ABSTRACT
The scientific method drives improvements in public health, but a strategy of obstructionism has impeded scientists from gathering even a minimal amount of information to address America’s gun violence epidemic. We argue that in spite of a lack of federal investment, large amounts of publicly available data offer scientists an opportunity to measure a range of firearm-related behaviors. Given the diversity of available data – including news coverage, social media, web forums, online advertisements, and Internet searches (to name a few) – there are ample opportunities for scientists to study everything from trends in particular types of gun violence to gun-related behaviors (such as purchases and safety practices) to public understanding of and sentiment towards various gun violence reduction measures. Science has been sidelined in the gun violence debate for too long. Scientists must tap the big media datastream and help resolve this crisis.

1. MISSED SCIENTIFIC OPPORTUNITIES
Many major improvements in the health and wellbeing of society derive from gathering epidemiological evidence about a disease and then applying the scientific method to test interventions that reduce harm or death from that disease. Just a brief sample of effective public health measures includes the widespread adoption of seatbelts in motor vehicles, the eradication of smallpox and routine vaccinations in general, the development of effective treatments for human immunodeficiency virus, and the prevention of smoking-attributable diseases through rigorous communication campaigns and other policy measures. Yet the power of science has essentially been locked out of one of the nation’s most pressing health crises, the nation’s seventh leading cause of mortality and the leading cause of death among persons aged 15 to 24: firearm injuries.

Two decades ago, as gun violence prevention research was accelerating at the Center for Disease Control and Prevention’s (CDC) National Injury Center, Congressman Jay Dickey (R-AR) amended the Omnibus Consolidated Appropriations Act of 1997 to read: “None of the funds made available for injury prevention and control at the [CDC] may be used to advocate or promote gun control.” In its wake, gun injury prevention funding within the CDC fell 96 percent to less than $100,000 annually (a sum that cannot cover the costs of even a modest traditional study). So under the thumb of the gun lobby was the CDC that the agency informed the National Rifle Association “as a courtesy” anytime investigators under their supervision studied issues related to gun violence, according to the New York Times.

The Dickey Amendment was duplicative and unnecessary – as a federal agency, the CDC is already prohibited from conducting advocacy – and in a January 2013 memorandum President Obama clarified this and requested the CDC and other science agencies within Health and Human Services resume research on gun violence prevention. But to date there is little new investment in gun violence research. The National Institutes of Health (the principal funder of public health research) has funded just two studies, and some health focused gun violence prevention research groups are relying on donations from the research staff themselves.

Limitations have also been established on a range of firearm-related data. Beginning in 2003, Congress enacted a series of amendments restricting access to and use of crime gun trace data collected by the nation’s law enforcement agencies and held by the Bureau of Alcohol, Tobacco, Firearms and Explosives. Those amendments continue to limit researchers’ ability to analyze the movement of guns from the large group of lawful purchasers to the much smaller population who use them in crime. Many states have also imposed restrictions, including exempting data on concealed carry permitting processes from public access.

Although the scientific community agrees that gun violence prevention can be enhanced and informed by public health science – even Congressman Dickey now supports gun violence prevention research – without funding to collect the necessary data or explicit restrictions from accessing data, they cannot begin scientific inquiry to inform gun violence prevention. To provide some sense of the consequence of this gap in information, just 234 reports on PubMed, public health’s publication database, include the...
term “gun violence” whereas, more than 88,000 studies indexed on PubMed make reference to “influenza,” with gun violence rivaling influenza as a cause of premature death [15]. The victims of gun violence cannot wait for legislation to unambiguously fund and support science that can inform prevention efforts.

Fortunately, we believe that public big media data can yield new scientific insights into gun violence today.

2. BIG MEDIA DATA CAN FILL THE GAP

Big media data— including large online media databases such as news aggregators, social media, and Internet searches—are yielding significant breakthroughs across science, particularly in public health [8, 25, 19, 30]. In disease surveillance, researchers have developed a variety of new sources that accurately track [29, 64] and forecast [59, 61, 63] influenza by analyzing data from Twitter [18]. Web Searches [32, 62], specialized apps [65] and Wikipedia [15, 54]. These sources have also been used to rapidly respond to emerging infectious diseases, such as dengue fever [7] and Zika [16, 42, 29], and Ebola [41, 20]. Big media data has yielded new insights into behavioral aspects of public health, including public responses to planned communication campaigns [2, 11, 6], awareness campaigns, [15, 8, 77, 35] and organic/spontaneous events, such as a celebrity health disclosure [30, 10, 22, 53, 12]. Online resources are also helping to fight the current opioid epidemic [36], as well as studies that track emerging drugs [58] and measure drug prevalence [58]. These efforts are perhaps most valuable when they provide insights about phenomenon for which—like gun violence—there are few existing high-quality data resources. These include work on a broad array of mental health areas: measuring the prevalence of common mental illnesses like psychological distress [7]; discovering seasonal patterns of mental illness [9], post traumatic stress disorder [25], depression [27, 24], and schizophrenia [17]; understanding eating disorder behaviors [21]; and predicting suicidal ideation [22, 37, 35, 17, 50]. That big media data provide insights into such a wide range of public health behaviors should give us hope that the same is possible for gun violence.

A significant benefit of big media databases is that they are already free, publicly accessible, and timely, with users creating large volumes of text in real-time. For instance, there are more than 70 million annual Twitter postings that include terms like “gun” or “guns,” [15] that touch on the full spectrum of gun-related behaviors (e.g., ownership, safety practices, unlicensed sales, etc.) and gun-related attitudes (e.g., understanding of current gun safety measures, support for additional safety measures, etc.). Anecdotally, these data already influence the public’s understanding of gun violence, as news reporting on many shootings rely on details gleaned from social media reports, like the Facebook Live broadcast of a police shooting in Minneapolis [16]. Big media data could be a proverbial silver bullet to interject science into gun violence prevention, providing investigators the data with which to implement rigorous studies.

A variety of private organizations are turning to strategies that generate, aggregate, and disseminate gun violence data to promote research. For instance, federal data on the occurrence of gun violence is limited and available only after significant delays. The nonprofit, nonpartisan Gun Violence Archive [4] aggregates incidents of gun violence from news reports, including both homicides and non-fatal shootings, and makes the data available in near real-time.

But big media data scientists can go beyond simply counting incidents to answer far more detailed questions. For example, shootings are routinely covered in traditional news media and these contemporaneous records include many additional details about the circumstances such as the weapon and magazine used; the name, relationship, and prior criminal history of the shooter, and other details. The news article itself may also be processed to reveal the reporting frame; how the shooting is presented in ways beyond the factual details. Understanding media framing is important for understanding the way the public perceives gun safety issues [44]. For instance, Everytown for Gun Safety maintains a database of unintentional shootings involving children, which shows that federal data vastly undercount the frequency of these tragic incidents, but more importantly the database shows that more than two-thirds of cases could have been prevented had the gun been stored responsibly.

Another powerful application of these data can be demonstrated through an examination of Internet search queries for guns [54]. Using Google Trends [54], a public index of Google search volume, we analyzed all gun-related searches emerging from the United States that included the terms “gun” or “guns” and related search trends to forty-eight of the most publicized of the more than one hundred thirty mass shootings in the same period [7] (Figure 2). Gun-related searches have been on the rise recently, and appear to be spiking around America’s spate of mass shootings since late 2012. For instance, gun-related searches reached record levels the day President Obama commented on the Sandy Hook Elementary shooting, and peaks also occurred in the days following the Sandy Hook Elementary shooting and notable mass shootings in San Bernardino and the Pulse nightclub in Orlando.

Moreover, we can measure the behaviors and thoughts of the public towards guns via the content of their searches—as we have done to describe the specific health concerns of the public in their own words [3] and link these to specific strategies for gun violence prevention. Figure 2 shows Google searches for gun shopping (including all queries with the terms “gun(s)” and “buy(s)" or “dealer(s)” or “shop(s)”), shopping for gun storage safes (“gun(s)” and “safe(s)”), and gun purchase background checks (“gun(s)” and “background(s)”) around the Sandy Hook Elementary shooting and President Obama’s subsequent speech. Searches for all three domains significantly ($p < 0.05$) and substantially deviated from search volumes during the prior year. Searches for gun safe-related terms (e.g., “buy a gun safe" and like queries)}
increased 195% (95%CI: 110 to 300), spiking the day of the Sandy Hook Elementary shooting and remained elevated for two months. Whereas searches for gun shopping terms and gun background check terms spiked both on the day of the shooting and subsequently when President Obama publicly commented on gun violence and the spikes were less durable.

3. FUTURE DIRECTIONS
As these case studies demonstrate, big media data have the potential to yield actionable insights for gun violence prevention research even without significant funding to support data collection, which is often the most costly component of the scientific process. Many unanswered questions about gun violence prevention can be explored this way, including those distinctly related to individual gun safety and public policy on gun safety.

Around individual gun safety the questions are only limited by the extent to which guns are discussed on big media channels. What other events trigger gun shopping? Is the public seeking out illicit gun paraphernalia such as after-market adapters that render weapons fully automatic? What safety aids (e.g., safes, trigger locks, etc.) are the public seeking? Unlike with traditional survey data, a big media data analysis of these questions has more face validity, where these behaviors are directly observed through Internet searches, social media posts, or microblogging. At the same time the structure of big media data, especially in social media, means the unit of analysis can be individuals (within established research ethics and privacy standards[23]), meaning gun safety data can be linked to demographic traits (e.g., sex, ethnicity, education, etc.) that are associated with individual social media accounts. Moreover, investigators could

Figure 1: Daily trends for all gun-related Google queries emerging from the United States (2006-2016).
Figure 2: Google searches for gun shopping, gun storage safes, and gun background checks increased after the Sandy Hook Elementary shooting.

study the co-occurrence of mitigating factors related to gun safety, such as mental illness and gun safety practices within individual social media account holders.

At the policy level we can assess the public’s understanding of existing and proposed gun safety regulation, summarize public feedback to inform the development of gun violence prevention regulations, and evaluate the extent to which regulations are being implemented or weakened by loopholes. For example, Everytown for Gun Safety tracked the emergence of a thriving online market for unlicensed gun sales without background checks which undermine existing gun safety laws by leaving an open door for convicted felons and domestic abusers to arrange gun sales and get armed with no questions asked. On a single website, they identified over 600,000 unique gun ads posted over a one-year-period, and the users’ geographic location and other information listed publicly in the ads provides opportunities to study the frequency and character of this heretofore poorly understood commerce.

Effective gun safety regulation is dependent on an informed public. Public health has successfully invested in educating the public on effective regulatory strategies to promote safety (e.g., tobacco, seatbelts, underage drinking). In addition to assessing public understanding of gun safety using big media data, researchers can also use these same channels to disseminate messages that better inform the public.

Ultimately big media data provides a pathway to use the scientific method to broadly inform gun violence prevention by leveraging billions of diverse data points. It is in everyone’s interest to have an empirically informed approach to gun violence prevention, with data driving us to the most effective strategies to improve public health. The speed of progress can be accelerated by data scientists and public health working together to harness big media data.

4. REFERENCES


