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

**Supplementary Materials**

Jennifer O. Ahweyevu†, Ngozi P. Chukwudebe†, Brittany M. Buchanan, Jingjing Yin,

Bishwa B. Adhikari, Xiaolu Zhou, Zion Tsz Ho Tse, Gerardo Chowell, Martin I. Meltzer,

Isaac Chun-Hai Fung\*

†Co-first authors; \*Corresponding author

**Outline**

Section A. Methods

Part I: Additional Texts

Part II: 2015 -2017 Data Collection and Documentation, Code book and Frequencies

Section B. Results

Part III: Main Result Tables Texts (continued)

Part III: Main Result Texts (continued)

Part V: Additional Result Tables

Section C. Supplementary Figures

**List of Acronyms:**

CDC: The Centers for Disease Control and Prevention

CI: confidence interval

GA: the state of Georgia in the United States of America

NCES: The National Center for Education Statistics

OSS: Online Systematic Search

USC: unplanned school closure

**Section A. Methods**

**Part I: Additional Texts**

**Text S1. Details of analysis**

*Data sources*

CDC researchers track USCs lasting a day or more via OSS using Google Alert, Google News, and LexisNexis (1). A list of keywords (“close,” “closed,” “closing,” “closings,” “closure,” “closures,” “closes,” or “dismiss”) are used daily by the researchers to search for and retrieve all likely USC events that have occurred in the previous 24 hours. Relevant USC events are recorded, verified through the school or school district website or social media pages, and saved. The Georgia USC data harvested by the CDC researchers were released to the authors for this research project.

*Dependent variables*

The dependent variables, with binary response categories, were:

1. “Schools with active Twitter coverage,”
2. “Schools identified by OSS as having at least one USC event among schools with active Twitter coverage,” and
3. “Schools identified by Twitter as having at least one USC event among schools with active Twitter coverage.”

Schools whose USC events that were not extracted by the CDC researchers (or from Twitter) were either schools that did not have an USC event (i.e., the true negative cases), or schools that had an USC event but did not captured by OSS or on Twitter (i.e., the false negative cases under either or both tools).

*Independent Variables*

**Total student population and student-teacher ratio:** The total number of students and the student-teacher ratio per school were available in the public schools’ data from the NCES website. The student-teacher ratio per school was determined by dividing the total number of students by the total number of teachers.

**Proportion of students that receive free/reduced price lunch:** In this study, the proportion of students that received free/ reduced price lunch was used as a proxy for poverty. Students enrolled in Georgia public schools are either eligible for free lunch or for reduced price lunch if they belong to a household with an income of ≤ 130% of the poverty threshold or > 130% but ≤ 185% of the poverty threshold respectively (2). In Georgia, in 2016-17 school year, students from family sizes of one and eight, earning less than $15,444 and 53,157 per annum were qualified for free lunch. Students from a family sizes of one to eight, earning from $15,445 to $21,978 and $53,158 to $75,647 per annum qualified for reduced price lunch (3).

**School locale:** According to the NCES revised school locale classification system, the major locale categories and definitions include:

1. City: A territory within an urbanized area and within a principal city;
2. Suburb: A territory outside a principal city and within an urbanized area;
3. Town: A territory within an urban cluster;
4. Rural: A census-defined rural territory (4).

**Statistical analysis**

Quantile-quantile plots were used to check the data normality of all the continuous variables (student population, student-teacher ratio and student proportion receiving free/reduced price lunch). The relationship between “schools with active Twitter coverage” and the covariates were evaluated using the chi-square test for the categorical variable (school locale), and the Welch two-sample t-test for the continuous variables.

The data subset of “schools with active Twitter coverage”, stratified by school year, was evaluated to determine the relationship between “schools identified by Twitter as having at least one USC event” (and “schools identified by OSS as having at least one USC event”) and the predictor variables.

**Sensitivity Analysis**

In the absence of a gold standard USC surveillance system, the evaluation of the sensitivity of Twitter and OSS was based on the assumption that all the true USC events that would have been captured by a hypothetical gold standard system would detect all USC identified by either one of both methods (see Table A1, A2 and A3).

**Table A1.** The distribution of USC events/schools with at least one USC detected by OSS and Twitter

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | **Detected by OSS** | |
|  |  | **Yes** | **No** |
| **Detected on Twitter** | **Yes** | A | B |
| **No** | C | D |

**Table A2.** The distribution of USC events detected by a hypothetical gold standard USC surveillance system and Twitter

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | **True USC event** | |  |
|  |  | **Yes** | **No** | **Total** |
| **Detected on Twitter** | **Yes** | True positive  (A+B) | False positive  (?) | Total detected on Twitter  (?) |
| **No** | False negative  (C) | True negative  (?) | Total missed on Twitter  (?) |
| **Total** | | Total True USC events  (A+B+C) | Total False USC events  (?) | Total  (?) |

**Table A3.** The distribution of USC events detected by a hypothetical gold standard USC surveillance system and OSS.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | **True USC event** | |  |
|  |  | **Yes** | **No** | **Total** |
| **Detected by OSS** | **Yes** | True positive  (A+C) | False positive  (?) | Total detected by OSS  (?) |
| **No** | False negative  (B) | True negative  (?) | Total missed by OSS  (?) |
| **Total** | | Total True USC events  (A+B+C) | Total False USC events  (?) | Total  (?) |

In addition, USC events identified by either method of USC surveillance were assumed to be true positives. Therefore, the sensitivity of Twitter and OSS for detecting USC events/schools with at least one USC were as follows:

1. Sensitivity (Twitter) =
2. Sensitivity (OSS) =

**Maps**

Maps were created via ArcGIS 10.4 (Esri, Redlands, CA, USA) and R version 3.2.3 (R Foundation for Statistical Computing, Vienna, Austria) illustrating the spatial distribution of OSS and Twitter coverage of Georgia public schools with at least one USC (Figure 2).

**Part II:** **2015-2017 Data Collection and Documentation, Code book and Frequencies**

Manual Codebook Creator: Ngozi Pearl Chukwudebe

Lead Manual Coders: Brittany Buchanan, Jennifer O. Ahweyevu and

Ngozi P. Chukwudebe

Project PI (and liaison with CDC): Isaac Chun-Hai Fung

Project Co-investigators: Jingjing Yin, Bishwa B. Adhikari, Xiaolu Zhou, Zion Tsz Ho Tse, Gerardo Chowell, Martin I. Melzer

Adapted from: Jackson et al. (5)

**Project Objective**

The aim of the study is to examine the potential of expanding the available data by analyzing Twitter posts on unplanned closures of Georgia public schools. By comparing USC data extracted via online systematic searches (OSSs) of news and other sources and Twitter, we will be able to demonstrate the potential for an integrative unplanned school closure (USC) surveillance system that provides a more comprehensive USC database.

1. **Procedure Data Collection, Processing and Editing**

**Component description**

The authors worked with six datasets, two of which were used in the statistical analysis (**events data** and **school USC data)**.

* **School data:** Georgia public school data spreadsheet, downloaded from the NCES website (<https://nces.ed.gov/>) in August 2017 that provides information on school characteristics such as school name, district, number of students, number of students on free lunch, number of student on reduced price lunch, and etc.
* **District data:** Georgia public school district data spreadsheet, downloaded from the NCES website in August 2017, that provides information on school district characteristics such as district name, number of students etc.
* **CDC USC events data:** A spreadsheet of USC events retrieved by researchers at the CDC via OSS using Google Alert, Google News, and LexisNexis, that provides information on all unscheduled dismissals of Georgia public schools/districts.
* **Twitter USC events data:** Twitter USC events file contains information on all unscheduled dismissals of Georgia public schools announced on Twitter retrieved by the authors.
* **Events data:** USC events spreadsheet with all USC events identified by CDC and by the authors on Twitter.
* **School USC data:** Contains data on school characteristics, Twitter coverage, Twitter and CDC identified closures, USC closure type, etc.

**Phase 1**

1. Download **school data** and **district data** from the NCES Website.
2. Search for Twitter handles for each school and district on the school and school district websites.
3. Verify Twitter handles by clicking on Twitter handle links on the school/district webpage.
4. If the webpage is not found **OR** a Twitter handle is not located on the webpage, search for the school/district name on Google with the word “Twitter” (ex. Georgia Southern Twitter).
5. Verify Twitter handle by checking the school address/location.
6. If the above methods are still unsuccessful, use the search function on the Twitter page (www.twitter.com) to search for the name of the school/district.
7. Only Twitter handles owned by the individual schools, school principal, school district, and school district superintendent are accepted.
8. Enter Twitter handles of schools (“IndTwHandle”) and districts (“DisTwiHandle”) in the **school data** and **district data** spread sheets from the NCES website respectively.

* “IndTwHandle” = Twitter handles owned by school accounts or school principals.
* “DisTwiHandle” = Twitter handles owned by school district accounts or superintendents.

1. Enter the school district Twitter handles in the **school data sheet**, such the district Twitter handles per district are applied to all schools within the district (“DisTwiHandle”).
2. Duplicate the **school data sheet** to create a **school USC data** sheet.
3. In the **school USC data** spreadsheet**,** create a new column/variable to indicate the type of Twitter account(s) per school (“TwiAccType”) from the information in the “DisTwiHandle” and the “DisTwiHandle” columns.

* “1” = District Twitter account only
* “2” = School Twitter account only
* “3” = Both district and school Twitter accounts
* “4” = None

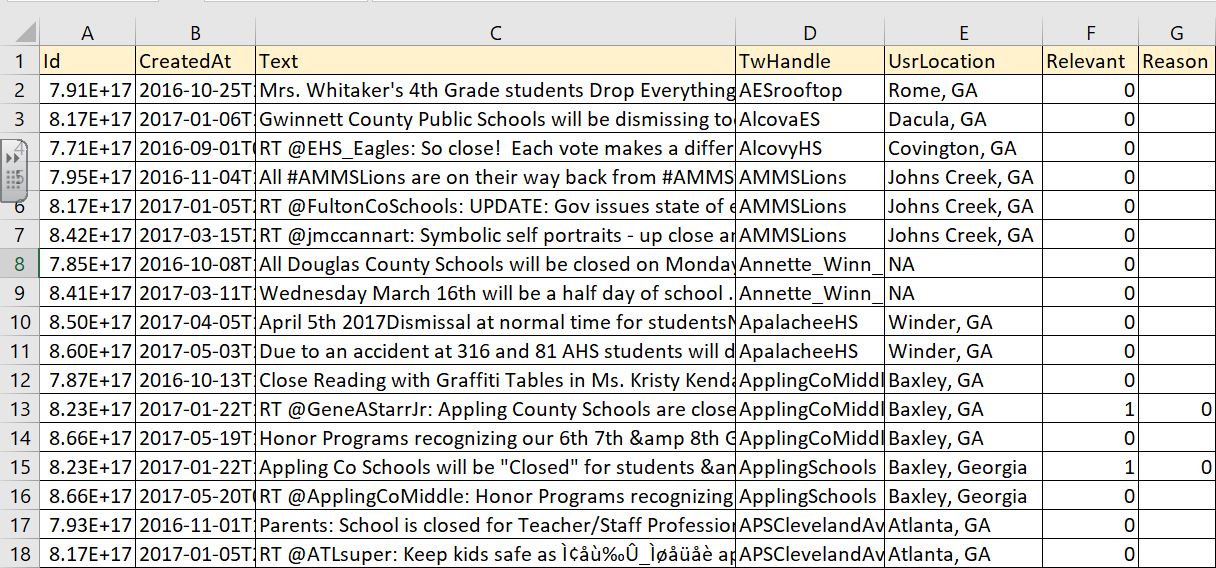
1. Recategorize the “Locale” variable from a 12-category to a 4-category variable.

* “City: Large” = “City”
* “City: Midsize” = “City”
* “City: Small” = “City”
* “Suburban: Large” = “Suburban”
* “Suburban: Midsize” = “Suburban”
* “Suburban: Small” = “Suburban”
* “Town: Distant” = “Town”
* “Town: Fringe” = “Town”
* “Town: Remote” = “Town”
* “Rural: Distant” = “Rural”
* “Rural: Fringe” = “Rural”
* “Rural: Remote” = “Rural”

1. Create a new variable for the proportion of students on free/reduced lunch

**Phase 2**

1. Submit Twitter handles for retrieval of tweets using Twitter API
2. Filter all likely tweets on school closures using a list of key words and phrases (“school closed”, “schools closed”, “schools are closed”, “schools will be closed”, “close”) created by the Community Interventions and Infection Control Unit at the CDC.
3. Save retrieved tweets in an Excel spreadsheet **(Twitter USC events data)** such as the one in the picture below.



1. Manually read through each tweet to identify unplanned school closure events. In a column labeled “relevant” and categorize each tweet as follows:
   * “1” = A relevant tweet of an unplanned school closure.
   * “0” = An irrelevant tweet.
2. According to the content of each relevant tweet, identify the reason for each USC event, categorizing the cause for each event, in a column labelled “Reason”, with the numbers “0 – 5” as illustrated below.
   * “0” = Weather related events
   * “1” = Event due to a facility issue(s)
   * “2” = Event due to teacher strike
   * “3” = Violence/threat related events
   * “4” = Event due to an illness of any kind
   * “5” = Other (USC tweets from which the reason for closure cannot be determined)

**Phase 3**

1. Filter USC events identified by the CDC 2017 (**CDC USC events data)** to retrieve unscheduled dismissals of Georgia public schools that occurred from August 2015 to June 2017 (inclusive).
2. Separate USC closure events in 2015-16 and 2016-17 school years in different Excel sheets.
3. Insert “Yes” in the “CDCClosure” column representing USC event captured by CDC
4. Duplicate the **CDC USC events data** spreadsheet to create the **Events data** spreadsheet.
5. Merge the USC information from the **Twitter USC events data** with the **CDC USC events data,** creating “CDCClosure” and “TwitterClosure” columns/variables with “Yes” or “No categories”.

**Phase 4**

1. In the **school USC data** sheet, apply all district closures identified by the CDC and Twitter to all the schools in the school district in the “CDCClosure” and “TwitterClosure” columns respectively, coding them as “Yes”.
2. In the “ClosureType” column enter the Type of closure “D” (District closure)
3. For each individual school closure identified by both the CDC and Twitter, enter the reason for closure, closure type, identification by CDC, identification by Twitter in the appropriate columns.
4. For schools with no USC event identified by CDC and/or Twitter, enter “No” in the appropriate column.
5. **Codebook and Frequencies**

**School characteristics variables**

The school characteristics variables were comprised of secondary data in a spreadsheet downloaded from the NCES website except the “STRatio” and “frlunchP” which were derivatives. The “Locality” variable was also collapsed from a 12-category to a 4-category variable.

**Locality - School locale**

Variable name: School locale

R Label: Locality

Variable type: Categorical

Dataset: School USC Data

|  |  | |  | | |
| --- | --- | --- | --- | --- | --- |
|  | **Code/Value** | **Value Description** | | **Count** | **Cumulative** |
| **All Georgia public schools** | City | Territory within an urbanized area and within a principal city | | 417 | 417 |
| Suburban | Territory outside a principal city and within an urbanized area | | 857 | 1,274 |
| Town | Territory within an urban cluster | | 288 | 1,562 |
| Rural | Census-defined rural territory | | 708 | 2,270 |
| **Schools with Twitter Covered** | City | Territory within an urbanized area and within a principal city | | 372 | 372 |
| Suburban | Territory outside a principal city and within an urbanized area | | 830 | 1,202 |
| Town | Territory within an urban cluster | | 171 | 1,373 |
| Rural | Census-defined rural territory | | 467 | 1,840 |

**STRatio – Student-teacher ratio**

Variable name: Ratio of Students to Teachers

R Label: STRatio

Variable type: Number

Dataset: School USC Data

| **Population** | **Code/Value** | | **Value Description** | **Count** | **Cumulative** |
| --- | --- | --- | --- | --- | --- |
| **All Georgia Public Schools** | | 1.8 – 206.7 | Range of values | 2,237 | 2,237 |
| NA | Missing | 62 | 2,299 |
| **Schools with Active Twitter Coverage** | | 1.8 - 136.4 | Range of values | 1,810 | 1,810 |
| NA | Missing | 54 | 1,864 |

**Students - Number of students**

Variable name: Number of students

R Label: Students

Variable type: Number

Dataset: School USC Data

| **Dataset** | **Code/Value** | **Value Description** | **Count** | **Cumulative** |
| --- | --- | --- | --- | --- |
| **All Georgia Public Schools** | 1 – 13,659 | Range of values | 2,265 | 2,265 |
| NA | Missing | 34 | 2,299 |
| **Schools with Active Twitter Coverage** | 1 – 13,659 | Range of values | 1,826 | 1,826 |
| NA | Missing | 38 | 1,864 |

**frlunchP -** **Proportion of students that receive free/reduced lunch**

Variable name: Proportion of students that receive free/reduced lunch

R Label: frlunchP

Variable type: Number

Dataset: School USC Data

| **Dataset** | **Code/Value** | **Value Description** | **Count** | **Cumulative** |
| --- | --- | --- | --- | --- |
| **All Georgia Public Schools** | 0 - 1 | Range of values | 2,250 | 2,250 |
| NA | Missing | 49 | 2,299 |
| **Schools with Active Twitter Coverage** | 0 - 1 | Range of values | 1,821 | 1,821 |
| NA | Missing | 43 | 1,864 |

**Twitter Coverage variables**

The Twitter coverage data comprised of primary data collected by the authors on the existence of school or district owned Twitter handles, and the use of these Twitter accounts.

**IndTwHandle – School Twitter handle**

**TwAccType - Twitter account type**

Variable name: Twitter account type

R Label: TwAccType

Variable type: Categorical

Dataset: School data

| **Code/Value** | **Value Description** | **Count** | **Cumulative** |
| --- | --- | --- | --- |
| 1 | District Twitter account | 1,226 | 1,226 |
| 2 | School Twitter account | 98 | 1,324 |
| 3 | Both | 540 | 1,864 |
| 4 | None | 435 | 2,299 |

**ActiveTw - Active Twitter coverage**

Variable name: Active Twitter coverage

R Label: ActiveTw

Variable type: Categorical

Dataset: School data and Events data

| **Code/Value** | **Value Description** | **Count** | **Cumulative** |
| --- | --- | --- | --- |
| 0 | No | 435 | 435 |
| 1 | Yes | 1,864 | 2,299 |

**Unplanned school closure variables**

The unplanned school closure variables comprised of variables representing the all the events, type of school closures and schools with USCs detected by OSS and Twitter.

**ClosType1 - Type of closure**

Variable name: Type of closure

R Label: ClosType1

Variable type: Categorical

Dataset: Events data

| **School Year** |  |  | **All GA public schools** | | **Schools with Twitter coverage** | |
| --- | --- | --- | --- | --- | --- | --- |
| **Code/Value** | **Value Description** | **Count** | **Cumulative** | **Count** | **Cumulative** |
| **2015-16** | D | District closure | 107 | 107 | 95 | 95 |
| I | School closure | 12 | 119 | 9 | 104 |
| **2016-17** | D | District closure | 148 | 148 | 96 | 96 |
| I | School closure | 23 | 171 | 21 | 117 |

**CDCClosure - Event detected by OSS**

Variable name: Event detected by OSS

R Label: CDCClosure

Variable type: Categorical

Dataset: Events data

| **School Year** |  |  | **All GA public schools** | | **Schools with Twitter coverage** | |
| --- | --- | --- | --- | --- | --- | --- |
| **Code/Value** | **Value Description** | **Count** | **Cumulative** | **Count** | **Cumulative** |
| **2015-16** | No |  | 19 | 19 | 19 | 19 |
| Yes |  | 88 | 107 | 51 | 70 |
| **2016-17** | No |  | 55 | 55 | 55 | 55 |
| Yes |  | 116 | 171 | 62 | 117 |

**TwitterClosure – Event detected on Twitter**

Variable name: Event detected on Twitter

R Label: TwitterClosure

Variable type: Categorical

Dataset: Events data

| **School Year** |  |  | **All GA public schools** | | **Schools with Twitter coverage** | |
| --- | --- | --- | --- | --- | --- | --- |
| **Code/Value** | **Value Description** | **Count** | **Cumulative** | **Count** | **Cumulative** |
| **2015-16** | No |  | 56 | 56 | 19 | 19 |
| Yes |  | 51 | 107 | 51 | 70 |
| **2016-17** | No |  | 84 | 84 | 30 | 30 |
| Yes |  | 87 | 171 | 87 | 117 |

**CDCClosure1 - School with ≥ 1 USC detected by OSS**

Variable name: School with ≥1 USC detected by OSS

R Label: CDCClosure1

Variable type: Categorical

Dataset: School USC data

|  |  |  | **All GA public schools** | | **Schools with Twitter coverage** | |
| --- | --- | --- | --- | --- | --- | --- |
| **School Year** | **Code/Value** | **Value Description** | **Count** | **Cumulative** | **Count** | **Cumulative** |
| **2015-16** | No |  | 1,770 | 1,770 | 1,448 | 1,448 |
| Yes |  | 529 | 2,299 | 416 | 1,864 |
| **2016-17** | No |  | 1,336 | 1,336 | 1,061 | 1,061 |
| Yes |  | 963 | 2,299 | 803 | 1,864 |

**TwitterClosure1 - School with ≥ 1 USC detected on Twitter**

Variable name: School with ≥1 USC detected on Twitter

R Label: TwitterClosure1

Variable type: Categorical

Dataset: School USC data

|  |  |  | **All GA public schools** | | **Schools with Twitter coverage** | |
| --- | --- | --- | --- | --- | --- | --- |
| **School Year** | **Code/Value** | **Value Description** | **Count** | **Cumulative** | **Count** | **Cumulative** |
| **2015-16** | No |  | 1,876 | 1,876 | 1,441 | 1,441 |
| Yes |  | 423 | 2,299 | 423 | 1,864 |
| **2016-17** | No |  | 1,425 | 1,425 | 990 | 990 |
| Yes |  | 874 | 2,299 | 874 | 1,864 |

**Section B. Results**

**Part III: Main Result Tables (Continued)**

**Table S1.** Adjusted relative risk (aRR) of (i) schools having an unplanned school closure (USC) identified via online systematic searches (OSS) and (ii) schools having active Twitter coverage, among all Georgia public schools (N=2,299).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 1. **Georgia public schools with USC identified via OSS** | | | | |
|  | **2015-16 school year** | | **2016-17 school year** | |
| **Predictor Variable** | **aRR (95% CI)** | **P-value** | **aRR (95% CI)** | **P-value** |
| Locality | Reference | | Reference | |
| City |
| Suburban | 0.6423 (0.4845 – 0.8418) | 0.0010 | 1.4168 (1.2653 - 1.5628) | <0.0001 |
| Town | 1.7327 (1.3958 – 2.0953) | <0.0001 | 1.5825 (1.4105 - 1.7387) | <0.0001 |
| Rural | 1.5540 (1.2804 – 1.8575) | <0.0001 | 1.3855 (1.2334 - 1.5329) | <0.0001 |
| Student Population | 1.0056 (0.9886 – 1.0207) | 0.4850 | 0.9885 (0.9768 - 0.9993) | 0.0478 |
| Free/Reduced Price Lunch Proportions | 1.0068 (0.9753 – 1.0393) | 0.6740 | 1.0478 (1.0277 - 1.0681) | <0.0001 |
| 1. **Georgia public schools with active Twitter coverage (2015-2017)** | | | | |
| **Predictor Variable** | **aRR (95% CI)** | | | **P-value** |
| Locality | Reference | | | |
| City |
| Suburban | 1.1388 (1.0838 - 1.1760) | | | <0.0001 |
| Town | 0.5028 (0.3881 - 0.6218) | | | <0.0001 |
| Rural | 0.5908 (0.4793 - 0.6984) | | | <0.0001 |
| Student Population | 1.0191 (1.0122 - 1.0261) | | | <0.0001 |
| Free/Reduced Price Lunch Proportions | 0.9907 (0.9798 - 1.0009) | | | 0.0777 |

CI: confidence intervals.

**Table S2.** Distribution of the unique USC events captured on Twitter compared to OSS.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| (a) All individual public schools (and school districts) in Georgia  (N=2,229) | | | | | | |
|  | **2015-16 school year** | | | **2016-17 school year** | | |
| Twitter | OSS | | | OSS | | |
| Captured | Missed | Total | Captured | Missed | Total |
| Captured | 32 | 19a | 51 | 32 | 55a | 87 |
| Missed | 56b | ? | ? | 84b | ? | ? |
| Total | 88 | ? | ? | 116 | ? | ? |
| (b) Individual public schools (and school districts) in Georgia with an active Twitter coverage (n=1,864) | | | | | | |
|  | **2015-16 school year** | | | **2016-17 school year** | | |
|  | OSS | | | OSS | | |
| Twitter | Captured | Missed | Total | Captured | Missed | Total |
| Captured | 32 | 19a | 51 | 32 | 55a | 87 |
| Missed | 19b | ? | ? | 30b | ? | ? |
| Total | 51 | ? | ? | 62 | ? | ? |

a Number missed by OSS but captured by Twitter.

b Number missed by Twitter but captured by OSS.

**Part IV: Main Result and Discussion Texts (Continued)**

**Text S2. Results**

Schools with active Twitter coverage were 50% less likely to be in towns (aRR= 0.5028; 95 CI%= 0.3881, 0.6218), 41% less likely to be in rural areas (aRR= 0.5908; 95 CI% = 0.4793, 0.6218), and 14% more likely to be in suburban areas (aRR= 1.1388; 95 CI% = 1.0838, 1.1760) when compared to city locations (Table S1[ii]).

Regarding the USC events that occurred in the 2015-16 school year, Twitter identified 51 USC events, of which 19 (37.25%) were missed by OSS. Conversely, OSS identified 88 unique USC events of which 56 (63.63%) were missed by Twitter. In the following school year, Twitter was able to identify 87 unique USC events, of which 55 (63.22%) were not captured by OSS. Likewise, OSS identified 116 unique USC events of which 84 (72.41%) were not identified by Twitter (Table S2a).

Of the 2,299 Georgia public schools, 529 (23.01%) and 963 (41.89%) schools had at least one USC announcement captured by OSS in the 2015-16 and 2016-17 school year respectively.  In the 2015-16 school year, compared to city schools, suburban schools were the least likely to have an USC identified by OSS (Suburban: aRR= 0.6423; 95%CI= 0.4845, 0.8418) while town schools and rural schools were more likely to have an USC identified by OSS (Town: aRR= 1.7327; 95%CI= 1.3958, 2.0953; Rural: aRR=1.5540; 95%CI= 1.2804, 1.8575) (Table S1[i]). In 2016-17 school year, city schools were the least likely to have an USC identified by OSS (Suburban: aRR=1.4168; 95%CI =1.2653, 1.5628; Town: aRR=1.5825; 95%CI=1.4105, 1.7387; Rural: aRR=1.3855; 95%CI=1.2334, 1.5329) (Table S1[i]).

After conducting the sensitivity analysis among **all** Georgia public schools, OSS was found to be more sensitive than Twitter in identifying USC events (2015-16: OSS = 82.24%; Twitter = 47.66%; 2016-17: OSS = 67.84%; Twitter = 50.88%) and identifying schools with at least one USC event (2015-16: OSS = 70.44%; Twitter = 56.32%; 2016-17: OSS = 83.40%; Twitter = 76.60%) (Table S8).

**Text S3. Discussion.**

From the perspective of a practitioner in a public health agency, whether Twitter can be a useful and important source of USC information may be related to a number of factors, such as whether the reason of USC (e.g., influenza outbreak) is identified, the length of USC, the ability of real-time data retrieval, the resources needed to retrieve Twitter data from all schools across the entire nation, the prevalence of Twitter use across schools in different states, and the validity of trend monitoring through Twitter. The adoption and the actual use of Twitter by school officials may change over time and may vary across different states. Variation in USC time trends as identified via Twitter may be confounded by variation in Twitter use over time. Likewise, the identification of USC announcements via OSS may be confounded by changes in the algorithms used by Google to order search results (such as Google PageRank algorithm) and press coverage of USC announcements over time. Since public health practitioners have no access to the proprietary algorithms used by Google, it is technically very difficult for public health researchers to assess the potential effect on the sensitivity and specificity of OSS over time. Furthermore, while technically it would have been possible for social media announcements to populate in the Google alert or Google news search, it was never noted in the OSS data sets for 2015-16 and 2016-17 school years.

We did not study the timing of the tweets. This was because the OSS data provided to us by CDC researchers was a retrospectively collected data set. We did not know when the CDC researchers picked up the news of the USC announcement. It could an hour later, or a day later, or a week later, for example. Likewise, we would not know when a tweet would have been picked up by a CDC researcher if Twitter became a routine source of information for them. A Twitter timestamp only indicated the time a tweet was posted. It did not indicate whether and when a tweet was read by a specific person.

We did not study other social media data sources as they are out of scope for this project. Unplanned USC announcements on Facebook have been studied in another study conducted by some other students of the senior author (unpublished data).

**Part V: Additional Result Tables**

**Table S3.** Distribution of unplanned school closure events identified via Twitter

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **2015-16 School Year** | | | **2016-17 School Year** | | |
|  | District Closures | Individual Closures | Total | District Closures | Individual Closures | Total |
| District Twitter account | 34 | 2 | 36 | 54 | 1 | 55 |
| School Twitter account | 10 | 5 | 15 | 14 | 18 | 32 |
| Total | 44 | 7 | 51\* | 68 | 19 | 87\* |

\* Total number of unplanned school closure events identified via Twitter.

**Table S4**.Distribution of unplanned school closure events identified via online systematic searches (OSS)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **2015-16 School Year** | | | **2016-17 School Year** | | |
| Schools with Active Twitter Coverage | District Closures | Individual Closures | Total | District Closures | Individual Closures | Total |
| Yes | 49 | 2 | 51 | 60 | 2 | 62 |
| No | 34 | 3 | 37 | 52 | 2 | 54 |
| Total | 83 | 5 | 88\* | 112 | 4 | 116\* |

\* Total number of unplanned school closure events identified via OSS.

**Table S5.** 2 x 2 tables of the district and individual closures captured by both online systematic searches (OSS) and Twitter in 2015-2016.

|  |  |
| --- | --- |
| **District Closures** | **Individual Closures** |
| |  |  |  |  | | --- | --- | --- | --- | | (a) | OSS Approach | |  | |  | Captured | Missed | Total | | Twitter |  |  |  | | Captured | 26 | 10 | 36 | | Missed | 57 | ? | ? | | Total | 83 | ? | ? |   District closures assumed to be all closures announced on district Twitter accounts.  Excludes district closures announced on school accounts (n = 10).  Includes school closures on district Twitter accounts (n = 2). | |  |  |  |  | | --- | --- | --- | --- | | (d) | OSS Approach | |  | |  | Captured | Missed | Total | | Twitter |  |  |  | | Captured | 0 | 26 | 26 | | Missed | 5 | ? | ? | | Total | 5 | ? | ? |   School closures assumed to be all closures announced on school Twitter accounts.  Excludes school closures announced on district accounts (n = 2).  Includes district closures on school Twitter accounts (n = 21). |
| |  |  |  |  | | --- | --- | --- | --- | | (b) | OSS Approach | |  | |  | Captured | Missed | Total | | Twitter |  |  |  | | Captured | 32 | 14 | 46 | | Missed | 51 | ? | ? | | Total | 83 | ? | ? |   District closures assumed to be all closures on district Twitter accounts and district closures on school Twitter accounts.  Includes school closures on district Twitter accounts (n = 2). | |  |  |  |  | | --- | --- | --- | --- | | (e) | OSS Approach | |  | |  | Captured | Missed | Total | | Twitter |  |  |  | | Captured | 0 | 28 | 28 | | Missed | 5 | ? | ? | | Total | 5 | ? | ? |   School closures assumed to be all closures on school Twitter accounts and school closures on district Twitter accounts.  Includes district closures on school Twitter accounts (n = 21). |
| |  |  |  |  | | --- | --- | --- | --- | | (c) | OSS Approach | |  | |  | Captured | Missed | Total | | Twitter |  |  |  | | Captured | 32 | 12 | 44 | | Missed | 51 | ? | ? | | Total | 83 | ? | ? |   District closures are all district closures announced on school and district Twitter accounts.  Excludes school closures on district Twitter accounts (n = 2). | |  |  |  |  | | --- | --- | --- | --- | | (f) | OSS Approach | |  | |  | Captured | Missed | Total | | Twitter |  |  |  | | Captured | 0 | 7 | 7 | | Missed | 5 | ? | ? | | Total | 5 | ? | ? |   School closures are all school closures announced on school and district twitter accounts.  Excludes district closures on school Twitter accounts (n = 21). |

Notes: 1. The columns refer to whether the closures were identified by the OSS method.

2. The rows refer to whether the closures were identified via Twitter.

3. School Twitter accounts and closures refer to accounts and closures per individual public school

4. District Twitter accounts and closures refer to accounts for and closures per public school district

**Table S6.** 2 x 2 tables of the district and individual closures captured by both online systematic searches (OSS) and Twitter in 2016-2017.

|  |  |
| --- | --- |
| **District Closures** | **Individual Closures** |
| |  |  |  |  | | --- | --- | --- | --- | | (a) | OSS Approach | |  | |  | Captured | Missed | Total | | Twitter |  |  |  | | Captured | 32 | 23 | 55 | | Missed | 80 | ? | ? | | Total | 112 | ? | ? |   District closures assumed to be all closures announced on district Twitter accounts.  Excludes district closures announced on school accounts (n = 14).  Includes school closures on district Twitter accounts (n = 1). | |  |  |  |  | | --- | --- | --- | --- | | (d) | OSS Approach | |  | |  | Captured | Missed | Total | | Twitter |  |  |  | | Captured | 0 | 54 | 54 | | Missed | 4 | ? | ? | | Total | 4 | ? | ? |   School closures assumed to be all closures announced on school Twitter accounts.  Excludes school closures announced on district accounts (n = 1).  Includes district closures on school Twitter accounts (n = 36). |
| |  |  |  |  | | --- | --- | --- | --- | | (b) | OSS Approach | |  | |  | Captured | Missed | Total | | Twitter |  |  |  | | Captured | 32 | 37 | 69 | | Missed | 80 | ? | ? | | Total | 112 | ? | ? |   District closures assumed to be all closures on district Twitter accounts and district closures on school Twitter accounts.  Includes school closures on district Twitter accounts (n = 1). | |  |  |  |  | | --- | --- | --- | --- | | (e) | OSS Approach | |  | |  | Captured | Missed | Total | | Twitter |  |  |  | | Captured | 0 | 55 | 55 | | Missed | 4 | ? | ? | | Total | 4 | ? | ? |   School closures assumed to be all closures on school Twitter accounts and school closures on district Twitter accounts.  Includes district closures on school Twitter accounts (n = 36). |
| |  |  |  |  | | --- | --- | --- | --- | | (c) | OSS Approach | |  | |  | Captured | Missed | Total | | Twitter |  |  |  | | Captured | 32 | 36 | 68 | | Missed | 80 | ? | ? | | Total | 112 | ? | ? |   District closures are all district closures announced on school and district Twitter accounts.  Excludes school closures on district Twitter accounts (n = 1). | |  |  |  |  | | --- | --- | --- | --- | | (f) | OSS Approach | |  | |  | Captured | Missed | Total | | Twitter |  |  |  | | Captured | 0 | 19 | 19 | | Missed | 4 | ? | ? | | Total | 4 | ? | ? |   School closures are all school closures announced on school and district twitter accounts.  Excludes district closures on school Twitter accounts (n = 36). |

Notes: 1. The columns refer to whether the closures were identified by the OSS method.

2. The rows refer to whether the closures were identified via Twitter.

3. School Twitter accounts and closures refer to accounts and closures per individual public school.

4. District Twitter accounts and closures refer to accounts and closures per public school district.

**Table S7.** Distribution of schools with at least one unplanned school closure (USC) announcement captured by Twitter compared to online systematic searches (OSS) among all Georgia public schools (N=2,299).

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **2015-16 School Year** | | | **2016-17 School Year** | | |
| Twitter | OSS | | | OSS | | |
| Captured | No USC announcement identified | Total | Captured | No USC announcement identified | Total |
| Captured | 201 | 222a | 423 | 696 | 178a | 874 |
| No USC announcement identified | 328b | 1,548 | 1,876 | 267b | 1,158 | 1,425 |
| Total | 529 | 1,770 | 2,299† | 963 | 1,336 | 2,299† |

†The total number of public schools in Georgia.

a Number missed by the OSS but captured by Twitter.

b Number missed by Twitter but captured by the OSS.

**Table S8.** Sensitivity of Twitter and OSS for identifying USC events and schools with at least one USC event among (i) all Georgia public schools and (ii) schools with active Twitter coverage.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Sensitivity (%)**  **(i) All Georgia public schools (n = 2,299)** | | | | |
|  | **2015-16** | | **2016-17** | |
|  | **Detected on Twitter** | **Detected by OSS** | **Detected on Twitter** | **Detected by OSS** |
| USC Events | 47.66 | 82.24 | 50.88 | 67.84 |
| Schools ≥ 1 USC | 56.32 | 70.44 | 76.60 | 83.40 |
| **Sensitivity (%)**  **(ii) Georgia public schools with active Twitter coverage (n = 1,864)** | | | | |
|  | **2015-16** | | **2016-17** | |
|  | **Detected on Twitter** | **Detected by OSS** | **Detected on Twitter** | **Detected by OSS** |
| USC Events | 72.86 | 72.86 | 74.36 | 52.99 |
| Schools ≥ 1 USC | 66.30 | 65.20 | 89.09 | 81.86 |

**Table S9.** Descriptive statistics of Georgia public schools, with or without an active Twitter account (2015-2017).

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Schools with an active Twitter account**  **N = 1840** | **¥Schools without any active Twitter account**  **N = 430** | **P value** |
| School Locale, n (%) |  |  | < 0.0001\* |
| City | 372 (20.22) | 45 (10.47) |  |
| Suburban | 830 (45.11) | 27 (6.28) |
| Town | 171 (9.29) | 117 (27.21) |
| Rural | 467 (25.38) | 241 (56.05) |
| Mean Student Population | 806.06 | 594.68 | < 0.0001\* |
| Mean Student-Teacher Ratio | 15.72 | 15.86 | 0.6763 |
| Mean Reduced Price/Free Lunch Proportion | 0.66 | 0.73 | < 0.0001\* |

Note: (1) 2,299 public schools in Georgia.

(2) The NCES data was missing the locality information of 29 schools.

¥ Schools whose Twitter accounts were inactive or did not have a Twitter account, and did not belong to a school district that had an active Twitter account.

**Table S10.** Descriptive statistics of Georgia public schools whose USC announcements were identified via OSS and schools with no identified USC announcements

|  |  |  |  |
| --- | --- | --- | --- |
| **2015 – 2016 school year** | | | |
|  | **Schools with announcements identified**  **N = 525** | **Schools with no identified announcements**  **N = 1,745** | **P value** |
| School Locale, n(%) |  |  | < 0.0001\* |
| City | 84 (16.00) | 333 (19.08) |
| Suburban | 111 (21.14) | 746 (42.75) |
| Town | 103 (19.62) | 185 (10.60) |
| Rural | 227 (43.24) | 481 (27.56) |
| Mean Student Population | 735.32 | 775.28 | 0.1312 |
| Mean Student-Teacher Ratio | 15.83 | 15.72 | 0.7169 |
| Mean Reduced Price/Free Lunch Proportion | 0.69 | 0.67 | 0.0615 |
| **2016 – 2017 school year** | | | |
|  | **Schools with announcements identified**  **N = 954** | **Schools with no identified announcements**  **N = 1,316** | **P value** |
| School Locale, n(%) |  |  | < 0.0001\* |
| City | 128 (13.42) | 289 (21.96) |
| Suburban | 362 (37.95) | 495 (37.61) |
| Town | 154 (16.14) | 134 (10.18) |
| Rural | 310 (32.49) | 398 (30.24) |
| Mean Student Population | 731.86 | 790.88 | 0.0053\* |
| Mean Student-Teacher Ratio | 15.92 | 15.62 | 0.2367 |
| Mean Reduced Price/Free Lunch Proportion | 0.70 | 0.65 | < 0.0001\* |

Note: (1) The NCES data was missing the locale information of 29 schools.

**Table S11.** Descriptive statistics of schools with USC events identified via Twitter and OSS among Georgia public schools with active Twitter coverage.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **2015 – 2016 school year** | | | | | | |
|  | **Schools with ≥ 1 USC Identified via OSS** | | **p-value** | **Schools with ≥ 1 USC Identified via Twitter** | | **p-value** |
| **Yes**  **N = 412** | **No**  **N = 1,428** | **Yes**  **N = 422** | **No**  **N = 1,418** |
| School Locale, n (%) |  |  | < 0.0001\* |  |  | < 0.0001\* |
| City | 83 (20.14) | 289 (20.24) |  | 31 (7.52) | 341 (24.05) |  |
| Suburban | 107 (25.97) | 723 (50.63) | 225 (53.32) | 605 (42.67) |
| Town | 69 (16.75) | 102 (7.14) | 56 (13.27) | 115 (8.11) |
| Rural | 153 (37.14) | 314 (21.99) | 110 (26.07) | 357 (25.18) |
| Total Student Population (Mean) | 776.05 | 814.74 | 0.2196 | 841.91 | 795.36 | 0.1397 |
| Mean Student-Teacher Ratio | 16.03 | 15.62 | 0.1374 | 15.81 | 15.69 | 0.6594 |
| Mean Reduced Price/Free Lunch Proportion | 0.67 | 0.65 | 0.2273 | 0.65 | 0.66 | 0.6006 |
| **2016 – 2017 school year** | | | | | | |
|  | **Schools with ≥ 1 USC Identified via OSS** | | **p-value** | **Schools with ≥ 1 USC Identified by Twitter** | | **p-value** |
| **Yes**  **N = 796** | **No**  **N = 1,044** | **Yes**  **N = 864** | **No**  **N = 976** |
| School Locale, n (%) |  |  | < 0.0001\* |  |  | 0.3475 |
| City | 128 (16.08) | 244 (23.37) |  | 189 (21.88) | 183 (18.75) |  |
| Suburban | 356 (44.72) | 474 (45.40) | 380 (43.98) | 450 (46.11) |
| Town | 102 (12.81) | 69 (6.61) | 83 (9.61) | 88 (9.02) |
| Rural | 210 (26.38) | 257 (24.62) | 212 (24.54) | 255 (26.13) |
| Total Student Population (Mean) | 751.97 | 847.45 | 0.0001\* | 779.11 | 829.95 | 0.0472\* |
| Mean Student-Teacher Ratio | 15.73 | 15.71 | 0.9198 | 15.82 | 15.63 | 0.4258 |
| Mean Reduced Price/Free Lunch Proportion | 0.69 | 0.63 | < 0.0001\* | 0.67 | 0.64 | 0.0093\* |

Note: (1) Schools have an active Twitter account.

(2) The NCES data was missing locality information of 24 schools.

**Table S12.** The crude relative risk of schools having active Twitter coverage among GA public schools (2015 – 2017, two school years combined).

|  |  |  |
| --- | --- | --- |
| **Predictor Variable** | **Relative Risk (95% CI)** | **P-value** |
| Student Population | 1.0259 (1.0202 - 1.0318) | <0.0001\* |
| Student-Teacher Ratio | 0.9994 (0.9965 - 1.0030) | 0.6790 |
| Free/Reduced Price Lunch Proportions | 0.9785 (0.9697 - 0.9868) | <0.0001\* |
| School Locale |  |  |
| City | Reference | |
| Suburban | 1.1607 (1.1194 - 1.1886) | <0.0001\* |
| Town | 0.5313 (0.4161 - 0.6486) | <0.0001\* |
| Rural | 0.6178 (0.5089 - 0.7208) | <0.0001\* |

**Table S13.** The crude relative risk of schools having an USC captured via OSS among GA public schools (2015 – 2016 school year).

|  |  |  |
| --- | --- | --- |
| **Predictor Variable** | **Relative Risk (95% CI)** | **P-value** |
| Student Population | 0.9872 (0.9702 – 1.0027) | 0.1270 |
| Student-Teacher Ratio | 1.0021 (0.9885 - 1.0175) | 0.7180 |
| Free/Reduced Price Lunch Proportions | 1.0260 (0.9977 - 1.0552) | 0.0740 |
| School Locale |  |  |
| City | Reference | |
| Suburban | 0.6517 (0.4978 – 0.8446) | 0.0009\* |
| Town | 1.7254 (1.3892 – 2.0883) | <0.0001\* |
| Rural | 1.5572 (1.2873 – 1.8567) | <0.0001\* |

**Table S14.** The crude relative risk of schools having an USC captured via OSS among GA public schools (2016 – 2017 school year).

|  |  |  |
| --- | --- | --- |
| **Predictor Variable** | **Relative Risk (95% CI)** | **P-value** |
| Student Population | 0.9859 (0.9750 - 0.9961) | 0.0091\* |
| Student-Teacher Ratio | 1.0050 (0.9968 - 1.0154) | 0.2636 |
| Free/Reduced Price Lunch Proportions | 1.0424 (1.0241 - 1.0609) | <0.0001\* |
| School Locale |  |  |
| City | Reference | |
| Suburban | 1.2964 (1.1500 - 1.4417) | <0.0001\* |
| Town | 1.5536 (1.3796 - 1.7134) | <0.0001\* |
| Rural | 1.3335 (1.1830 - 1.4813) | <0.0001\* |

**Table S15.** The crude relative risk of schools having an USC identified via OSS among GA public schools with an active Twitter coverage (2015 – 2016 school year).

|  |  |  |
| --- | --- | --- |
| **Predictor Variable** | **Relative Risk (95% CI)** | **P-value** |
| Student Population | 0.9889 (0.9706 – 1.0052) | 0.2160 |
| Student-Teacher Ratio | 1.0110 (0.9951 – 1.0285) | 0.1670 |
| Free/Reduced Price Lunch Proportions | 1.0192 (0.9883 – 1.0511) | 0.2270 |
| School Locale |  |  |
| City | Reference | |
| Suburban | 0.5780 (0.4365 – 0.7584) | <0.0001\* |
| Town | 1.8070 (1.4061 – 2.2391) | <0.0001\* |
| Rural | 1.4677 (1.1806 – 1.7920) | <0.0001\* |

**Table S16.** The crude relative risk of schools having an USC identified via OSS among GA public schools with an active Twitter coverage (2016 – 2017 school year).

|  |  |  |
| --- | --- | --- |
| **Predictor Variable** | **Relative Risk (95% CI)** | **P-value** |
| Student Population | 0.9783 (0.9666 - 0.9896) | 0.0002\* |
| Student-Teacher Ratio | 1.0005 (0.9892 - 1.0116) | 0.9200 |
| Free/Reduced Price Lunch Proportions | 1.0436 (1.0244 - 1.0630) | <0.0001\* |
| School Locale |  |  |
| City | Reference | |
| Suburban | 1.2064 (1.0605 - 1.3524) | 0.0057 |
| Town | 1.5774 (1.3807 - 1.7518) | <0.0001\* |
| Rural | 1.2549 (1.0934 - 1.4139) | 0.0020\* |

**Table S17.** The crude relative risk of schools having an USC identified via Twitter search among GA public schools with an active Twitter coverage (2015 – 2016 school year).

|  |  |  |
| --- | --- | --- |
| **Predictor Variable** | **Relative Risk (95% CI)** | **P-value** |
| Student Population | 1.0102 (0.9960 – 1.0249) | 0.1540 |
| Student-Teacher Ratio | 1.0035 (0.9853 - 1.0192) | 0.6610 |
| Free/Reduced Price Lunch Proportions | 0.9920 (0.9626 – 1.0223) | 0.6000 |
| School Locale |  |  |
| City | Reference | |
| Suburban | 2.4004 (1.9836 – 2.8332) | <0.0001\* |
| Town | 2.6794 (2.1650 – 3.1555) | <0.0001\* |
| Rural | 2.1765 (1.7318 – 2.6487) | <0.0001\* |

**Table S18.** The crude relative risk of schools having an USC identified via Twitter search among GA public schools with an active Twitter coverage (2016 – 2017 school year).

|  |  |  |
| --- | --- | --- |
| **Predictor Variable** | **Relative Risk (95% CI)** | **P-value** |
| Student Population | 0.9906 (0.9805 - 0.9999) | 0.0579 |
| Student-Teacher Ratio | 1.0041 (0.9940 - 1.0159) | 0.4350 |
| Free/Reduced Price Lunch Proportions | 1.0232 (1.0057 - 1.0409) | 0.0094\* |
| School Locale |  |  |
| City | Reference | |
| Suburban | 0.8942 (0.7701 - 1.0231) | 0.107 |
| Town | 0.9520 (0.7665 - 1.1443) | 0.623 |
| Rural | 0.8861 (0.7487 - 1.0297) | 0.119 |

**Table S19.** Distribution of schools with and without at least one unplanned school closure (USC) announcement captured by Twitter compared to online systematic search (OSS) among 1864 individual Georgia public schools with active Twitter coverage.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | 2015-16 | | 2016-17 | |
|  | n (% of sub-category) | % of schools with active Twitter coverage | n (% of sub-category) | % of schools with active Twitter coverage |
| **Schools with USC announcements identified: Total** | **638 (100)** | **34.23** | **981 (100)** | **52.63** |
| *City* | 87 (13.64) |  | 190 (19.37) |  |
| *Suburban* | 285 (44.67) |  | 405 (41.28) |  |
| *Town* | 80 (12.54) |  | 115 (11.72) |  |
| *Rural* | 182 (28.53) |  | 261 (26.61) |  |
| *Unknown locale* | 4 (0.63) |  | 10 (1.02) |  |
|  |  |  |  |  |
| **Via OSS and Twitter: Subtotal** | **201 (100)** | **10.78** | **696 (100)** | **37.34** |
| *City* | 27 (13.43) |  | 127 (18.25) |  |
| *Suburban* | 48 (23.88) |  | 331 (47.56) |  |
| *Town* | 45 (22.39) |  | 70 (10.06) |  |
| *Rural* | 80 (39.80) |  | 161 (23.13) |  |
| *Unknown locale* | 1 (0.50) |  | 7 (1.01) |  |
|  |  |  |  |  |
| **Via OSS only: Subtotal** | **215 (100)** | **11.53** | **107 (100)** | **5.74** |
| *City* | 56 (26.05) |  | 1 (0.93) |  |
| *Suburban* | 59 (27.44) |  | 25 (23.36) |  |
| *Town* | 24 (11.16) |  | 32 (29.91) |  |
| *Rural* | 73 (33.95) |  | 49 (45.79) |  |
| *Unknown locale* | 3 (1.40) |  | 0 (0) |  |
|  |  |  |  |  |
| **Via Twitter only: Subtotal** | **222 (100)** | **11.91** | **178 (100)** | **9.55** |
| *City* | 4 (1.80) |  | 62 (34.83) |  |
| *Suburban* | 178 (80.18) |  | 49 (27.53) |  |
| *Town* | 11 (4.95) |  | 13 (7.30) |  |
| *Rural* | 29 (13.06) |  | 51 (28.65) |  |
| *Unknown locale* | 0 (0) |  | 3 (1.69) |  |
|  |  |  |  |  |
| **Schools without USC announcements identified: Total** | **1226 (100)** | **65.77** | **883 (100)** | **47.37** |
| *City* | 285 (23.25) |  | 182 (20.61) |  |
| *Suburban* | 545 (44.45) |  | 425 (48.13) |  |
| *Town* | 91 (7.42) |  | 56 (6.34) |  |
| *Rural* | 285 (23.25) |  | 206 (23.33) |  |
| *Unknown locale* | 20 (1.63) |  | 14 (1.59) |  |
|  |  |  |  |  |
| **Schools with active Twitter coverage - Total** | **1864** | **100** | **1864** | **100** |

**Table S20.** On exclusion of USCs due to Tropical Storm Hermine and Hurricane Matthew, distribution of schools with and without at least one unplanned school closure (USC) announcement captured by Twitter compared to online systematic search (OSS) among 1864 individual Georgia public schools with active Twitter coverage.

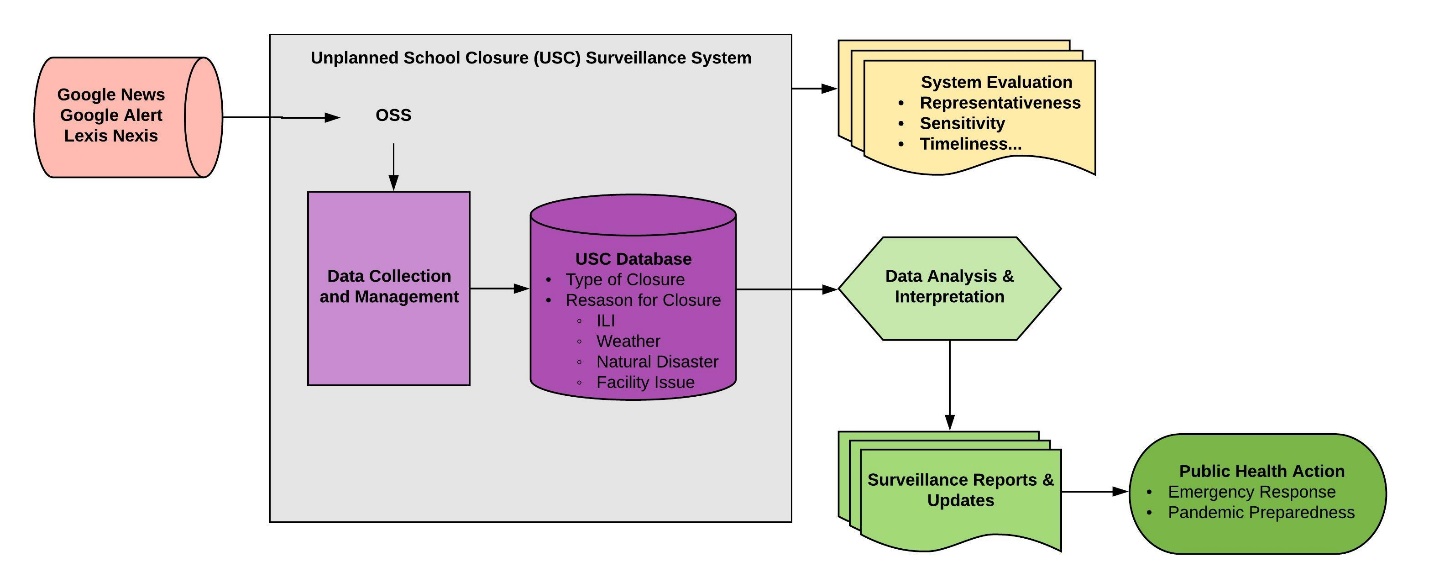
|  |  |  |
| --- | --- | --- |
|  | 2016-17 | |
|  | n (% of sub-category) | % of schools with active Twitter coverage |
| **Schools with USC announcements identified: Total** | 732 (100) | 39.27 |
| *City* | 105 (14.34) |  |
| *Suburban* | 358 (48.91) |  |
| *Town* | 67 (9.15) |  |
| *Rural* | 196 (26.78) |  |
| *Unknown locale* | 6 (0.82) |  |
|  |  |  |
| **Via OSS and Twitter: Subtotal** | 451 (100) | 24.20 |
| *City* | 34 (7.54) |  |
| *Suburban* | 287 (63.64) |  |
| *Town* | 40 (8.87) |  |
| *Rural* | 88 (19.51) |  |
| *Unknown locale* | 2 (0.44) |  |
|  |  |  |
| **Via OSS only: Subtotal** | 112 (100) | 6.01 |
| *City* | 4 (3.57) |  |
| *Suburban* | 26 (23.21) |  |
| *Town* | 22 (19.64) |  |
| *Rural* | 59 (52.68) |  |
| *Unknown locale* | 1 (0.89) |  |
|  |  |  |
| **Via Twitter only: Subtotal** | 169 (100) | 9.07 |
| *City* | 67 (39.64) |  |
| *Suburban* | 45 (26.63) |  |
| *Town* | 5 (2.96) |  |
| *Rural* | 49 (28.99) |  |
| *Unknown locale* | 3 (1.78) |  |
|  |  |  |
| **Schools without USC announcements identified: Total** | 1132 (100) | 60.73 |
| *City* | 267 (23.59) |  |
| *Suburban* | 472 (41.70) |  |
| *Town* | 104 (9.19) |  |
| *Rural* | 271 (23.94) |  |
| *Unknown locale* | 18 (1.59) |  |
|  |  |  |
| **Schools with active Twitter coverage: Total** | 1864 | 100 |

**Table S21.** On exclusion of USCs due to Tropical Storm Hermine and Hurricane Matthew, adjusted relative risk (aRR) of schools having an USC identified via Twitter search, among 1864 Georgia public schools with active Twitter coverage.

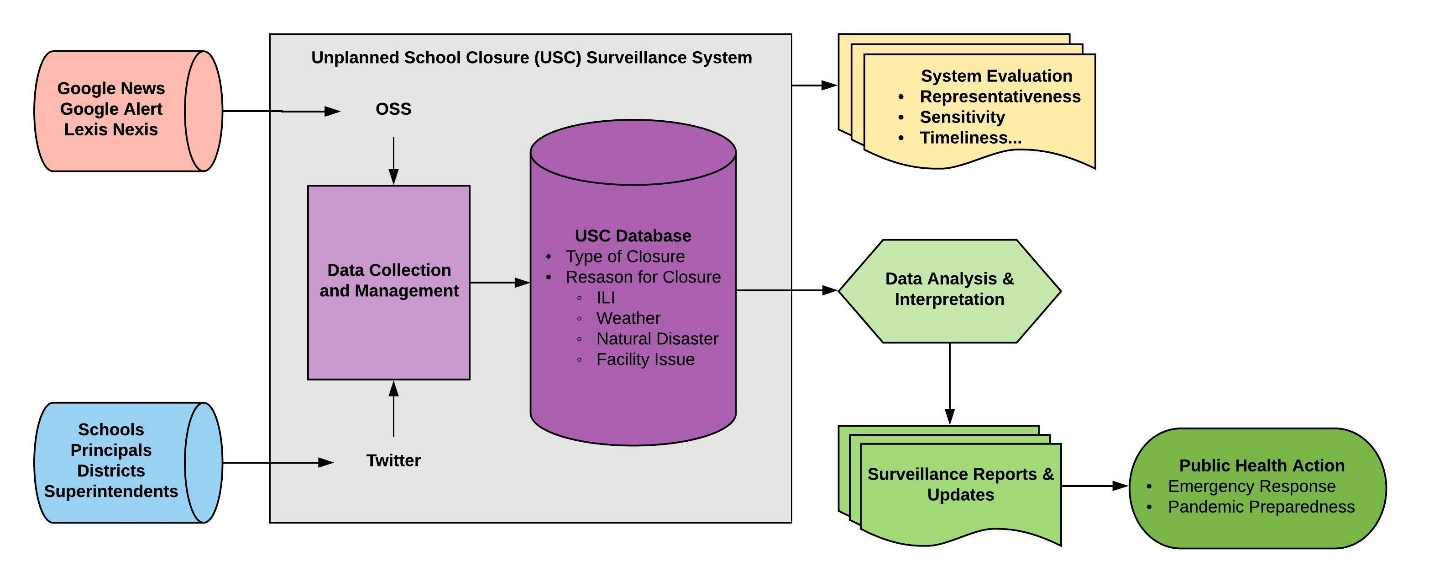
|  |  |  |
| --- | --- | --- |
| **Schools with USC(s) identified via Twitter Search among Georgia Public Schools with active Twitter coverage** | | |
|  | **2016-17 School Year** | |
| **Predictor Variable** | **aRR (95% CI)** | **P-value** |
| School Locale | Reference | |
| City |
| Suburban | 1.4452 (1.2405 - 1.6552) | <0.0001 |
| Town | 0.9904 (0.7357 - 1.2755) | 0.9453 |
| Rural | 1.0718 (0.8701 - 1.2931) | 0.5012 |
| Student Population | 0.9869 (0.9724 - 1.0002) | 0.0688 |
| Free/Reduced Price Lunch Proportions | 0.9984 (0.9736 - 1.0235) | 0.8977 |

**Section C. Supplementary Figures**

**Figure S1. Existing Methods of Online Systematic Searches (OSS) for Unplanned School Closure Surveillance in the USA.**

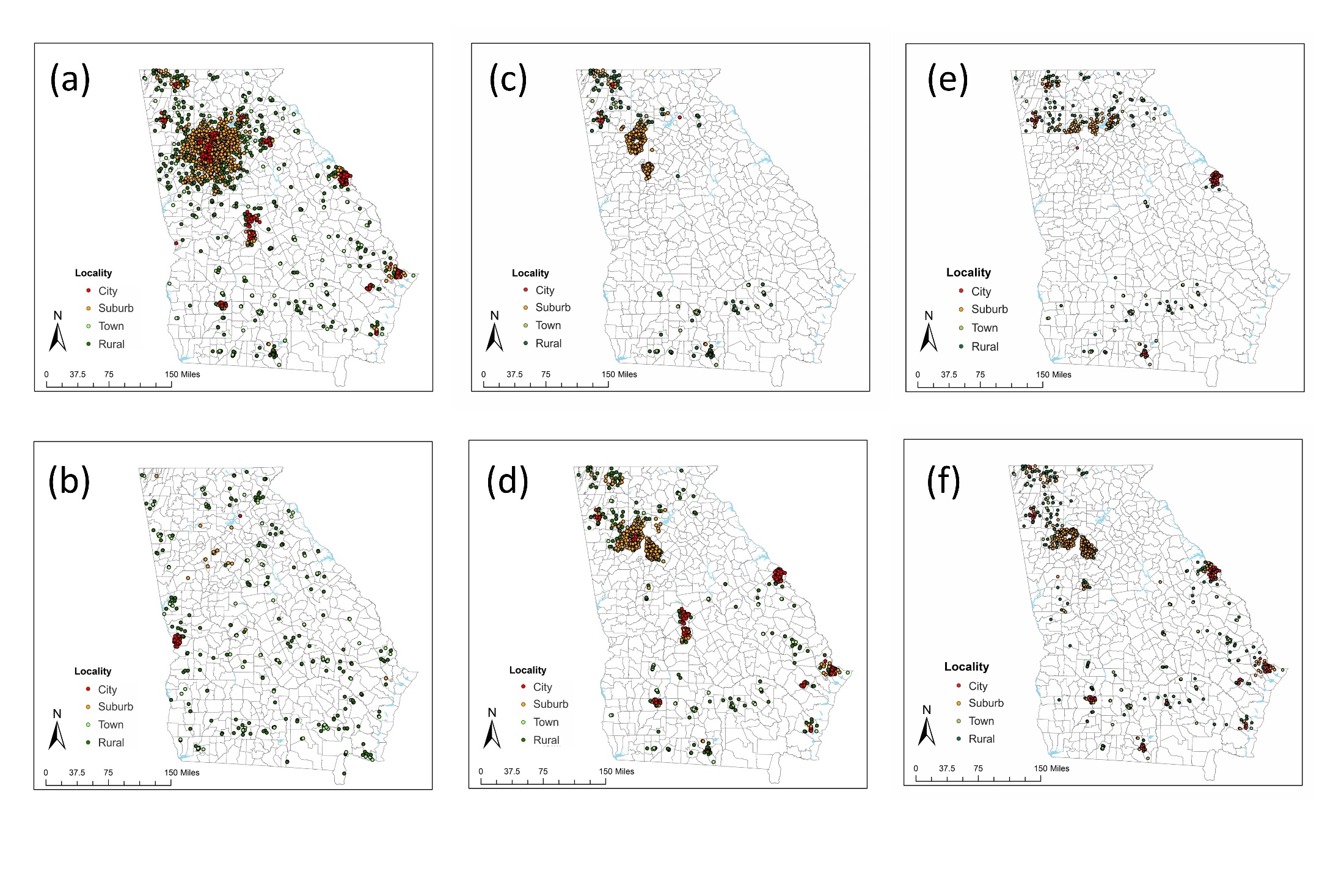
****

**Figure S2. Proposed Additions of Twitter Data to Unplanned School Closure Surveillance in the USA.**

****

**Figure S3. Flowchart of Identification of Public Schools in Georgia, USA, with or without Active Twitter Coverage.**

E:\Files\Documents\GSU_work\Research\Digital\Projects\29_School\Manuscript-GA\Figure1_20180807.tif

**Figure S4. Map of Georgia Depicting School Locations.** The dots represent individual public schools in Georgia and the color of the dot depends on the locality of the school. Panels represent schools: (a) With active Twitter coverage. (b) Without active Twitter coverage. (c) With active Twitter coverage and used their Twitter accounts to announce at least one unplanned school closure in 2015-16 school year. (d) With active Twitter coverage and used their Twitter accounts to announce at least one unplanned school closure in 2016-17 school year. (e) With active Twitter coverage and with at least one unplanned school closure captured by OSS in 2015-16 school year. (f) With active Twitter coverage and with at least one unplanned school closure captured by OSS in 2016-17 school year.

References

1. Wong KK, Shi J, Gao H, Zheteyeva YA, Lane K, Copeland D, et al. Why is school closed today? Unplanned K-12 school closures in the United States, 2011-2013. PloS one. 2014;9(12):e113755.

2. McFarland J, Hussar B, de Brey C, Snyder T, Wang X, Wilkinson-Flicker S, et al. The Condition of Education 2017 (NCES 2017- 144). 2017.

3. Georgia Department of Education. Georgia Announces 2016-17 Free and Reduced-Price Meal Income Guidelines 2017 [Available from: <http://www.gadoe.org/External-Affairs-and-Policy/communications/Pages/PressReleaseDetails.aspx?PressView=default&pid=446>.

4. The National Center for Education Statistics. Rural education in America: Definitions [Available from: <https://nces.ed.gov/surveys/ruraled/definitions.asp>.

5. Jackson AM, Mullican LA, Tse ZTH, Yin J, Zhou X, Kumar D, et al. Unplanned closure of public schools in Michigan, 2015 – 2016: Cross-sectional study on rurality and digital data harvesting. Journal of School Health. 2020:In press.