



Published in final edited form as:

J Adolesc Health. 2019 July ; 65(1): 94–100. doi:10.1016/j.jadohealth.2018.12.025.

Temporal and Geographic Patterns of Social Media Posts About an Emerging Suicide Game

Steven A. Sumner, M.D., M.Sc.^{a,*}, Stacey Galik, M.L.I.S.^b, Jennifer Mathieu, Ph.D.^b, Megan Ward, Ph.D.^b, Thomas Kiley, M.S.^b, Brad Bartholow, Ph.D.^a, Alison Dingwall, Ph.D.^b, Peter Mork, Ph.D.^b

^aDivision of Violence Prevention, National Center for Injury Prevention and Control, U.S. Centers for Disease Control and Prevention (CDC), Atlanta, Georgia

^bThe MITRE Corporation, McLean, Virginia

Abstract

Purpose: Rates of suicide are increasing rapidly among youth. Social media messages and online games promoting suicide are a concern for parents and clinicians. We examined the timing and location of social media posts about one alleged youth suicide game to better understand the degree to which social media data can provide earlier public health awareness.

Methods: We conducted a search of all public social media posts and news articles on the Blue Whale Challenge (BWC), an alleged suicide game, from January 1, 2013, through June 30, 2017. Data were retrieved through multiple keyword search; sources included social media platforms Twitter, YouTube, Reddit, Tumblr, as well as blogs, forums, and news articles. Posts were classified into three categories: individual “pro”-BWC posts (support for game), individual “anti”-BWC posts (opposition to game), and media reports. Timing and location of posts were assessed.

Results: Overall, 95,555 social media posts and articles about the BWC were collected. In total, over one-quarter (28.3%) were “pro”-BWC. The first U.S. news article related to the BWC was published approximately 4 months after the first English language U.S. social media post about the BWC and 9 months after the first U.S. social media post in any language. By the close of the study period, “pro”-BWC posts had spread to 127 countries.

Conclusions: Novel online risks to mental health, such as prosuicide games or messages, can spread rapidly and globally. Better understanding social media and Web data may allow for detection of such threats earlier than is currently possible.

Keywords

suicide; self-harm; mental health; social media; online

*Address correspondence to: Steven A. Sumner, M.D., M.Sc., Centers for Disease Control and Prevention, National Center for Injury Prevention and Control, Division of Violence Prevention, 4770 Buford Highway, NE, Mailstop F-63, Atlanta, GA 30341. hvo5@cdc.gov (S.A. Sumner).

Conflicts of interest: The authors have no conflicts of interest relevant to this article to disclose.

Publisher's Disclaimer: The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

Suicide and nonfatal self-harm are rapidly growing problems among adolescents and young adults. Over the past decade, from 2006 to 2016, rates of suicide among youth aged 10e24 years in the U.S. have increased by 39.1% [1]. Furthermore, rates of emergency department (ED) visits for nonfatal self-inflicted injuries in the U.S. have also increased precipitously. A recent study assessing data from 2001 to 2015 demonstrated that ED visits for self-harm increased 50.6% over this period, with significant increases from 2008 forward [2]. Concerningly, the rise in ED visits for self-harm was greatest among the youngest femalesdfor example, from 2009 to 2015, ED visits for self-harm among females aged 10e14 years increased over 189% from 109.8 to 317.7 per 100,000 population [2]. Although these studies cannot distinguish between self-harm with suicidal intent and nonsuicidal self-harm, such findings indicate that injurious behavior is nonetheless increasing rapidly. Beyond the U.S., suicide also remains a global problem. Although the most recently published cross-national trend analyses do not include data past 2012, youth suicide rates in nearly half of high-income nations are higher than rates in the U.S. [3].

The causes of elevated rates of suicide among youth are complex; however, a growing body of research is exploring novel risk factors for suicidal behaviors, including exposure to harmful content online. At a macro-level, cross-national analyses have linked the prevalence of Internet usage to suicide rates [4], and survey data at the individual-level have detected associations between electronic device and social media use with depressive symptoms and suicide-related outcomes [5]. Although further research is needed to establish causal associations and it remains unknown which types of online content may be harmful and the magnitude of such an effect, one particular mechanism for such associations is supported by Social Cognitive Theory. Social Cognitive Theory establishes that behavior is learned through a social context or social environment, involving observation and interaction with others; suicide-related behaviors are one such health behavior that is influenced by social learning from others[6]. Indeed, attention to social learning is particularly important in adolescent populations as research has indicated that brain development in this stage of life confers a heightened sensitization to social evaluation and peer norms [7]. Unfortunately, suicide-related behaviors can spread among youth and be difficult for clinicians and public health practitioners to detect before morbidity and mortality occurs. For example, contagion or clustering of suicide in school settings is a well-documented phenomenon that results in frequent public health agency field investigations [8]. Related to this phenomenon, health practitioners have expressed concerns about online content that may promote imitative self-harming behavior among vulnerable youth [9,10] or online groups where self-directed violence is encouraged [11,12].

Suicide games are one potential mental health threat that can spread, be challenging to identify early, and need additional study [13,14]. Early research using traditional epidemiologic methods such as death reports and surveys have examined fatalities resulting from youth games related to self-asphyxiation, establishing some degree of population health risk from such games [15,16]. Suicide games may emerge and evolve rapidly and are difficult to study using available clinical data sources and tools.

Although a growing body of literature by mental health researchers has studied aspects of self-harmerelated content online [17–23], fundamental challenges to public health practice

and public health surveillance remain to be solved. Specifically, health authorities are often only alerted to some emerging health threats from media reports. Unfortunately, by this time, significant morbidity or mortality may have occurred. If emerging risks or threats to health could be identified sooner, enhanced opportunities to prevent injury and death could occur.

Thus, in this research, we investigate whether novel data sources, such as social media and Web data, can potentially shed valuable light on emerging health threats before awareness by traditional media. In this study, we examine in detail an alleged game encouraging self-harm and suicide, known as the Blue Whale Challenge (BWC). In brief, the BWC purports to encourage youth to participate in a series of tasks or challenges over a 50-day period that gradually increase in their self-harming nature, culminating in a final suicide challenge [24]. It is important to note that the risk for suicidal behavior associated with the BWC is currently not known, and causal studies are lacking; reports of harm from the BWC currently exist largely from statements by family members [25], and systematic epidemiologic study is needed. However, we selected the BWC for examination as it exemplified a potential threat to health which ultimately received widespread attention. Although it is difficult to understand the scope, authenticity, and impact of many such emerging threats as they cannot be easily studied using traditional health data, a study of the BWC allows for an examination of a potential threat to health that generated significant concern, if not least from a proliferation of self-harm-related content. Our specific research questions and aims in this study were to (1) describe how much earlier social media posts about the BWC emerged before identification of the game by media reports to understand the potential additional time that could be afforded for prevention response through earlier detection, (2) characterize whether such social media posts about the BWC were indicating support or opposition for the game to better understand the attitude and nature of sentinel posts about this health threat, and (3) explore the geographic origin of posts about the BWC and the geographic spread of such posts to better understand the reach of such content.

Methods

Data sources

We conducted a retrospective review of publically available social media and Web posts on the BWC from January 1, 2013, through June 30, 2017. This period was selected as June 30, 2017, was approximately the date of initiation of this retrospective study. January 1, 2013, was selected to allow for over 4 years of review, a period of time hypothesized to be sufficient to detect early emergence of the game. All posts examined were fully public facing, and all information was available to any Internet user. Data sources included social media platforms with fully publicly available postings and comprised Twitter, YouTube, Reddit, and Tumblr as well as public blogs and forums. Nonpublic data were not included. The estimated prevalence of usage for each platform by U.S. teens aged 13-17 years is 85%, 32%, 9%, and 7% for YouTube, Twitter, Tumblr, and Reddit, respectively [26]. Social media data across these sites were queried using Crimson Hexagon, a leading commercially available social media analytics tool used widely in such research [27,28]. In addition to social media platforms, Crimson Hexagon indexes news media.

Search strategy

To collect relevant postings, we created a search strategy of keywords or phrases that are consistent with BWC terminology. An initial set of keywords and phrases was identified through a review of news articles about the BWC. The search strategy was refined based on social media post results. Using an iterative process, we added new terminology and related hashtags. Terms ultimately incorporated in the search strategy included greater than 20 keywords/phrases used by participants to identify related messages. Various exclusion terms were included to remove off-topic posts. Keywords are not published herein to minimize the risk of harm; further details on search strategy are available upon request from the authors. Social media posts are defined as any individual message that contained one of the keywords; thus, a post could be either an original message that mentioned the BWC or a reply/comment on another message provided the reply/comment contained one of the BWC keywords.

Categories of interest

All BWC posts retrieved from the keyword search strategy were then classified into four main categories of interest using human annotation and Crimson Hexagon's built-in, proprietary machine learning classifier. These categories of interest were defined as (1) media warnings and reports, (2) user postings that were "pro"-BWC, (3) user postings that were "anti"-BWC, and (4) off-topic posts that were then discarded from analysis. Media warnings and reports were defined as messages coming from formal news media outlets and government entities, including international news organizations. An example message includes "CBSNEWS reports 'Blue Whale Challenge' urges young people to end their lives." "Pro"-BWC posts were defined as posts with individual expressions of curiosity or interest in the game, suggesting possible openness and risk of participation or posts explicitly mentioning participation or encouraging others people's participation. For example, a "pro"-BWC post (edited to provide anonymity) states "I'm ready to do it #blue-whalechallenge." "Anti"-BWC posts were defined as messages originating from individual users that warned others about the game or attempted to share protective or preventive information; such posts could include sharing a news article but the warning originated from an individual user. An example "anti"-BWC post is "Stay safe and PLEASE don't play #bluewhalegame." The categories of "pro-" and "anti"-BWC posts were selected to distinguish the attitude of user-level posts about the BWC. Specifically, we sought to understand whether early posts about suicide games mostly comprise people interested in or participating in the game or represent comments from a community of users who are trying to warn others or share preventive messages. These categories may be useful in informing prevention response as it indicates the status of community norms around the game during early emergence.

The Crimson Hexagon machine learning classifier was initially trained on 155 posts that were manually annotated by two human reviewers who achieved nearly perfect interrater agreement ($Kappa = .96$) owing to the clear distinction between post types (i.e., posts in support of the game use language that is distinctly different from posts opposing the game). Once the algorithm was trained and deployed to classify all posts in the study, validation of the classification algorithm was performed in which analysts manually coded a random

sample of 200 posts that the algorithm had categorized, revealing an 86% accuracy of the classifier ($Kappa = .78$), which compares favorably to related research [29].

Data analysis and visualization

Posts about the BWC were then analyzed to examine the platform of origin, change in volume over time, and language. In addition, where possible, the country of origination of the post was assessed; approximately one third of posts (32,371) were able to be geo-located to a country by Crimson Hexagon using geospatial tags present or location information provided in account/post information. This was the only variable for which missing data were present; all posts included text and date/time information. Data were examined to identify the first public post from an individual user about the BWC as well as the first post from a media outlet from the top 15 countries generating posts about the BWC. Our analysis of retrospective publicly available media postings was granted Institutional Review Board exemption.

Results

From January 1, 2013, through June 30, 2017, 95,555 social media and Web posts about the BWC were collected. Most posts were from Twitter (50.5%), Tumblr (28.0%), and YouTube (12.3%; Table 1). Approximately, 53% of posts were estimated to be from females, and 47% from male users by the Crimson Hexagon software.

The first objective of this research was to describe how much earlier social media posts about the BWC emerged before identification of the game by media reports to better understand potential lead time for prevention activities afforded by early detection using social media data. To answer this question, Figure 1 displays a global timeline of the dates of the first social media post about the BWC by country, for each of the 15 countries having the most posts about the BWC. Additional events are provided for the U.S. on the timeline. The first U.S. social media post occurred on December 8, 2015, on the platform Twitter and was authored in Russian (although search terms for the study were largely in English, posts in other languages could be retrieved as not all hashtags were necessarily language-specific, such as #F57). The first U.S. social media post authored in English was detected on May 17, 2016, from a messaging forum known as 4chan. The first U.S. news article related to the BWC was published on September 13, 2016, nearly 4 months after the first U.S. social media post authored in English and more than 9 months after the very first U.S. social media post about the BWC. From a global perspective, we note that the first observed social media message about the BWC was posted in Russia on March 1, 2014. Subsequently, messages about the BWC next appear in Indonesia, the U.S., France, Italy, Kenya, and Ukraine.

The second main aim of our research was to characterize whether social media posts about the BWC were indicating support or opposition for the game to better understand the attitude and nature of sentinel posts about this health threat. Globally, of the 95,555 total posts retrieved, 27,032 posts (28.3%) were classified as posts from users who were “pro-”BWC, expressing interest, curiosity, or participation in the game. Overall, 49,670 posts (52.0%) were classified as posts from users who were “anti-”BWC, warning against

participation in the game. Finally, 18,853 posts (19.7%) were classified as news articles or from media sources.

Figure 2 depicts a timeline of posts by source from January 1, 2017, through June 30, 2017, the period of time when posts about the BWC become relatively frequent. User posts on social media that were “pro-”BWC were more frequent than posts of any other type in the early part of the timeline. In early March 2017, user posts against the BWC grow in frequency after a spike in user posts “pro-”BWC. Around this same period, media reports begin to appear at a relatively low level and do not surge until May 2017 when there is a large amount of content about the BWC being discussed.

Our third research aim was to explore the geographic origin of posts about the BWC and the geographic spread of such posts to better understand the reach of such content. Figure 3 displays the global spread of the BWC by plotting the location of “pro-”BWC posts over time. Over the study period, posts reflecting promotion or participation in the game spread from a single country to 127 countries in a relatively short time span.

Discussion

This study aimed to explore the utility of examining social media and Web data to provide early insights into potentially emerging risks for suicide. We detected that social media posts about the BWC preceded identification by news media by several months and that posts supporting participation in the BWC spread globally, suggesting utility in the analysis of social media data to enable more rapid public health responses to emerging online risks.

There were several noteworthy findings from this examination of online data. First, traditional public health surveillance for trends and risk factors related to mental health generally does not use real-time data. Indeed, official public health surveillance reports generated from death, ED, or survey data are often delayed by 1–2 years, limiting rapid prevention responses. Emerging trends, which are identified through media reports, may represent a more timely identification; however, health threats may be ongoing for some time before they rise to attention in the media. Thus, our findings that social media data can potentially assist identification of emerging health threats several months earlier are notable for public health agencies.

Early detection may offer opportunities for health professionals to focus more on prevention activities rather than reaction to a widespread problem. Prior research on social media has indicated that there are robust online communities where youth share and occasionally promote self-harm content [17,30]. Attempts to moderate or block such content to protect youth is easily subverted through the creation of new, related keywords to help such groups re-establish contact and sharing online [31], suggesting that the only feasible and effective approach may be to better identify emerging mental health and suicide-related risks and develop more effective, online prevention strategies. Although further research is needed, such approaches could include the delivery of programs such as cognitive behavioral therapy or skill-based training through online modalities [32] or the use of strategies which promote positive peer support [33]. The “anti-”BWC posts detected in our study also offer ideas for

improved public health response. Early research has identified the types of messages that may be most protective against suicidality [34], and public health professionals can work to disseminate best practices in messaging to enable bystanders to most effectively address harmful norms. As Social Cognitive Theory suggests, although harmful beliefs and behaviors related to suicide may be learned in online communities [11], such communities also have potential through social learning to increase adoption of positive attitudes and behaviors, such as the promotion of help-seeking, stigma reduction, or evidence-based treatment [35].

The global spread of the BWC that we detected further underscores the interconnectedness of individuals. Threats related to self-harm and suicide can spread rapidly across nations, just as emerging infectious diseases can rapidly cross borders. Indeed, while studying the global nature of infectious diseases, some researchers have studied spread in awareness, fear, and apprehension about infectious diseases using social media, indicating that such data hold utility in more fully assessing threats to population level health [36].

Finally, given the importance of early detection of online risks as identified in this research, one critical question is how such capabilities may be possible, particularly when new threats may not be initially known by name. Although research in the field of emerging risk detection is still developing, identification of new methods and precipitants of self-harm may be possible through studying changes in co-occurrence of new terminology with traditional self-harm phrases or images of self-harm over time. Advanced methods, such as assessing contextual or semantic similarity of emerging words through complex natural language processing algorithms, have also been demonstrated as promising in other areas of injury prevention [37].

Some important limitations of this research should be noted. First, given the retrospective nature of the study, some online content that was originally authored may no longer be present and retrievable if it was deleted or otherwise not able to be indexed. This limitation would most likely affect posts indicating support or participation in the BWC. Second, it is also important to note that our search strategy focused on social media platforms with fully publicly available data. Hence, not all platforms (such as Facebook, Instagram, or the Russian platform VK) were included in this study. It is possible that including additional platforms would detect an even larger volume and earlier occurrence of social media posts about the BWC. Temporal information in this study should be viewed as best available estimates only. Third, although media reports have linked the BWC to morbidity and mortality experienced by youth, it is impossible to fully assess the veracity, impact, and extent of the BWC, as traditional public health data sources do not gather the requisite information. Furthermore, the alleged online game operates through encouraging suicide-related behaviors, which may not be specific to the game. This underscores the importance of assessment of new data sources to try to examine such events. Nonetheless, even if the extent and nature of the BWC cannot be known, this analysis holds merit as Internet phenomena, however real, can lead to actual health consequences by stimulating a proliferation of self-harm related messages and encouraging imitative behavior among vulnerable populations. Also, although this study takes a stated focus on youth, the actual ages of users posting about the BWC cannot be known. Although surveys indicate that social

media usage among individuals aged 13–17 years and 18–24 years may be several fold higher than older age groups, it is likely that some older individuals are among those who are affected, and additional study on risk and behaviors by age is needed [26]. Fourth, we note that successful geolocation of posts to a country, which was performed by the Crimson Hexagon software, relied on the presence of geo-enabled features on a social media post or location description in profile information. Thus, not all posts were able to be geo-coded, which may affect results to an unknown degree. In addition, although this study classified individual-level posts about the BWC, we were not able to assess for whether such posts appeared to originate from automated sources, which are a growing concern with regard to health content and remains an area for future research [38]. And finally, although this study suggests promise in use of social media data for improved public health monitoring, significant future research is needed to operationalize and automate such a detection system for prospective use that does not rely on prespecified keywords. Furthermore, the accuracy of such a system for detecting and classifying posts into levels of severity of messages is not known and remains to be assessed.

The implications of this study are considerable; we demonstrate the global and distributed nature of potential online threats to mental health, which merits increased attention in light of rapidly rising suicide and self-harm rates [2]. Furthermore, our findings reveal that early detection of suicide-related and mental health threats may be possible using social media and Web data, and further research to best enable these capabilities is needed. In fact, this study underscores the importance of these data sources as key elements of forward-thinking public health practice as such novel threats to mental health simply cannot be examined using only traditional data. Improved attention to innovative approaches to identify health threats may play an important role in reversing recent increases in self-harm and suicide.

Funding Source

This study was funded by the Centers for Disease Control and Prevention.

References

- [1]. Centers for Disease Control and Prevention, National Center for Injury Prevention and Control. Web-based injury statistics query and reporting system, <http://www.cdc.gov/injury/wisqars/index.html>. 2017 Accessed July 16, 2018.
- [2]. Mercado MC, Holland K, Leemis RW, et al. Trends in emergency department visits for nonfatal self-inflicted injuries among youth aged 10 to 24 years in the United States, 2001–2015. *JAMA* 2017;318:1931–3. [PubMed: 29164246]
- [3]. Roh B-R, Jung EH, Hong HJ. A comparative study of suicide rates among 10–19-year-olds in 29 OECD countries. *Psychiatry Invest* 2018;15:376.
- [4]. Shah A The relationship between general population suicide rates and the Internet: A cross-national study. *Suicide Life Threat Behav* 2010;40:146–50. [PubMed: 20465349]
- [5]. Twenge JM, Joiner TE, Rogers ML, Martin GN. Increases in depressive symptoms, suicide-related outcomes, and suicide rates among US adolescents after 2010 and links to increased new media screen time. *Clin Psychol Sci* 2018;6:3–17.
- [6]. Lester D Suicide as a learned behavior. Springfield, IL: Charles C Thomas, Publisher; 1987.
- [7]. Somerville LH. The teenage brain: Sensitivity to social evaluation. *Curr Dir Psychol Sci* 2013;22:121–7. [PubMed: 24761055]

- [8]. Davidson LE, Rosenberg ML, Mercy JA, et al. An epidemiologic study of risk factors in two teenage suicide clusters. *JAMA* 1989;262:2687–92. [PubMed: 2810601]
- [9]. Ayers JW, Althouse BM, Leas EC, et al. Internet searches for suicide following the release of 13 Reasons Why. *JAMA Intern Med* 2017;177: 1527–9. [PubMed: 28759671]
- [10]. O'Brien KHM, Knight JR, Harris SK. A call for social responsibility and suicide risk screening, prevention, and early intervention following the release of the Netflix series 13 Reasons Why. *JAMA Intern Med* 2017;177: 1418–9. [PubMed: 28759674]
- [11]. Colombo GB, Burnap P, Hodorog A, Scourfield J. Analysing the connectivity and communication of suicidal users on twitter. *Computer Commun* 2016; 73:291–300.
- [12]. Daine K, Hawton K, Singaravelu V, et al. The power of the web: A systematic review of studies of the influence of the internet on self-harm and suicide in young people. *PLoS One* 2013;8:e77555. [PubMed: 24204868]
- [13]. Egge MK, Berkowitz CD, Toms C, Sathyavagiswaran L. The choking game: A cause of unintentional strangulation. *Pediatr Emerg Care* 2010;26:206–8. [PubMed: 20216282]
- [14]. Linkletter M, Gordon K, Dooley J. The choking game and YouTube: A dangerous combination. *Clin Pediatr* 2010;49:274–9.
- [15]. Centers for Disease Control Prevention. Unintentional strangulation deaths from the “choking game” among youths aged 6–19 years—United States, 1995–2007. *MMWR: Morbidity mortality weekly Rep* 2008;57: 141–4.
- [16]. Centers for Disease Control Prevention. “Choking game” awareness and participation among 8th graders—Oregon, 2008. *MMWR: Morbidity mortality weekly Rep* 2010;59:1–5.
- [17]. Lewis SP, Heath NL, St Denis JM, Noble R. The scope of nonsuicidal self-injury on YouTube. *Pediatrics* 2011;127:e552–7. [PubMed: 21339269]
- [18]. Cash SJ, Thelwall M, Peck SN, et al. Adolescent suicide statements on MySpace. *Cyberpsychol Behav Soc Netw* 2013;16:166–74. [PubMed: 23374167]
- [19]. Luxton DD, June JD, Fairall JM. Social media and suicide: A public health perspective. *Am J Public Health* 2012;102:S195–200. [PubMed: 22401525]
- [20]. Dyson MP, Hartling L, Shulhan J, et al. A systematic review of social media use to discuss and view deliberate self-harm acts. *PLoS One* 2016;11: e0155813. [PubMed: 27191728]
- [21]. Birnbaum ML, Ernala SK, Rizvi AF, et al. A collaborative approach to identifying social media markers of schizophrenia by employing machine learning and clinical appraisals. *J Med Internet Res* 2017;19:e289. [PubMed: 28807891]
- [22]. Jashinsky J, Burton SH, Hanson CL, et al. Tracking suicide risk factors through Twitter in the US. *Crisis* 2014;35:51–9. [PubMed: 24121153]
- [23]. De Choudhury M, Counts S, Horvitz E. Social media as a measurement tool of depression in populations. *Proceedings of the 5th Annual ACM Web Science Conference Paris, France, May 2–4, 2013*:47–56.
- [24]. Sousa D, Filho J, Bezerra Cavalcanti R, et al. The impact of the ‘Blue Whale game in the rates of suicide: Short psychological analysis of the phenomenon. *Int J Soc Psychiatry* 2017;63:796–7. [PubMed: 28936914]
- [25]. Timm-Garcia J, Hartung K. Family finds clues to teen’s suicide in blue whale paintings. CNN. Available at: <https://www.cnn.com/2017/07/17/health/blue-whale-suicide-game/index.html>. Accessed November 21, 2018.
- [26]. Anderson M, Jiang J. Teens, Social media & technology 2018. Pew Research Center - Internet and Technology Available at: <http://www.pewinternet.org/2018/05/31/teens-technology-appendix-a-detailed-tables/>. Accessed November 21, 2018.
- [27]. McClellan C, Ali MM, Mutter R, et al. Using social media to monitor mental health discussions evidence from Twitter. *J Am Med Inform Assoc* 2017; 24:496–502. [PubMed: 27707822]
- [28]. Crimson Hexagon. Available at: <https://www.crimsonhexagon.com/>. Accessed July 9, 2018.
- [29]. O’Dea B, Wan S, Batterham PJ, et al. Detecting suicidality on Twitter. *Internet Interventions* 2015;2:183–8.

- [30]. Lewis SP, Heath NL, Sornberger MJ, Arbuthnott AE. Helpful or harmful? An examination of viewers' responses to nonsuicidal self-injury videos on YouTube. *J Adolesc Health* 2012;51:380–5. [PubMed: 22999839]
- [31]. Chancellor S, Pater JA, Clear T, et al. #thyghgapp: Instagram content moderation and lexical variation in pro-eating disorder communities. Paper Presented at: Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing San Francisco, February 27–March 02, 2016.
- [32]. Clarke AM, Kuosmanen T, Barry MM. A systematic review of online youth mental health promotion and prevention interventions. *J Youth Adolesc* 2015;44:90–113. [PubMed: 25115460]
- [33]. Ali K, Farrer L, Gulliver A, Griffiths KM. Online peer-to-peer support for young people with mental health problems: A systematic review. *JMIR Ment Health* 2015;19:e19.
- [34]. De Choudhury M, Kiciman E. The language of social support in social media and its effect on suicidal ideation risk. Paper Presented at: Proceedings of the International AAAI Conference on Weblogs and Social Media San Jose, California, May 7–12, 2016.
- [35]. Stone DM, Holland KM, Bartholow BN, et al. Preventing suicide: A technical package of policies, programs, and practice. Atlanta: Centers for Disease Control and Prevention; 2017.
- [36]. Fung IC-H, Tse ZTH, Cheung C-N, et al. Ebola and the social media. *The Lancet* 2014;384(9961):2207.
- [37]. Ding T, Roy A, Chen Z, et al. Analyzing and retrieving illicit drug-related posts from social media. Paper Presented at: Bioinformatics and Biomedicine (BIBM), 2016 IEEE International Conference Shenzhen, China, December 15–18, 2016.
- [38]. Broniatowski DA, Jamison AM, Qi S, et al. Weaponized health communication: Twitter bots and Russian trolls amplify the vaccine debate. *Am J Public Health* 2018;0:e1–7.

IMPLICATIONS AND CONTRIBUTION

Robust posts about online suicide games are present before identification of such risks by traditional news sources, and such games can spread rapidly and globally. Improved study of social media and Web data may offer opportunities for early detection and prevention of youth mental health risks.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

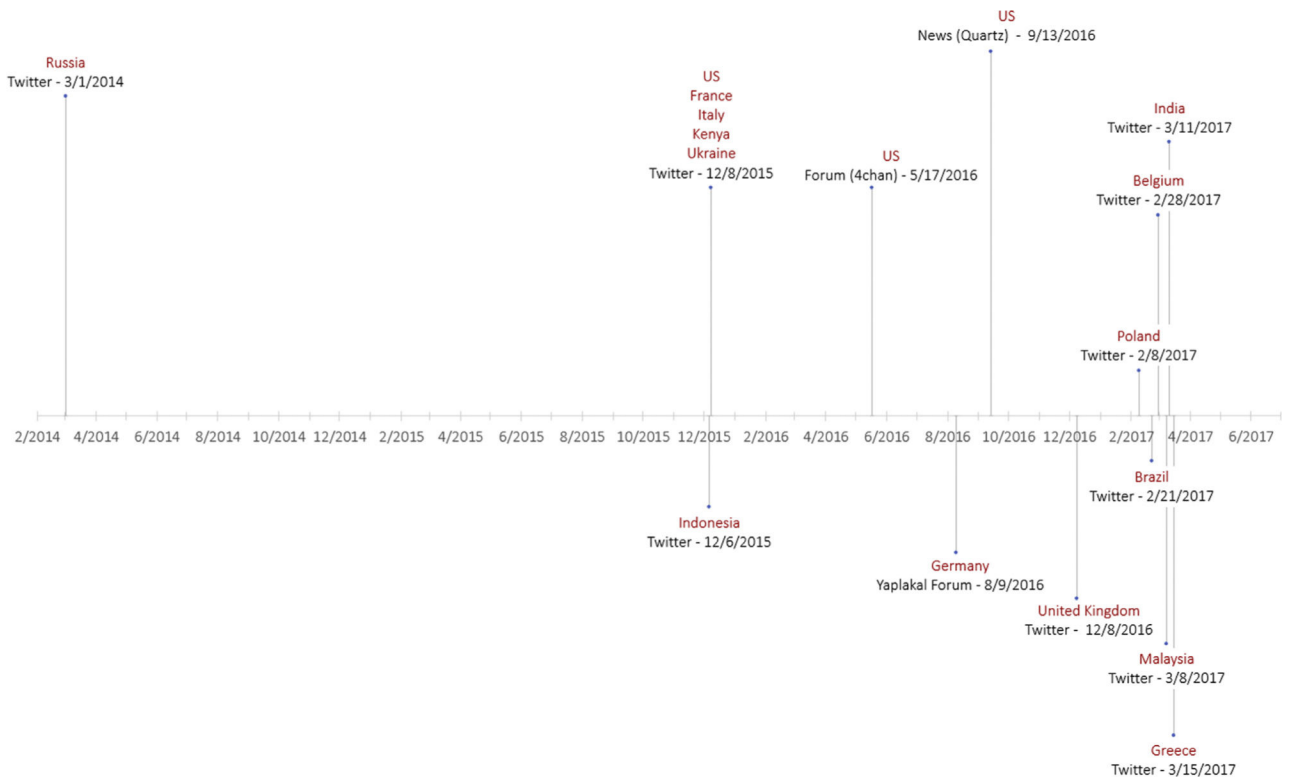


Figure 1. Timeline of first social media posts about Blue Whale Challenge by country. For the U.S., the timeline displays the date of first U.S. social media posting in any language (December 8, 2015), the first U.S. social media posting in English (May 17, 2016), and the first U.S. news article about Blue Whale Challenge (September 13, 2016) for context.

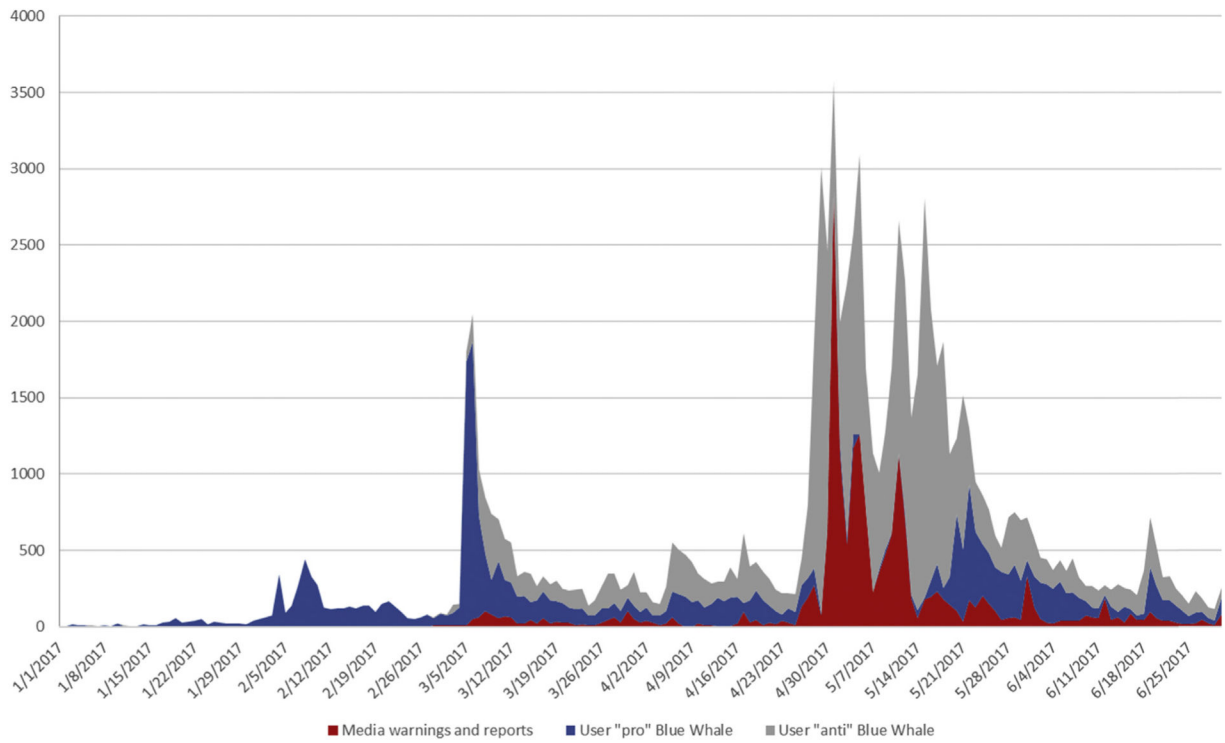


Figure 2.
Number of posts about Blue Whale Challenge by type and date, January 1, 2017, through June 30, 2017. Graph shown is a stacked area chart all categories are visible if present during a given period.

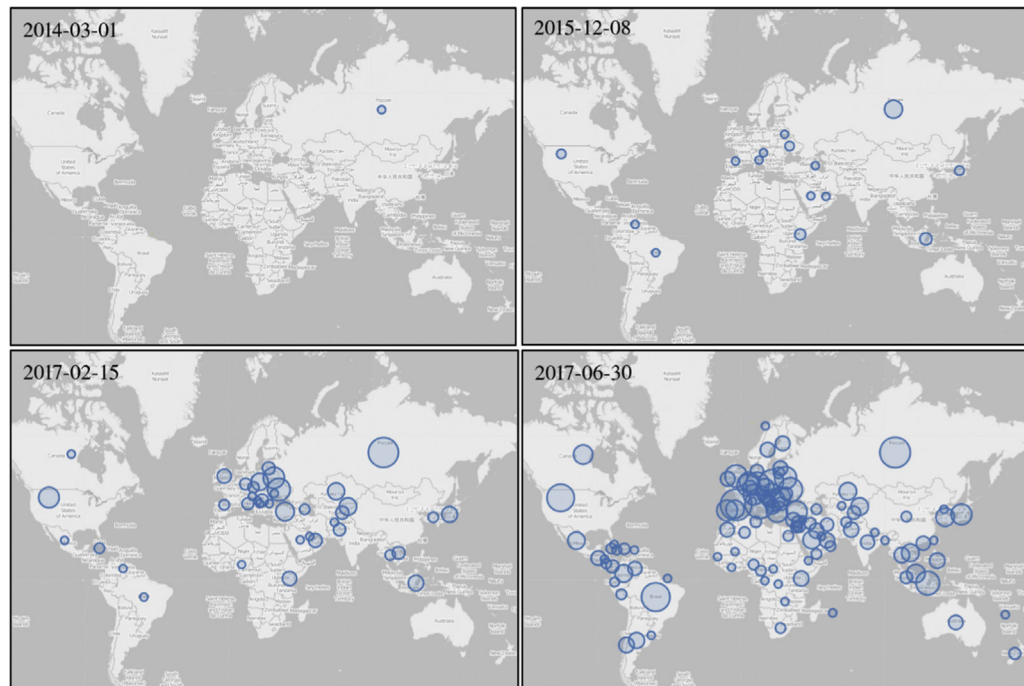


Figure 3. Global spread of “pro” Blue Whale Challenge posts over time and by country of the world. This figure shows the global spread of user-level posts that are “pro” Blue Whale Challenge over time. Panels show various time points throughout the study period and are meant to help visualize growth in the global volume in messages. Circle sizes at each time point represent the cumulative volume of “pro” Blue Whale posts up to that time point from the beginning of the study period (January 1, 2013) and are scaled proportional to the number of posts per country.

Table 1

Social media, Web, and news postings about the Blue Whale Challenge, January 1, 2013, through June 30, 2017

Source	N	Percentage of total
Twitter	48,256	50.5
Tumblr	26,736	28.0
YouTube	11,799	12.3
News	5,807	6.1
Forums	1,849	1.9
Blogs	821	0.9
Reddit	264	0.3
Other	23	0.0
Total	95,555	100%

Other category includes a small number of comments posted on blogs or news articles.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript