

Improving Foodborne Complaint and Outbreak Detection Using Social Media, New York City



Foodborne Illness

- Illness caused by the consumption of contaminated food
- Many different pathogens
- Vastly underdiagnosed and not reported to health departments
- Estimated 48 million episodes of foodborne illness in United States each year¹
- Eating food prepared outside of the home increases risk²

¹Scallan E, Griffin PM, Angulo, FJ, Tauxe RV, Hoekstra, RM. Foodborne Illness Acquired in the United States-Unspecified Agents. Emerging Infectious Diseases. 2011; 17(1): 16-22.

²Jones TF, Angulo, FJ. Eating in Restaurants: A Risk Factor for Foodborne Disease? Clinical Infectious Diseases. 2006; 43: 1324-1328.

Foodborne Illness Complaints in New York City (NYC)

- 25,000 restaurants and 15,000 food retailers
- Over 8.5 million residents; 78% report going out to eat \geq once per week³
- NYC's non-emergency information service - 311
 - Allows residents to submit various types of complaints, including food poisoning
 - Submitted by phone or online
- NYC Department of Health and Mental Hygiene (DOHMH) receives ~3,500 restaurant associated complaints each year
- Approx. 30 restaurant related outbreaks each year

³New York City Health and Nutrition Examination Survey: New York City Department of Health and Mental Hygiene. <http://nychanes.org/data/>
Accessed 10 July 2016.

Unreported Foodborne Illness Complaints

- 2011: Outbreak investigation identified multiple complaints on the restaurant review site, Yelp.com that were not submitted via 311
- Nine month pilot project launched in 2012
 - Found 468 reviews consistent with foodborne illness; identified 3 previously unreported outbreaks
 - Many reviewers unaware of 311
- Collaborated with Yelp and Columbia University to obtain daily feed of Yelp reviews and develop a machine learning program to classify and identify reviews pertaining to foodborne illness

⁴Harrison C, Jorder M, Stern H, Stavinsky F, Reddy V, Hanson H, Waechter H, Lowe L, Gravano L, Balter S. Using Online Reviews by Restaurant Patrons to Identify Unreported Cases of Foodborne Illness – New York City, 2012-2013. Morbidity and Mortality Weekly Report. 2014; 63(20): 441-445.

Yelp Integration

- Nightly process to pull data from Yelp via Amazon Web Services; Google geocoding used to select only for NYC restaurants
- Program developed by Columbia assigns a “sick score” indicating the likelihood that a review pertains to foodborne illness; ranges from 0-1.0 and is based on if the review indicates:
 - Food poisoning or symptoms of foodborne illness (vomit, diarrhea, etc.)
 - Multiple people became sick
 - Incubation period

Yelp Integration

- All reviews enter the Foodborne Illness Tracker application; those with sick score ≥ 0.5 manually classified by DOHMH staff
- Messages sent to those identified as complaints, requesting they contact DOHMH for interview
- Feedback data sent to Columbia

Yelp Example

8. Complaint **2225895**

User Name:
Restaurant

Date Created:
Comment Date:
Review Text:

Rating:
Sick Score:
View on Yelp
Setup OEI Interview:

0.886

**Pros: Good food and friendly service! Takes
Convenient location! Cons: Ate an early dinner
going to see Hamilton. Food poisoning hit during
leave during intermission because I couldn't
sick and shaking. Really bummed I missed the second
Hamilton! Maybe only go if you don't have \$300 theater
tickets that night. Heed my mistake.**

Is this a case of foodborne disease: Yes
Did it occur within the last month: Unknown
Did two or more people get sick: Yes
Action item: Urgent
Are there severe symptoms: Yes
What type of message was sent: Interview message

mm/dd/yyyy
✕Cancel Save

Start Investigation Add to Investigation Annotate

Yelp Response

	Number	Percent
Reviews	19720	100
Foodborne	9998	50.7
Messaged	5554	55.6
Responded	1315	23.7
Interviewed	682	12.3

Data for Yelp reviews from July 1, 2012- February 28, 2018

- Since 2012, 10 outbreaks identified using Yelp

Unreported Foodborne Illness Complaints

- Even with the inclusion of Yelp, still likely not receiving all complaints of restaurant associated foodborne illness in NYC
- Other jurisdictions have reported success in identifying foodborne illness complaints using Twitter
- We sought to:
 - Identify tweets indicating foodborne illness
 - Validate tweets using an online survey
 - Integrate completed surveys into foodborne illness complaint system

Twitter Integration

- Use publicly available Twitter application program interface (API)
- Data received every two hours via a targeted API query that searches for keywords pertaining to foodborne illness
- Tweets assigned a sick score (0-1.0); solely based on indication of illness, food poisoning, symptoms, etc.
- Location obtained via metadata; based on user account registration

Twitter Integration

- Tweets with sick score ≥ 0.5 manually reviewed by DOHMH staff
- Survey link tweeted back to users complaining of foodborne illness
- Completed surveys qualify as complaints
- Feedback data sent to Columbia

@NYCFoodborne

Home Moments Notifications Messages Search Twitter Tweet

FOODBORNE NYC Health

TWEETS 1,756 FOLLOWING 29 FOLLOWERS 34 LIKES 25 MOMENTS 0 Edit profile

NYC Foodborne
@NYCFoodborne
Investigating reports of food poisoning on Twitter, Yelp, and @nyc311. Stay #healthy New York!
New York, NY
www1.nyc.gov/site/doh/health...
Joined March 2015

Tweets Tweets & replies Media

You Retweeted
nycHealthy @nycHealthy · 1h
It's #HepatitisAwarenessMonth. Regularly washing your hands with soap can help prevent hepatitis A: on.nyc.gov/2qu9oyn #TipTuesday

Your Tweet activity
Your Tweets earned 1,368 impressions over the last week
View your top Tweets

Twitter Example


1. Complaint **2240122**

Tweet Text: **Got a bad case of food poisoning on Friday. How long before I'm not afraid to eat food again?**

Tweet Date:
Date detected:
User:
Tweet Geo-Location: Unknown
User location: Bronx, NY **Bronx, NY**

View on Twitter: [Twitter](#) **0.612**

Sick Score: 0.612

Setup OEI Interview: [Click here](#) 

Did someone get sick:
NYC-located: UNKNOWN
Notes:
Last annotated By:
On: NEVER


[Contact User](#) [Link to Restaurant](#) [Fill-in Survey](#) [Annotate](#)

2. Complaint **2239279**

Tweet Text: **I felt sick to my stomach hearing about Kim's story. She had 2 young children. Rich or not. Totally devastating. Folks were callus. Smdh.**

Tweet Date:
Date detected:
User:
Tweet Geo-Location:
User location:
View on Twitter: [Twitter](#)

Sick Score: 0.974

Setup OEI Interview: [Click here](#) 

Did someone get sick:
NYC-located: UNKNOWN
Notes:
Last annotated By:
On: NEVER

[Contact User](#) [Link to Restaurant](#) [Fill-in Survey](#) [Annotate](#)

We found your tweet about possible food poisoning and we would like to ask you some questions.

If you think you have food poisoning in New York City, please complete this form. The info will be sent to the NYC Dept. of Health and Mental Hygiene so they can take any necessary action.

What happened?

Please be as descriptive as possible, letting us know what food items you consumed, what time of day, what was your first symptom, and for how long you were sick.

* What restaurant was it?

e.g. Restaurant Name

* Where was the restaurant located?

e.g. 42-09 28th St, Queens, NY 11101

(Please provide address, if known, or cross streets)

* What happened?

e.g. I went to [Restaurant Name] on [Date] at [Time] with 4 friends. I ate the chicken salad sandwich and tomato soup. The next morning I experienced severe diarrhea and vomiting and felt sick for about 24 hours.

What day did you order food from the restaurant?

mm/dd/yyyy



How can the City reach you?

First name

Jane

Last name


Doe

Email

janedoe@example.com

Phone

212-555-5555

 Submit your report

Twitter Survey Response

	Number	Percent
Tweets reviewed	13928	100
Foodborne and NYC	2646	19.0
Survey sent	2044	14.7
Survey completed	34	1.7
Interview completed	13	38.2

-602

Data for tweets identified 11/29/16-1/10/2018

- Of 34 completed surveys:
 - 25 (73.5%) reported foodborne illness associated with NYC restaurant; none were reported via 311/Yelp
 - 15 (60%) provided complete contact information
 - 13 (86.7%) completed interviews

Twitter Public Response

	Number	Percent
Tweets sent	2044	100
Likes	106	5.2
Retweets	15	0.7
Replies	74	3.6
Survey link clicks	378	18.5
Profile views	258	12.6
Detail expansions	1173	57.4

Data for tweets identified 11/29/16-1/10/2018

Engagement Promotion

- Survey response rate very low (1.7%)
- Implemented changes to promote engagement and survey completion
 - Tweet more frequently to increase Twitter presence
 - Edit response message
 - Add infographic

Edit Response Message

- Please complete this survey to report illness to the NYC Health Dept: <https://t.co/L60xqSb7nE>
- **Sorry that you're sick!** Please complete this survey to report illness to the NYC Health Dept: <https://t.co/TRB0I54LIw>
- Hope you're feeling better! **Help the NYC Health Dept. prevent this from happening again:** <https://t.co/CfqbmkdXiy>
- Hope you are feeling better! Help the NYC Health Dept. prevent **#foodpoisoning** from happening again: <https://t.co/Lzb0NC3ngd>

Add Infographic



The NYC Dept. of Health found your tweet about possible food poisoning and would like to ask you a few questions.

If you think you have food poisoning in New York City, please complete this form. The info will be sent to the NYC Dept. of Health and Mental Hygiene so we can take any necessary action to prevent others from becoming sick.

Daily Report

- SAS is used to generate and email a daily report; aggregates data from 311, Yelp, and Twitter surveys
- Complaints are matched based on restaurant name and location
- Restaurants with multiple complaints within a 30 day period are flagged for investigation

Implementation Requirements

- Supported by grants from Alfred P. Sloan Foundation and National Science Foundation
- Collaboration with Columbia University Department of Computer Science
- Additional Staff
 - Programmer/developer
 - Epidemiologist
 - Student intern
 - Public health inspector
- Coordination with IT, ongoing maintenance

Challenges

- Technical difficulties
 - Firewall issues, access to data blocked by IT security
 - Public facing server with increased IT security required for survey
- Twitter
 - Geolocation data no longer provided in public API
 - Anti-spamming measures
 - Difficult to contact, less willing to collaborate
- Noise/background
- Engagement

Next Steps

- Evaluate improved classifier that was recently implemented for Yelp reviews
- Use Twitter feedback data to improve sensitivity and specificity of classifier
- Incorporate additional data sources
- Evaluate methods to increase response rates

Questions?

Katelynn Devinney, MPH
Bureau of Communicable Disease
NYC Department of Health and Mental Hygiene
kdevinney@health.nyc.gov