datasets are available, or if they do not meet a minimum standard of quality, then information is not included in the data visualization. In all cases, the Igarapé Institute solicits and stores information on homicide from all possible sources available. However, the legal attribution of the death can change over time. For example, deaths that may have been attributed to manslaughter, suicide, or a police self-defense can later be re-categorized as homicide. As such, it is important to routinely update/revise datasets as events are processed in national and state criminal justice systems.

Uppsala Conflict Data Program (UCDP)

The UCDP collects global data on organized violence (collective violence) in three categories: state-based armed conflict, non-state armed conflict, and one-sided violence. The categories are mutually exclusive so they can be aggregated to an overall measure of organized violence, including an overall count of violent deaths. All the data are collected as georeferenced events data (date/location).

News sources are mainly used to code UCDP data (70–80 percent), as well as IGO/NGO reports, truth commission reports, historical accounts, information from area experts, etc. News sources often contain reports on the individual events but less on context. IGO/NGO reports, conversely, provide the necessary context to interpret the events to give additional information on individual events, and aggregated figures with which to compare the sum of individual events. The approach of UCDP has always been to report death figures that can be confirmed with some level of certainty. Death figures are therefore either near the real figure or on the conservative side. UCDP uses available expertise and contextual information to judge what would be the most likely figure for any given event, and then report a best estimate of the number of deaths, together with a low and a high estimate. The high estimate should be understood as the upper limit of the range of plausible estimates—not as the highest number reported. Some reported numbers may be much higher, but judged to be implausible. Although the numbers given by the UCDP data are thus conservative, they are sufficiently consistent over time and across cases to map trends and differences across cases.

UCDP includes all organized violence that generates at least 25 deaths in a calendar year. Thus, deaths from non-organized violence and deaths from violence between organized groups generating less than 25 deaths in a calendar year are not included.[32] UCDP also does not include extrajudicial killings if they occur inside government facilities.[33]

Aside from the data that are released to the public, the UCDP database system also contains multiple violent events coded as "unclear," which are not made public (see Purgatories in part 4 below). Because the database is constantly revised with new information, data previously coded as unclear can be moved into publicly available datasets as they become available.

Situations with large-scale organized crime violence groups operating in the same area pose challenges. It is rare that organized crime gangs inflict 25 or more deaths against governments, civilians, or other gangs in one year, but when this happens it is often difficult to ascertain which group perpetrated the violence, leading to unclear events. Moreover, when it is possible to ascertain perpetrators, there may be some overlap between UCDP data and homicide data.

The Global Violent Deaths Database (GVD)

The Small Arms Survey's GVD database methodology is based on a unified approach to lethal violence and the conviction that prevention of all forms of violence and violent deaths is necessary to achieve "peaceful and inclusive societies," as envisaged in the 2030 Sustainable Development agenda.[34] National or other country-level indicators are used to

32 UCDP holds that this violence makes up a very small portion of all violence.

33 Extrajudicial killings inside government facilities are excluded because of difficulties in obtaining reliable information on deaths and causes of deaths inside government facilities.

34 See Claire Mc Evoy and Gergely Hideg, *Global Violent Deaths 2017: Time to Decide* (Geneva: Small Arms Survey, 2017); and Gergely Hideg and Anna Alvazzi del Frate, *Darkening Horizons: Global Violent Deaths Scenarios, 2018–30,* Briefing Paper (Geneva: Small Arms Survey, May 2019). Methodology for the database was developed through the Global Burden of Armed Violence (GBAV) study and accessible at http://www.genevadeclaration.org/fileadmin/docs/GBAV3/GBAV3-Methodological-Annexe.pdf.

track changes in lethal violence worldwide. The database contains data dating to 2004 and is updated yearly.

GVD is a multi-source database that consolidates several international and national data sources providing records of casualties, including a wide range of lethal violence. GVD uses—among others—data from GReVD consortium members ACLED, GTD, and UCDP.

The GVD's "unified approach" allows for counting all violent deaths under one composite indicator. Violence takes different manifestations in varied locations, which tends to add more ambiguity to conflict and non-conflict contexts. Typically, lethal violence is characterized by factors such as the context or the intentions of the perpetrator, which help identify the level of organization and the motivation behind the violent acts, broadly resulting in a distinction between organized (collective) and interpersonal (individual) violence, and between conflict (politically motivated) and criminal (economically motivated) violence. GVD highlights how these detailed distinctions and classifications still do not eliminate the complexities and ambiguities of recording violent deaths data.

GVD builds on national and cross-national specialized datasets to produce a global lethal violence estimate. Data is derived from multiple sources, most of which are "incident reporting" mechanisms. According to GVD, incident reporting includes passive surveillance of multiple people reported to have died in violent events through hospital, mortuary, police, or criminal justice data collection. The incident reporting systems usually result in three different types of data sources: one for criminal justice statistics, one for public health data, and one for direct conflict deaths.[35]

GVD Estimates on Violent Deaths

GVD estimates are based on two major series for lethal events data: *intentional homicide data* and *direct conflict death data*. Through the construction of the corresponding datasets (GVD Homicide Database and GVD Conflict Database) two estimates are produced (GVD Homicide Estimates[36] and GVD Conflict Deaths Estimates[37]). The GVD database provides national-level data. On the top of these two estimates, the GVD elaborates a global estimate that fills the gap for the categories of legal interventions and unintentional homicides.

The GVD Homicide Database includes data for 223 countries and territories dating to 2004. GVD provides disaggregated data on the gender of homicide victims and the use of firearms. GVD Homicide Database relies on both national (e.g., national statistical offices, ministries of interior, health, and justice, and the national police), regional and international sources (e.g., UNODC, WHO, Institute for Health Metrics and Evaluation (IHME), and Eurostat), as well as non-governmental organizations, media reports, and other observatories. GVD database contains absolute values of homicides per country. Additionally, it includes data on homicides at the subnational level (province, municipal, and city level).

35 Geneva Declaration Secretariat, *Global Burden of Armed Violence: Lethal Encounters* (Cambridge: 2011).

36 Where possible, these can be disaggregated by (1) homicide victims by sex, (2) homicide by firearm, and (3) homicide by firearm by sex.

37 Where possible, the GVD database disaggregates conflict deaths by sex and instrument.

The main sources of homicide data that GVD relies on are traditional sources of data on intentional homicides that are produced by the criminal justice and public health systems that get disseminated by governmental agencies such as national statistical offices. Among other data-providing institutions are national and international organizations like Eurostat or UNODC. Additionally, GVD makes use of the various observatories on violence, crime, and conflict that also provide data on intentional homicides at the national and local level.

The GVD Conflict Deaths Database includes data on conflict-affected countries as well as data from a set of countries that are affected by violence and insecurity.[38] The GVD produces comprehensive estimates that combine the two datasets with sources of data on other killings, such as legal interventions, terrorism, etc. The Small Arms Survey has identified the lack of consistent datasets on these forms of lethal violence as a significant gap. These types of violent deaths are often missing in either homicide data or conventional armed conflict datasets.

Disaggregation by Gender, Age, and Other Characteristics

It is well documented that most perpetrators and victims of violence are male and that most perpetrators are youth (15–29).[39] This notwithstanding, very little of the currently available data on violent deaths can be disaggregated by gender, age, or any other identifying characteristics of either perpetrators or victims. None of the leading databases on conflict deaths currently disaggregates at this level. It is likely that the Office of the UN High Commissioner for Human Rights (OHCHR) data on conflict deaths associated for indicator 16.1.2 can be disaggregated at this level, but at the expense of coverage for incidents that do not include these identifying characteristics. Such disaggregation will be necessary for policymakers attempting to monitor progress on SDG 16 for the 2030 Agenda. It is not sufficient to merely monitor the overall increase or decrease in violent deaths; even if homicide and conflict deaths decline, it may not result in declines for all groups. If, for example, there is a decline in males as victims of violent death but little or no change in the rate of females as victims of violent death, there may be evidence of structural issues with prevention.[40] Without disaggregation of this kind, global violence trends cannot be fully understood. Given these challenges and coverage, it is also not possible to determine whether violent deaths are undercounted for groups of interest (income, sex, age, race, ethnicity, migratory status, disability and geographic location, following UNStats guidance). Event time and location information are necessary to match counts of violent deaths to identity and characteristics of individual entries in the registry.

---

38 For inclusion criteria for collection of conflict deaths, see draft version "Methodological Annexe: Annexe of the *Global Burden of Armed Violence 2015: Every Body Counts*," http://www.genevadeclaration.org/fileadmin/docs/GBAV3/GBAV3-Methodological-Annexe.pdf.

39 See GBAV 2011 and GBAV 2015, particularly on disaggregation by age.

40 Jacqui True, "Winning the Battle but Losing the War on Violence: A Feminist Perspective on the Declining Global Violence Thesis." *International Feminist Journal of Politics* 17, no. 4 (2015): 554–572.

PART 3 A Common Framework for Understanding Gaps in Coding of Violent Deaths

To better understand the challenges of coding violent deaths, the GReVD consortium has developed a framework to describe the process by which a violent death might be coded. This 5x9 framework, which combines five channels for coding (first section) and nine possible stages of coding (second section), is illustrated in Figure 2. The challenges associated with coding are further described in the section using the 5x9 Framework to Assess Gaps (below).

FIVE CHANNELS FOR CAPTURING VIOLENT DEATHS

The raw data on violent deaths are predominantly recorded through five different channels: (1) administrative data from state institutions and intergovernmental bodies, (2) media reporting, (3) monitoring by expert groups/observatories, (4) representative surveys, and (5) direct reporting by the public (including through social media). Most of the data on homicides come from administrative records, whereas most data on conflict-related deaths are compiled through media reports. What all these disparate sources have in common is that they record available and observable information.[41] Yet for multiple reasons, which this section details, data produced by these sources are largely incomplete, biased, and frequently of a political nature (see Box 2: Political Challenges of Counting Violent Deaths). The data only cover a proportion of the total number of violent deaths, which results in a disparity between the recorded number of violent deaths and the real—but unknown—number of violent deaths.

41 Todd Landman and Anita Gohdes, "A Matter of Convenience," in *Counting Civilian Casualties: An Introduction to Recording and Estimating Nonmilitary Deaths in Conflict*, ed. Taylor B. Seybolt, Jay D. Aronson, and Baruch Fischhoff (New York: Oxford University Press, 2013).

BOX 2: POLITICAL CHALLENGES OF COUNTING VIOLENT DEATHS

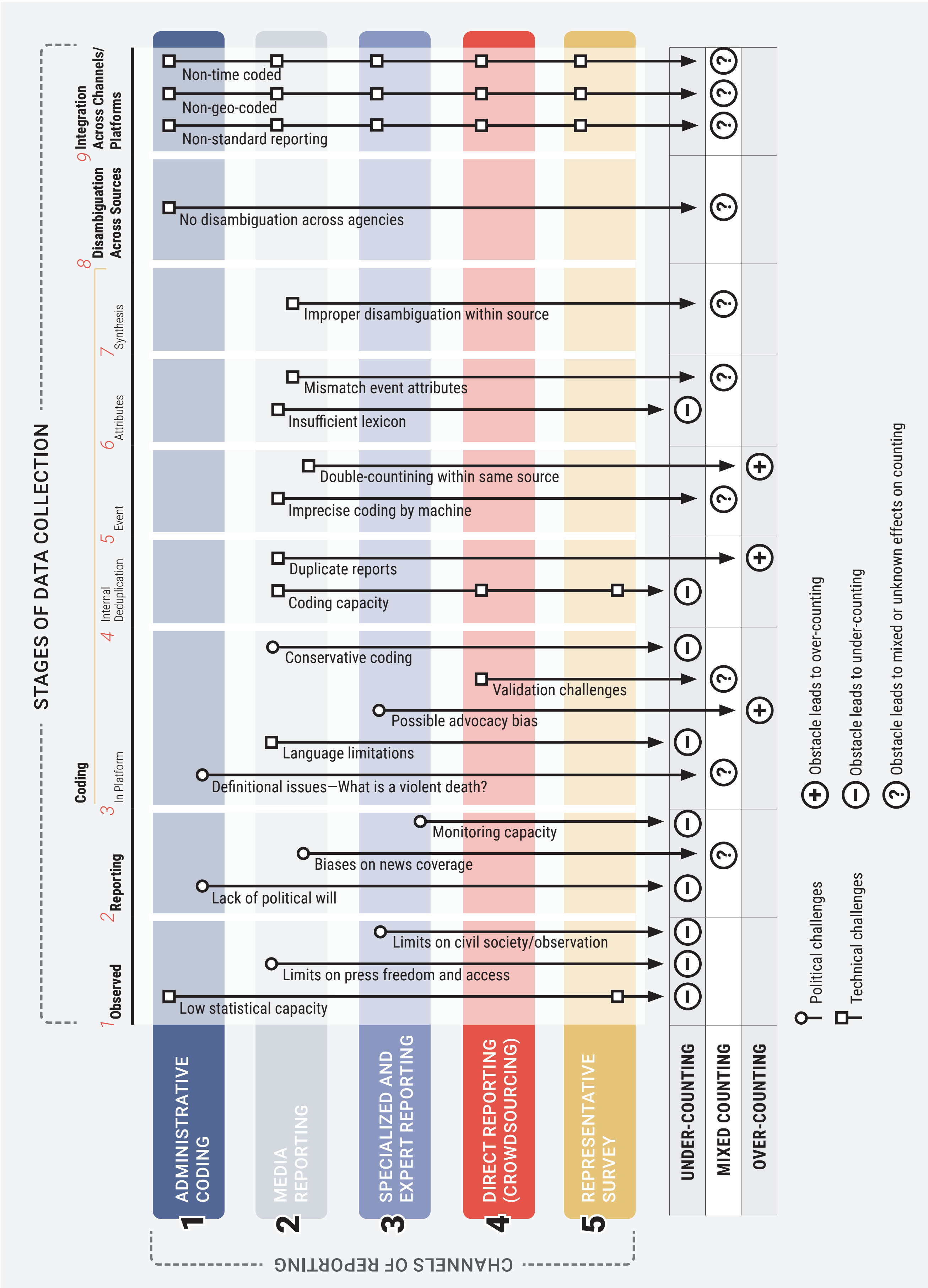
Counting and classifying violent deaths is not a politically neutral exercise. Choosing what to count inevitably makes a political claim about what not to count. Figures are highly influential, and the ability to manipulate counts—by inflating, downplaying, or censoring numbers—can be used and abused by those in power. Key challenges for counting violent deaths include who generates and uses these data, and for what purposes.

State institutions generate much of the existing mortality data (see Channel 1: Administrative Data). Although this is an important function of governments and part of the common administrative procedure of recording and documenting vital statistics and criminal justice data, data can be presented to form specific narratives. Counting and manipulating casualties might be done to serve operational purposes (e.g., the highly inflated figures presented by the U.S. Defense Department on the number of communist fighters killed in the Vietnam war as an indicator of success[42]). Battlefield reporting in the 1991 Gulf War, conversely, was largely censored in an attempt to prevent the recording of civilian casualties that might negatively affect the perceived legitimacy of the intervention.[43] Mortality data for conflicts can also influence the type of narrative created by a government around the nature of the conflict and its role in it. Media (see Channel 2: Media Reporting) and nongovernmental organizations (NGOs) are other important sources recording casualties (see Channel 3: Specialized Reporting), which may also suffer from bias. Human rights organizations and NGOs may have advocacy aims. As a result, contexts and narratives of a conflict are often only partly reported to the public.

42 Keith Krause, "Bodies Count: The Politics and Practices of War and Violent Death Data," *Human Remains and Violence: An Interdisciplinary Journal* 3, no. 1 (2017).
43 Ibid.

FIGURE 2

Understanding the 5 x 9 Framework



Channel 1: Administrative Coding

Administrative data on violent deaths are captured by two main sources: the criminal justice system and/or the public health system. These are illustrated in row 1 of Figure 2. There are important differences in the data produced by these two institutions.

Criminal justice institutions classify (violent) deaths primarily based on intent. Incident-based reports produced by state agencies' part of the criminal justice system, such as the police and the military, often provide information on the weapon used, perpetrator, time, and place, but may not provide information on victims. Court data in published form may be less amenable to coding of event data because of reporting formats. Nevertheless, court judgements, where available, may provide verification for other data. Backlogs in processing criminal cases can mean impunity for violent deaths and inaccurate data.

Public health institutions on the other hand, classify deaths based on the type of injury that was the cause of death (e.g., gunshot wound). These data often include victim and cause of death, but rarely include data on perpetrators.[44] Furthermore, when data are coded off of death certificates, they may include very little information on the event.[45] As a result, preventability, motivation or intent, factors which may be necessary to determine whether a death should be counted as a violent death, may be unavailable.

Different state institutions or administrations often also have different counting rules for violent deaths. Many jurisdictions do not count police killings as criminal deaths, including in Nigeria, Kenya, and Jamaica.[46] In the United States, such reporting is voluntary. In some jurisdictions, up to 20 percent of violent deaths are caused by the police but not captured in homicide statistics.[47] Moreover, administrative data compiled by state institutions are predominantly produced for internal monitoring rather than for public release. When data are released, they are often in the aggregate, rather than specific, individual event.[48] This can make it nearly impossible to reconcile such administrative data with event data (see Stage 9 in Coding Violent Deaths, below).

Globally, there are large disparities between countries and regions in capacity and political will for reporting violent deaths. Monitoring and recording capacities are correlated with overall levels of economic and institutional development.[49] According to WHO and the World Bank, most developing nations are unable to accurately record deaths. Ninety-one percent of African states lack credible mortality data.[50] In 70 countries, neither the criminal justice

44 John Sloboda and Elizabeth Minor, *Paper 3: The Range of Sources in Casualty Recording* (London: Oxford Research Group, 2012).

45 Asha Z. Ivey-Stephenson, Janet M. Blair, and Alex E. Crosby. "Efforts and Opportunities to Understand Women's Mortality due to Suicide and Homicide Using the National Violent Death Reporting System," *Journal of Women's Health* 27, no. 9 (2018): 1073–1081.

46 Keith Krause, "Challenges to Counting and Classifying Victims of Violence in Conflict, Post-Conflict, and Non-Conflict Settings," in *Counting Civilian Casualties: An Introduction to Recording and Estimating Nonmilitary Deaths in Conflict*, ed. Taylor B. Seybolt, Jay D. Aronson, and Baruch Fischhoff (New York: Oxford University Press, 2013).

47 Ibid.

48 Megan Price and Patrick Ball, "Big Data, Selection Bias, and the Statistical Patterns of Mortality in Conflict," *SAIS Review of International Affairs* 34, no. 1 (2014).

49 Krause, *Counting Civilian Casualties: An Introduction to Recording and Estimating Nonmilitary Deaths in Conflict*.

50 World Bank and WHO, *Global Civil Registration and Vital Statistics: Scaling up Investment Plan 2015–2024* (World Bank Group, 2014).

nor the public health systems produce statistical information on homicide.[51] Nevertheless, even in developed countries, data on violent deaths are irregular. For example, in 2018, only 40 out of 50 U.S. states participated in the National Violent Death Reporting System.[52]

Administrative data are particularly problematic in conflict-affected countries. Armed conflict can lead to breakdowns of the mechanisms that normally make death notifications available to a nation's statistical authority as well as weakened or collapsed public health and criminal justice institutions.[53] Consequently, official data in conflict-affected countries are often missing. This further complicates attempts to reconcile violent deaths from multiple channels in these countries.

Finally, as mentioned earlier, global aggregates suffer compounded complications due to differences in how national criminal justice systems classify killing based on intent. Homicide, manslaughter, unintentional homicide, and even vehicular deaths may be present in administrative criminal records, jointly or separately, depending of the statistical traditions of a particular country.

Channel 2: Media Reporting

Media reports are the most common sources for conflict datasets, including violent deaths due to conflict. Although the quality of media reporting may vary by context—depending on transparency and freedom—and availability, it has multiple strengths in terms of information necessary for coding (e.g., date, place, number killed, means, perpetrator).[54]

Media reporting may be subject to selection bias and may underreport or overreport certain types of events due to the characteristics of the event. The probability of reporting a violent death may depend on where the event happened, media access to conflict-affected areas, race of the victim or perpetrator, and the perceived newsworthiness of an event. Violent deaths in remote, rural areas are less frequently reported than killings in urban areas. Bias against rural areas is influenced by factors such as inadequate road infrastructure, lack of electricity, distance of the event relative to news agencies, and their principal readership. Local news outlets in general have better coverage of rural and hard-to-reach areas than international media; however, using such sources for monitoring violent deaths is often restricted due to language barriers and the cost of collection (see Ways forward for GReVD in part 4 below). Journalists are also at times deliberately denied access to conflict sites by governments or NSAs. In other instances, journalists wishing to maintain working relationships with state actors are less likely to report on violence in which the perpetrators are identified as state agents.[55] There are few conflict regions, however, reported to be completely inaccessible to journalists.[56]

---

51 UNODC, *Global Study on Homicide 2013* (United Nations, 2014).

52 Centers for Disease Control and Prevention, "CDC's National Violent Death Reporting System Now Includes All 50 States," Press Release (September 5, 2018), https://www.cdc.gov/media/releases/2018/p0905-national-violent-reporting-system.html.

53 Ewa Tabeau and Jan Zwierzchowki, "A Review of Estimation Methods for Victims of the Bosnian War and the Khmer Rouge Regime," in *Counting Civilian Casualties: An Introduction to Recording and Estimating Nonmilitary Deaths in Conflict*, ed. Taylor B. Seybolt, Jay D. Aronson, and Baruch Fischhoff (New York: Oxford University Press, 2013); and Salama, *Counting Casualties*.

54 Sloboda and Minor, *Paper 3: The Range of Sources in Casualty Recording*.

55 Christian Davenport and Patrick Ball, "Views to a Kill: Exploring the Implications of Source Selection in the Case of Guatemalan State Terror, 1977–1995," *Journal of Conflict Resolution* 46, no. 3 (2002).

56 Sloboda and Minor, *Paper 3: The Range of Sources in Casualty Recording*.

In addition to geographical and other access factors, the probability that a given event is reported is related to the scale of the event. The probability of coverage and the number of sources covering the event increases when events are large and particularly violent, and when there is no other major news story occurring at the same time. Conversely, events that involve only one victim are less likely to be documented.[57] The effects this has on monitoring violent deaths are that events that are covered by a single source are less likely to be picked up by databases gathering recordings on conflict-related deaths. Nevertheless, such generalization needs to be treated with caution as documenting violence can be a dangerous task and an increase in violence can inversely lead to a decrease in reported violence.[58]

Armed conflicts and other violent events are generally considered to be relatively newsworthy and therefore likely to be reported.[59] Such events are considered particularly newsworthy when newspapers are for-profit and express an interest in protest and/or human rights.[60] Nevertheless, newsworthiness tends to fluctuate: media "fatigue" may result in under-documentation later in a conflict, or when other newsworthy stories may limit the amount of time and space available to cover victims of a specific conflict.[61] Although the interest of international media in a specific armed conflict gradually declines over time, it increases when a great power military or UN-led forces intervene.[62] News reports are less useful in enumerating violent deaths resulting from criminal activity and everyday interpersonal violence, particularly in countries where violent crime is relatively common.

Fatality accounts risk suffering from description bias. It is not only *that* an event is reported that is of importance, but also *how* an event is reported.[63] Difficulties arise in distinguishing between perpetrators and victims, particularly when armed actors camouflage themselves among the civilian population. Media reporting on conflict-related fatalities is many times subject to government or rebel propaganda. The issue of description bias is especially severe when media outlets receive information directly or indirectly from conflict actors. Nevertheless, a large share of events would be missed if such reports were excluded, which would in turn further exacerbate selection bias.[64]

Channel 3: Specialized and Expert Reporting

Armed violence monitoring systems—specialized surveillance systems that collect data from multiple sources, often focusing on a particular kind of violent death—also record violent deaths. These monitoring systems continually and systematically collect and analyze

57 Price and Ball, "Big Data, Selection Bias."

58 Davenport and Ball, "Views to a Kill."

59 Magnus Öberg and Margareta Sollenberg, "Gathering Conflict Information Using News Resources," *Understanding Peace Research: Methods and Challenges* (Abingdon: Routledge, 2011).

60 Davenport and Ball, "Views to a Kill."

61 Price and Ball, "Big Data, Selection Bias."

62 Sven Chojnacki et al., "Event Data on Armed Conflict and Security: New Perspectives, Old Challenges, and Some Solutions," *International Interactions* 38, no. 4 (2012).

63 Ibid.

64 Ibid.

data on armed violence, and include crime or violence observatories and early warning systems. Expert and niche monitoring sources include organizations such as NGOs focusing on human rights violations, drone strikes, or landmines. Such monitoring systems provide a specialized understanding of specific issues or sub-themes of violence and are integral to monitoring violent deaths, because they record data that are often not available through media streams or they may have access to sources other than the media. They are also particularly useful in monitoring and recording violent deaths resulting from criminal activity in countries where violent crime is common and thus not considered newsworthy by the media.

Crime and violence observatories, which are mostly found in Latin America and Africa, have three basic functions: collect data, analyze data, and disseminate data publicly, which may include advocacy. In terms of their thematic focus, they can be either general observatories (monitoring violence, security, crime, etc.,) or specialize in specific thematic issues (landmines, youth violence, sexual violence, extrajudicial killings, organized crime, etc.).

Conversely, conflict early warning systems focus on systematic data collection and analysis that identify where armed conflict could erupt or aggravate. They are usually set up in the latent stages of a perceived potential armed conflict to foster preventative action. Because observatories and early warning systems mainly derive their data from criminal justice and vital registration statistics, hospital and morgue records, NGO and media reports, and household surveys, it can be argued that these kinds of monitoring systems are biased in favor of settings with functioning governmental registration systems or healthy media coverage.[65] Furthermore, these channels have many of the same strengths and weaknesses of the other channels described here.

In addition to observatories and early warning systems, international NGOs, such as Amnesty International and Human Rights Watch, and local civil society organizations operating in conflict-affected countries often monitor and report violence. Data can include recording of date, time, place, identity of victims, and circumstances of death. Where these actors can operate independently, they can provide information about state activity otherwise unreported. However, as a result, human rights NGOs may cover specific types of regimes, and ignore others.[66] Moreover, the quality of NGO data varies considerably; data (particularly aggregated data) may be unsourced or imprecise and can be difficult to integrate with other data without risk of double-counting.

Channel 4: Direct Reporting (Crowdsourcing)

Crowdsourcing has been used to collect information on violent incidents from the affected population, increasingly through new technologies such as text messages, email, Twitter, Facebook, and other social media. Micropublishing incidence of violence can further use advanced technology to filter, process, and georeference information received, making data available and mapped within minutes or hours of an incident.[67] These reports can be

65 Elisabeth Gilgen and Lauren Tracey, *Contributing Evidence to Programming: Armed Violence Monitoring Systems* (Geneva Declaration, 2011).

66 Davenport and Ball, "Views to a Kill."

67 Sloboda and Minor, *Paper 3: The Range of Sources in Casualty Recording*.

searched for and collected by casualty recorders. Increasingly, press and media organizations are incorporating social media outputs into their own data collection and dissemination processes.[68]

Using social media for micropublishing is valuable in conflicts characterized by restrictions on the free movement of journalists and other data-gatherers (e.g., human rights monitors) on the ground.[69] However, direct reporting using these technologies is limited to places where mobile internet technology is widespread. Furthermore, because documents are derived from different individuals, it can be highly variable in content and quality, creating challenges for filtering and aggregating the information. Algorithms that extract casualties from social media need to simultaneously address noise unique to social media source and the significant difficulties encountered in event extraction from traditional sources.[70] Social media reports are also incredibly difficult to verify; unlike news reporting, social media reports do not pass through any verification schemes. Immediacy and speed may sacrifice accuracy and analysis. Social media may also be used to spread government or NSA disinformation using bots and propaganda.

Channel 5: Representative Surveys

Survey methods can also serve as a substitute for monitoring violence. The approach typically involves randomized sampling of households in conflict-affected zones to obtain basic data on family size, adult and child mortality rates, and reported causes of death. Random sample surveys, significantly less expensive than censuses, are also useful when demographic records are not maintained or have been discontinued during a conflict period. Nevertheless, undertaking high-quality population surveys in conflict zones is extremely difficult and places both interviewers and interviewees at high risk. Hence, such surveys are often conducted in the post-conflict period once the intensity of violence is reduced.[71]

In conflict zones and other humanitarian crises, indirect mortality is also frequently estimated using retrospective mortality surveys (RMS). The approach often involves random or semi-random cluster sampling of the national and/or area-specific population. The main advantage is the ability to rapidly assess mortality in areas where prospective surveillance does not exist. Moreover, RMS methods have been standardized through an inter-agency humanitarian initiative (Working Group for Mortality Estimation in Emergencies). However, RMS also have significant limitations. They may not capture the true medical causes of death because information cannot be independently verified. It often can be difficult to distinguish between violent and non-violent causes, and logistical challenges may frustrate RMS because data are often politically sensitive.[72] As a result of these limitations, there may not be enough high-quality surveys on which to base global estimates. However, given

68 Salama, *Counting Casualties*.

69 Ibid.

70 Specific difficulties with extracting events from social media reports include duplication challenges caused by the sheer amount of data to sift through, with many people talking about the same event, and determining the authenticity of social media reports, which suffer from propaganda and bots, among others.

71 Jana Asher, "Using Surveys to Estimate Casualties Post-Conflict," in *Counting Civilian Casualties: An Introduction to Recording and Estimating Nonmilitary Deaths in Conflict*, ed. Taylor B. Seybolt, Jay D. Aronson, and Baruch Fischhoff (New York: Oxford University Press, 2013).

72 Robert Muggah, *Counting Conflict Deaths: Options for SDG 16.1* (2015).

that robust population-based surveys are seen by many researchers as the best route for determining the scale and distribution of conflict deaths, greater investment in surveys may be necessary to fill data gaps.[73] Using representative surveys, however, to supplement other data sources raises the prospect for double-counting where there is overlap (see Stage 9 in Coding Violent Deaths, below).

STAGES IN CODING VIOLENT DEATHS

Coding violent deaths so they can be counted and aggregated, from any of the channels described above, poses significant data processing challenges. In addition to multiple unreported deaths, the sheer number of reported deaths can be overwhelming to a coding process. The above channels—particularly media reporting—contain far more mentions of violent death than could be accurately monitored at global scales by human readers alone. This implies a need for algorithmic methods. A pilot study by the Cline Center for Advanced Social Research assessed off-the-shelf technologies and potential software development strategies. They found that existing lexico-syntactic—that is, grammar and dictionary-based—software is poorly suited to casualty detection, even when modified to incorporate quantity detection methods. It found that the most promising, feasible approach lies in enhancing the scalability and accuracy of semantic role-labeling platforms that detect casualty-related quantities.[74]

Even with perfect human and machine coding, any text-based approach, however, involves several stages of inference that affect the potential for all purely algorithmic solutions. In principle, one might count casualties with perfect accuracy by text-mining media sources given absolute confidence in each of the links in the chain that connects reality to reports, and then to algorithmic output. This requires making heroic assumptions about both collecting and processing the content. First, it means assuming that at least one outlet reports every relevant event. Second, every relevant outlet is believed to be monitored, and all relevant documents are collected. Third, relevant documents must be passed into a process that identifies events without error. Finally, the attributes of these extracted events must be correctly characterized and disambiguated. Errors in the content collection process mean that irrelevant material may be processed, generating false events. Relevant reports may also be missed, causing an undercount of violent deaths. In the event extraction process, numerical attributes like casualty counts or geospatial coordinates may be incorrectly assigned. Errors cascade through the channel, so that earlier errors have downstream effects for all latter stages. For example, overreporting high-profile events makes textual de-duplication harder, and dramatically increases the number of documents processed by event

73 Geneva Declaration Secretariat, *Global Burden of Armed Violence* (Small Arms Survey, 2008).

74 This section summarizes findings and analysis from the Cline Center, as reported by Scott Althaus, Dan Shalmon, Buddy Peyton, and Loretta Auvil, complemented by findings from the GReVD workshops.

For conclusions, see Scott L. Althaus and Dan A. Shalmon, *Promises and Pitfalls in Using Computational Strategies for Deriving Accurate and Timely Data on Violent Deaths around the World,* Report, (Champaign, IL: Cline Center for Advanced Social Research, 2018).

For detailed analysis and technical assessment, see Loretta Auvil, Buddy Peyton, and Dan A. Shalmon, *Technical Report for Promises and Pitfalls in Using Computational Strategies for Deriving Accurate and Timely Data on Violent Deaths around the World* (Tech.). (Champaign, IL: Cline Center for Advanced Social Research, 2018).

To request a copy, please contact shalmon2@illinois.edu.

and attribute extraction algorithms. Each such coding creates an opportunity for error. This, in turn, increases difficulty in the final stage: disambiguating extracted events and casualty counts. The cumulative effect of these errors can create significantly biased estimates of violent death rates.

The complexity set out above reflects the many challenges of moving the event through a coding process. This is further complicated by trying to integrate findings across the five data coding channels. To better understand this complexity and organize the gaps in our knowledge, the consortium identified nine stages of coding, relevant for human and machine coding.

NINE STAGES IN CODING VIOLENT DEATHS

STAGE 1: *A violent death is either **observed**, or it is not.* At the very least, for a violent death to be recorded, it must first be observed by a party able and willing to disseminate the information. Various factors contribute to the risk that a death is unobserved, including cover up by perpetrators, limitations on press freedom or NGO access, or simply that there are no witnesses.

STAGE 2: *A lethal violent event is either **reported** through some outlet somewhere in the world, or it is not.* Because of biases inherent in news reporting, some violent events will be overreported, whereas others go unreported. Similarly, due to insufficient technical capacity and lack of political will, not all deaths will be recorded by specialized state institutions or NGOs.

Errors in these first and second stages are mostly beyond the control of researchers. However, it may be possible to mitigate their effects by estimating the propensity for selection bias or failure to retrieve relevant data in certain countries.

STAGE 3: *Reports describing events are recorded into a research database (or **platform**), or they are not.* Extreme-scale monitoring and data collection—in particular, monitoring the global information environment—is technically and logistically challenging.

STAGE 4: *Within the research database, **duplicate** copies of roughly the same report are detected and removed, or they are not.* This problem often arises, for example, when local news outlets publish articles produced by syndicated transnational newswires. Depending on the newsworthiness of the event, syndication can generate multiple, nearly identical stories. This challenge is aggravated by the propensity of local outlets to make minor modifications to newswire content. Moreover, both locally produced and syndicated stories are re-published and updated as reporters acquire more information. A similar process applies to social media and specialized reporting. NGOs and citizens tend to produce multiple reports on high-profile events, echoing trusted sources, and update their descriptions over time. The implication is that any algorithmic solution must be able to screen out nearly identical content.

STAGE 5: *Sentences containing **events** in each unique news report are correctly identified, and relevant events are extracted, or they are not.* Once duplicates have been removed, the next stage involves finding the subset of text that contains relevant events and converting unstructured natural language into structured event-level records. Document classification and event span identification are challenging tasks and produce both false positives and false negatives. Errors in this stage include document classification and event identification failures.

Documents can be identified, for example, using a combination of keyword or extract textual feature queries and machine learning-based classification. Once documents are selected, events must be extracted. The Cline Center pilot study found that off-the-shelf event data platforms fail to identify many relevant events. One reason for this failure is that they read sentence-by-sentence and typically assume that one pair of interacting agents in one sentence is equivalent to one event. Given that design, a sentence like “Boko Haram killed 100 people in three bombings last week” will generate a single event. Omissions are particularly likely when events are described in multiple text spans (sentences). Unfortunately, cross-sentence distribution of event information is very common—after naming an agent once, we tend to refer to them with a pronoun like “he” “she,” or “they,” for example. These systems may also make categorization errors, like classifying violent riots as peaceful protests, or vice versa.

STAGE 6: *Salient event **attributes** contained in event-laden sentences—such as time, place, casualty counts, etc.—are either extracted accurately, extracted with error, or not extracted at all.* This stage includes categorizing the event in greater detail, extracting casualty counts, and identifying the time and location of events so they can be linked to a reporting period and a country. This is a complex task, because natural language is often vague, and text must be transformed into numbers and coordinates in a spreadsheet.

Accurately extracting event attributes requires converting ambiguous descriptions into clear, structured data. Expressions like “yesterday” and “the same town” should be transformed into specific dates and locations, and phrases like “up to a dozen killed or injured” into the appropriate number of entries (12). Current-generation event extraction systems often fail to correctly identify locations and dates.[75] Currently, only semantic systems—new technologies that have not yet been deployed in a conflict monitoring project—are designed to identify casualties.

STAGE 7: *The salient attributes within a single report are either correctly disambiguated, **synthesized**, and integrated into complete event records, or they are not.* A document may contain multiple sentences describing an event, but no single sentence describes all its attributes simultaneously, explicitly, and without ambiguity or contradiction. To generate accurate event records, the spans of text that contain attributes must be found and stitched together. This requires resolving complex patterns of co-reference, extracting a semantically accurate count of distinct events, and resolving discrepancies in the reported attributes. For example, there may be conflicting casualty counts from multiple sources, or updated location information inconsistent with initial reports. In documents describing multiple events, a sentence may refer to a single event, multiple distinct events, or may aggregate events

75 Scott Althaus, Buddy Peyton, and Dan A. Shalmon, *Spatial and Temporal Dynamics of Boko Haram Activity across Six Event Generation Channels: Testing a New Approach to Event Data Validation*. Presented at the APSA Annual Meeting & Exhibition (Boston, MA: Cline Center for Advanced Social Research, 2018); and S. J., Lee, H. Liu, and M. D. Ward, “Lost in Space: Geolocation in Event Data,” *Political Science Research and Methods* (2018): 1–18.

to describe their cumulative effect or a trend. Assembling this information into accurate, unique event records is time-consuming for trained human analysts, and exceedingly difficult for machines.

STAGE 8: *Events extracted from different sources are integrated and* ***disambiguated*** *so that each real-world incident is represented by one and only one event record, or they are not.* This challenge is slightly different, because it arises after events have been extracted from documents. It therefore depends critically on the accuracy of all prior steps. It involves cross-document disambiguation and integration. That is, it requires reasoning across documents to merge redundant records while preserving distinct ones and resolving discrepant codings of the same real-world occurrence. Consider a single highly visible atrocity reported differently in multiple documents. They may describe a wide range of casualty counts. A well-documented example of this is the 2015 Baga Massacre, which is described in the Cline Center's case study.[76] Ideally, an algorithm would find all the redundant records of the atrocity, and combine them into a single event, converting conflicting counts into a single "best" casualty estimate. This stage would also integrate redundant and incomplete records into a single record that maximizes available event information.[77]

This stage distinguishes "knowledge extraction" from "information extraction." Extracting events from a given textual input is an information extraction task and entails minimal interpretation. One can extract information from many sources with perfect accuracy but extract false knowledge from them. For example, a platform might accurately process each individual report of an atrocity, producing many redundant events, each with an accurately extracted casualty figure. A method that naively converts these data into a fact—a description of multiple, distinct atrocities—would generate flawed knowledge.

It requires considerable effort and skill for a human to analyze multiple documents and event records to find overlapping information and identify the number of truly distinct events. Doing so typically involves implicit or background knowledge. For example, an analyst logically considers proximate or similar-sounding locations, dates, or actors. It is worth pointing out, however, that although human analysis tends to be considered the "gold standard," for most text annotation tasks, there is an upper limit on inter-annotator agreement at around 85–90 percent. Simplified, it is incredibly difficult to get multiple people to interpret the same document identically.

STAGE 9: *Given a singular event record for each real-world incident in at least one database, multiple databases are either correctly* ***integrated*** *and disambiguated, or they are not.* Even if a platform codes Stages 1–8 perfectly, there will still be inconsistencies across databases. Multiple monitoring projects exist, which may record the same event inconsistently because they rely on distinct sources and coding protocols.[78] For example, depending on sources and coding rules, a single real world event could be recorded as occurring on different days,

---

76 Kelly M. Greenhill, "Nigeria's Countless Casualties—The Politics of Counting Boko Haram's Victims," *Foreign Affairs (Snapshot)* (February 9, 2015), https://www.foreignaffairs.com/articles/africa/2015-02-09/nigerias-countless-casualties.

77 For example, two documents might refer to the same small-scale attack. One might describe casualties and location and omit information about the perpetrator. A second might describe the perpetrator but omit the casualty count. Merging these two redundant and incomplete records is the optimal result.

78 Karsten Donnay et al., "Integrating Conflict Event Data," *Journal of Conflict Resolution* 63, no. 4 (2018).

in different locations, and/or with a different number of casualties, and/or inconsistently categorized.[79] The difficulties here are similar to the within-database integration challenge in Stage 8. Reasoning across databases to create unique, accurate records requires examining multiple attributes; potentially including the type, scale, and outcomes of the events and the actors involved.[80]

Errors in the third through to ninth stages may be detected and remedied by researchers. However, errors in these stages represent significant engineering challenges at the leading edge of data science innovation. By creating a common research infrastructure and methodology drawing on the knowledge of leading experts in the field, GReVD offers a unique opportunity to resolve these technical challenges.

USING THE 5X9 FRAMEWORK TO ASSESS GAPS

Databases like the UCDP, ACLED, and GTD pick up signals of violent deaths occurring through different channels (primarily media sources), which then progress through the nine stages of inference (see Five Channels for Capturing Violent Deaths and Stages in Coding Violent Deaths, above). Depending on the channel, errors can occur in multiple ways, resulting in both undercounting or overcounting, as well as mixed and unknown effects. In this section, we walk through the coding process for every stage of each channel to demonstrate how errors might occur.

Administrative coding of violent deaths: Low statistical capacity and the aggregation of data are main reasons for why a violent death is not observed through this channel. Statistical capacities of public institutions to record violent deaths correlate with overall levels of economic and institutional development, and thus vary tremendously across countries. Conflict-affected countries are particularly affected, as monitoring and reporting structures often break down and public health and criminal justice systems are weakened. Often details on events are lost in the administrative processes that eventually release aggregated data.[81] Even when the state has the capacity to produce high-quality mortality statistics, lack of political will can result in deaths not being reported. States, particularly if they are involved in hostilities, may find it politically sensitive to publish or even collect data on conflict-related deaths (see Box 2: Political Challenges of Counting Violent Deaths).[82] Many times violent deaths go unreported due to fear of reporting. Armed groups may use terror to intimidate communities and assert control, beyond that of the authorities. In some cases, armed groups dismember or mutilate victims both to avoid detection and intimidate communities. Efforts to avoid detection can lead to undercounting of violent deaths.

Even in less contentious contexts there may be political motivations behind recording violent deaths or not. Law enforcement may have a vested interest in identifying challenges to

79 For example, a single real-world massacre may be described as "terrorism" in one dataset, "one-sided violence" in another, and a "politically motivated attack" in a third.

80 Donnay et al., "Integrating Conflict Event Data."

81 See Channel 1 Administrative Coding in part 3. One way to address the issue of aggregated administrative data in the future is to code at the individual record level to ensure consistency at the registry level; this may be part of capacity building in the future (see Conclusions).

82 Salama, *Counting Casualties*; and Mc Evoy and Hideg, *Global Violent Deaths 2017*.

public order.[83] In the United States, for example, law enforcement agencies are not required to report on police killings. In fact, they have an interest in treating such incidents as justifiable homicides, which may or may not be included in intentional homicide statistics. The voluntary nature of the reporting system means that significant numbers of killings by police are not included in the official numbers.[84] Moreover, the disparities in the definition of violent deaths are likely to result in mixed counting effects, with some deaths in some contexts being overcounted, whereas in others, such as in the example above, undercounted.

Administrative coding of violent deaths is less affected in the remaining inference stages. However, in Stage 8, Disambiguation across Multiple Sources, there is the potential risk that records across agencies are not disambiguated, leading to overcounting of events. For example, national forensic institutes and the police should produce matching reports on homicides. If these are not disambiguated, the same deaths would be double-counted. Nevertheless, important issues around data completeness and generation processes warrant caution. For example, in Colombia, although the police and the national forensic institute have different procedures and are involved at different stages of the investigative process, they should in theory report identical numbers, as both institutions play a central role in every homicide investigation. However, significant differences in the numbers reported by the two institutions have been found.[85]

Media reporting of violent deaths: Limits on press freedom and other access restrictions are critical causes of unobserved violent deaths in media-based channels (see Figure 2). The biases inherent in news reporting tend to produce both overcounts and undercounts (Stage 2)—certain types of events are reported repeatedly, whereas others are ignored. Local media outlets are more likely to cover deaths that international media ignore, which could remedy the undercount of deaths occurring in, for example, rural and isolated areas. However, selection bias exists also in local media, which renders certain types of violence, by certain actors or certain locations, effectively invisible. Moreover, local reports are less likely to be retrieved into a research database due to technical and language barriers (Stage 3).

To support data generation efforts like GReVD, text must be collected, cleaned, indexed, and classified to identify documents that contain lethal events (see Box 3: Example of the Challenges to Machine Coding). Performing each of these tasks in multiple languages is a significant challenge. Most Natural Language Processing (NLP) algorithms, which are used in automated source monitoring and document classification processes, tend to be language-specific. These barriers arise at both the level of language and dialect—even in English-only efforts, local dialects can affect the output of NLP tools. Monitoring the global news environment is further complicated by sources that lack websites, impose paywalls, or create other barriers to monitoring.

83 Davenport and Ball, "Views to a Kill."
84 Mc Evoy and Hideg, *Global Violent Deaths 2017*.
85 Jule Krüger et al., "It Doesn't Add Up," in *Counting Civilian Casualties: An Introduction to Recording and Estimating Nonmilitary Deaths in Conflict*, ed. Taylor B. Seybolt, Jay D. Aronson, and Baruch Fischhoff (New York: Oxford University Press, 2013).

Even accessible sites may suffer from poor design, which makes extracting high-quality text difficult. A text processing platform must address obstacles like errant html scripts, banners, ads, comments sections, and captions, all of which can make it difficult to effectively deploy NLP algorithms. That is, textual noise makes Stages 4–7 more challenging, because each of the processes involved requires close reading and detailed analysis of linguistic features across sentences and documents.

BOX 3: EXAMPLE OF THE CHALLENGES OF MACHINE CODING

Two stages are particularly challenging for machine event coding: event identification and attribution extraction (Stages 5 and 6, respectively). The complexity of the task is illustrated in this excerpt of a Reuters article processed in the Cline Center pilot study:

> Suicide bombers struck two bus stations in different parts of northern Nigeria on Tuesday, killing at least 26 people...In the first, a suicide bomber rushed onto a bus in the northeastern town of Potiskum before setting off a blast that destroyed the bus and killed 16 people... On Sunday, a girl with explosives strapped to her killed five people outside a market in the same town... In Tuesday's second attack, two suicide bombers in a car struck a major bus station in the north's main city of Kano, killing at least 10 people.[86]

Initially, effective event extraction requires "span identification." That is, the pieces of text that contain the event-related information must be found. The excerpt above includes numerous ellipses, which exclude text unrelated to the lethal attacks. A machine must perform a similar form of summarization—for example, excluding sentences (as has been done here) like "Nigeria's neighbours are being targeted also as they join in the battle," because it does not contain a lethal event, even though "battle" and "targeted" are words often associated with potentially lethal violence.

Event-span identification is challenging. As shown in the example above, news reports often include multiple events in a complex linguistic structure of references spread across multiple sentences and paragraphs. When this happens, it is easy for algorithms to mix up attributes of different events in ways that create "false positive" events composed of the details of the same underlying incident. They may also produce "false negatives" when details from different events are mistakenly merged into a single composite incident span.

In the Cline Center pilot study, one of the systems used only detected one attack in this document. Because there were at least three bombings, this represents two omitted events (false negatives). Another system, by contrast, identified more events than three, producing multiple false positives.

Once event-related text has been found, algorithms (or humans) may characterize it incorrectly. For example, a potentially lethal attack may be described as a nonviolent protest, excluding it from the database. Off-the-shelf systems often rely on dictionaries of potential perpetrators and targets and do not code events with unknown agents. If their dictionaries are not periodically updated, or adapted to new regional domains, they will fail to detect events.[87]

Event span detection is only the first step—both the existence of an event and its salient attributes must be extracted (Stages 5 and 6, respectively). Attributes typically include time, place, casualty counts, etc. Unfortunately, current-generation systems often fail to correctly identify locations and dates, which are necessary for linking extracted events to a reporting period and country.[88] This is difficult for machines, in part, because humans rely on logical reasoning and

86 Joe Hemba and Nnekule Ikemfuna, *Suicide Bombers Kill at Least 26 across North Nigeria*, Reuters (February 25, 2015), https://www.reuters.com/article/us-nigeria-violence/suicide-bombers-kill-at-least-26-across-north-nigeria-idUSKBN0LT0ZM20150225.

87 This is a time-consuming and challenging task. This imposes some challenges on any effort to scale dictionary-based systems to match the need for timely GReVD data.

88 A previous Cline Center study found that location data were missing entirely in more than 35 percent of machine-generated potentially lethal conflict events. In a subset of machine-generated events examined by trained analysts, between 35 and 70 percent of country-level locations were missing or incorrectly identified. See Althaus, Peyton, and Shalmon, *Spatial and Temporal Dynamics of Boko Haram Activity*.

In a similar test, Lee et. al. found accuracy rates between 31 and 67 percent. S. J. Lee, H. Liu, M. D. Ward, "Lost in Space."

background information to identify these attributes. A machine must be able to replicate these capacities to do the job correctly. None of the computational methods currently used by conflict monitoring projects can extract casualty-related attributes. Even when existing tools are combined with a quantity detection algorithm, they performed very poorly in the Cline Center study.[89]

Consider the Reuters excerpt quoted above. To accurately extract dates, locations, and the number of distinct events, a reader must correctly distinguish related but distinct events across multiple sentences, then reason using prior knowledge to compare temporal expressions to a publication date.[90] For example, a human knows that "the same town" refers to "Potiskum," mentioned earlier in the sentence. They also know that "Sunday" likely refers to the Sunday immediately preceding the Tuesday at the beginning of the document. A dictionary-based algorithm that reads sentence-by-sentence struggles with this sort of implicit temporal and spatial information. Newer systems—like the semantic platform evaluated by the Cline Center team—use machine learning to replicate the logical inferences made by human readers.

Detecting and extracting casualty counts, which is central to GReVD, entails four steps, none of which is simple.

- First, at least one death must be identified, that is, a quantity must be detected in natural language. This quantity may be implicit, for example, a suicide bomber's death is usually not counted in the casualty number reported in the media, but they are killed in the attack, nonetheless. The quantity may also be vague, such as "at least a dozen" or "hundreds."
- Second, the algorithm must determine that the quantity terms indicate casualties. In the excerpt above, one algorithm used in the Cline Center study detected six quantities. They are: "two," "16," "at least 26," "five," "two," and "at least 10." Only three of these are actually casualty figures—the others refer to the number of targets and bombers, respectively.
- Third, particular events must be correctly linked with specific casualty figures. Here, the "16" casualty count should be linked to a suicide attack in Potiskum. Ideally, the system would know—as a human does—that when 16 targeted people are killed by a successful suicide attack, the total casualty count is actually 17.
- Fourth, casualty-counting language must be converted to numbers in a spreadsheet, which becomes particularly challenging when numbers are implicit or vague.

Finally, the event attributes within a single report must be correctly disambiguated and integrated into complete, accurate event records (Stage 7).

In this case, three of the casualty counts—"at least 26," "16," and "at least 10" are actually nested—that is, 26 is the sum of "10" and "at least 16." An ideal algorithm must be able to determine that the document contains two casualty-generating events, and that a sentence that contains a relevant casualty count—"Suicide bombers struck two bus stations...killing at least 26 people"—does not actually describe any one really-existing event. This stage would also involve resolving other ambiguous co-references—called anaphora—for example, recognizing that the attack on the market "in the same town" occurred in "Potiskum."[91]

Once events have been extracted from documents and processed through nine stages detailed above, records extracted from different news reports must be integrated and disambiguated. Each real-world incident should be represented by only one event record (Stage 8). This challenge is slightly different from the others, because it occurs after events have been extracted from natural language. It depends critically on the accuracy of prior steps. The same kinds of ambiguities encountered within a single document also occur across reports. For example, articles describing the same attacks as the Reuters excerpt (above) may use different casualty numbers, describe geographical locations differently, and provide varying levels of detail. Effectively disambiguating and converting extracted events into unique "facts" involves cross-document inference, which is a difficult problem in computational linguistics.

89 Between 75 and 88 percent of the extracted quantities could not be matched to a casualty count. Casualties extracted using emerging semantic methods were associated with a true casualty figure nearly 60 percent of the time, but they still failed to detect most casualty counts. See Auvil, Peyton, and Shalmon, *Technical Report for Promises and Pitfalls,* 33–34

90 This is a difficult task in computational linguistics. See Q. Ning, B. Zhou, Z, Feng, H. Peng, and D. Roth, *Cogcomptime: A Tool for Understanding Time in Natural Language,* Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing: System Demonstrations (November 2018): 72–77).

91 For computational work on reasoning about quantities and anaphoric references, respectively, see S., Roy, T., Vieira, and D. Roth, "Reasoning about Quantities in Natural Language," *Transactions of the Association for Computational Linguistics* 3 (2015): 1–13; and K. Clark and C. D. Manning, *Improving Coreference Resolution by Learning Entity-Level Distributed Representations*, in Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Vol. 1: Long Papers, 2016): 643–653.

Due to mass syndication of wire service reports, duplication issues are of concern for media-based platforms. Near-duplicate detection requires comparison of all potential duplicate-pairs (Stage 4). This is a computationally demanding task because these stories often have different headlines and varying amounts of original content. Failing to detect duplicate reports tends to increase the number of redundant event records, and overcounts of violent deaths for a given place and time. Once duplicate documents are removed, sentences containing event information in the news report must be correctly identified, extracted, and used to code events.

Direct reporting of violent deaths (crowdsourcing): Direct reporting (crowdsourcing and social media) are relatively new technologies and thus the specific errors pertaining to this channel as reports progress through the various stages are hard to accurately ascertain. Although social media has an important role in in filling gaps that administrative data and news reports leave open, these data have several limitations relative to news media data for use in estimating violent death totals. The problem of identifying duplicate reports of the same event is exponentially more complicated in social media data than in news media data because there is so much more social media data to sift through, and so many people are posting about the same events. However, the limitations of social media data go well beyond this problem. Language used in social media is more sparse, diverse, and location-specific than in news media, which greatly complicates the use of algorithmic solutions. The diversity and location-specificity of social media language is also less stable over time than for news media language, so that any algorithmic solutions would need to be continually re-engineered to keep up with rapidly evolving patterns of slang and context-dependent semantic references. Social media data are also beset with emerging privacy issues and changing terms of data access, which can affect validity and sustainability of coding platforms built on social media. Finally, in many countries ravaged by political violence, social media communications are used to spread governmental disinformation about politically relevant events with bots and propaganda that may be difficult to detect algorithmically from authentic reports of lethal violence. Temporary shutdown of social media services can also affect coding.

Specialized and expert reporting of violent deaths: As with administrative coding, errors in this channel mainly concentrate in the first three stages. In terms of a violent death being observed or not, the main obstacle with specialized and expert reporting relates to the restricted operational space of NGOs carrying out this kind of activity. On the other hand, where these agencies do report on violent deaths, there is a risk of potential advocacy bias.

Representative survey findings on violent deaths: Representative surveys will always be limited in scope and are, as previously discussed, extremely difficult to conduct in areas of high levels of violence, such as conflict affected countries. Surveys are expensive and cannot identify specific events, but they can be useful with sufficient statistical capacity to identify violent deaths underrepresented or unidentified in other sources. The UNODC Global Study on Homicide draws from a representative survey it conducted together with the National Bureau of Statistics of Nigeria, which included a module on homicide. The results of the survey provided strong evidence that the level of lethal violence in Nigeria is likely to be higher than commonly assumed.[92] Nevertheless, while surveys are useful for pointing researchers to where violent deaths might be occurring, they do not account for the specific coding of actual deaths.

INTEGRATION ACROSS ALL FIVE CHANNELS

Finally, integration across all channels is affected by the lack of standardized reporting, events not being geocoded or with low geo-precision, and the lack of time coding.[93] As previously described, the same event may be reported across channels and datasets inconsistently because they rely on distinctive sources and coding protocol.[94] When integrating across multiple channels/platforms, complications can arise when an event can be coded as occurring on different days, in different locations, with a different number of casualties, and being labeled as a different type of conflict activity. Integrating multiple data sets requires making decisions about whether differences mean that entries capture unique events or instead reflect uncertainty or variation in measurement.[95]

When aggregating and reconciling overlaps in conflict datasets, and integrating with homicide data, it is important to consider both duplication and disparities based on differences in coding procedures. Disparities are often based on differences in coding rules. For example, ACLED has an “atomic event” that occurs on a specific day, location, in a categorizable way, and between one or more agents. As detailed above, there are many cases where multiple similar events occur in the same space and within the same day and coded in multiple datasets. Therefore, not all cases of multiple events date- and location-stamped are necessarily duplicates. In another case, one dataset may define civilian targeting in only one way, or only if attacks are perpetrated by specific actors, whereas another may define civilian targeting in multiple ways (e.g., as violence against civilians but also specific “remote violence” events in which civilians are the target). There are also differences in conceptual definitions, such as what constitutes a conflict event for the purpose of respective datasets, in geographical coverage and in access to different source material (access to different corpora).

92 UNODC, Global Study on Homicide 2019 (Vienna, 2019).

93 For a discussion of discrepancies between reporting sources, see Karsten Donnay and Vladimir Filimonov, “Views to a War: Systematic Differences in Media and Military Reporting of the War in Iraq.” *EPJ Data Science* 3, no. 1 (2014): 25.

94 Kristine Eck, “In Data We Trust? A Comparison of UCDP, GED and ACLED Conflict Events Datasets,” *Cooperation and Conflict* 47, no. 1 (2012); and Karsten Donnay and Ravi Bhavnani, “The Cutting Edge of Research on Peace and Conflict,” in *Peace and Conflict 2016* (Routledge, 2016): 20–34, https://www.researchgate.net/publication/310951486_The_Cutting_Edge_of_Research_on_Peace_and_Conflict for further discussion on data discrepancies.

95 Donnay et al., “Integrating Conflict Event Data.”

PART 4 Ways Forward for GReVD

This report has summarized multiple gaps in current methodology for counting violent deaths. The goal of the GReVD consortium is to remedy these gaps, to eventually provide a registry with a single entry for every violent death, identified by location and time, and where possible, by type of violence and characteristics of actors, eventually including victims and perpetrators (where possible). It will do this by creating a shared research ontology, infrastructure, and methodology, convening leading experts and institutions working on these issues, to develop better estimates over the next decade. Based on this assessment of the gaps in our knowledge and the challenges in monitoring violent deaths, the next objective will be to create a shared, "imperfect but good enough," methodology for solving these problems for the near term, while building better solutions for the long-term. Next steps for GReVD can be broken into three lines of action:

- Improving **coding** for partners, including human in the loop, machine in the loop, and machine reading processes.
- Increased **precision** through shared ontologies and better estimation methods with partners.
- Building and filling the **registry**, including a fully integrated database drawing on consortium partners, better integration with administrative sources, and new source development.

IMPROVING CODING, INCLUDING BOTH MACHINE CODING AND HUMANS IN THE LOOP

A pilot study by the Cline Center for Advanced Social Research assessed off-the-shelf technologies and potential software development strategies. They found that existing lexico-syntactic software—that is, grammar and dictionary-based software—is poorly suited to casualty detection, even when modified to incorporate quantity detection methods. The study found that the most promising, feasible approach lies in enhancing the scalability and accuracy of semantic role-labeling platforms that detect casualty-related quantities. As we described in the framework section above, coding complex events like violent deaths is currently more difficult for machines, because humans rely on logical reasoning and background information to identify salient attributes. Still, it is acknowledged that machines will be necessary to process the immense (and increasing) amount of raw data being made available; investment is needed now in the machine coding systems that will help to code that data.

This is not to say that machine coding doesn't already offer some promising leads. The GReVD workshops identified synergies on how machine coding could improve current human coding processes. GTD already uses algorithms to filter relevant and useful source articles to increase efficiency of human coders. Preliminary results from the Cline Center study suggested that machines could potentially be used to determine which sources report identical "facts," which could be used to simplify human labor and help to improve human coding by prioritizing articles to read first. Workshop participants discussed using document summarization algorithms to improve upstream deduplication efforts. This could substantially reduce redundant downstream human coding. Some initial tests demonstrated efficiency improvements of 10–15 percent. These and other "human in the loop" innovations can improve coding for consortium members.

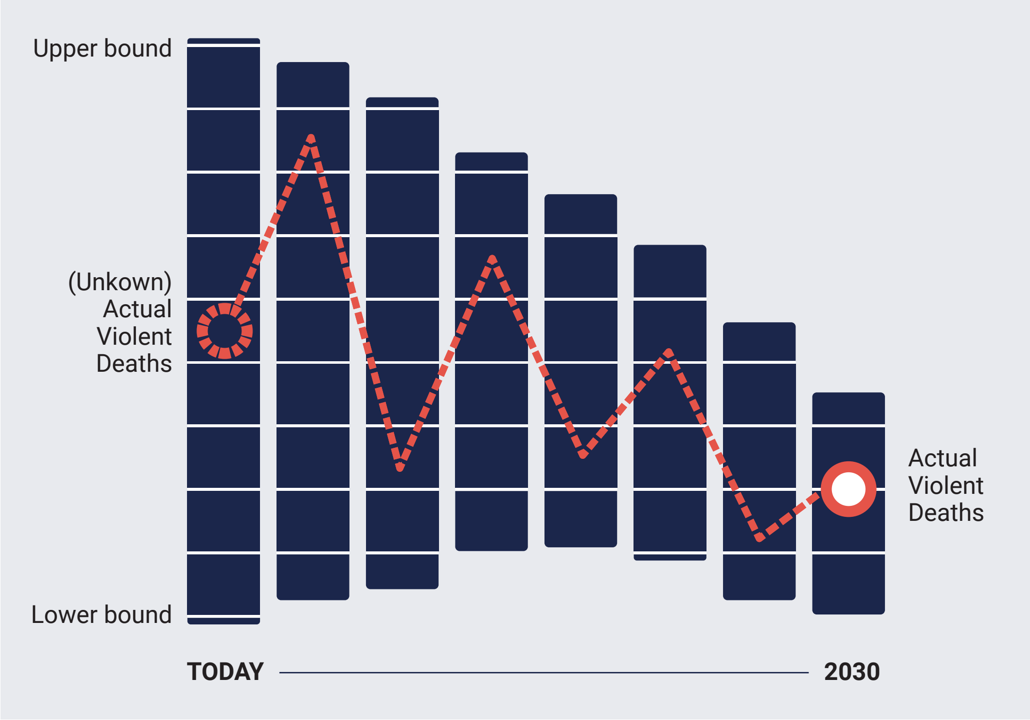
The consortium is piloting a consultative process to improve machines-in-the-loop coding processes for the GTD. If successful and with lessons learned from this pilot, other coding advances can be made available to other GReVD consortium partners and others who work on coding or reporting violent deaths.

INCREASING PRECISION

All current reported numbers of violent deaths are estimates. As discussed in the first two parts of this report, the amount of uncertainty around these estimates can be attributed to multiple reasons—definitional issues, coding challenges, and political will. The GReVD consortium will contribute to increased precision by reducing uncertainty of these estimates. This uncertainty is represented in Figure 3 by the height of the bar between lower bound and upper bound.

FIGURE 3

Reducing Uncertainty over Time and the Lower and Upper Bounds



Source: Authors based on the Geneva Workshop of the GReVD consortium (February 2019)

The upper bound of violent deaths is the difference between all deaths and all non-violent deaths. This is the “counting down” method—subtracting all deaths that are not conceivably violent from the total number of global deaths. Preliminary estimates suggest that this number would be approximately 1.5–2 million violent deaths per annum, depending on which categories of WHO mortality data are included or excluded in the definition of violent death.

Counting up is conservative because it includes all deaths that are known and can be documented. There are many violent deaths (including the categories described above) that go unreported or underreported through media or administrative sources every year. The GVD number of 589,000 violent deaths in 2017 is a lower bound, counting up estimate, and is the best estimate we have. Therefore, the GVD estimate is the working GReVD consortium lower bound estimate.

Ontology and a Common Research Infrastructure

A first step for the GReVD consortium is to build a shared ontology to resolve definitional issues between datasets and to transcend reporting channels to resolve many of the "gap" and "overlap" issues identified in this report. This will be vital if machines are to be increasingly integrated into the coding process (see next section) because off-the-shelf systems often rely on pre-defined dictionaries of potential perpetrators and targets, and cannot code events with unknown agents. If their dictionaries are not periodically updated, or are adapted into new regional domains, they will fail to detect events. This shared ontology can be used by peace researchers beyond those in the consortium to provide a unifying definition of violent deaths through 2030. The shared ontology can also be a foundation for a common research infrastructure because it can be used to identify current definitional issues, which can be assigned to task teams to solve coding issues by violent death type. Finally, the shared ontology may facilitate negotiation of the consortium collectively, for access to larger source corpus that can be used by all consortium partners.

Known and Unknown Unknowns

Violent deaths are defined differently across coding processes, disciplines, and institutions. As a consequence, violent deaths may be observed but not coded through a particular data coding process. Databases that define armed conflict according to specific intensity levels may discard numerous incidents because specific thresholds were not met. Additionally, violent deaths may not be coded due to technical or political constraints. The consortium refers to observations that enter the coding process but are not coded for whatever reason as "known unknowns." In many cases, consortium members take these uncoded observations and place them in a holding file, called "purgatories" (see Purgatories below). One promising research area for future work is mining purgatories from all consortium members to fill gaps in coverage because the GReVD ontology on violent deaths will be larger than individual member definitions.

Unknown numbers of violent deaths are never identified—that is, they never enter the coding stages. This may be due to geographical constraints on recording, fear of reporting crimes in certain contexts, or simply that the death was not observed. By definition, these violent deaths are referred to by the consortium as "unknown unknowns." One way of identifying and remediating these gaps is through statistical modeling and estimating unknown unknowns, building off of approaches like the capture method (see Appendix 1: Capture Methods). Note that multiple system estimation models (MSE) and other statistical estimation methods would not contribute to filling the registry directly, because it is just a tool for estimating unknown population sizes; however, it could be used to identify where further research should be focused to identify likely missing entries for the registry.

Other methods being considered include modeling the most likely places for missing observations, applying survey methods to identify areas with large incidence of unknown unknowns (see UNODC Global Study on Homicides, 2019), and unpacking coding purgatories from consortium members (and other sources that code violent deaths, which may involve expanding the consortium).

Building a single comprehensive, global dataset for violent deaths will require not only integrating existing resources, but also supplementing them with new data (unknowns) to fill the larger and most important gaps in existing monitoring systems.

Purgatories—An Example of Known Unknowns

In addition to issues of underreporting discussed earlier, restrictive coding criteria exclude a proportion of violent deaths observed. This is largely the result of insufficient information available about the victim and perpetrator(s) around the circumstances of the death. If a violent death cannot be sufficiently attributed, according to the specific coding criteria of a database, it is most likely excluded—hence, not counted. In such cases, it enters into a holding file the consortium calls "purgatory." These data purgatories, with observed violent deaths that cannot be coded, are a very real example of known unknowns and a first order priority for GReVD in reducing uncertainty.

Databases have their own specific coding criteria and deal with these so-called purgatories differently. For the UCDP, which collects data on organized violence pertaining to state-based armed violence, non-state armed conflict, and one-sided violence, the fluidity of armed groups may generate purgatories. When there are multiple organizations in an area, or where it is known that organizational affiliation is fluid and that groups often split up and merge, it is hard to determine with certainty that deaths are an outcome of organized violence. Examples are violence in El Salvador, Honduras, Guatemala, Brazil, and Mexico.

Similarly, in order for an event to be recorded in the GTD it must be documented by at least one primary (as opposed to secondary) source that is independent (generally free of influence from the government, political perpetrators, or corporations), and that routinely reports externally verifiable content. "Events that are only documented by distinctly biased or unreliable sources are not included in the GTD. Note that particular scarcity of high-quality sources in certain geographic areas results in conservative documentation of attacks in those areas in the GTD."[96] This means that certain deaths that take place may not be counted, either because there is an intractable degree of ambiguity about the event or because limited resources for data collection mean that analysts are not able to conduct additional research to make sure that every ambiguous case has been exhaustively sourced.

One way forward is for GReVD consortium members to use each other's purgatories, using the GReVD common research infrastructure to facilitate cooperation. Another possibility is to use machine-coding to mine existing purgatories and program machines to identify uncoded events that have been overlooked in other coding processes.

BUILDING AND FILLING THE REGISTRY

In addition to improving coding methods, improving precision, and reducing uncertainty around defining violent deaths, the consortium will build the actual registry database. This

96 START, "Codebook."

is no simple matter because it requires a front-end interface that allows users to access geocoded data, and researchers to access the dataset, as well as a backend that connects to original source material, including the original databases of the GReVD consortium members (ACLED, GTD, and UCDP).

This line of action begins with proper specification and hosting requirements for the registry for a minimum of 15 years, based on the Washington and Geneva workshops, current concepts, and coding challenges. It allows for filtering using a consortium-defined ontology and allows users to identify filter and sorting criteria. The main data-producing members of the consortium will be consulted on the design process of the registry, and once online, the registry will ensure their contributions are acknowledged through consortium endorsed standards and principles.

Conclusion

Developing improved ontologies and common research infrastructure will be iterative, requiring consortium members to improve on the methodology through joint research over the coming years. Furthermore, machines increasingly will be integrated into coding processes, both to improve human coding and to eventually bridge coding processes with human-in-the-loop systems. Opportunities exist to mine purgatories and access increasingly larger source corpus.

As a next step, the consortium will build on current approaches to continue to refine the upper and lower bound of estimates of violent deaths. The upper bound—based on a subtractive method that takes all deaths and removes those that are indisputably *not* violent deaths—will be refined with guidance from those who regularly work with health statistics. The lower bound estimate will use an additive approach —built upon the current GVD methodology[97]—adding together what is known from current datasets (resolving duplicate coding) and increasing coverage of coding sources where possible. Owing to the discrepancies in definitions and quality of data, the range for the upper bound and lower bound estimates, drawing on the subtractive and additive approaches, respectively, will be wide at first. Future improvements on the methodology, standard-setting, and in reporting and monitoring will contribute to decreasing the range of these bounds, converging upon the "real" number of violent deaths as the methodology for the registry improves.

As the GReVD methodologies and data improve, estimates will be updated annually, including retroactively so that over time, comparators that are consistent with the current methodology and data are available to users of GReVD. A common research infrastructure will be built to overcome the challenges described in this report through task team approaches to identifying problems and solving these challenges.

97 2019 edition with 2017 estimates, see Hideg and Alvazzi del Frate, *Darkening Horizons*.

APPENDIX 1 Capture Methods

Statistical models have been developed that can be used to estimate the size of an unknown population, thereby helping to address issues related to known unknowns and unknown unknown violent deaths. Dual and multiple system estimation (MSE) models have been used to this end because they have been developed for estimating a population of unknown size in situations where probability sampling is insufficient or infeasible. These methods have been used for estimating casualties in armed conflicts, specifically in support of historical verification and truth commission processes, criminal justice tribunals, and national human rights campaigns by NGOs. MSE methods have been used to reveal invisible dimensions of lethal violence, allowing researchers not only to call into question established but unverified figures, but also to provide a new and reliable estimates of total deaths.[98] More recently, MSE have also been used for studying documentation dynamics—specifically the extent to which human rights observers have been able to document killings in the armed conflict in Syria to help interpret the relationship between documentation and conflict dynamics.[99]

MSE methods require that multiple (at least two) intersecting but incomplete lists exist that partially capture the population of interest. These lists are then used to detect inclusion or capture patterns in order to quantify the probability that a death will be missed. In short, MSE models attempt to estimate the number of cases that were not included in lists that partially enumerate a closed population. Nevertheless, MSE comes with multiple strict assumptions, some which pose problems for estimating unknown numbers in conflict and post-conflict contexts.

The first, and also least problematic, assumption is that all samples must refer to the same closed system of observation. Because estimating conflict-related deaths is largely done retrospectively, the population cannot migrate in or out of the area and time period for which one seeks to estimate all deaths that occurred.

The second assumption demands that the observations reported in more than one source must be perfectly matched (i.e., that there are no duplications). However, many records of war victims—when they exist—frequently contain errors, which impede reliable and full matching of the sources and in turn risks overestimating the undercount.[100]

The third assumption requires that every observation must have the same probability of being recorded as any other, and that this be true for all lists. This assumption is mostly likely to be violated with conflict and post-conflict data because lists of conflict-related deaths are

98 This includes Kosovo: Patrick Ball, Wendy Betts, Fritz Scheuren, Jana Dudukovich, and Jana Asher, "Killings and Refugee Flow in Kosovo March–June 1999," *Washington, DC: A Report to the International Criminal Tribunal for the Former Yugoslavia* (2002); Bosnia: Patrick Ball, E. Tabeau, and P. Verwimp (2007); Guatemala: Davenport and Ball, "Views to a Kill"; Peru: Patrick Ball et. al. (2003), and Colombia: Tamy Guberek, D. Guzman, M. Price, K. Lum, and P. Ball (2010).

99 Megan Price, Anita Gohdes, and Patrick Ball, "Documents of War: Understanding the Syrian Conflict," *Significance* 12, no. 2 (2015).

100 Jan Zwierzchowski and Ewa Tabeau, *The 1992–95 War in Bosnia and Herzegovina: Census-Based Multiple System Estimation of Casualties' Undercount*, paper presented at the the Global Cost of Conflict (Berlin, February 1–2, 2010), https://pdfs.semanticscholar.org/5b8c/480e42c4c50976fdd51ec90ae4becb5a0060.pdf?_ga=2.95824348.1388251583.1555396473-596354133.1555396473.

rarely random samples of the true number of deaths.[101] It is highly plausible that that undiscovered deaths differ fundamentally from discovered ones, in this way generating capture heterogeneity. As seen, deaths are more likely to be captured in urban areas, in areas easily accessible, and in events that involve more victims. There are, however, ways of controlling for this by stratifying by space and time, effectively seeking out within-list capture probabilities that are heterogeneous at the aggregate level but can still be usefully treated as homogenous within strata.

The fourth assumption holds that the sources documenting the observations must be independent in their recording efforts. This assumption is also frequently violated in the case of systems capturing conflict-related deaths. The various actors who are producing lists of deaths (e.g., human rights activists, the police, the military,) tend to operate either simultaneously or successively. They may draw from the same sources, collect data from separate but overlapping populations, or even draw information from each other (e.g., referring cases to, copying from, exchanging with or consulting the other). List dependence may also be related to the issue of capture heterogeneity. Because of list dependence, the capture probability of a given death into one system (e.g., a truth commission) is higher if that death was also captured by another system (e.g., NGOs) because the two systems are somehow related.[102]

Having detailed data and multiple lists provides a richer set of information from which to draw inferences, as every additional list greatly increases the number of capture patterns and relies on weaker assumptions. However, it comes at the price of high technical complexity. Considerable statistical expertise is needed to understand the assumptions and limitations of the methods and to apply them correctly, which in turn raises sizeable communication challenges, that puts at risk the clear dissemination and discussion of results. Credibility losses associated with communication challenges can reduce the value of capture methods in estimating politically sensitive numbers like fatalities.

101 Jule Krüger and Kristian Lum, *An Exploration of Multiple Systems Estimation for Empirical Research with Conflict-Related Deaths*, Visions in Methodology Conference, University of Kentucky, May (2015).

102 Ibid.

List of Acronyms

ACLED	Armed Conflict Location and Event Data Project
CMR	crude mortality rates
CMU-CREATE Lab	The Community Robotics, Education and Technology Empowerment Lab at Carnegie Mellon University
COW	Correlates of War
DALYs	disability-adjusted life years
GVD	Global Violent Deaths Database
GBD	Global Burden of Disease
GHDx	Global Health Data Exchange
GHO	Global Health Observatory
GReVD	Global Registry of Violent Deaths
GTD	Global Terrorism Database
HALE	healthy life expectancy
HLPF	High-Level Political Forum
IAEG-SDG	Inter-Agency and Expert Group on SDG Indicators
ICCS	International Classification of Crime for Statistical Purposes
IHME	Institute for Health Metrics and Evaluation
MELTT	Matching Event Data by Location, Time and Type
MMR	maternal mortality ratio
MSE	Multiple Systems Estimation
NGO	non-governmental organization
NSA	non-state actor (specifically non-state armed actor in this analysis)
NLP	Natural Language Processing
PRIO	Peace Research Institute Oslo
RMS	Retrospective Mortality Surveys
SAS	Small Arms Survey
SDG	Sustainable Development Goals
SIPRI	Stockholm International Peace Research Institute
UCDP	Uppsala Conflict Data Program

UMD	University of Maryland, College Park
UN	United Nations
UNODC	United Nations Office on Drugs and Crime
WHO	World Health Organization
YLDs	years lived with disability
YLLs	years of life lost

Bibliography

Asher, Jana. “Using Surveys to Estimate Casualties Post-Conflict.” In *Counting Civilian Casualties: An Introduction to Recording and Estimating Nonmilitary Deaths in Conflict.* Edited by Taylor B. Seybolt, Jay D. Aronson, and Baruch Fischhoff. New York: Oxford University Press, 2013.

Chojnacki, Sven, Christian Ickler, Michael Spies, and John Wiesel. “Event Data on Armed Conflict and Security: New Perspectives, Old Challenges, and Some Solutions.” *International Interactions* 38, no. 4 (2012): 382–401.

Davenport, Christian and Patrick Ball. “Views to a Kill: Exploring the Implications of Source Selection in the Case of Guatemalan State Terror, 1977–1995.” *Journal of Conflict Resolution* 46, no. 3 (2002): 427–50.

Donnay, Karsten, Eric T. Dunford, Erin C. McGrath, David Backer, and David E. Cunningham. “Integrating Conflict Event Data.” *Journal of Conflict Resolution* 63, no. 4 (2018).

Eck, Kristine. “In Data We Trust? A Comparison of UCDP, GED and ACLED Conflict Events Datasets.” *Cooperation and Conflict* 47, no. 1 (2012): 124–41.

Gilgen, Elisabeth and Lauren Tracey. *Contributing Evidence to Programming: Armed Violence Monitoring Systems.* Geneva Declaration, 2011.

Global Health Estimates 2016: Disease Burden by Cause, Age, Sex, by Country and by Region, 2000–2016. Geneva: World Health Organization, 2018.

Jensen, Michael. “Discussion Point: The Benefits and Drawbacks of Methodological Advancements in Data Collection and Coding: Insights from the Global Terrorism Database (GTD).” https://www.start.umd.edu/news/discussion-point-benefits-and-drawbacks-methodological-advancements-data-collection-and-coding.

Krause, Keith. “Challenges to Counting and Classifying Victims of Violence in Conflict, Post-Conflict, and Non-Conflict Settings.” In *Counting Civilian Casualties: An Introduction to Recording and Estimating Nonmilitary Deaths in Conflict*, edited by Taylor B. Seybolt, Jay D. Aronson, and Baruch Fischhoff. New York: Oxford University Press, 2013.

———. “Bodies Count: The Politics and Practices of War and Violent Death Data.” *Human Remains and Violence: An Interdisciplinary Journal* 3, no. 1 (2017): 90–115.

Krüger, Jule, Patrick Ball, Megan Price, and Amelia Hoover Green. “It Doesn’t Add Up.” In *Counting Civilian Casualties: An Introduction to Recording and Estimating Nonmilitary Deaths in Conflict*, edited by Taylor B. Seybolt, Jay D. Aronson, and Baruch Fischhoff. New York: Oxford University Press, 2013.

Krüger, Jule and Kristian Lum. *An Exploration of Multiple Systems Estimation for Empirical Research with Conflict-Related Deaths*. Visions in Methodology Conference, University of Kentucky, May 2015.

Kumar, T.S. Mohan, Tanuj Kanchan, K. Yoganarasimha, and G. Pradeep Kumar. “Profile of Unnatural Deaths in Manipal, Southern India 1994–2004.” *Journal of Clinical Forensic Medicine* 13, no. 3 (2006): 117–20.

Landman, Todd and Anita Gohdes. “A Matter of Convenience.” In *Counting Civilian Casualties: An Introduction to Recording and Estimating Nonmilitary Deaths in Conflict*, edited by Taylor B. Seybolt, Jay D. Aronson, and Baruch Fischhoff. New York: Oxford University Press, 2013.

Mc Evoy, Claire and Gregely Hideg. *Global Violent Deaths 2017*. Switzerland: Small Arms Survey, 2017.

Muggah, Robert. "Counting Conflict Deaths: Options for SDG 16.1." https://igarape.org.br/wp-content/uploads/2015/10/counting-conflict-deaths-muggah-2015.pdf.

Öberg, Magnus and Margareta Sollenberg. "Gathering Conflict Information Using News Resources." *Understanding Peace Research: Methods and Challenges*. Abingdon: Routledge (2011): 47–73.

Price, Megan and Patrick Ball. "Big Data, Selection Bias, and the Statistical Patterns of Mortality in Conflict." *SAIS Review of International Affairs* 34, no. 1 (2014): 9–20.

Price, Megan, Anita Gohdes, and Patrick Ball. "Documents of War: Understanding the Syrian Conflict." *Significance* 12, no. 2 (2015): 14–19.

Salama, Hana. *Counting Casualties: Operationalising SDG Indicator 16.1.2 in Libya.* Switzerland: Small Arms Survey, 2018.

Secretariat, Geneva Declaration. *Global Burden of Armed Violence*. Geneva: Small Arms Survey, 2008.

Secretariat, Geneva Declaration. *Global Burden of Armed Violence: Every Body Counts*. Geneva: Small Arms Survey, 2015.

Secretariat, Geneva Declaration. *Global Burden of Armed Violence: Lethal Encounters*. Cambridge, 2011.

Seybolt, Taylor B., Jay D. Aronson, and Baruch Fischhoff. "Introduction." In *Counting Civilian Casualties: An Introduction to Recording and Estimating Nonmilitary Deaths in Conflict*, edited by Taylor B. Seybolt, Jay D. Aronson, and Baruch Fischhoff. New York: Oxford University Press, 2013.

Sloboda, John and Elizabeth Minor. *Paper 3: The Range of Sources in Casualty Recording*. London: Oxford Research Group, 2012.

Smit, Paul R., Rinke R. de Jong, and Catrien C.J.H. Bijleveld. "Homicide Data in Europe: Definitions, Sources, and Statistics." In *Handbook of European Homicide Research*. Springer, 2012: 5–23.

Tabeau, Ewa and Jan Zwierzchowki. "A Review of Estimation Methods for Victims of the Bosnian War and the Khmer Rouge Regime." In *Counting Civilian Casualties: An Introduction to Recording and Estimating Nonmilitary Deaths in Conflict*, edited by Taylor B. Seybolt, Jay D. Aronson, and Baruch Fischhoff. New York: Oxford University Press, 2013.

UNODC. *Global Study on Homicide 2013*. United Nations, 2014.

UNODC. *International Classification of Crime for Statistial Purposes.* UNODC, 2015.

UNODC. *Global Study on Homicide 2019*. United Nations, 2019.

Vaghela, Prithvirajsinh. "Profile of Unnatural Deaths in Bhuj (Gujarat): A Retrospective Study." *National Journal of Integrated Research in Medicine* 3, no. 2 (2012).

World Bank and WHO. *Global Civil Registration and Vital Statistics: Scaling up Investment Plan 2015–2024.* World Bank Group, 2014.

Zwierzchowski, Jan and Ewa Tabeau, "The 1992–95 War in Bosnia and Herzegovina: Census-Based Multiple System Estimation of Casualties' Undercount." Conference Paper for the International Research Workshop on "The Global Costs of Conflict," Berlin, February1–2, 2010. https://pdfs.semanticscholar.org/5b8c/480e42c4c50976fdd51ec90ae4becb5a0060.pdf?_ga=2.95824348.1388251583.1555396473-596354133.1555396473.

About the GReVD Consortium

To address the issues in this report, the consortium established a long-term goal to create a registry with a single entry for every violent death, identified by location and time, disaggregated by type of violence and characteristics of actors where possible. The consortium will achieve this through improving methodologies for monitoring and coding violent deaths, improving estimates by jointly working together, and establishing a common research infrastructure and ontology, thereby contributing to an improved, shared understanding of violent deaths based on a set of common definitions that can be used freely by anyone.

This Gaps Report is the first step in this research initiative. It sets out the many definitional issues and methodological concerns, including a survey of data gaps, data comparability issues, and double-counting issues that complicate the counting of violent deaths. It contributes to this research area with a 5x9 framework for understanding these issues, which organizes the main methodological challenges around five channels by which data are coded and nine stages in coding violent death data. As a gaps report, this survey does not contain solutions, but concludes with some possible ways forward for reconciling the global datasets into what would eventually become a single registry (while maintaining the independence of the data collection efforts in their own right). Findings are based on a thorough stocktaking of the current literature and the outcome of three workshops with the consortium, convened by SIPRI and the Brookings Institution: two in Washington, DC, and one hosted by the Small Arms Survey in Geneva.[103] The consortium partners have also provided specific input in the form of multiple background papers, and inputs and edits to this report (although errors remain those of the authors). The report draws heavily from the results of a trial using machine learning techniques to describe political violence in Nigeria by the Cline Center for Advanced Social Research (December 2018).

CONSORTIUM MEMBERS INCLUDE THE FOLLOWING ORGANIZATIONS:

Armed Conflict and Location Database (ACLED) (https://www.acleddata.com/) is a disaggregated conflict collection, analysis and crisis mapping project. ACLED collects the dates, actors, types of violence, locations, and fatalities of all reported political violence and protest events across Africa, South Asia, South East Asia, the Middle East, Europe, and Latin America. Political violence and protest include events that occur within civil wars and periods of instability, public protest, and regime breakdown. ACLED's aim is to capture the forms, actors, dates, and locations of political violence and protest as it occurs across states. The ACLED team conducts analysis to describe, explore, and test conflict scenarios, and makes both data and analysis open to freely use by the public.

The Brookings Institution (www.brookings.edu) is a nonprofit organization devoted to independent, in-depth research. It brings together more than 300 leading experts in government and academia from all over the world who provide high quality research, policy recommendations, and analysis on a full range of public policy issues. Research topics cover foreign policy, economics, development, government, and metropolitan policy. It traces its beginnings to

103 This research for producing the Gaps Report, including the workshops, were supported by UKAID.

1916, when a group of reformers founded the Institute for Government Research—the first private organization devoted to analyzing public policy issues at national level.

The Center for International Development and Conflict Management (CIDCM) at the University of Maryland (UMD) (https://www.start.umd.edu) has developed the tool, "Matching Event Data by Location, Time and Type" (MELTT) in 2015 by a team of START-affiliated researchers. MELTT aims to aid in the unbiased understanding of political instability through the integration of event data sets typical of an unstable political climate, ranging from protests to political violence to terrorist incidents.

The Center for Peace and Security Studies (cPASS) at the University of California, San Diego (UCSD) (http://cpass.ucsd.edu) conducts rigorous, data-driven research on international affairs and U.S. foreign policy. The focus at cPASS is on new and emerging modes of interstate conflict. Research at cPASS applies innovative thinking and diverse methodologies (experiments, deductive modeling, statistical analysis, case studies, "big data") to traditional security issues made more dynamic and difficult by increased complexity.

Cline Center for Advanced Social Research at University of Illinois (https://clinecenter.illinois.edu) connects computational expertise in the data sciences with subject matter expertise in the social sciences and humanities to address pressing societal problems around the world. It equips and empowers social scientists, humanists, and data scientists to take up key challenges that threaten human flourishing in the 21st century— including climate change, civil unrest, sustainability, inequality, security, and public health—by applying advanced computational techniques at extreme scales to discover innovative solutions hidden in unstructured data.

The Community Robotics, Education and Technology Empowerment Lab (CREATE Lab) at the Carnegie Mellon University (CMU) (https://cmucreatelab.org) explores socially meaningful innovation and deployment of robotic technologies. Specifically, it aims to (1) empower a technologically fluent generation through experiential learning opportunities in and outside of school, and (2) empower everyday citizens and scientists with affordable environmental sensing and documentation instruments, and powerful visualization platforms for sense-making and sharing of gathered scientific data to promote evidence-based decision making, public discourse, and action.

Global Terrorism Database (GTD) at START at University of Maryland (UMD) (https://www.start.umd.edu/gtd) is an open-source database including information on terrorist events around the world from 1970 through 2017 (with additional annual updates planned for the future). The GTD includes systematic data on domestic as well as transnational and international terrorist incidents that have occurred during this time period and now includes more than 180,000 cases. The database is maintained by the National Consortium for the Study of Terrorism and Responses to Terrorism (START). It is also the basis for other terrorism-related measures, such as the Global Terrorism Index (GTI) published by the Institute for Economics and Peace.

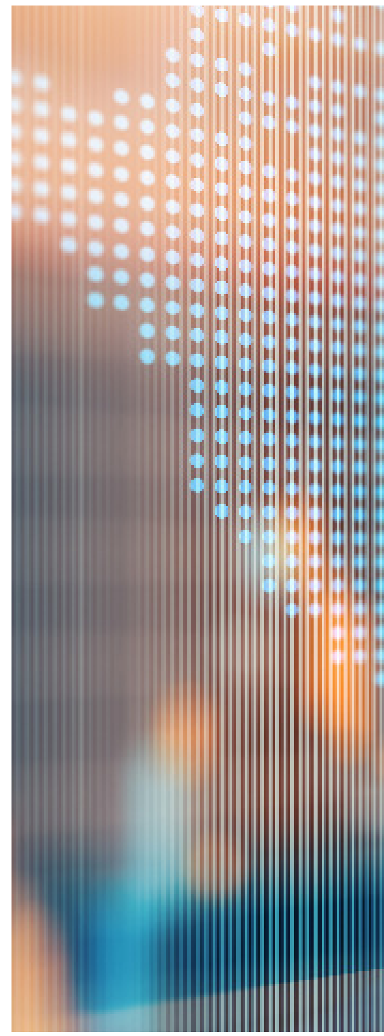
The Igarapé Institute (https://igarape.org.br/en) is an independent think and do tank, formed in 2011 and devoted to integrating security, justice, and development agendas. The Institute's goal is to propose innovative solutions to complex social challenges through research, new technologies, influence in public policies and articulation. The Institute currently works with five overarching themes: (1) national and global drug policies, (2) citizen security, (3) building peace, (4) safer cities, and (5) cybersecurity. Based in Rio de Janeiro, Brazil, it produces the Homicide Monitor.

The Peace Research Institute Oslo (PRIO) (https://www.prio.org) is an independent foundation, established in 1959. It conducts research on the conditions for peaceful relations between states, groups, and people. It seeks to understand the processes that bring societies together or split them apart, exploring how conflicts erupt and how they can be resolved; how different kinds of violence affect people; and how societies tackle crises—and the threat of crisis.

Small Arms Survey (SAS) (http://www.smallarmssurvey.org) is a global center of excellence whose mandate is to generate impartial, evidence-based, and policy-relevant knowledge on all aspects of small arms and armed violence. It is an international source of expertise, information, and analysis on small arms and armed violence issues, and acts as a resource for governments, policymakers, researchers, and civil society. Located in Geneva, Switzerland, SAS is an associate program of the Graduate Institute of International and Development Studies. The Global Violent Deaths (GVD) project provides expertise that can be used to assess progress made in achieving peaceful, just, and inclusive societies through reductions in violent deaths and illicit arms flows. The Small Arms Survey updates its GVD database annually. It includes data on homicides, direct conflict deaths, and other violent deaths from 223 countries/territories from 2004, disaggregated by sex and by instrument.

Stockholm International Peace Research Institute (SIPRI) (https://www.sipri.org/about) is an independent international institute dedicated to research into conflict, armaments, arms control, and disarmament. Established in 1966, SIPRI provides data, analysis, and recommendations, based on open sources, to policymakers, researchers, media and the interested public. Based in Stockholm, SIPRI is regularly ranked among the most respected think tanks worldwide.

Uppsala Conflict Data Programme (UCDP), the University of Uppsala (https://ucdp.uu.se) provides data on organized violence and the oldest ongoing data collection project for civil war, with a history of almost 40 years. Its definition of armed conflict has become the global standard of how conflicts are defined and studied. UCDP produces high-quality data, which are systematically collected, have global coverage, are comparable across cases and countries, and have long time series that are updated annually.



GReVD

A Global Registry of Violent Deaths

Stockholm International Peace Research Institute
SIPRI, Signalistgatan 9
SE-169 72 Solna
Sweden

+46 8 655 97 00

www.GReVD.org