Brief Report

Reactions to foodborne *Escherichia coli* outbreaks: A text-mining analysis of the public's response

Elizabeth M. Glowacki PhD [a,*], Joseph B. Glowacki MS [b], Arnold D. Chung MA [c], Gary B. Wilcox PhD [c]

[a] *Department of Communication Studies, College of Arts, Media, and Design, Northeastern University, Boston, MA*
[b] *Philadelphia College of Osteopathic Medicine, Philadelphia, PA*
[c] *Stan Richards School of Advertising and Public Relations, Moody College of Communication, University of Texas at Austin, Austin, TX*

Key Words:
Foodborne illness
Social media
Twitter
Health communication
Food safety
Public health

Foodborne illnesses caused by bacteria are being reported at an increasing rate in the United States. We performed a text-mining analysis to look at nearly 13,000 tweets from two foodborne *Escherichia coli* outbreaks in 2018. Concerns from the public included staying informed about contaminated lettuce, recognizing signs of infection, and holding responsible farms accountable. At the end of the second outbreak, comments were focused on assessing symptoms, using the traceback process to locate outbreak sources, and calling for better food labeling practices.

© 2019 Association for Professionals in Infection Control and Epidemiology, Inc. Published by Elsevier Inc. All rights reserved.

There is evidence to suggest that foodborne disease outbreaks and illnesses may be increasing in the United States.[1] The Centers for Disease Control and Prevention (CDC) defines a "foodborne disease outbreak" as when 2 or more people get the same illness from the same contaminated food or drink.[2] The CDC also estimates that 48 million people get sick, 128,000 people are hospitalized, and 3,000 people die from foodborne diseases each year in the United States.[3] One key pathogen that has been consistently responsible for many foodborne illnesses and outbreaks is *Escherichia coli*. There are many forms of *E coli*, and whereas most are harmless and have a mutual relationship with the human body, there are certain types of *E coli* (known as strains) that can be particularly harmful.[3] Enterohemorrhagic *E coli* (EHEC) are strains capable of producing toxins, called Shiga toxin, and are therefore called Shiga toxin-producing *E coli* (STEC). *E coli* O157:H7 is a common serotype (subgroup) of EHEC in the United States and has been responsible for multiple foodborne disease outbreaks.[4]

E coli can cause a variety of serious complications, including bloody diarrhea, urinary tract infections, neonatal meningitis, and even kidney failure.[3] Additionally, treating certain forms and manifestations of *E coli*, such as foodborne *E coli* O157:H7, can be particularly challenging because antibiotics have been shown to be ineffective.[3] Treatment largely consists of awareness and avoidance of causative agents and supportive care through fluids and rest.[3] Given these difficulties, the public needs to be informed of preventive measures to avoid consuming foods contaminated with *E coli*, as well as steps for treating *E coli*.

Social media platforms such as Twitter can play a critical role in information dissemination, especially when the public is faced with a rapidly spreading health threat.[5] Twitter can also be used to monitor the public's questions, concerns, and responses to health threats such as *E coli*. Many people, including health experts, use Twitter as a source for up-to-date information and to gauge the public's reactions to public health crises such as the Ebola virus, the Zika virus, and opioid abuse.[5-7] However, few studies have looked at the public's awareness of and reactions to *E coli*. Considering that foodborne illnesses continue to be a growing problem and that there are ample opportunities for the spread of misinformation about this topic, public health officials need to identify public reactions to foodborne illness outbreaks as a stepping point for effective information dissemination.[8] Therefore, the goal of the current study is to examine the public's responses to 2 foodborne *E coli* outbreaks during 2018. We used a text-mining analysis to identify concerns and reactions to these outbreaks.

METHODS

Data acquisition

Tweets and retweets were collected using NUVI software (Brickfish, Chicago, Illinois) in conjunction with Twitter's Search

* Address correspondence to Elizabeth M. Glowacki, PhD, Department of Communication Studies, College of Arts, Media, and Design, Northeastern University, 360 Huntington Ave, Boston, MA 02115.
E-mail address: e.glowacki@northeastern.edu (E.M. Glowacki).
Conflicts of interest: None to report.

Table 1
Topics and descriptions—outbreaks 1 and 2

Topic ID	Topic	No. of tweets	Description
First outbreak			
1	Ecoli, + ecoli outbreak, + salad, + eat, romaine lettuce	556	Focus on consuming romaine lettuce, concerns about having already consumed contaminated lettuce
2	Coli,romaine, + eat, +salad, linked	489	Information about number of cases, scope of the outbreak, and the nationwide recall
3	Arizona, romaine lettuce, Yuma, + link, + romaine	488	Mentions of outbreak sources (Arizona and California), avoiding lettuce from these areas, confirming locations
4	State, people, + sicken, linked, CDC	432	Comments about spreading to multiple states, blaming food producers responsible, questions about *Escherichia coli* and its potentially deadly nature
5	E-coli, + e-coli outbreak, romaine lettuce, + eat, + romaine	367	More information about being careful with lettuce consumption
6	E-coli, tainted, likely, stores, longer	354	CDC waring the public to avoid romaine lettuce, concerns about feeling sick and knowing how to identify *E coli* infection
7	First, eating, + die, + eat romaine lettuce, person	263	More recommendations to avoid affected lettuce, outbreaks in multiple states, speculation about food contamination
8	Dead, five, + coli outbreak, Arizona, coli	228	Discussion about what *E coli* is and how it's caused, number of deaths from *E coli*
9	O157,:h7, multistate, + infection, + food	189	Focus on specific strain of *E coli* responsible for foodborne illness
10	Warning, + prompt, lettuce warning, Georgia, ecoli	169	Additional warnings about spread of *E coli* and avoiding lettuce
11	E coli case, double, NY, + case, + link	161	Comments about No. of cases, CDC's warnings to the public
12	American shopper, + shopper, American, + avoid, + coli outbreak	102	Knowing what to avoid when shopping for food
Second outbreak			
1	Romaine, + lettuce, + romaine lettuce, + eat, + outbreak	1,604	Mixed messages about lettuce safety—some sources say lettuce is safe to eat and others say the public should still avoid certain kinds of lettuce
2	Water, + Trump, + year, news, + cause	1,072	Impact of government and policies on public food safety, assuring the public government agencies are doing their jobs
3	CDC, romaine, + warn, + eat, coli contamination	894	Food safety updates from the CDC, encouraging public to check labels, still be careful about consuming romaine
4	Farm, + recall, + produce, + link, additional produce	884	California farms responsible have been identified, outbreaks spreading to different parts of the country
5	Food, + shutdown, + inspection, health	827	Concerns about impact of government shutdown on food safety and food inspections, Canadian Food Inspection Agency issuing warnings
6	Infection, CDC, coli infection investigation, + investigation, + add	758	Additional investigations into farms, consumers should check labels, CDC expects additional illnesses
7	Farm, + infect, + county, at least one, Barbara	586	Identifying water reservoirs and irrigation as *E coli* source on California farms, food labels should indicate origin of product
8	Santa, FDA, + source, + farm, + county	528	Pinpointing source of outbreak, explanation of traceback process
9	Leaf, farming, bros, inc, + green leaf lettuce	526	Not all lettuce cleared, new nationwide food recalls announced include cauliflower
10	Year, + customer, + recall, monte, del	520	Del Monte foods recalled out of an abundance of caution, includes canned corn
11	Linked, lettuce, likely, coli, California	402	New recall of infected romaine lettuce affects the United States and Canada
12	Cauliflower, red, romaine e, + sell, green	397	Red and green leaf lettuce recalled, knowing what symptoms of infection of look for, recommendations for food safety

ID, identification.

application programming interface.[9] Because our focus was on the public concerns related to *E coli*, key words included variations of the word "*E coli*" combined with what would be used with any of the following words:

advisory, bacteria, bacterium, CDC, Centers for Disease Control, prevention, ehec, enterohemorrhagic e. coli serotype o157:h7, FDA, food reservoirs, food safety, foodborne illness, foodborne transmission, hemolytic uremic syndrome, hus, o157:h7, outbreak, phac, Public Health Agency of Canada, romaine lettuce, shiga toxin, shiga toxin-producing e. coli, stec, United States Food and Drug Administration.

This study used textual analytics to identify topics and extract meanings contained in unstructured textual data. The Twitter messages were captured during 2 time periods, April 9, 2018 (the date

on which multiple people had first been identified as infected) through June 29, 2018, and October 8, 2018 through January 9, 2019 (the date on which the CDC issued its final update and declared the outbreak over). These time periods were selected because they capture the 2 major *E coli* outbreaks in the United States during 2018. Retweets/duplicate tweets were removed from both time periods. A total of 3,798 messages was used in the analysis of period 1 data, and 8,998 messages were used in the analysis of period 2 data.

Text analytics

The public-generated tweet content was analyzed using Text Miner (version 12.1; SAS, Cary, North Carolina). Text parsing is a function of SAS, which provides the ability to extract, filter, and store text information. Tweets are then assembled into related topics for introspection and insights from the unstructured data.[10]

Using text parsing, each message was divided into individual words. These words were listed in a frequency matrix, and words that did not contribute to a topic, such as auxiliary verbs and conjunctions, were excluded from the analysis. Next, a text filter node was used to exclude words that appeared in < 4 messages. The words initially included (and excluded) in the analysis were visually inspected to ensure accuracy and identify unrecognizable symbols for exclusion. With the inclusion criteria set, a text topic node was used to combine terms into 8-12 topic groups. After examining each of the created topics, the 12-topic solution most clearly illustrated the main themes and produced the final topic groups. Finally, the researchers inspected both the individual tweets and retweets of the final topic groups to interpret the themes.

RESULTS

Table 1 shows the 12 mutually exclusive topics generated from the public's tweets during the first foodborne *E coli* outbreak (April 2018), and the 12 mutually exclusive topics generated from the public's tweets during the second major outbreak (October 2018). The most significant concerns (illustrated by the number of tweets) were consuming contaminated romaine lettuce, locating the source of origin, worry surrounding a lack of knowledge about *E coli,* inconsistent messaging about lettuce safety, holding the responsible farms and food companies accountable, the usefulness of the trace-back process, and the need for clear food labeling. Reactions during the second time period were more focused on providing credible information and described specific steps to take to prevent and treat illnesses from *E coli*.

DISCUSSION

The public's comments during 2 foodborne *E coli* outbreaks suggest that there is concern regarding the symptoms and spread of *E coli*. These concerns specifically centered on not being informed in time of a contaminated lettuce recall, uncertainty about symptoms of an *E coli* infection, and receiving mixed messaged messages about food safety. There were noticeable differences in the nature of the content posted during the 2 time periods such that the tweets from the second outbreak were more focused on providing specific, concrete steps for taking action to address the issue. This included discussions about the traceback method, the importance of labels on food packaging, and raising questions about accountability for contaminated food sources.

One of the most prevalent topics to emerge was worry regarding the possibility that one might have already consumed contaminated lettuce. This was accompanied by concerns and questions about the number of states affected and discussions about the possible death rate. Organizations such as the CDC and the Food Safety and Inspection Service may play a role in alleviating feelings of worry and uncertainty by reaching out to the public more frequently on social media sites such as Twitter. Although the *E coli* outbreaks described here may not be as severe in nature, interacting with the public in this way provide with efficaciousness and reassurance, especially since

there were questions raised at the end of the second outbreak about the impact of the government shutdown on food inspections.

There are limitations to this study that should be addressed. Twitter only serves as 1 window into social media commentary, and multiple sources of information should be considered when evaluating discussions about a public health issue. Twitter has the potential to be an effective means of communication, but has not yet been fully harnessed by public health researchers and the government. Further, individuals who use Twitter are not necessarily representative of the larger population. Individuals who take the time to discuss *E coli* on a public forum may have different attitudes about their susceptibility to foodborne outbreaks. Finally, future studies would benefit from distinguishing between what experts like the CDC and the US Department of Health and Human Services are posting on social media and how their reactions to an *E coli* outbreak differ from those posted by laypeople or those who have little knowledge on the subject.

CONCLUSIONS

Our findings suggest that social media sites such as Twitter can be used as a tool by public health agencies that wish to identify concerns about foodborne disease outbreaks. Monitoring these conversations can help public health officials respond to the public's uncertainty about *E coli* and possibly help stop the spread of misinformation. Foodborne illnesses continue to be a public health challenge, but can be addressed by disseminating accurate, up-to-date information. Public health officials can use the data made available to them by their constituents via this channel to advocate for recommended practices such as clear food labels and product recalls. Identifying the public's perceptions of the severity and their susceptibility to infections from contaminated foods can be helpful for mitigating widespread concern.

References

1. Centers for Disease Control and Prevention. Foodborne diseases active surveillance network (FoodNet). Available from: https://www.cdc.gov/foodnet/reports/prelim-data-intro-2017.html. Accessed February 6, 2019.
2. Centers for Disease Control and Prevention. Foodborne disease outbreak 2011 case definition. Available from: https://wwwn.cdc.gov/nndss/conditions/foodborne-disease-outbreak/case-definition/2011/. Accessed February 4, 2019.
3. Centers for Disease Control and Prevention. *E. coli (Escherichia coli)*. Available from: https://www.cdc.gov/ecoli/index.html. Accessed February 4, 2019.
4. Clearly TG. The role of Shiga-toxin-producing *Escherichia coli* in hemorrhagic colitis and hemolytic uremic syndrome. Semin Pediatr Infect Dis 2004;15:260-5.
5. Lazard AJ, Scheinfeld E, Bernhardt JM, Wilcox GB, Suran M. Detecting themes of public concern: a text mining analysis of the Centers for Disease Control and Prevention's Ebola live Twitter chat. Am J Infect Control 2015;43:1109-11.
6. Glowacki EM, Lazard AJ, Wilcox GB, Mackert M, Bernhardt JM. Identifying the public's concerns and the Centers for Disease Control and Prevention's reactions during a health crisis: an analysis of a Zika live Twitter chat. Am J Infect Control 2016;44:1709-11.
7. Glowacki EM, Glowacki JB, Wilcox GB. A text-mining analysis of the public's reactions to the opioid crisis. Subst Abus 2018;39:129-33.
8. Council to Improve Foodborne Outbreak Response. Guidelines for foodborne disease outbreak response. Available from: http://cifor.us/products/guidelines. Accessed February 6, 2019.
9. NUVI. Available from: http://www.nuvi.com. Accessed January 10, 2019.
10. Chakraborty G, Pagolu M, Garla S. Text mining and analysis: practical methods, examples, and case studies using SAS. Cary (NC): SAS Institute; 2013.