

# Evaluating Twitter as a Data Source for Foodborne Illness Outbreak Detection in New York City

Katelynn Devinney, MPH

Foodborne Disease Epidemiologist

New York City Department of Health and Mental Hygiene

# 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 year<sup>1</sup>
- Eating food prepared outside of the home increases risk<sup>2</sup>

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

<sup>2</sup>Jones 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)

- 24,000 restaurants and 15,000 food retailers
- Over 8.5 million residents; 78% report going out to eat  $\geq$  once per week<sup>3</sup>
- 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) identifies foodborne illness complaints from Yelp reviews
- DOHMH receives ~5,000 restaurant associated complaints each year
- Approx. 30 restaurant related outbreaks annually

<sup>3</sup>New 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

- Even with the inclusion of Yelp reviews, 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
- Collaboration with Columbia University Department of Computer Science
- Supported by grants from Alfred P. Sloan Foundation and National Science Foundation

# Twitter Integration

- Publicly available Twitter application program interface (API)
- Data received every two hours via targeted API query that searches for keywords pertaining to foodborne illness
- Text mining and machine learning program assigns sick score to tweets (0-1)
- Location obtained via metadata; based on user account registration and selects only for those potentially in NYC

# Twitter Integration

- Data enters Foodborne Illness Tracker application
- Tweets meeting threshold sick score of 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**

**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 1,756 FOLLOWING 29 FOLLOWERS 34 LIKES 25 MOMENTS 0

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](http://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

View on Twitter: [Twitter](#) 0.612

Sick Score: 0.612

Setup OEI Interview: [Click here](#)

Did someone get sick:   
NYC-located:   
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:   
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	12085	100
Foodborne and NYC	2299	19.0
Survey sent	1785	14.8
Survey completed	27	1.5
Interview completed	10	37.0

*Note: A red bracket indicates a difference of -514 between the 'Foodborne and NYC' count (2299) and the 'Survey sent' count (1785).*

Data for tweets identified 11/29/16-9/30/2017

- Of 27 completed surveys:
  - 20 (74.7%) reported foodborne illness associated with NYC restaurant; none were reported via 311/Yelp
  - 11 (55%) provided complete contact information
    - 10 (90.9%) completed interviews

# Public Response

	Number	Percent
Tweets sent	1785	100
Likes	93	5.2
Retweets	12	0.7
Replies	65	3.6
Survey link clicks	349	19.6
Profile views	232	13.0
Detail expansions	1084	60.7

Data for tweets identified 11/29/16-9/30/2017

# Engagement Promotion

- Survey response rate very low (1.5%)
- 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.

# 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

- Implement changes to increase survey response rates
  - Account verification
  - Add question
  - Provide additional information
- Use feedback data to improve sensitivity and specificity of classifier
- Continue to evaluate usefulness
- Incorporate additional data sources



# Acknowledgements

- Vasudha Reddy
- Sharon Balter
- HaeNa Waechter
- Lan Li
- Faina Stavinsky
- Daniel O'Hallorhan
- Jazmin Fontenot
- Chandrajeet Padhy
- David Howell
- Adile Bekbay
- Mark Ozman
- Suneethakumari Polepalle
- Sharon Greene
- Team Salmonella
- Luis Gravano
- Daniel Hsu
- Thomas Effland
- Lampros Flokas
- Alfred P. Sloan Foundation
- National Science Foundation
- Fund for Public Health NYC
- FoodCORE
- EHS-Net

Thank you!

Questions: [kdevinney@health.nyc.gov](mailto:kdevinney@health.nyc.gov)

# Extra Slides

# Yelp Integration

- Collaboration with Columbia University; 2012 pilot
- Nightly process pulls data from Yelp for NYC restaurants
- Text mining and machine learning program classifies reviews; assigns “sick score” indicating likelihood that review pertains to foodborne illness
- Reviews meeting threshold level sick score are evaluated and manually classified by DOHMH staff
- Messages sent to those identified as complaints, requesting contact DOHMH for interview
- Feedback data sent to Columbia