Can Big Media Data Revolutionize Gun Violence Prevention?

John W. Ayers
San Diego State University
San Diego, CA, USA
ayers.john.w@gmail.com

Benjamin M. Althouse
Santa Fe Institute
Santa Fe, NM
althouse@santafe.edu

Eric C. Leas
UC San Diego
La Jolla, CA, USA
eleas@ucsd.edu

Ted Alcorn
Everytown for Gun Safety
New York, New York
talcorn@everytown.org

Mark Dredze
Johns Hopkins University
Baltimore, MD, USA
mdredze@cs.jhu.edu

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 [48] and the leading cause of death among persons aged 15 to 24: firearm injuries [50].

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." [1] 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) [43]. 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 [40].

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 [55] President Obama clarified this and requested the CDC and other science agencies within Health and Human Services resume research on gun violence prevention [51]. 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 [60].

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.[1] 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 [31, 49, 14] – 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

---

[1] http://everytownresearch.org/reports/access-denied/

term "gun violence"[2] whereas, more than 88,000 studies indexed on PubMed make reference to "influenza,"[3] with gun violence rivaling influenza as a cause of premature death [48]. 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 [5, 28, 19, 30]. In disease surveillance, researchers have developed a variety of new sources that accurately track [26, 64] and forecast [59, 61, 63] influenza by analyzing data from Twitter [18], Web Searches [32, 62], specialized apps [65] and Wikipedia [45, 34]. These sources have also been used to rapidly respond to emerging infectious diseases, such as dengue fever [4] 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, [13, 8, 57, 33] and organic/spontaneous events, such as a celebrity health disclosure [39, 10, 52, 53, 12]. Online resources are also helping to fight the current opioid epidemic [36], as well as studies that track emerging drugs [38] 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 [47]; understanding eating disorder behaviors [21]; and predicting suicidal ideation [22, 37, 35, 17, 56]. 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 [46]. 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[5], 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* [6], 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 fourty-eight of the most publicized of the more than one hundred thirty mass shoootings in the same period [7] (Figure 1). 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)

---

[2] http://www.ncbi.nlm.nih.gov/pubmed?term=%22gun%20violence%22

[3] http://www.ncbi.nlm.nih.gov/pubmed/?term=influenza

[4] http://www.gunviolencearchive.org/

[5] http://www.everytownresearch.org/notanaccident

[6] http://www.google.com/trends

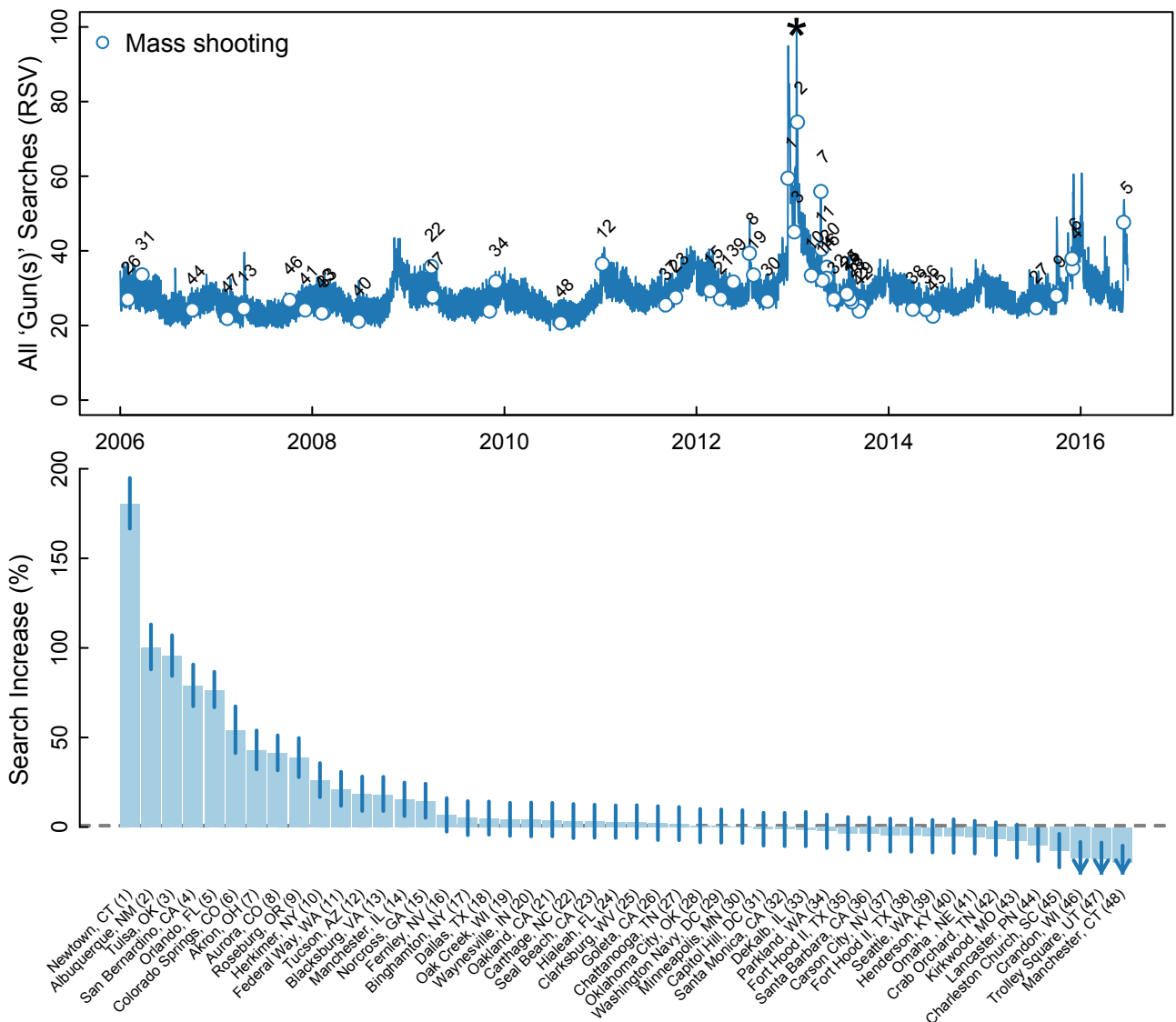[7] http://everytownresearch.org/mass-shootings/

Figure 1: Daily trends for all gun-related Google queries emerging from the United States (2006-2016).

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
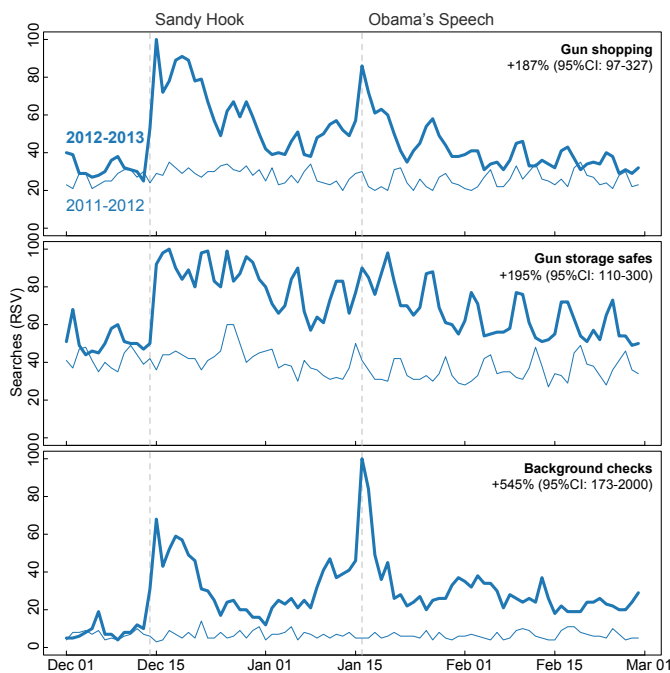
Figure 2: Google searches for gun shopping, gun storage safes, and gun background checks increased after the Sandy Hook Elementary shooting.

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,[8] 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.[9]

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 [49].

---

[8] https://everytownresearch.org/reports/point-click-fire/

[9] https://everytownresearch.org/reports/business-as-usual

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

[1] *PUBLIC LAW 104–208—SEPT. 30, 1996: Omnibus Consolidated Appropriations Act, 1997.* 104th Congress, 1996.

[2] J.-P. Allem, P. Escobedo, K.-H. Chu, D. W. Soto, T. B. Cruz, and J. B. Unger. Campaigns and counter campaigns: reactions on twitter to e-cigarette education. *Tobacco control*, 2016.

[3] B. M. Althouse, J.-P. Allem, M. A. Childers, M. Dredze, and J. W. Ayers. Population health concerns during the United States' great recession. *American Journal of Preventive Medicine*, 46(2):166–170, February 2014.

[4] B. M. Althouse, Y. Y. Ng, and D. A. Cummings. Prediction of dengue incidence using search query surveillance. *PLoS Negl Trop Dis*, 5(8):e1258, 2011.

[5] J. Ayers, B. Althouse, and M. Dredze. Could behavioral medicine lead the web data revolution? *JAMA*, 311(14):1399–1400, 2014.

[6] J. W. Ayers and B. M. Althouse. "tips from former smokers" can benefit from considering all available data. *American journal of preventive medicine*, 49(6):e133, 2015.

[7] J. W. Ayers, B. M. Althouse, J.-P. Allem, M. A. Childers, W. Zafar, C. Latkin, K. M. Ribisl, and J. S. Brownstein. Novel surveillance of psychological distress during the great recession. *Journal of affective disorders*, 142(0):323–330, 12 2012.

[8] J. W. Ayers, B. M. Althouse, J.-P. Allem, D. E. Ford, K. M. Ribisl, and J. E. Cohen. A novel evaluation of world no tobacco day in latin america. *Journal of medical Internet research*, 14(3):e77, 2012.

[9] J. W. Ayers, B. M. Althouse, J.-P. Allem, J. N. Rosenquist, and D. E. Ford. Seasonality in seeking mental health information on google. *American journal of preventive medicine*, 44(5):520–525, 2013.

[10] J. W. Ayers, B. M. Althouse, M. Dredze, E. C. Leas, and S. M. Noar. News and internet searches about human immunodeficiency virus after Charlie Sheen's disclosure. *JAMA Intern Med*, 2016.

[11] J. W. Ayers, B. M. Althouse, and S. Emery. Changes in internet searches associated with the" tips from former smokers" campaign. *American journal of preventive medicine*, 48(6):e27–9, 2015.

[12] J. W. Ayers, B. M. Althouse, S. M. Noar, and J. E. Cohen. Do celebrity cancer diagnoses promote primary cancer prevention? *Preventive medicine*, 58:81–84, 2014.

[13] J. W. Ayers, J. L. Westmaas, E. C. Leas, A. Benton, Y. Chen, M. Dredze, and B. Althouse. Leveraging big data to improve health awareness campaigns: A novel evaluation of the Great American Smokeout. *JMIR Public Health and Surveillance*, 2016.

[14] C. L. Barry, E. E. McGinty, J. S. Vernick, and D. W. Webster. After Newtown — public opinion on gun policy and mental illness. *New England Journal of Medicine*, 368(12):1077–1081, 2013.

[15] A. Benton, B. Hancock, G. Coppersmith, J. W. Ayers, and M. Dredze. After Sandy Hook Elementary: A year in the gun control debate on Twitter. In *Bloomberg Data for Good Exchange*, 2016.

[16] I. I. Bogoch, O. J. Brady, M. Kraemer, M. German, M. I. Creatore, M. A. Kulkarni, J. S. Brownstein, S. R. Mekaru, S. I. Hay, E. Groot, et al. Anticipating the international spread of Zika virus from Brazil. *Lancet (London, England)*, 387(10016):335–336, 2016.

[17] S. R. Braithwaite, C. Giraud-Carrier, J. West, M. D. Barnes, and C. L. Hanson. Validating machine learning algorithms for Twitter data against established measures of suicidality. *JMIR mental health*, 3(2):e21, 2016.

[18] D. A. Broniatowski, M. J. Paul, and M. Dredze. National and local influenza surveillance through Twitter: an analysis of the 2012-2013 influenza epidemic. *PloS one*, 8(12):e83672, 2013.

[19] J. S. Brownstein, C. C. Freifeld, and L. C. Madoff. Digital disease detection — harnessing the web for public health surveillance. *New England Journal of Medicine*, 360(21):2153–2157, 2009.

[20] M. Carter. How Twitter may have helped Nigeria contain Ebola. *BMJ*, 2014.

[21] S. Chancellor, T. Mitra, and M. De Choudhury. Recovery amid pro-anorexia: Analysis of recovery in social media. In *Conference on Human Factors in Computing Systems (CHI)*, pages 2111–2123. ACM, 2016.

[22] M. D. Choudhury, E. Kiciman, M. Dredze, G. Coppersmith, and M. Kumar. Discovering shifts to suicidal ideation from mental health content in social media. In *Conference on Human Factors in Computing Systems (CHI)*, 2016.

[23] M. Conway. Ethical issues in using Twitter for public health surveillance and research: developing a taxonomy of ethical concepts from the research literature. *Journal of medical Internet research*, 16(12):e290, 2014.

[24] G. Coppersmith, M. Dredze, and C. Harman. Quantifying mental health signals in Twitter. In *ACL Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality*, pages 51–60, 2014.

[25] G. Coppersmith, C. Harman, and M. Dredze. Measuring post traumatic stress disorder in Twitter. In *ICWSM*, 2014.

[26] A. Culotta. Towards detecting influenza epidemics by analyzing Twitter messages. In *First workshop on social media analytics*, pages 115–122. ACM, 2010.

[27] M. De Choudhury, S. Counts, E. J. Horvitz, and A. Hoff. Characterizing and predicting postpartum depression from shared facebook data. In *Computer supported cooperative work (CSCW)*, pages 626–638. ACM, 2014.

[28] M. Dredze. How social media will change public health. *IEEE Intelligent Systems*, 27(4):81–84, 2012.

[29] M. Dredze, D. A. Broniatowski, and K. M. Hilyard. Zika vaccine misconceptions: A social media analysis. *Vaccine*, 2016.

[30] G. Eysenbach. Infodemiology and infoveillance. *American Journal of Preventive Medicine*, 40(5):S154–S158, 2011.

[31] S. Frattaroli, D. W. Webster, and G. J. Wintemute. Implementing a public health approach to gun violence prevention: The importance of physician engagement. *Annals of Internal Medicine*, 158(9):697–698, 2013.

[32] J. Ginsberg, M. H. Mohebbi, R. S. Patel, L. Brammer, M. S. Smolinski, and L. Brilliant. Detecting influenza epidemics using search engine query data. *Nature*, 457(7232):1012–1014, 2009.

[33] R. W. Glynn, J. C. Kelly, N. Coffey, K. J. Sweeney, and M. J. Kerin. The effect of breast cancer awareness month on internet search activity - a comparison with awareness campaigns for lung and prostate cancer. *BMC Cancer*, 11(1):1–9, 2011.

[34] K. S. Hickmann, G. Fairchild, R. Priedhorsky, N. Generous, J. M. Hyman, A. Deshpande, and S. Y. Del Valle. Forecasting the 2013–2014 influenza season using wikipedia. *PLoS Comput Biol*, 11(5):e1004239, 2015.

[35] J. Jashinsky, S. H. Burton, C. L. Hanson, J. West, C. Giraud-Carrier, M. D. Barnes, and T. Argyle. Tracking suicide risk factors through Twitter in the us. *Crisis*, 2014.

[36] T. Katsuki, T. K. Mackey, and R. Cuomo. Establishing a link between prescription drug abuse and illicit online pharmacies: analysis of Twitter data. *Journal of medical Internet research*, 17(12), 2015.

[37] M. Kumar, M. Dredze, G. Coppersmith, and M. D. Choudhury. Shifts in suicidal ideation manifested in social media following celebrity suicides. In *Conference on Hypertext and Social Media*, 2015.

[38] J. E. Lange, J. Daniel, K. Homer, M. B. Reed, and J. D. Clapp. Salvia divinorum: effects and use among youtube users. *Drug and alcohol dependence*, 108(1):138–140, 2010.

[39] E. C. Leas, B. M. Althouse, M. Dredze, N. Obradovich, J. H. Fowler, S. M. Noar, J. Allem, and J. W. Ayers. Big data sensors of organic advocacy: The case of Leonardo DiCaprio and climate change. *PLoS One*, 2016.

[40] M. Luo. N.R.A. stymies firearms research, scientists say, January 25 2011. [Online].

[41] M. S. Majumder, S. Kluberg, M. Santillana, S. Mekaru, and J. S. Brownstein. 2014 ebola outbreak: Media events track changes in observed reproductive number. *PLoS Currents Outbreaks*, 2015.

[42] M. S. Majumder, M. Santillana, S. R. Mekaru, D. P. McGinnis, K. Khan, and J. S. Brownstein. Utilizing nontraditional data sources for near real-time estimation of transmission dynamics during the 2015-2016 Colombian Zika virus disease outbreak. *JMIR public health and surveillance*, 2(1):e30, 2016.

[43] Mayors Against Illegal Guns. Access denied: How the gun lobby is depriving police, policy makers, and the public of the data we need to prevent gun violence. 2013.

[44] E. E. McGinty, D. W. Webster, and C. L. Barry. Effects of news media messages about mass shootings

on attitudes toward persons with serious mental illness and public support for gun control policies. *American Journal of Psychiatry*, 170(5):494–501, 2013.

[45] D. J. McIver and J. S. Brownstein. Wikipedia usage estimates prevalence of influenza-like illness in the United States in near real-time. *PLoS Comput Biol*, 10(4):e1003581, 2014.

[46] E. C. McLaughlin. Woman streams aftermath of fatal officer-involved shooting, July 8 2016. [Online].

[47] M. Mitchell, K. Hollingshead, and G. Coppersmith. Quantifying the language of schizophrenia in social media. In *Association for Computational Linguistics Workshop on Computational Linguistics and Clinical Psychiatry*, page 11, 2015.

[48] A. Mokdad, J. Marks, D. Stroup, and J. Gerberding. Actual causes of death in the United States, 2000. *JAMA*, 291(10):1238–1245, 2004.

[49] D. Mozaffarian, D. Hemenway, and D. S. Ludwig. Curbing gun violence: lessons from public health successes. *JAMA*, 309(6):551–552, 2013.

[50] National Center for Health Statistics. *Health, United States, 2015, With Special Feature on Racial and Ethnic Health Disparities*. Hyattsville, MD, 2016.

[51] National Institutes of Health. Notice of legislative mandates in effect for fy2012 (not-od-12-034), January 20 2012. [Online].

[52] S. M. Noar, B. M. Althouse, J. W. Ayers, D. B. Francis, and K. M. Ribisl. Cancer information seeking in the digital age effects of angelina jolie's prophylactic mastectomy announcement. *Medical Decision Making*, 35(1):16–21, 2015.

[53] S. M. Noar, K. M. Ribisl, B. M. Althouse, J. F. Willoughby, and J. W. Ayers. Using digital surveillance to examine the impact of public figure pancreatic cancer announcements on media and search query outcomes. *J Natl Cancer Inst Monogr*, 2013(47):188–194, 2013.

[54] S. V. Nuti, B. Wayda, I. Ranasinghe, S. Wang, R. P. Dreyer, S. I. Chen, and K. Murugiah. The use of google trends in health care research: a systematic review. *PloS one*, 9(10):e109583, 2014.

[55] B. Obama. Presidential memorandum – engaging in public health research on the causes and prevention of gun violence, January 16, 2013. [Online].

[56] B. O'Dea, M. Larsen, P. Batterham, A. Calear, and H. Christensen. Talking suicide on Twitter: Linguistic style and language processes of suicide-related posts. *European Psychiatry*, 33:S274, 2016.

[57] L. Owen and B. Youdan. 22 years on: the impact and relevance of the uk no smoking day. *Tobacco Control*, 15(1):19–25, 2006.

[58] M. J. Paul, M. S. Chisolm, M. W. Johnson, R. G. Vandrey, and M. Dredze. Assessing the validity of online drug forums as a source for estimating demographic and temporal trends in drug use. *Journal of Addiction Medicine*, 2016.

[59] M. J. Paul, M. Dredze, and D. Broniatowski. Twitter improves influenza forecasting. *PLOS Currents Outbreaks*, 2014.

[60] R. R. Tale of 2 agencies: CDC avoids gun violence research but nih funds it. *JAMA*, 315(16):1689–1692,

2016.

[61] M. Santillana, A. T. Nguyen, M. Dredze, M. J. Paul, E. O. Nsoesie, and J. S. Brownstein. Combining search, social media, and traditional data sources to improve influenza surveillance. *PLoS Comput Biol*, 11(10):e1004513, 2015.

[62] M. Santillana, D. W. Zhang, B. M. Althouse, and J. W. Ayers. What can digital disease detection learn from (an external revision to) Google Flu Trends? *American Journal of Preventive Medicine*, 47(3):341–347, September 2014.

[63] J. Shaman and A. Karspeck. Forecasting seasonal outbreaks of influenza. *Proceedings of the National Academy of Sciences*, 109(50):20425–20430, 2012.

[64] A. Signorini, A. M. Segre, and P. M. Polgreen. The use of Twitter to track levels of disease activity and public concern in the US during the influenza A H1N1 pandemic. *PloS one*, 6(5):e19467, 2011.

[65] M. S. Smolinski, A. W. Crawley, K. Baltrusaitis, R. Chunara, J. M. Olsen, O. Wójcik, M. Santillana, A. Nguyen, and J. S. Brownstein. Flu near you: crowdsourced symptom reporting spanning 2 influenza seasons. *American journal of public health*, 105(10):2124–2130, 2015.