

Original Paper

Using Web-Based Search Data to Study the Public's Reactions to Societal Events: The Case of the Sandy Hook Shooting

Nir Menachemi^{1,2*}, BS, MPH, PhD; Saurabh Rahurkar^{2*}, BDS (India), DrPH; Mandar Rahurkar^{3*}, PhD

¹Richard M. Fairbanks School of Public Health, Health Policy and Management, Indiana University-IUPUI, Indianapolis, IN, United States

²Regenstrief Institute, Center for Biomedical Informatics, Indianapolis, IN, United States

³CareChime, Mountain View, CA, United States

* all authors contributed equally

Corresponding Author:

Saurabh Rahurkar, BDS (India), DrPH

Regenstrief Institute

Center for Biomedical Informatics

1101 West Tenth Street

Indianapolis, IN, 46202

United States

Phone: 1 317 274 9338

Fax: 1 317 274 9305

Email: rahurkar@regenstrief.org

Abstract

Background: Internet search is the most common activity on the World Wide Web and generates a vast amount of user-reported data regarding their information-seeking preferences and behavior. Although this data has been successfully used to examine outbreaks, health care utilization, and outcomes related to quality of care, its value in informing public health policy remains unclear.

Objective: The aim of this study was to evaluate the role of Internet search query data in health policy development. To do so, we studied the public's reaction to a major societal event in the context of the 2012 Sandy Hook School shooting incident.

Methods: Query data from the Yahoo! search engine regarding firearm-related searches was analyzed to examine changes in user-selected search terms and subsequent websites visited for a period of 14 days before and after the shooting incident.

Results: A total of 5,653,588 firearm-related search queries were analyzed. In the after period, queries increased for search terms related to "guns" (+50.06%), "shooting incident" (+333.71%), "ammunition" (+155.14%), and "gun-related laws" (+535.47%). The highest increase (+1054.37%) in Web traffic was seen by news websites following "shooting incident" queries whereas searches for "guns" (+61.02%) and "ammunition" (+173.15%) resulted in notable increases in visits to retail websites. Firearm-related queries generally returned to baseline levels after approximately 10 days.

Conclusions: Search engine queries present a viable infodemiology metric on public reactions and subsequent behaviors to major societal events and could be used by policymakers to inform policy development.

(*JMIR Public Health Surveill* 2017;3(1):e12) doi:[10.2196/publichealth.6033](https://doi.org/10.2196/publichealth.6033)

KEYWORDS

Internet; search engine; firearms; health policy; information seeking behavior; public health informatics; gun control debate

Introduction

Nearly 9 out of every 10 Americans have Internet access at home [1] and Web browsing accounts for an average of 23 hours per week that includes activities such as communication, entertainment, news, shopping, and social networking [1,2]. Importantly, searching the Web for information using search engines far surpasses most other types of activities with over

91% of US adults contributing to Web traffic of this nature [3]. Consequently, Web searches generate a vast amount of data in the form of users' search queries which capture their information-seeking preferences (eg, what they search for) and behavior (eg, what sites they visit). Analysis of this information—a form of infodemiology [4]—could be used to improve our understanding of various issues which in turn can inform policy development.

Infodemiology is an emerging discipline that focuses on analyzing electronic information from the Internet (eg, search queries, social media, and so on) in order to provide information on public health and policy [4]. Previous infodemiology literature has examined Web search query data to evaluate various public health and health care research questions. For example, several studies used Web search query data to identify influenza outbreaks ahead of conventional population detection methods in the United States [5-13] and abroad [14-20], as well as other public health surveillance [21-24]. Researchers have also analyzed search queries for the detection and prevention of adverse drug events, or any other drug related complications [25-27]. Finally, Web-based search logs have been utilized to predict health care utilization and costs following information seeking on search engines [22,28-30]. To our knowledge, no study has examined search data to better understand the public's sentiments, reactions, and behaviors to major societal events.

We consider major societal events consistent with the "social crises" definition from the crisis management literature. These events are characterized by the severe consequences of the incident, low probability of incident occurrence, and the informational and situational uncertainty that occur among members of the public [31,32]. These situations are inevitably accompanied by collective anxiety, improvised group behaviors, and adaptive collaboration among the public [32-36]. Public mass shooting events share these characteristics; they have a low probability of occurrence, they are followed by lack of reliable information regarding details and consequences of the event, and generate heightened anxiety and public outcry in response to the situation.

The purpose of this paper was to analyze search query data in the context of a major societal event. We decided to study the Sandy Hook Elementary Shooting incident that occurred on December 14, 2012, in Newton Connecticut to determine whether such data can be used to better understand the public's reactions to such an event. The act of a lone gunman causing the deaths of 20 children and 6 adults received national and international attention, prompting renewed public interest in gun issues [37]. We are interested in understanding how firearm-related information seeking (eg, looking up relevant laws, learning about advocacy) and Web-based behavior (eg, visits to firearm-related retailers) changed immediately after the incident. Understanding these trends will provide insights into how Americans responded to the incident which can enhance societal debates and inform policy development related to firearms.

Methods

Data Source and Preparation

We examined deidentified data from Yahoo! search engine queries in a 28-day period before (14-day) and after (14-day) the Sandy Hook shooting incident. Our population consisted of all users of the Yahoo! search engine located in the United States (including Puerto Rico and Mariana Islands) that queried firearm-related searches during the study period. The majority of the information consumed on the Web starts as search queries entered by the user. The choices made by the user in the form

of websites they click from the list of populated search results present a much more comprehensive picture of a user's information needs. Our goal was to use the search query data to evaluate patterns of information seeking regarding firearms and to evaluate broadly the changes in intent based on differences in the content (retail, news, education, and so on) and sources (commercial entities, noncommercial organizations, government entities, and so on) of information sought.

From the complete Yahoo! search query database, we identified all firearm-related queries from November 31, 2012 to December 28, 2012. Queries were text strings consisting of single words or phrases that users typed into the search engine; we identified these using keywords that would match partially or completely with words in the queries. Firearm-related search queries were identified by using keywords in the following categories: Gun type (gun, firearm, handgun, rifle, pistol, revolver, and shotgun), ammunition (ammunition, ammo, and bullets), law related (Brady Act, second amendment of the US constitution), and shooting. In order to choose keywords in each category, we examined Web-based trends of firearm-related search queries for December 2012 using Google Trends. We did this by first examining simpler queries (eg, handgun), and the 10 most correlated searches for these queries. This was repeated recursively with each of the correlated queries until we found no new or correlated searches. This gave us a set of 247 queries that were related to firearms. We wanted our keywords to have the ability to identify these 247 queries as well as any other searches that may be firearm related. Thus our keywords consisted of single words which could identify most firearm-related searches based on complete or partial matches with user queries. As such, our analysis included users' actual search queries that included keywords in any of the 4 categories.

In addition, we also analyzed the uniform resource locators (URLs) that each individual user clicked from the search results generated by their search queries. First, we identified the domain for each URL that the user clicked; for example, if the user clicked the URL "http://en.wikipedia.org/wiki/Second_Amendment," the domain was identified as "wikipedia.org." Next, we categorized these URLs based on the top-level domain (TLD) into commercial entities (.com), noncommercial organizations (.org), government entities (.gov, .state.us, and .mil), educational institutions (.edu), and others (country specific, .pro, .tv, and so on). Including TLDs in our analysis allows us to infer the nature of the organization; for instance, TLDs such as .gov, .mil, and .edu have legal restrictions which prevent them from being used by organizations other than government, military, and educational entities. Moreover, search ranking algorithms are unlikely to place URLs from entities with erroneously used TLDs higher in the search results. These factors allow the use of TLDs to categorize the nature of organizations fairly reliably. Next, each domain was categorized as retail (websites for the purchase of guns, ammunition, and gun accessories; including gun shows), news (websites of newspapers, news channels etc), educational (websites, regardless of TLD, that host information regarding gun safety, gun laws, gun maintenance, and may include websites of gun advocacy groups), showbiz (websites of movies, television

shows, music videos, and so on) or “other” which included all remaining uncategorized websites.

The TLD and the content describe different characteristics of the same website and thus examining them together provides a richer understanding of the information seeking patterns. As such we created a variable that assigns a class to each website in the dataset derived from its content category and TLD. Thus, a website with retail content hosted by a commercial entity would be classified as “retail content, .com.” Finally, we created a variable to capture all of the websites owned or affiliated with the National Rifle Association (NRA) as listed on the NRA’s website [38]. Such websites were classified as gun rights advocacy groups. The NRA website also identifies other sites that it categorizes as “antigun lobbying organizations” [39]. We categorized these websites as gun control advocacy groups.

To evaluate the association between the Sandy Hook incident and the nature of information sought, we first examined the distributions of various characteristics of the domains visited by users following the search query (category of keyword, top-level domain, category of the website’s content, and advocacy view of the websites visited). Next, we investigated differences in website characteristics in the period before and after the shooting incident using the website classes. We also examined the percentage change in website visits for each of the characteristics relative to the total websites in the before period to those in the after period. Additionally, we examined the percent change in website visits for each of the characteristics in the after period to the website visit for the same characteristics in the before period.

Finally, it is possible that observed changes in information-seeking behavior over time may be due to the presence of secular or temporal trends and not as a result of the Sandy Hook shooting incident. For example, given that our study period overlapped with the holiday shopping season, one might expect an increase in Web-based shopping activity that can include increases in firearm-related searches, independent of the Sandy Hook incident. To differentiate the shopping activity related increase in search activity from that related to the shooting incident, we included a control query that would be agnostic to the trend due to the Sandy Hook incident but sensitive to the temporal trends of the holiday season. Thus, a query for “bicycle” (and related synonyms) was used as a control search term.

Limitations

The following limitations must be noted. First, given that Yahoo! search accounted for about 12% of the US search engine market

share in December 2012 [40], we recognize that caution must be used before generalizing to the entire US population. Additionally, the Web pages visited by the users may also be associated with result-ranking algorithms which vary by search engines. Since 2011, Yahoo! search is powered by Bing [41] and whereas the exact algorithms are proprietary, evidence suggests that Bing emphasizes keywords (search strings) in ranking search results [42]. Second, our analysis was focused on the query-level (ie, website visited after each search) and not the user level which may include several queries in a given search episode. Third, approximately 30% of all observations consisted of a large number of unique domains occurring with a low frequency and thus could not be classified. Nevertheless, these domains individually accounted for less than 1% of all observations and thus their effect on the findings is likely minimal. Finally, our work represents an exploratory study to examine whether search data can be used for a new purpose. Thus, the existing body of literature provided little guidance on the methods or approaches to analyzing such data. We recognize that future studies may identify additional techniques for analyzing similarly complex data.

Results

A total of 5,653,588 firearm-related search queries were identified by our keywords in a 28-day period before (14-day) and after (14-day) the Sandy Hook shooting incident. By each search query category (see Table 1), the majority (59.62%; 3,370,523/5,653,588) focused on a gun type (eg, queries with the term pistol, shotgun, or rifle) with the rest focused on the shooting incident (22.47%; 1,270,122/5,653,588), ammunition (16.88%; 954,363/5,653,588), or law related searches (1.04%; 58,580/5,653,588). Based on TLD, users were most likely to visit websites of commercial entities (.com: 88.03%; 4,976,990/5,653,588) followed by noncommercial organizations (.orgs: 6.63%; 374,863/5,653,588) and government entities (.gov, state.us, mil: 1.06%; 59,939/5,653,588). Users most frequently clicked on links that brought them to retail websites (30.33%; 1,714,504/5,653,588), followed by news websites (23.38%; 1,321,706/5,653,588), educational websites (20.32%; 1,148,897/5,653,588), and showbiz websites (2.09%; 118,174/5,653,588). A total of 66,581 websites that users visited could be classified as those of gun rights (68.86%; 45,848/66,581) or gun control (31.14%; 20,733/66,581) advocacy groups. Finally, our control search query for bicycle synonyms yielded 597,859 individual observations during the same study period.

Table 1. Characteristics of the search query data. Source: Authors' analysis of Yahoo! search queries for December 2012.

Variables	Proportion n (%)
Search keywords	
Firearm-related (n=5,653,588)	
Gun type	3,370,523 (59.62)
Shooting incidents	1,270,122 (22.47)
Ammunition	954,363 (16.88)
Law related	58,580 (1.04)
Counterfactual (n=597,859)	
Bicycle	597,859 (100.00)
Top-level domain	
Firearm-related (n=5,653,588)	
Commercial entities	4,976,990 (88.03)
Noncommercial organizations	374,863 (6.63)
Government entities	59,939 (1.06)
Educational institutions	9419 (0.17)
Other	232,377 (4.11)
Category	
Firearm-related (n=5,653,588)	
Retail	1,714,504 (30.33)
News	1,321,706 (23.38)
Educational	1,148,897 (20.32)
Showbiz	118,174 (2.09)
Other or uncategorized	1,350,307 (23.88)
Stance on gun control	
Firearm-related (n=66,581)	
Gun control advocacy group	20,733 (31.14)
Gun rights advocacy group	45,848 (68.86)

Bivariate relationships between user search queries and the class of websites visited based on content and TLD are presented in [Table 2](#). In all categories there was an increase in firearm-related search queries in the period after the shooting. Gun type searches which were the most common firearm-related query showed the least relative change after the shooting incident with a 50.06% increase in the proportion of user searches. In contrast, the law category of search queries after the shooting incident had a 535.47% increase in the proportion of searches although it was the least searched. Although users searching for gun types (+61.02%) or ammunition (+173.15%) were more likely to visit

retail content on commercial entity websites after the shooting incident, a greater proportion (+1054.37%) visited news content on commercial entity websites for shooting incident searches. Law-related searches, however, had a greater proportion of visits to websites with educational content from noncommercial organizations (+702.70%), commercial entities (+484.20%), and educational institutions (+593.97%). Importantly, when examining changes to bicycle-related search terms (the counterfactual) in the before and after period, we observed a relatively modest decrease in overall searches (-8.64%).

Table 2. Changes in search patterns before and after the Sandy Hook school shooting incident (December 14, 2012). Source: Authors' analysis of Yahoo! search queries for December 2012.

Search query	n	Before period ^a	After period ^a	Delta %	Cumulative %
Gun type					
Retail content, .com	965,795	370,002 (38.31%)	595,793 (61.69%)	61.02%	28.70%
News content, .com	595,689	340,883 (57.22%)	254,806 (42.78%)	-25.25%	46.40%
Educational content, .com	560,295	196,435 (35.06%)	363,860 (64.94%)	85.23%	63.05%
Educational content, .org	184,979	46,886 (25.35%)	138,093 (74.65%)	194.53%	68.54%
Other content, .com	774,034	293,009 (37.85%)	481,025 (62.15%)	64.17%	91.54%
Total	3,365,359	1,345,833 (39.99%)	2,019,526 (60.01%)	50.06%	100%
Shooting incidents					
News content, .com	648,233	51,678 (7.97%)	596,555 (92.03%)	1054.37%	51.45%
Educational content, .com	94,150	26,941 (28.61%)	67,209 (71.39%)	149.47%	58.93%
Educational content, .org	54,830	7911 (14.43%)	46,919 (85.57%)	493.09%	62.28%
Showbiz related content, .com	44,135	15,732 (35.65%)	28,403 (64.35%)	80.54%	66.78%
Other content, .com	319,892	98,564 (30.81%)	221,328 (69.19%)	124.55%	92.18%
Total	1,259,817	236,050 (18.74%)	1,023,767 (81.26%)	333.71%	100%
Ammunition					
Retail related content, .com	609,246	163,272 (26.80%)	445,974 (73.20%)	173.15%	63.85%
Educational content, .com	95,687	31,256 (32.66%)	64,431 (67.34%)	106.14%	73.88%
News content, .com	61,025	10,849 (17.78%)	50,176 (82.22%)	362.49%	80.28%
Other content, .com	100,711	36,104 (35.85%)	64,607 (64.15%)	78.95%	90.83%
Total	954,158	268,670 (28.16%)	685,488 (71.84%)	155.14%	100%
Laws					
Educational content, .org	25,429	2817 (11.08%)	22,612 (88.92%)	702.70%	43.41%
Educational content, .com	12,473	1823 (14.62%)	10,650 (85.38%)	484.20%	64.70%
Educational content, .edu	3422	431 (12.59%)	2991 (87.41%)	593.97%	70.54%
News content, .com	2663	383 (14.38%)	2280 (85.62%)	495.30%	75.09%
Other content, .com	8978	1502 (16.73%)	7476 (83.27%)	397.74%	90.41%
Total	58,580	7965 (13.60%)	50,615 (86.40%)	535.47%	100%
Bicycle					
Retail related content, .com	142,209	75,551 (53.13%)	66,658 (46.87%)	-11.77%	24.08%
Educational content, .com	87,964	45,550 (51.78%)	42,414 (48.22%)	-6.88%	38.98%
Other content, .com	266,220	137,068 (51.49%)	129,152 (48.51%)	-5.78%	84.06%
Other content, .other	23,605	12,254 (51.91%)	11,351 (48.09%)	-7.37%	88.06%
Other content, .org	19,957	10,233 (51.28%)	9724 (48.72%)	-4.97%	91.44%
Total	590,530	308,603 (52.26%)	281,927 (47.74%)	-8.64%	100%

^aIndicates \pm 14 days from the Sandy Hook event.

Table 3 presents the bivariate relationship between the time period and the advocacy view (gun rights vs gun control) of the websites visited by users following search queries for each category of firearm-related searches. Search results increased in all categories in the period after the shooting incident (range: 79.94%-418.39%). The majority of search queries for gun type (67.92%; 40,069/58,998), shooting incident (84.55%;

2490/2945), and ammunition (93.16%; 2316/2486) resulted in users visiting websites of gun rights advocacy groups, whereas those searching for laws were more likely to visit websites of gun control advocacy groups (54.79%; 1179/2152). On the whole, Web-based users had an increase of between 285.71% and 660.58% of visiting gun control advocacy group websites after the shooting incident.

Table 3. Bivariate relationship between time period and the advocacy view of the websites visited by users for each category of firearm-related searches. Source: Authors' analysis of Yahoo! Search queries for December 2012.

Search query	n (%)	Before ^a	After ^a	Delta %
Gun type				
Gun rights	40,069 (67.92)	28.29%	71.71%	153.53%
Gun control	18,929 (32.08)	14.97%	85.03%	467.93%
Total	58,998	24.01%	75.99%	216.42%
Shooting incidents				
Gun rights	2490 (84.55)	39.04%	60.96%	56.17%
Gun control	455 (15.45)	17.58%	82.42%	368.75%
Total	2945	35.72%	64.28%	79.94%
Ammunition				
Gun rights	2316 (93.16)	36.44%	63.56%	74.41%
Gun control	170 (6.84)	20.59%	79.41%	285.71%
Total	2486	35.36%	64.64%	82.82%
Laws				
Gun rights	973 (45.21)	21.69%	78.31%	261.14%
Gun control	1179 (54.79)	11.62%	88.38%	660.58%
Total	2152	16.17%	83.83%	418.39%

^aIndicates ± 14 days from the Sandy Hook event.

Figure 1 presents the trend data graphed in the before and after period for firearm-related and bicycle-related searches for 4 categories of TLDs. As can be seen, Web traffic as a result of firearm-related search queries saw a sharp increase corresponding to the Sandy Hook shooting incident for domains of commercial entities, educational institutions, government entities, and noncommercial organizations. Additionally, depending on the TLD, a relatively smaller peak in Web traffic is seen at days 6 and 11 before the shooting incident following firearm-related searches, with the greatest increase seen for

.com domains. Conversely, bicycle-related searches during the same period appear relatively unchanged. Figure 2 presents the trend data graphed in the before and after period for firearm-related search queries for advocacy view. Websites of both gun control and gun rights advocacy groups saw a sharp increase in traffic corresponding to the shooting incident following firearm-related searches. The traffic decreased slowly for both over the after period with slight increase in traffic at day 11.

Figure 1. Websites visited following firearm-related and bicycle-related search queries by top-level domain. Source: Authors' analysis of Yahoo! Search queries for December 2012.

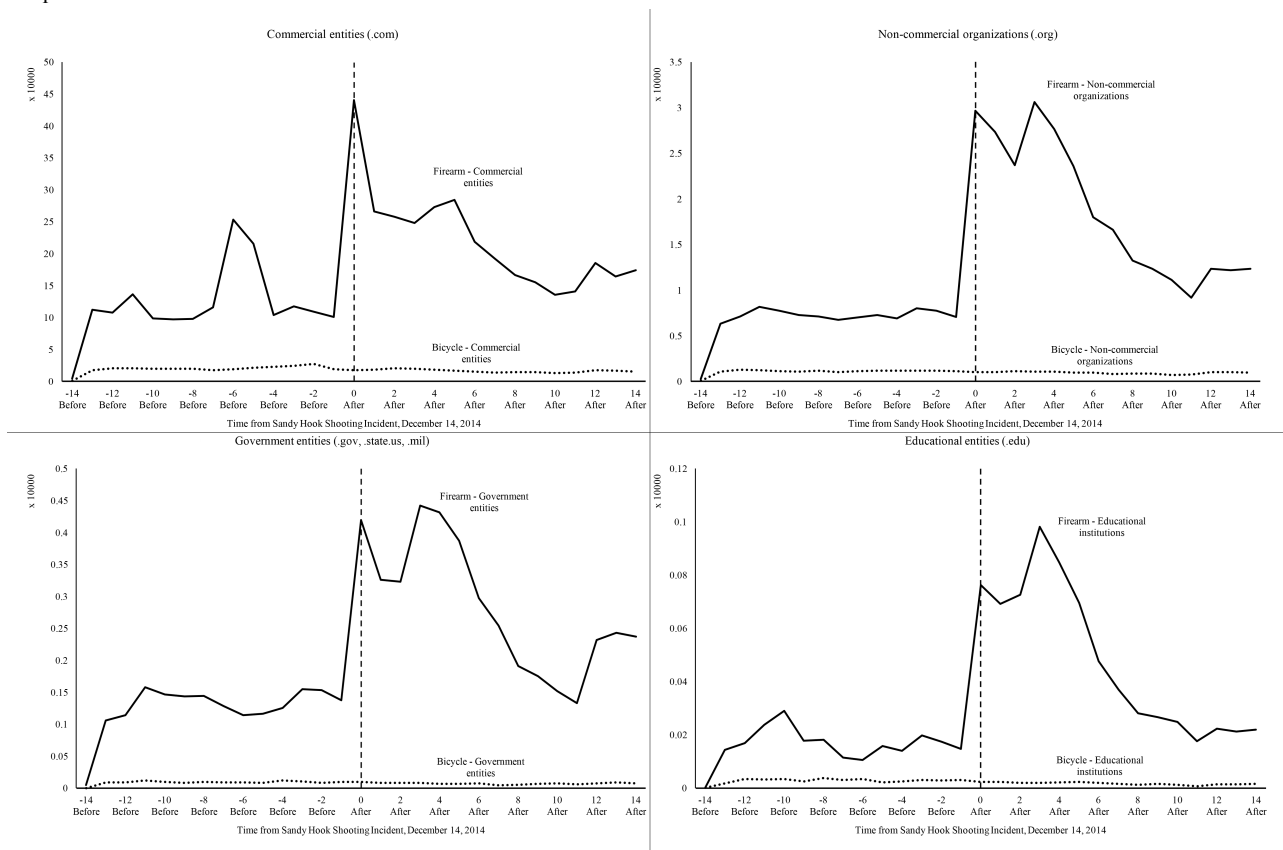
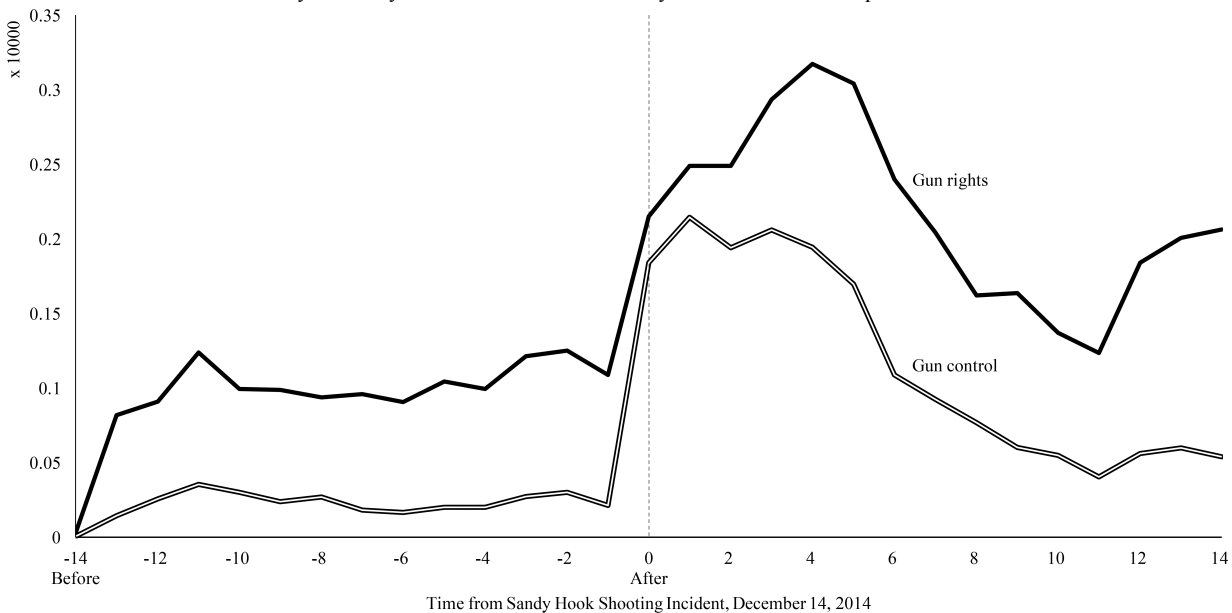


Figure 2. Trends in website visits by advocacy view. Source: Authors' analysis of Yahoo! search queries for December 2012.



Discussion

Principal Findings

One of the key findings of our analysis was that firearm-related searches more than doubled immediately after the Sandy Hook shooting incident in contrast to the control searches for “bicycle” which showed a small change with a decrease in the number of searches in the after period. This finding suggests that

Web-based user search queries capture the immediate change in public interest following events of the nature of Sandy Hook shooting and can thus potentially serve as real-time indicators of the public psyche.

Overall, retail websites were the most visited websites following searches for gun types and ammunition. A salient finding was that gun type and ammunition searches had a 2-fold to 3-fold increase after the shooting incident. Furthermore, although it

may seem natural to expect a greater interest in news articles following a major societal event, retail website visits had the highest and second highest increase after the shooting incident for gun type and ammunition searches, respectively. This finding may possibly be the result of a heightened interest in purchasing firearms and/or ammunition for one's protection against the apparent public safety concerns raised by the mass shooting [43]. Additionally, it is possible that some individuals may anticipate an increase in regulatory control over access to firearms as a ramification of the Sandy Hook incident and as such prompt purchase of firearms before any such legislative action is passed.

Furthermore, there was a 6-fold increase in law-related firearm queries in the period immediately after the shooting. Importantly, these were the least likely searched terms in the before period and noted the greatest percent increase in the after period. This increased interest may be due—in part—to the purchase-related search or inquiry conducted by the potential firearm and ammunition buyers discussed above, not to mention the renewed interest in the gun-policy debate after Sandy Hook. Interestingly, most users seeking law-related information were interested in educational information and chose websites of noncommercial organizations, commercial entities, or educational institutions. From the advocacy perspective, more people visited websites of gun rights groups than did the gun control groups. However, despite gun control websites forming a lower proportion of all websites supporting an advocacy stance, they experienced the greatest percent increase from before to the after period. This trend was seen in all categories of firearm-related searches, with Web traffic to gun control advocacy groups exhibiting between almost 4-fold and 6-fold increase in the after period.

In addition to the trends discussed previously, a key feature of user searches and the subsequent URL clicks was that in all categories users were far less likely to choose content from a government entity. For example, even though the majority of the law-related searches are directed toward educational content, users are more likely to choose noncommercial organizations (including gun control or gun rights advocacy groups), commercial entities, and educational institutions as their preferred sources of information. The nature of advocacy groups is such that they exist to influence stakeholder decision to align with their agenda and therefore, the resulting conflict of interest may be an impediment to providing unbiased information. Thus, it is also likely that users seeking information about gun laws may obtain this information from websites of advocacy groups.

Our analysis of user search query data presents several key implications from a policy perspective. First, as stated above, user search queries present a valuable real-time indicator of the attitudes of the population as shown by the effect of the Sandy Hook shooting incident. In fact, the spike seen 6 days before the Sandy Hook event corresponds to 2 news stories: one on December 7, when supermarket employees found a handgun in frozen meat [44] and another on December 9, when a 7-year-old boy was fatally shot in the parking lot of a gun store [45]. Similarly, the spike seen around day 12 corresponds with the much publicized advocacy speech given by a prominent American sportscaster on television [46]. These spikes highlight user search queries as a timely measure of the public's reaction

to societal events. The time period immediately after a major event is characterized by heightened awareness and information-seeking behavior that may not be representative of public action during normal states (eg, buying firearms at twice the regular prices [43]). Indeed, Oh et al note that "rumormongering" is common after major societal events including shooting events [32]. On the one hand, this may indicate that policymakers should consider the timing of their actions noting that while a societal event can trigger interest in a topic, it ironically may not be the best time to debate major tenets of policy change. On the other hand, some observed behavior may be due to fears arising from misinformation. For example, the increased purchase-related queries in our findings corroborate increased firearm sales due to fear of increased gun control legislation [43].

Second, it is possible that people are accessing information sources with either commercial or advocacy-related interest, at the same time being far less likely to choose content from government and educational institution websites. This may be because websites of government and educational institutions rank lower in the search results compared with those of commercial and advocacy interest groups. Although search engine optimization (SEO) may play a role in the higher ranking of commercial and advocacy interest websites, it is also possible the information presented by government and educational institutions may be less accessible. This may be due to suboptimal website design, jargon-filled language, poor SEO, lack of up-to-date information, and so on. Policy efforts should focus on providing reliable information as well as improved dissemination of this information by government institutions. Government entities may collaborate with educational institutions toward the creation of information portals focused on dissemination of accurate, timely, and high-quality information that is easy to understand. Furthermore, resources allocated toward making the public aware of these portals as well as on SEO may ensure that these websites rank higher in search results and thus visited more often.

Finally, the increased interest generated by the shooting incident appears to start tending toward normal levels around day 10, eventually returning to the levels before the shooting. This indicates that the increased interest generated due to incidents such as Sandy Hook presents a short window in which to form the public's opinion. As discussed previously, this may not present the best opportunity to engage in public debate due to the increased anxiety and fear following these events. Whether this fear was driven by the need to protect oneself or the possibility of losing the right to purchase a firearm, it is unlikely that political sentiment for policy change will be easy to accomplish when fear is driving some stakeholder's perspectives. Instead, policymakers should consider preemptively addressing some of the anticipated fears by implementing targeted campaigns that focus on specific groups of individuals. A recent US study reported that 3 percent of the US population owns nearly half of all firearms in the country with an average of 17 firearms each [43,47-49]. The median firearm ownership, however, remains at 1 to 2 firearms per owner. These individuals are likely to indulge in firearm purchases [43] after events such as the Sandy Hook shooting.

Furthermore, given that personal protection against other people remains the most prevalent reason for firearm ownership in the US [47], mass shooting events may also motivate those on the fence to purchase firearms. As such, targeted campaigns that focus on these groups of individuals in order to allay fears and reduce reactionary purchase of firearms may help achieve some policymaker's goals of lower rates of firearm ownership.

Conclusions

Our findings enabled us to identify directions for future research; web browsing choices and attitudes toward firearms may be affected by numerous other factors. As such, it may be valuable

to examine the differences between attitudes toward firearms based on state characteristics such as political affiliation, socioeconomic status, and gun ownership. It may also be interesting to look at ordered queries nested within each deidentified user based on the order in which the user clicked each URL to provide richer data on users' search intent. Search query data presents a valuable infodemiology metric of near real-time analysis of peoples' attitudes and responses to major societal events. We believe future studies can employ the use of other search query datasets possibly with active user participation to examine the impact of society events over a longer period of time.

Acknowledgments

This work was done when the author, Mandar Rahurkar was employed at Yahoo! Labs. The authors would like to thank Suju Rajan at Yahoo! Labs for reviewing the manuscript and supporting this research.

Conflicts of Interest

None declared.

References

1. Purcell K, Brenner J, Rainie L. Pew Internet Research. 2012 Mar 09. Search engine use 2012 URL: http://www.pewinternet.org/files/old-media/Files/Reports/2012/PIP_Search_Engine_Use_2012.pdf [accessed 2016-05-25] [WebCite Cache ID 6hlw9efLi]
2. eMarketer. eMarketer. 2013 Jul 02. Social Usage Involves More Platforms, More Often URL: <http://www.emarketer.com/Article/Social-Usage-Involves-More-Platforms-More-Often/1010019> [accessed 2016-05-24] [WebCite Cache ID 6hkEp7KKi]
3. Purcell K. Pew Research Center. SearchEmail Still Top the List of Most Popular Online Activities URL: http://www.pewinternet.org/files/old-media/Files/Reports/2011/PIP_Search-and-Email.pdf [accessed 2016-05-25] [WebCite Cache ID 6hlw7LKBJ]
4. Eysenbach G. Infodemiology and infoveillance: framework for an emerging set of public health informatics methods to analyze search, communication and publication behavior on the Internet. *J Med Internet Res* 2009;11(1):e11 [FREE Full text] [doi: [10.2196/jmir.1157](https://doi.org/10.2196/jmir.1157)] [Medline: [19329408](https://pubmed.ncbi.nlm.nih.gov/19329408/)]
5. Bentley RA, Ormerod P. A rapid method for assessing social versus independent interest in health issues: a case study of 'bird flu' and 'swine flu'. *Soc Sci Med* 2010 Aug;71(3):482-485. [doi: [10.1016/j.socscimed.2010.03.042](https://doi.org/10.1016/j.socscimed.2010.03.042)] [Medline: [20494504](https://pubmed.ncbi.nlm.nih.gov/20494504/)]
6. Brownstein JS, Freifeld CC, Madoff LC. Digital disease detection--harnessing the web for public health surveillance. *N Engl J Med* 2009 May 21;360(21):2153-5, 2157 [FREE Full text] [doi: [10.1056/NEJMp0900702](https://doi.org/10.1056/NEJMp0900702)] [Medline: [19423867](https://pubmed.ncbi.nlm.nih.gov/19423867/)]
7. Dugas AF, Hsieh Y, Levin SR, Pines JM, Mareiniss DP, Mohareb A, et al. Google Flu Trends: correlation with emergency department influenza rates and crowding metrics. *Clin Infect Dis* 2012 Feb 15;54(4):463-469 [FREE Full text] [doi: [10.1093/cid/cir883](https://doi.org/10.1093/cid/cir883)] [Medline: [22230244](https://pubmed.ncbi.nlm.nih.gov/22230244/)]
8. Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS, Brilliant L. Detecting influenza epidemics using search engine query data. *Nature* 2009 Feb 19;457(7232):1012-1014. [doi: [10.1038/nature07634](https://doi.org/10.1038/nature07634)] [Medline: [19020500](https://pubmed.ncbi.nlm.nih.gov/19020500/)]
9. Hulth A, Rydevik G, Linde A. Web queries as a source for syndromic surveillance. *PLoS One* 2009;4(2):e4378 [FREE Full text] [doi: [10.1371/journal.pone.0004378](https://doi.org/10.1371/journal.pone.0004378)] [Medline: [19197389](https://pubmed.ncbi.nlm.nih.gov/19197389/)]
10. Pelat C, Turbelin C, Bar-Hen A, Flahault A, Valleron AJ. More diseases tracked by using Google Trends. *Emerg Infect Dis* 2009 Aug;15(8):1327-1328 [FREE Full text] [doi: [10.3201/eid1508.090299](https://doi.org/10.3201/eid1508.090299)] [Medline: [19751610](https://pubmed.ncbi.nlm.nih.gov/19751610/)]
11. Polgreen PM, Chen Y, Pennock DM, Nelson FD. Using internet searches for influenza surveillance. *Clin Infect Dis* 2008 Dec 1;47(11):1443-1448 [FREE Full text] [doi: [10.1086/593098](https://doi.org/10.1086/593098)] [Medline: [18954267](https://pubmed.ncbi.nlm.nih.gov/18954267/)]
12. Seifter A, Schwarzwald A, Geis K, Aucott J. The utility of "Google Trends" for epidemiological research: Lyme disease as an example. *Geospat Health* 2010 May;4(2):135-137. [doi: [10.4081/gh.2010.195](https://doi.org/10.4081/gh.2010.195)] [Medline: [20503183](https://pubmed.ncbi.nlm.nih.gov/20503183/)]
13. Zhou X, Ye J, Feng Y. Tuberculosis surveillance by analyzing Google trends. *IEEE Trans Biomed Eng* 2011 Aug;58(8). [doi: [10.1109/TBME.2011.2132132](https://doi.org/10.1109/TBME.2011.2132132)] [Medline: [21435969](https://pubmed.ncbi.nlm.nih.gov/21435969/)]
14. Cho S, Sohn CH, Jo MW, Shin S, Lee JH, Ryoo SM, et al. Correlation between national influenza surveillance data and google trends in South Korea. *PLoS One* 2013;8(12):e81422 [FREE Full text] [doi: [10.1371/journal.pone.0081422](https://doi.org/10.1371/journal.pone.0081422)] [Medline: [24339927](https://pubmed.ncbi.nlm.nih.gov/24339927/)]
15. Seo D, Jo M, Sohn CH, Shin S, Lee J, Yu M, et al. Cumulative query method for influenza surveillance using search engine data. *J Med Internet Res* 2014 Dec 16;16(12):e289 [FREE Full text] [doi: [10.2196/jmir.3680](https://doi.org/10.2196/jmir.3680)] [Medline: [25517353](https://pubmed.ncbi.nlm.nih.gov/25517353/)]

16. Woo H, Cho Y, Shim E, Lee J, Lee C, Kim SH. Estimating influenza outbreaks using both search engine query data and social media data in South Korea. *J Med Internet Res* 2016;18(7):e177 [FREE Full text] [doi: [10.2196/jmir.4955](https://doi.org/10.2196/jmir.4955)] [Medline: [27377323](https://pubmed.ncbi.nlm.nih.gov/27377323/)]
17. Kang M, Zhong H, He J, Rutherford S, Yang F. Using Google Trends for influenza surveillance in South China. *PLoS One* 2013;8(1):e55205 [FREE Full text] [doi: [10.1371/journal.pone.0055205](https://doi.org/10.1371/journal.pone.0055205)] [Medline: [23372837](https://pubmed.ncbi.nlm.nih.gov/23372837/)]
18. Kelly H, Grant K. Interim analysis of pandemic influenza (H1N1) 2009 in Australia: surveillance trends, age of infection and effectiveness of seasonal vaccination. *Euro Surveill* 2009 Aug 6;14(31) [FREE Full text] [Medline: [19660248](https://pubmed.ncbi.nlm.nih.gov/19660248/)]
19. Valdivia A, Lopez-Alcalde J, Vicente M, Pichiule M, Ruiz M, Ordobas M. Monitoring influenza activity in Europe with Google Flu Trends: comparison with the findings of sentinel physician networks - results for 2009-10. *Euro Surveill* 2010;15(29) [FREE Full text] [Medline: [20667303](https://pubmed.ncbi.nlm.nih.gov/20667303/)]
20. Eysenbach G. Infodemiology: tracking flu-related searches on the web for syndromic surveillance. *AMIA Annu Symp Proc* 2006:244-248 [FREE Full text] [Medline: [17238340](https://pubmed.ncbi.nlm.nih.gov/17238340/)]
21. Zheluk A, Quinn C, Hercz D, Gillespie JA. Internet search patterns of human immunodeficiency virus and the digital divide in the Russian Federation: infoveillance study. *J Med Internet Res* 2013;15(11):e256 [FREE Full text] [doi: [10.2196/jmir.2936](https://doi.org/10.2196/jmir.2936)] [Medline: [24220250](https://pubmed.ncbi.nlm.nih.gov/24220250/)]
22. Yom-Tov E, White RW, Horvitz E. Seeking insights about cycling mood disorders via anonymized search logs. *J Med Internet Res* 2014;16(2):e65 [FREE Full text] [doi: [10.2196/jmir.2664](https://doi.org/10.2196/jmir.2664)] [Medline: [24568936](https://pubmed.ncbi.nlm.nih.gov/24568936/)]
23. Ling R, Lee J. Disease monitoring and health campaign evaluation using Google search activities for HIV and AIDS, stroke, colorectal cancer, and Marijuana use in Canada: a retrospective observational study. *JMIR Public Health Surveill* 2016 Oct 12;2(2):e156 [FREE Full text] [doi: [10.2196/publichealth.6504](https://doi.org/10.2196/publichealth.6504)] [Medline: [27733330](https://pubmed.ncbi.nlm.nih.gov/27733330/)]
24. Foroughi F, Lam AK, Lim MS, Saremi N, Ahmadvand A. "Googling" for cancer: an Infodemiological assessment of online search interests in Australia, Canada, New Zealand, the United Kingdom, and the United States. *JMIR Cancer* 2016 May 04;2(1):e5. [doi: [10.2196/cancer.5212](https://doi.org/10.2196/cancer.5212)]
25. Simmering JE, Polgreen LA, Polgreen PM. Web search query volume as a measure of pharmaceutical utilization and changes in prescribing patterns. *Res Social Adm Pharm* 2014;10(6):896-903. [doi: [10.1016/j.sapharm.2014.01.003](https://doi.org/10.1016/j.sapharm.2014.01.003)] [Medline: [24603135](https://pubmed.ncbi.nlm.nih.gov/24603135/)]
26. White RW, Tatonetti NP, Shah NH, Altman RB, Horvitz E. Web-scale pharmacovigilance: listening to signals from the crowd. *J Am Med Inform Assoc* 2013 May 1;20(3):404-408 [FREE Full text] [doi: [10.1136/amiajnl-2012-001482](https://doi.org/10.1136/amiajnl-2012-001482)] [Medline: [23467469](https://pubmed.ncbi.nlm.nih.gov/23467469/)]
27. Yom-Tov E, Gabrilovich E. Postmarket drug surveillance without trial costs: discovery of adverse drug reactions through large-scale analysis of web search queries. *J Med Internet Res* 2013;15(6):e124 [FREE Full text] [doi: [10.2196/jmir.2614](https://doi.org/10.2196/jmir.2614)] [Medline: [23778053](https://pubmed.ncbi.nlm.nih.gov/23778053/)]
28. White RW, Horvitz E. Web to world: predicting transitions from self-diagnosis to the pursuit of local medical assistance in web search. *AMIA Annu Symp Proc* 2010;2010:882-886 [FREE Full text] [Medline: [21347105](https://pubmed.ncbi.nlm.nih.gov/21347105/)]
29. Agarwal V, Zhang L, Zhu J, Fang S, Cheng T, Hong C, et al. Impact of predicting health care utilization via web search behavior: a data-driven analysis. *J Med Internet Res* 2016 Sep 21;18(9):e251 [FREE Full text] [doi: [10.2196/jmir.6240](https://doi.org/10.2196/jmir.6240)] [Medline: [27655225](https://pubmed.ncbi.nlm.nih.gov/27655225/)]
30. White RW, Horvitz E. From health search to healthcare: explorations of intention and utilization via query logs and user surveys. *J Am Med Inform Assoc* 2014;21(1):49-55 [FREE Full text] [doi: [10.1136/amiajnl-2012-001473](https://doi.org/10.1136/amiajnl-2012-001473)] [Medline: [23666794](https://pubmed.ncbi.nlm.nih.gov/23666794/)]
31. Runyan RC. Small business in the face of crisis: identifying barriers to recovery from a natural disaster. *JCCM* 2006 Mar;14(1):12-26. [doi: [10.1111/j.1468-5973.2006.00477.x](https://doi.org/10.1111/j.1468-5973.2006.00477.x)]
32. Oh O, Agrawal M, Rao HR. Community intelligence and social media services: a rumor theoretic analysis of tweets during social crises. *MIS Q* 2013 May;37(2):407-426.
33. Bharosa N, Lee J, Janssen M. Challenges and obstacles in sharing and coordinating information during multi-agency disaster response: propositions from field exercises. *Inf Syst Front* 2009 May 9;12(1):49-65. [doi: [10.1007/s10796-009-9174-z](https://doi.org/10.1007/s10796-009-9174-z)]
34. Janssen M, Lee J, Bharosa N, Cresswell A. Advances in multi-agency disaster management: Key elements in disaster research. *Inf Syst Front* 2009 May 9;12(1):1-7. [doi: [10.1007/s10796-009-9176-x](https://doi.org/10.1007/s10796-009-9176-x)]
35. Majchrzak A, Jarvenpaa S, Hollingshead A. Coordinating expertise among emergent groups responding to disasters. *Organization Science* 2007 Feb;18(1):147-161. [doi: [10.1287/orsc.1060.0228](https://doi.org/10.1287/orsc.1060.0228)]
36. Kendra J, Wachtendorf T. Elements of resilience after the World Trade Center disaster: reconstituting New York City's Emergency Operations Centre. *Disasters* 2003 Mar;27(1):37-53. [Medline: [12703151](https://pubmed.ncbi.nlm.nih.gov/12703151/)]
37. Barron J. NYTimes.: New York Times; 2012 Dec 15. Children Were All Shot Multiple Times With a Semiautomatic, Officials Say URL: http://www.nytimes.com/2012/12/16/nyregion/gunman-kills-20-children-at-school-in-connecticut-28-dead-in-all.html?_r=0 [accessed 2017-02-27] [WebCite Cache ID [6hkcJTAeI](https://www.webcitation.org/6hkcJTAeI)]
38. membership.nrahq. 2014 Feb 13. NRA Websites URL: <https://membership.nrahq.org/othersites.asp> [accessed 2017-02-27] [WebCite Cache ID [6hkWQqHnv](https://www.webcitation.org/6hkWQqHnv)]

39. NRAILA. 2014 Aug 20. NRA-ILA: Anti-Gun Lobbying Organizations URL: <https://www.nraila.org/articles/20140820/anti-gun-lobbying-organizations> [accessed 2017-02-27] [WebCite Cache ID 6oayrUYoC]
40. comscore. 2014 Apr 15. comScore Releases March 2014 U.S. Search Engine Rankings URL: http://www.comscore.com/Insights/Press-Releases/2014/4/comScore-Releases-March-2014-U.S.-Search-Engine-Rankings?cs_edgescape_cc=US [accessed 2017-02-27] [WebCite Cache ID 6hkcsDL5p]
41. Radhakrishnan K. Ysearchblog. 2011 Oct 18. Search Alliance Global Algo Transition Update URL: <http://web.archive.org/web/20140705105032/http://www.ysearchblog.com/2011/10/18/search-alliance-global-algo-transition-update/> [accessed 2017-02-27] [WebCite Cache ID 6ob0RubaV]
42. Webmaster. Blogs.bing. 2011 Apr 13. Collection of SEO related documents from the Bing Ecosystem URL: <https://web.archive.org/web/20160618042243/http://blogs.bing.com/webmaster/2011/04/13/collection-of-seo-related-documents-from-the-bing-ecosystem/> [accessed 2017-02-27] [WebCite Cache ID 6jme0CGzb]
43. Beckett L. Thetrace.: The Trace; 2016 Sep 20. Meet America's Gun Super-Owners — With An Average of 17 Firearms Each URL: <https://web.archive.org/web/20161011153019/https://www.thetrace.org/2016/09/gun-super-owners-harvard-survey/> [accessed 2017-02-27] [WebCite Cache ID 6lBKELcff]
44. Pfeiffer E. news.Yahoo. 2012 Dec 07. Loaded pistol found in package of frozen meat URL: <http://web.archive.org/web/20130203193717/http://news.yahoo.com/blogs/sideshow/loaded-pistol-found-package-frozen-meat-193134816.html> [accessed 2017-02-27] [WebCite Cache ID 6hkccovVL]
45. Levitz J. WSJ.: Wall Street Journal; 2012 Dec 09. Boy Killed Accidentally Outside Gun Store URL: <http://www.wsj.com/news/articles/SB10001424127887324339204578169493535248354> [accessed 2016-05-24] [WebCite Cache ID 6hkc9iC7o]
46. Strauss C. USAToday.: USA Today; 2012 Dec 03. Bob Costas gives anti-gun speech on 'Sunday Night Football' URL: <http://www.usatoday.com/story/gameon/2012/12/03/nfl-costas-speech-reaction/1742313/> [accessed 2017-02-27] [WebCite Cache ID 6hkZ0DfSz]
47. Azrael D, Hepburn L, Hemenway D, Miller M. The Stock and Flow of US Firearms: Results from the 2015 National Firearms Survey. In: Implications for Regulation and Enforcement Conference Hub. Paper presented at: The Underground Gun Market: Russell Sage Foundation; 2016 Apr 29 Presented at: RSF Journal Conference: The Underground Gun Market; April 28, 2016; New York, NY URL: http://www.russellsage.org/sites/all/files/RSF_Journal/Cook_Pollack/Azrael_et_al.pdf
48. NPR.: NPR; 2016 Sep 20. Nearly Half Of Guns In U.S. Owned By 3 Percent Of Population, Study Finds URL: <http://www.npr.org/2016/09/20/494765559/nearly-half-of-guns-in-u-s-owned-by-3-percent-of-population-study-finds> [accessed 2016-02-27] [WebCite Cache ID 6lBHUeQoH]
49. Beckett L. TheGuardian.: The Guardian; 2016 Sep 19. Gun inequality: US study charts rise of hardcore super owners URL: <https://www.theguardian.com/us-news/2016/sep/19/us-gun-ownership-survey> [accessed 2016-10-11] [WebCite Cache ID 6lBHIVrp3]

Abbreviations

- NRA:** National Rifle Association
SEO: search engine optimization
TLD: top level domain
URL: uniform resource locator

Edited by G Eysenbach; submitted 08.06.16; peer-reviewed by D Walker, Y Kwon; comments to author 04.08.16; revised version received 29.10.16; accepted 03.02.17; published 23.03.17

Please cite as:

Menachemi N, Rahrurkar S, Rahrurkar M

Using Web-Based Search Data to Study the Public's Reactions to Societal Events: The Case of the Sandy Hook Shooting

JMIR Public Health Surveill 2017;3(1):e12

URL: <http://publichealth.jmir.org/2017/1/e12/>

doi: [10.2196/publichealth.6033](https://doi.org/10.2196/publichealth.6033)

PMID: [28336508](https://pubmed.ncbi.nlm.nih.gov/28336508/)

©Nir Menachemi, Saurabh Rahrurkar, Mandar Rahrurkar. Originally published in JMIR Public Health and Surveillance (<http://publichealth.jmir.org>), 23.03.2017. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete

bibliographic information, a link to the original publication on <http://publichealth.jmir.org>, as well as this copyright and license information must be included.