HHS Public Access

Author manuscript

Disaster Med Public Health Prep. Author manuscript; available in PMC 2022 April 01.

Published in final edited form as:

Disaster Med Public Health Prep. 2021 October ; 15(5): 568–572. doi:10.1017/dmp.2020.65.

Using Twitter to Track Unplanned School Closures: Georgia Public Schools, 2015–17

Jennifer O. Ahweyevu, MPH[#], **Ngozi P. Chukwudebe, MPH**[#], **Brittany M. Buchanan, MPH**, **Jingjing Yin, PhD**

Department of Biostatistics, Epidemiology and Environmental Health Sciences, Jiann-Ping Hsu College of Public Health, Georgia Southern University, Statesboro, Georgia

Bishwa B. Adhikari, PhD,

Health Economics and Modeling Unit, Division of Preparedness and Emerging Infections, National Center for Emerging and Zoonotic Infectious Diseases, Centers for Disease Control and Prevention, Atlanta, Georgia

Xiaolu Zhou, PhD,

Department of Geology and Geography, College of Science and Mathematics, Georgia Southern University, Statesboro, Georgia; Department of Geography, Texas Christian University, Fort Worth, Texas

Zion Tsz Ho Tse, PhD,

School of Electrical and Computer Engineering, College of Engineering, The University of Georgia, Athens, Georgia; Department of Electronic Engineering, University of York, Heslington, York, United Kingdom

Gerardo Chowell, PhD,

Department of Population Health Sciences, School of Public Health, Georgia State University, Atlanta, Georgia

Martin I. Meltzer, PhD,

Health Economics and Modeling Unit, Division of Preparedness and Emerging Infections, National Center for Emerging and Zoonotic Infectious Diseases, Centers for Disease Control and Prevention, Atlanta, Georgia

Isaac Chun-Hai Fung, PhD

Department of Biostatistics, Epidemiology and Environmental Health Sciences, Jiann-Ping Hsu College of Public Health, Georgia Southern University, Statesboro, Georgia; Health Economics and Modeling Unit, Division of Preparedness and Emerging Infections, National Center for Emerging and Zoonotic Infectious Diseases, Centers for Disease Control and Prevention, Atlanta, Georgia

[#] These authors contributed equally to this work.

Correspondence and reprint requests to Isaac Chun-Hai Fung, Department of Biostatistics, Epidemiology & Environmental Health Sciences, Jiann-Ping Hsu College of Public Health, P.O. Box 7989, Georgia Southern University, Statesboro, GA 30460-7989 (cfung@georgiasouthern.edu).

Supplementary material

To view supplementary material for this article, please visit https://doi.org/10.1017/dmp.2020.65

Abstract

Objectives: To aid emergency response, Centers for Disease Control and Prevention (CDC) researchers monitor unplanned school closures (USCs) by conducting online systematic searches (OSS) to identify relevant publicly available reports. We examined the added utility of analyzing Twitter data to improve USC monitoring.

Methods: Georgia public school data were obtained from the National Center for Education Statistics. We identified school and district Twitter accounts with 1 or more tweets ever posted ("active"), and their USC-related tweets in the 2015–16 and 2016–17 school years. CDC researchers provided OSS-identified USC reports. Descriptive statistics, univariate, and multivariable logistic regression were computed.

Results: A majority (1,864/2,299) of Georgia public schools had, or were in a district with, active Twitter accounts in 2017. Among these schools, 638 were identified with USCs in 2015–16 (Twitter only, 222; OSS only, 2015; both, 201) and 981 in 2016–17 (Twitter only, 178; OSS only, 107; both, 696). The marginal benefit of adding Twitter as a data source was an increase in the number of schools identified with USCs by 53% (222/416) in 2015–16 and 22% (178/803) in 2016–17.

Conclusions: Policy-makers may wish to consider the potential value of incorporating Twitter into existing USC monitoring systems.

Keywords

digital health; public health surveillance; social distancing; social media

Recommendations and decisions to conduct unplanned school closures (USCs) are made by school and public health officials based on available information regarding the threats to community and student safety.[1] Reasons for such closures include severe weather, natural disasters, and facility problems.[2] For example, preemptive school closures and dismissals may be recommended as community-level, nonpharmaceutical interventions to slow influenza pandemic transmission in the community.[3] The rationale for such preemptive closing is that schools serve as amplification points for influenza transmission, as children exhibit higher transmission rates than adults.[4,5] Therefore, USC monitoring helps track outbreak response and control, and inform future pandemic preparedness efforts.[2,6]

Representativeness, accuracy, cost-effectiveness, and timeliness are important attributes to consider in developing and evaluating public health surveillance systems.[7,8] With the increasing popularity of social media,[9] community-based surveillance incorporating social media monitoring may represent access to affordable, relevant real-time data.[10] Researchers at the Centers for Disease Control and Prevention (CDC) have demonstrated the utility of "online systematic searches" (OSS) for USC monitoring. This strategy consists of systematic scans and searches of online media and data, using Google Alert, Google News, and LexisNexis[2] (Supplemental Figure S1). In addition, a study on USCs in Michigan, demonstrated how monitoring posts on Twitter can enhance USC surveillance[11] (Supplemental Figure S2). We expand the literature on the potential use of Twitter to monitor USCs by reporting on a comparison of the coverage of USCs in Georgia public

schools. For this purpose, we compared tweets and OSS-based reports of USCs for 2015–16 and 2016–17 school years.

METHODS

The list of 2299 Georgia public schools and 232 school districts were downloaded from the National Center for Education Statistics (NCES) website[12] in August 2017. Based on the NCES data and using the Google and Twitter search functions with additional checking of school websites, schools with a Twitter account were manually identified in August 2017. The same procedure was applied to the list of school districts to identify district-owned Twitter accounts. Only official school and school district Twitter accounts were included. All publicly available tweets posted by these accounts were downloaded using the Twitter Search Application Programming Interface (API) on October 9, 2017. Schools that had ever posted a tweet on their page or were under school districts that had ever posted a tweet on their page, were coded as having active Twitter coverage (Supplemental Figure S3).

Using a list of keywords (provided by CDC researchers as in their OSS method) to identify tweets likely to include information on USC, all likely tweets on school closures were filtered from the tweets downloaded, and then stratified by school year. Relevant tweets were then manually coded to identify USCs and type of closure (school or school district closure). All duplicates were identified and eliminated manually. Twitter USC announcements were then matched to all the schools in the NCES data.

CDC researchers provided OSS-acquired USC data. (See Supplemental Materials for details.) We filtered the data to retrieve USC events announced by Georgia public schools and school districts, from August 2015 to June 2017. USC data obtained from OSS and Twitter were compared with individual school as a unit of analysis. We compared the frequency of schools with 1 or more USC announcements in a year to schools without USC announcements.

Statistical Analysis

Statistical analyses were performed in R version 3.4.3, by means of RStudio version 1.1.383. Statistical significance was determined a priori at $\alpha = 0.05$. To compare USCs determined by Twitter with that by OSS, some characteristics (such as school locale, total student population, student-teacher ratio, reduced price/free lunch proportion) of schools with at least 1 USC determined by Twitter and OSS, respectively, were summarized by frequencies if the character variable is categorical (such as school locale) or by sample means if the character variable is continuous (such as student population). These descriptive statistics were then compared using Chi-square test for frequencies and Welch 2-sample t-test for sample means. Univariate and multivariable logistic regression were performed to determine whether certain school characteristics/factors were associated with USCs announced by Twitter and by OSS, respectively. We used city as the reference category for the school locale variable, facilitating comparison with our results from Michigan.[11] Using the R package “sjstat”, the crude and adjusted odds ratios were converted to crude and adjusted relative risks (aRR).

This study was approved by the Institutional Review Board of Georgia Southern University (H15083) under the B2 exempt category, as the social media data analyzed are considered publicly observable behavior.

RESULTS

Of the 2299 Georgia public schools, 1864 (81%) schools had active Twitter coverage (627 schools had an active Twitter account, and 1237 schools were under districts with an active Twitter account). The proportion of schools with active Twitter coverage varied by geographic location: 89% (372/417) for city, 97% (830/857) for suburban, 59% (171/288) for town, and 66% (467/708) for rural, and 83% (24/29) for schools missing locale data (Supplemental Figure S4; Supplemental Table S9). Schools with active Twitter coverage had a significantly higher mean student population (806 students vs 595 students; $P < 0.0001$) and lower mean proportion of students receiving free/reduced-price lunch (66% vs 73%; $P < 0.0001$) than those without (Supplemental Table S9).

Among the 1864 Georgia public schools with active Twitter coverage, 638 (34%) and 981 (53%) with 1 or more USCs were identified in 2015–16, and 2016–17, respectively. In 2015–16, among the 638 schools with USCs, 222 were identified by Twitter alone, 215 by OSS alone, and 201 by both. In 2016–17, among the 981 schools with USCs, 178 were identified by Twitter alone, 107 by OSS alone, and 696 by both. Given that OSS is the current method adopted by CDC researchers, the marginal benefit of adding Twitter as a data source was an increase in the number of schools with USCs identified by 53% (222/416) in 2015–16 and 22% (178/803) in 2016–17 (Table 1).

Stratified by locality, in 2015–16, in cities or suburban areas, 182 schools with 1 or more USCs were identified by Twitter alone, 115 by OSS alone, and 75 by both; in town or rural areas, 40 were identified by Twitter alone, 97 by OSS alone, and 125 by both. In 2016–17, in cities or suburban areas, 111 schools with 1 or more USCs were identified by Twitter alone, 26 by OSS alone, and 458 by both; in town and rural areas, 64 were identified by Twitter alone, 81 by OSS alone, and 231 by both (Supplemental Table S19).

In 2015–16 school year, compared with city schools, suburban schools were the least likely to have an USC identified by OSS, while in 2016–17 school year, city schools were the least likely to have an USC identified by OSS (Table 2[i]). In 2015–16 school year, schools in all other locales, when compared with city schools, were more likely to have at least 1 USC announced on Twitter in 2015–16, but no differences were observed in 2016–17 (Table 2[ii]). However, additional analysis excluding tweets on USCs due to Tropical Storm Hermine (August 31 to September 2, 2016) and Hurricane Matthew (October 4 to 10, 2016), found that suburban schools were the most likely to have an USC identified by Twitter in 2016–17. (See Supplemental Tables S20 and S21.)

DISCUSSION

Among public schools in Georgia with active Twitter coverage for the 2015–16 and 2016–17 school years, using both Twitter and OSS resulted in higher identification of schools with

USCs compared with using either method alone. Furthermore, we found that school locale was an important predictor of identifying USCs by means of Twitter and OSS.

Twitter has a high penetration among Georgia public schools. A total of 81% of Georgia public schools had active Twitter coverage compared with the national average Twitter penetration of 24% among US adults.[13] Among schools with active Twitter coverage, slightly more schools with at least 1 USC announcement were identified by means of Twitter than OSS (7 more in 2015–16; 73 more in 2016–17). In 2015–16, suburban schools were the least likely to be identified by OSS in reference to city schools, but this was not the case for Twitter. This is noteworthy because most of the schools in Georgia are in the suburban locale. While this does not translate to Twitter being a better alternative, it provided additional USC data that the OSS did not pick up and vice versa.

The 2016–17 school year saw an increase in the amount of schools with at least 1 USC. When compared with the previous school year, there was a change in the relationship between the predictor variables and the ability to detect schools with at least 1 USC. There was an increase in the probability of schools with at least 1 USC in the suburban locale to be identified by means of OSS, while schools were equally as likely to be detected by Twitter regardless of locale. Based on our additional analytical findings (Supplemental Table S20 and S21), we infer that these changes were chiefly due to Tropical Storm Hermine and Hurricane Matthew that struck the southeastern United States on September 2 and October 7, 2016.[14,15] Due to the severe weather, several schools closed in preparation for the storm and to deal with the aftermath.

Spatial generalizability was assessed by comparing our results with those reported from Michigan.[11] We found that, in both Michigan and Georgia, geographical locations of schools are predictors of whether schools with USCs are identified by either method. In Michigan, city schools had the highest risk of being identified by Twitter, but they had the lowest risk of being identified in Georgia. We hypothesize that this difference could be attributed to geographic differences in the *actual* use of Twitter to announce USCs. Because the direction of association between rurality and schools with USCs being identified by means of OSS is the same in both states, we believe this difference may be related to Twitter behavior of school officials in both states. There are 2 levels to this behavior: The adoption of Twitter by schools and school districts and their use of Twitter to announce USCs. Although Georgia schools (81%) have higher Twitter coverage than schools in Michigan (58%), city and suburban schools have higher Twitter coverage than rural and town schools in both states. This rules out the adoption of Twitter as a potential cause, leaving us with the use of Twitter to announce USCs.

There are several limitations associated with our study. First, it is a cross-sectional study. Retrospective Twitter data, up to 3,200 tweets per Twitter account as per the limit of Twitter API, was retrieved in 1 day. Twitter coverage was assumed constant over the 2 school years. We did not investigate the specific date on which a school or school district joined Twitter. If some schools or school districts joined only in 2016–17, it would not have been possible for them to make Twitter USC announcements in 2015–16. Second, human errors were possible even with 2 coders reading the tweets. Finally, Twitter use is neither uniform across time

nor geography. When and where tweets are less frequent, the performance of the Twitter method may suffer. Future research can extend the analysis across more years or other states to establish a comprehensive picture on how the use of Twitter to announce USCs may vary. Future studies can also collect information on how long schools stay closed during an USC event.

In conclusion, 81% of Georgia public schools have active Twitter coverage. Through Twitter, 222 and 178 additional Georgia public schools with at least 1 USC were identified in 2015–16 and 2016–17 school years. Twitter was able to fill in a gap in OSS coverage in the suburban locale. Altogether, our study demonstrates that Twitter remains a complementary USC data source across 2 years in Georgia. If future results are consistent with our findings, Twitter has the potential to become an important supplement to OSS, creating a more robust USC database. (See discussion in Supplemental Text S3.) This would improve the ability for public health professionals to make better informed decisions. It would also help identify how well public health measures are followed during a public health emergency.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors thank Dr. Amra Uzicanin, Dr. Hongjiang Gao, Dr. Yenlik Zheteyeva, and Ashley Jackson, MPH, of the Community Interventions for Infection Control Unit, Division of Global Migration and Quarantine, National Center for Emerging Infectious Diseases, Centers for Disease Control and Prevention, for sharing their unplanned school closure online systematic search data. The authors are grateful to the late Dr. Toby Merlin of the Division of Preparedness and Emerging Infections for his leadership and institutional support. The authors thank Mr. Dharamendra Kumar of the University of Georgia for Twitter data retrieval.

Funding

Dr. Isaac Chun-Hai Fung acknowledges salary support from the National Center for Emerging and Zoonotic Infectious Diseases, Centers for Disease Control and Prevention (18IPA1808820; 19IPA1908208).

Disclaimer

The findings and conclusions of this report are those of the authors and do not necessarily represent the views of the Centers for Disease Control and Prevention or the United States Government.

REFERENCES

1. Community Preventive Services Task Force. Emergency preparedness and response: school dismissals to reduce transmission of pandemic influenza 2012. https://www.thecommunityguide.org/findings/emergency-preparedness-and-response-school-dismissals-reduce-transmission-pandemic-influenza.
2. Wong KK, Shi J, Gao H, et al. Why is school closed today? Unplanned K-12 school closures in the United States, 2011–2013. PLoS One. 2014;9(12):e113755. [PubMed: 25463353]
3. Qualls N, Levitt A, Kanade N, et al. Community mitigation guidelines to prevent pandemic influenza - United States, 2017. MMWR Recomm Rep. 2017;66(1):1–34.
4. Fowlkes A, Dasgupta S, Chao E, et al. Estimating influenza incidence and rates of influenza-like illness in the outpatient setting. Influenza Other Respir Viruses. 2013;7(5):694–700. [PubMed: 22984820]
5. Nesti MM, Goldbaum M. Infectious diseases and daycare and preschool education. J Pediatr (Rio J). 2007;83(4):299–312. [PubMed: 17632670]

6. Zheteyeva Y, Gao H, Shi J, Uzicanin A, eds. Exploring Usability of School Closure Data for Influenza-Like Illness Surveillance. ISDS Annual Conference Proceedings 2015. Online J Public Health Inform. 2016;8. 10.5210/ojphi.v8i1.6496. Accessed April 18, 2020.
7. German RR, Lee LM, Horan JM, Milstein RL, Pertowski CA, Waller MN. Updated guidelines for evaluating public health surveillance systems: recommendations from the Guidelines Working Group. MMWR Recomm Rep. 2001;50(Rr-13):1–35; quiz CE1–7.
8. Groseclose SL, Buckeridge DL. Public health surveillance systems: recent advances in their use and evaluation. Ann Rev Public Health. 2017;38(1):57–79. [PubMed: 27992726]
9. Ewaldsson U 5 trends shaping innovation in ICT. Ericsson Technology Review. 2016. https://www.ericsson.com/en/reports-and-papers/ericsson-technology-review/articles/technology-trends-2016. Accessed May 2, 2020.
10. Salathe M, Freifeld CC, Mekaru SR, Tomasulo AF, Brownstein JS. Influenza A (H7N9) and the importance of digital epidemiology. N Engl J Med. 2013;369(5):401–404. [PubMed: 23822655]
11. Jackson AM, Mullican LA, Tse ZTH, et al. Unplanned closure of public schools in Michigan, 2015–2016: cross-sectional study on rurality and digital data harvesting. J School Health. 2020: 10.1111/josh.12901.
12. National Center for Education Statistics. Data tools https://nces.ed.gov/datatools/. Accessed April 18, 2020.
13. Pew Research Center. Social Media Fact Sheet 2018 http://www.pewinternet.org/fact-sheet/social-media/. Accessed April 18, 2020.
14. Stewart SR. Hurricane Matthew. National Hurricane Center; 2017. Accessed April 18, 2020.
15. National Weather Service. Tropical Storm Hermine - September 2016 2016. https://www.weather.gov/chs/TropicalStormHermine-Sep2016. Accessed April 18, 2020.

TABLE 1

Number (%) of Public Schools With and Without USC Announcements Identified in 2015–17, Among 1864 Public Schools in Georgia With Active Twitter Coverage

No. of public schools in Georgia as of 2017		2299 (100%)
No. of public schools in Georgia with active Twitter coverage as of 2017		1864 (81.08%)
	2015–16	**2016–17**
Schools with USC announcements identified via OSS and Twitter	201 (10.78%)	696 (37.34%)
Schools with USC announcements identified via OSS only	215 (11.53%)	107 (5.74%)
Schools with USC announcements identified via Twitter only	222 (11.91%)	178 (9.55%)
Schools with USC announcements identified	638 (34.23%)	981 (52.63%)
Schools without USC announcements identified	1226 (65.77%)	883 (47.37%)
Schools with active Twitter coverage	1864 (100%)	1864 (100%)

Abbreviations: OSS, online systematic searches; USC, unplanned school closure.

TABLE 2

aRR of (i) Schools Having a USC Identified via OSS and (ii) Schools Having a USC Identified via Twitter Search, Among 1864 Georgia Public Schools With Active Twitter Coverage

(i) Schools With USC(s) Identified via OSS Among Georgia Public Schools With Active Twitter Coverage

	2015–16 School Year		2016–17 School Year	
Predictor Variable	**aRR (95% CI)**	***P*-Value**	**aRR (95% CI)**	***P*-Value**
School locale				
City	Reference		Reference	
Suburban	0.5605 (0.4170 – 0.7453)	<0.0001	1.3358 (1.1830 – 1.4839)	<0.0001
Town	1.8302 (1.4260 – 2.2634)	<0.0001	1.6218 (1.4290 – 1.7894)	<0.0001
Rural	1.4497 (1.1583 – 1.7799)	0.0017	1.3196 (1.1546 – 1.4790)	0.0002
Student population	1.0050 (0.9868 – 1.0211)	0.5522	0.9796 (0.9669 – 0.9918)	0.0014
Free/reduced price lunch proportions	1.0002 (0.9660 – 1.0356)	0.9907	1.0434 (1.0224 – 1.0646)	<0.0001

(ii) Schools With USC(s) Identified via Twitter Search Among Georgia Public Schools With Active Twitter Coverage

	2015–16 School Year		2016–17 School Year	
Predictor Variable	**aRR (95% CI)**	***P*-Value**	**Predictor Variable**	**aRR (95% CI)**
School locale				
City	Reference		Reference	
Suburban	2.4172 (1.9922 – 2.8520)	<0.0001	0.9405 (0.8081 – 1.0766)	0.3898
Town	2.7039 (2.1953 – 3.1689)	<0.0001	0.9767 (0.7881 – 1.1703)	0.8129
Rural	2.2009 (1.7522 – 2.6721)	<0.0001	0.9062 (0.7645 – 1.0534)	0.2105
Student population	1.0037 (0.9880 – 1.0188)	0.6620	0.9906 (0.9798 – 1.0004)	0.0761
Free/reduced price lunch proportions	1.0002 (0.9660 – 1.0356)	0.9907	1.0434 (1.0224 – 1.0646)	<0.0001

Abbreviations: aRR, adjusted relative risk; OSS, online systematic searches; USC, unplanned school closure.