

Published in final edited form as:

Prev Med. 2014 October ; 67: 264–269. doi:10.1016/j.ypmed.2014.08.003.

Online Reports of Foodborne Illness Capture Foods Implicated in Official Foodborne Outbreak Reports

Elaine O. Nsoesie, PhD^{1,2,*}, Sheryl A. Gordon, MSc¹, and John S. Brownstein, PhD^{1,2,3}

¹Children's Hospital Informatics Program, Boston Children's Hospital, Boston, Massachusetts, 02215, USA

²Department of Pediatrics, Harvard Medical School, Boston, Massachusetts, 02215, USA

³Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montreal, Quebec, H3A 1A2, Canada

Abstract

Objective—Traditional surveillance systems only capture a fraction of the estimated 48 million yearly cases of foodborne illness in the United States. We assessed whether foodservice reviews on Yelp.com (a business review site) can be used to support foodborne illness surveillance efforts.

Methods—We obtained reviews from 2005–2012 of 5824 foodservice businesses closest to 29 colleges. After extracting recent reviews describing episodes of foodborne illness, we compared implicated foods to foods in outbreak reports from the U.S. Centers for Disease Control and Prevention (CDC).

Results—Broadly, the distribution of implicated foods across five categories was as follows: aquatic (16% Yelp, 12% CDC), dairy-eggs (23% Yelp, 23% CDC), fruits-nuts (7% Yelp, 7% CDC), meat-poultry (32% Yelp, 33% CDC), and vegetables (22% Yelp, 25% CDC). The distribution of foods across 19 more specific food categories was also similar, with spearman correlations ranging from 0.60 to 0.85 for 2006–2011. The most implicated food categories in both Yelp and CDC were beef, dairy, grains-beans, poultry and vine-stalk.

Conclusions—Based on observations in this study and the increased usage of social media, we posit that online illness reports could complement traditional surveillance systems by providing near real-time information on foodborne illnesses, implicated foods and locations.

© 2014 The Authors. Published by Elsevier Inc. All rights reserved.

Corresponding author: Elaine O. Nsoesie, Boston Children's Hospital, 1 Autumn St. #14, Boston, MA. 02215, onelaine@vt.edu, Phone: 857-218-5108, Fax: NA.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Conflict of interest

The authors declare no conflict of interest.

Financial disclosure

The authors of this paper report no financial disclosures.

Keywords

foodborne illness; foodborne diseases; disease surveillance; social media; gastroenteritis; population surveillance

INTRODUCTION

An estimated 48 million people experience foodborne illness in the United States each year (1). Most foodborne illnesses are associated with acute gastroenteritis (defined as diarrhea and vomiting) (2), but affected individuals can also experience abdominal cramps, fever and bloody stool (3, 4). Although there are several surveillance systems for foodborne illnesses at the local, state and territorial level, these systems only capture a fraction of the foodborne illness burden in the United States mainly due to few affected individuals seeking medical care and lack of reporting to appropriate authorities (3). One way to improve surveillance of foodborne illnesses is to utilize nontraditional approaches to disease surveillance (5).

Nontraditional approaches have been proposed to supplement traditional systems for monitoring infectious diseases such as influenza (6, 7) and dengue (8). Examples of nontraditional data sources for disease surveillance include social media, online reports and micro-blogs (such as Twitter) (6–9). These approaches have been recently examined for monitoring reports of food poisoning and disease outbreaks (5, 10). However, only one recent study by New York City Department of Health and Mental Hygiene in collaboration with researchers at Columbia University (28) has examined foodservice review sites as a potential tool for monitoring foodborne disease outbreaks.

Online reviews of foodservice businesses offer a unique resource for disease surveillance. Similar to notification or complaint systems, reports of foodborne illness on review sites could serve as early indicators of foodborne disease outbreaks and spur investigation by proper authorities. If successful, information gleaned from such novel data streams could aid traditional surveillance systems in near real-time monitoring of foodborne related illnesses.

The aim of this study is to assess whether crowdsourcing via foodservice reviews can be used as a surveillance tool with the potential to support efforts by local public health departments. Our first aim is to summarize key features of the review dataset from Yelp.com. We study reviewer-restaurant networks and assess degree distributions by state and year to identify and eliminate reviewers whose extensive reviewing might have a strong impact on the data. Furthermore, we identify and further investigate report clusters (greater than two reports in the same year). Our second aim is to compare foods implicated in outbreaks reported to the U.S. Centers for Disease Control and Prevention (CDC) Foodborne Outbreak Online Database (FOOD) to those reported on Yelp.com. Attribution of foodborne illness and disease to specific food vehicles and locations is important for the monitoring and estimation of foodborne illness, which is necessary for public policy and regulatory decisions (11–14).

METHODS

Data Sources

Yelp—Yelp.com is a business review site created in 2004. Data from Yelp has been used to evaluate the correlation between traditional hospital performance measures and commercial website ratings (24), and the value of forecasting government restaurant inspection results based on the volume and sentiment of online reviews (25). We obtained data from Yelp containing de-identified reviews from 2005 to 2012 of 13,262 businesses closest to 29 colleges in fifteen states (Table A.1). 5,824 (43.9%) of the businesses were categorized as Food or Restaurant businesses.

CDC—We also obtained data from CDC's Foodborne Outbreak Online Database (FOOD) (26) to use as a comparator. FOOD contains national outbreak data voluntarily submitted to the CDC's foodborne disease outbreak surveillance system by public health departments in all states and U.S. territories. The data comprises information on the numbers of illnesses, hospitalizations, and deaths, reported food vehicle, species and serotype of the pathogen, and whether the etiology was suspected or confirmed. Note, outbreaks not identified, reported, or investigated might be missing or incomplete in the system. For each of the fifteen states represented in the Yelp data, we extracted data from FOOD in which reported illness was observed between January 2005 and December 2012.

Analysis

Keyword Matching—We constructed a keyword list based on a list of foodborne diseases from the CDC and common terms associated with foodborne illnesses (such as diarrhea, vomiting, and puking) (Table A.2). Each review of a business listed under Yelp's food or restaurant category (Table A.5) was processed to locate mentions of any of the keywords. 4,088 reviews contained at least one of the selected keywords. We carefully read and selected reviews meeting the classification criteria (discussed in the next section) for further analysis.

Classification Criteria—We focused on personal reports and reports of alleged eyewitness accounts of illness occurring after food consumption (see Table 1 for examples). We concentrated on recent accounts of foodborne illness and eliminated episodes in the distant past, such as childhood experiences. For each relevant review, we documented the following information, if reported: date of illness, foods consumed, business reviewed, and number of ill individuals.

Bias and Cluster Analysis—Data bias could be introduced by false reviews from disgruntled former employees and competitors. Yelp has a process for eliminating such reviews. We therefore focused on identifying bias introduced by individuals with a large number of negative reviews compared to the median in the dataset using network analysis and visualization. If a reviewer had significantly more reports than the median, we would investigate the impact of including and excluding this individual from the analysis. We also identified and investigated restaurants with more than two foodborne illness reports in the same year, since most restaurants appeared to have one or two reports, and because the CDC

defines a foodborne disease outbreak as more than one case of a similar illness due to consumption of a common food (4, 27).

Comparison of food vehicles—We extracted food vehicles mentioned in the FOOD outbreak reports and the Yelp data according to the CDC convention of categorizing and grouping implicated foods (15, 16). Broadly, the taxonomy consisted of three major categories: aquatic animals, land animals and plants. These categories were hierarchically distributed into subcategories as shown in Figure 2. Initially, we grouped the data into five major categories: aquatic, dairy-eggs, fruits-nuts, meat-poultry, and vegetables. Based on observations from this grouping, we further analyzed nineteen more specific categories, capturing all the major food groups. The nineteen categories consisted of fish, crustaceans, mollusks, dairy, eggs, beef, game, pork, poultry, grains-beans, fruits-nuts, fungi, leafy, root, sprout, vine-stalk, shellfish, vegetables, and meat. The aquatic, shellfish, vegetables and meat consisted of all foods that belonged to these categories but could not be assigned to the more specific categories such as leafy, crustaceans, poultry, etc. We excluded the oils-sugars category since most meals include natural or processed oils and/or sugars.

Foods implicated in foodborne illness were either categorized as *simple* or *complex*. Simple foods consisted of a single ingredient (e.g., lettuce) or could be classified into a single category (e.g., fruit salad). Complex foods consisted of multiple ingredients that could be classified into more than one commodity (e.g., pizza). For example, if pizza were implicated in an alleged foodborne illness report, we documented three food categories: grains-beans (crust), vine-stalk (tomato sauce), and dairy (cheese). If a report included a food item not easily identifiable (such as a traditional dish), we used Google search engine to locate the main ingredients in a typical recipe (e.g., meat, vegetable, aquatic, etc.) and categorized the food accordingly.

To compare foods implicated by Yelp and the CDC, we focused on reports from 2006 to 2011, because the 2012 Yelp data were incomplete. We ranked the nineteen food categories separately for Yelp and FOOD, according to the frequency with which each food category was implicated per year. Food categories with the same frequency were assigned the average of their rankings. Correlations of the ranked food categories were assessed using Spearman's rank correlation coefficient, ρ . Analyses were performed in SAS 9.1.3 (SAS Institute, Inc., Cary, NC).

RESULTS

Data

De-identified reviews of 13,262 businesses closest to 29 U.S. colleges in fifteen states (Table A.1) were obtained from Yelp.com. Of the 13,262 businesses included in the dataset, 5,842 (43.9%) were classified as foodservice businesses. The data included all reviews from 2005 to 2012. The volume of yearly reviews and reports of foodborne illness increased linearly from 2005–2011, with the majority of data observed between 2009 and 2012 (see Figure 1).

538 (9.2%) foodservices had at least one alleged foodborne illness report resulting in 760 reports with mentions of foodborne diseases and terms commonly associated with foodborne illness (such as diarrhea, vomiting etc.). Each review containing at least one of the foodborne illness-related terms was carefully read to extract information on date of illness, foods consumed, business reviewed and number of ill individuals. Most individuals mentioned being sick recently, but only 130 (17.1%) indicated the actual date of illness. 12 (1.58%) individuals with an alleged illness mentioned visiting a doctor or being hospitalized, and 80 (10.5%) reports indicated that more than one individual experienced illness. Since each review includes the restaurant information, the data can be visualized and also used by public health authorities for further investigation. Similar visualizations can be constructed for various cities and locations based on data availability.

We also studied the characteristics of reviewers who submitted reports of foodborne illness to identify any “super-reporters”. The highest number of reports by a single individual was four and the median number of reports was one. Since most reviewers (99.5%) had only one or two reports of alleged illness, we did not need to perform the bias analysis outlined in the Methods or eliminate any reviewers from the analysis.

We disaggregated the data by state and found that California (n = 319), Massachusetts (n=109) and New York (n=57) had the most illness reports. Since the data were generated based on colleges, those in sparsely populated regions might have fewer restaurants and therefore fewer reviews. We observed six clusters of more than two illness reports implicating the same business between 2007 and 2012, however, in most cases, reports were observed in different years. The six restaurants were located in California (four), Georgia (one) and Massachusetts (one). Per Yelp, one of the restaurants has closed. Restaurant inspection reports (see Table A.3) for four of the restaurants suggested at least one food violation in the last four years. These violations included: contaminated equipment, improper holding temperature, and cleanliness of food and nonfood contact surfaces.

Implicated Foods

557 (73.3%) Yelp foodborne illness reports and 1,574 (47.4%) CDC FOOD outbreak reports included the foods consumed prior to illness. Of the 1,574 CDC outbreak reports, 383 (24.3%) identified the contaminated ingredient. Foods were categorized based on the CDC’s convention of categorizing and grouping implicated foods (see Figure 2) (15, 16). We initially focused on five major food categories: aquatic, dairy-eggs, fruits-nuts, meat-poultry, and vegetables (see Figure 3). The distribution of implicated foods across these categories was extremely similar with identical proportions observed for the dairy-eggs (23%), and fruits-nuts (7%) categories. The other food categories had a 1% to 4% difference between Yelp and CDC. We then further disaggregated the data by year and focused on nineteen specific categories based on Figure 2. Rankings of the frequency of the nineteen food categories (shown in Table A.4) were positively correlated, with a mean of 0.78. The correlations for 2006 through 2011 were 0.60, 0.85, 0.85, 0.80, 0.77, and 0.79, respectively, with $p < 0.01$ for each year. We also present the proportion of foods within each category in Table 2. Lastly, we focused on illness reports from 2009 through 2011 since the most illness reports were noted during this period, as previously stated. The most frequently implicated

groups for 2009–2011 were beef (6.30% Yelp, 9.12% CDC), dairy (11.67% Yelp, 13.30% CDC), grains-beans (29.19% Yelp, 19.73 % CDC), poultry (9.37% Yelp, 9.57% CDC) and vine-stalk (8.14% Yelp, 10.16 % CDC).

DISCUSSION

In this study, we assessed reports of foodborne illness in foodservice reviews as a possible data source for disease surveillance. We observed that reports of foodborne illness on Yelp were sometimes extremely detailed, which could be useful for monitoring foodborne illness and outbreaks. We also located clusters of reports for particular restaurants, some of which had health safety violations related to food handling and hygiene. This suggests that tracking reviews in near real-time could reveal clusters useful for outbreak detection. Most importantly, we found that foods implicated in foodborne illness reports on Yelp correlated with foods implicated in reports from the CDC. This could be useful for identifying food vehicles for attribution and estimation of the extent of foodborne illness. Additionally, institutions and foodservices are considered principal locations for foodborne outbreaks (3), and studies suggest that Americans are increasingly consuming food outside the home (17, 18), which could lead to increased exposure to pathogens associated with foodborne illness. Approximately 44% and 3-4% of outbreaks contained in the CDC FOOD dataset were suspected or confirmed to be associated with restaurants and schools, respectively. A better understanding of foods and locations typically implicated in reports of foodborne illness is therefore needed in order to improve surveillance and food safety.

Although this data source could be useful for monitoring foodborne illness, there are several limitations in the data and the analysis. First, the incubation periods differ for different foodborne diseases, which can lead to misleading reports on time and source of infection. Second, some reports are delayed by several weeks or months, which could be challenging for surveillance. Individuals could be encouraged to report symptoms in near real-time or indicate the suspected date of infection. Third, the zero percentages in Table 2 could be due to missing data from the Yelp.com reviews and/or from the CDC reports and should therefore be treated with caution. As a result, the reported correlations could also be affected by missing data, in addition to other factors (such as the scheme used in categorizing and grouping foods). Fourth, the term list used in extracting foodborne illness reports are limited to typical symptoms of gastroenteritis and foodborne diseases, thereby missing some terms and slang words that could be used to describe foodborne illness. In future studies, we will develop a more comprehensive list that includes additional terms to better capture reports of foodborne illness. Fifth, the data are limited to businesses closest to specific colleges implying only a sample of foodservices in each state were included in the dataset thereby limiting the conclusions that can be drawn from the comparison with the FOOD data, which although limited is aimed at statewide coverage of disease outbreaks. Sixth, the number of restaurants serving particular food items could influence the distribution of implicated foods across the food categories. For example, cities in the central part of the U.S. might be more likely to serve meat-poultry products compared to aquatic products. Consequently, individuals are more likely to be exposed to foodborne pathogens present in foods that are more regularly served, which could partially explain the implications of these foods in foodborne illness reports. Lastly, the CDC warns that the data in FOOD are incomplete.

However, this is the best comparator available for this analysis at a national scale. More detailed state or city-level analyses could further refine the evaluation of this online data source. The lack of near real-time reports of foodborne outbreaks at different geographical resolutions reinforces the need for alternative data sources to supplement traditional approaches to foodborne disease surveillance.

In addition, data from [Yelp.com](https://www.yelp.com) can be combined with data from other review sites, micro-blogs such as Twitter and crowdsourced websites such as Foodborne Chicago (<https://foodborne.smartchicagoapps.org>) to improve coverage of foodborne disease reports. Furthermore, although this study is limited to the United States, foodborne diseases are a global issue with outbreaks sometimes spanning multiple countries. We could therefore use a similar approach to assess and study trends and foods implicated in foodborne disease reports in other countries.

CONCLUSIONS

Social media and similar data sources provide one approach to improving food safety through surveillance (19). One major advantage of these nontraditional data sources is timeliness. Detection and release of official reports of foodborne disease outbreaks could be delayed by several months (20), while reports of foodborne illnesses on social media can be available in near real-time. A study by Pelat et al (21) illustrated that searches for gastroenteritis were significantly correlated with incidence of acute diarrhea from the French Sentinel Network. Other studies leveraging data from social media (such as Twitter) have been able to track reports of foodborne illnesses and identify clusters suggesting outbreaks (22, 23). Most individuals who experience foodborne illnesses do not seek medical care but might be willing to share their experiences using social media platforms. By harnessing the data available through these novel sources, automated data mining processes can be developed for identifying and monitoring reports of foodborne illness and disease outbreaks. Continuous monitoring, rapid detection, and investigation of foodborne disease outbreaks are crucial for limiting the spread of contaminated food products and for preventing reoccurrence by prompting changes in food production and delivery systems.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work is supported by a research grant from the National Library of Medicine, the National Institutes of Health (5R01LM010812-03).

Role of funding source

The funding source had no role in the design and analysis of the study, and writing of the manuscript.

References

1. CDC Estimates of Foodborne Illness in the United States. [Accessed July 16, 2013] Centers for Disease Control and Prevention. Available at: <http://www.cdc.gov/foodborneburden/>

2. Lucado J, Mohamoud S, Zhao L, Elixhauser A. Infectious enteritis and foodborne illness in the United States, 2010. *Statistical Brief #150*. 2013
3. McCabe-Sellers BJ, Beattie SE. Food safety: emerging trends in foodborne illness surveillance and prevention. *J Am Diet Assoc*. 2004; 104:1708–1717. [PubMed: 15499359]
4. Daniels NA, et al. Foodborne disease outbreaks in United States schools. *Pediatr Infect J*. 2002; 21:623–628.
5. Brownstein JS, Freifeld CC, Madoff LC. Digital disease detection--harnessing the Web for public health surveillance. *N Engl J Med*. 2009; 360:2153–2155. 2157. [PubMed: 19423867]
6. Aramaki, E.; Maskawa, S.; Morita, M. Proceedings of the Conference on Empirical Methods in Natural Language Processing. Association for Computational Linguistics, Edinburgh, United Kingdom; 2011. Twitter catches the flu: Detecting influenza epidemics using twitter. In; p. 1568-1576.
7. Yuan Q, et al. Monitoring influenza epidemics in China with search query from Baidu. *PLoS One*. 2013; 8:e64323. [PubMed: 23750192]
8. Chan EH, Sahai V, Conrad C, Brownstein JS. Using web search query data to monitor dengue epidemics: a new model for neglected tropical disease surveillance. *PLoS Negl Trop Dis*. 2011; 5:e1206. [PubMed: 21647308]
9. Madoff LC. ProMED-mail: an early warning system for emerging diseases. *Clin Infect Dis*. 2004; 39:227–232. [PubMed: 15307032]
10. Wilson K, Brownstein JS. Early detection of disease outbreaks using the Internet. *CMAJ*. 2009; 180:829–831. [PubMed: 19364791]
11. Nyachuba DG. Foodborne illness: is it on the rise? *Nutr Rev*. 2010; 68:257–269. [PubMed: 20500787]
12. Kuchenmuller T, et al. Estimating the global burden of foodborne diseases--a collaborative effort. *Euro Surveill*. 2009; 14 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19422776>.
13. Scallan E, Mahon BE, Hoekstra RM, Griffin PM. Estimates of illnesses, hospitalizations and deaths caused by major bacterial enteric pathogens in young children in the United States. *Pediatr Infect J*. 2013; 32:217–221.
14. Woteki CE, Kineman BD. Challenges and approaches to reducing foodborne illness. *Annu Rev Nutr*. 2003; 23:315–344. [PubMed: 12626692]
15. Painter JA, et al. Recipes for foodborne outbreaks: a scheme for categorizing and grouping implicated foods. *Foodborne Pathog Dis*. 2009; 6:1259–1264. [PubMed: 19968563]
16. Painter JA, et al. Attribution of foodborne illnesses, hospitalizations, and deaths to food commodities by using outbreak data, United States, 1998–2008. *Emerg Infect Dis*. 2013; 19:407–415. [PubMed: 23622497]
17. Nielsen SJ, Siega-Riz AM, Popkin BM. Trends in food locations and sources among adolescents and young adults. *Prev Med*. 2002; 35:107–113. [PubMed: 12200094]
18. Poti JM, Popkin BM. Trends in energy intake among US children by eating location and food source, 1977–2006. *J Am Diet Assoc*. 2011; 111:1156–1164. [PubMed: 21802561]
19. Newkirk RW, Bender JB, Hedberg CW. The potential capability of social media as a component of food safety and food terrorism surveillance systems. *Foodborne Pathog Dis*. 2012; 9:120–124. [PubMed: 22217109]
20. Bernardo MT, et al. Scoping review on search queries and social media for disease surveillance: A chronology of innovation. *J Med Internet Res*. 2013; 15:e147. [PubMed: 23896182]
21. Pelat C, Turbelin C, Bar-Hen A, Flahault A, Valleron A. More diseases tracked by using Google trends [letter]. *Emerg Infect Dis*. 2009; 15 Available at: <http://wwwnc.cdc.gov/eid/article/15/8/09-0299.htm>.
22. Sadilek, A.; Brennan, SP.; Kautz, HA.; Silenzio, V. Proceedings of the First AAAI Conference on Human Computation and Crowdsourcing, HCOMP, November 7–9, 2013. Palm Springs, CA, USA: AAAI 2013; 2013. nEmesis: Which restaurants should you avoid today?. ISBN 978-1-57735-607-3
23. Ordun C, et al. Open Source Health Intelligence (OSHINT) for foodborne illness event characterization. *Online J Public Health Inform*. 2013; 5 Available at: <http://ojphi.org/ojs/index.php/ojphi/article/view/4442>.

24. Bardach NS, Asteria-Peñaloza R, Boscardin WJ, Dudley RA. The relationship between commercial website ratings and traditional hospital performance measures in the USA. *BMJ Qual Saf.* 2013; 22:194–202.
25. Kang JS, Kuznetsova P, Choi Y, Luca M. Using text analysis to target government inspections: evidence from restaurant hygiene inspections and online reviews (Harvard Business School). 2013 Available at: <http://ideas.repec.org/p/hbs/wpaper/14-007.html>.
26. CDC Foodborne Outbreak Online Database. Foodborne Outbreak Online Database. Available at: <http://wwwn.cdc.gov/foodborneoutbreaks>.
27. Jones TF, Rosenberg L, Kubota K, Ingram LA. Variability among states in investigating foodborne disease outbreaks. *Foodborne Pathog Dis.* 2013; 10:69–73. [PubMed: 23249418]
28. Harrison C, et al. Using Online Reviews by Restaurant Patrons to Identify Unreported Cases of Foodborne Illness — New York City, 2012–2013. *MMWR.* 2014; 63(20):441–445. [PubMed: 24848215]

Highlights

- Foodborne illness reports on Yelp were sometimes extremely detailed
- Illness reports included details such as date of illness and foods consumed
- Foods implicated in Yelp were also implicated in CDC foodborne outbreak reports
- Restaurants with health safety violations were implicated in report clusters

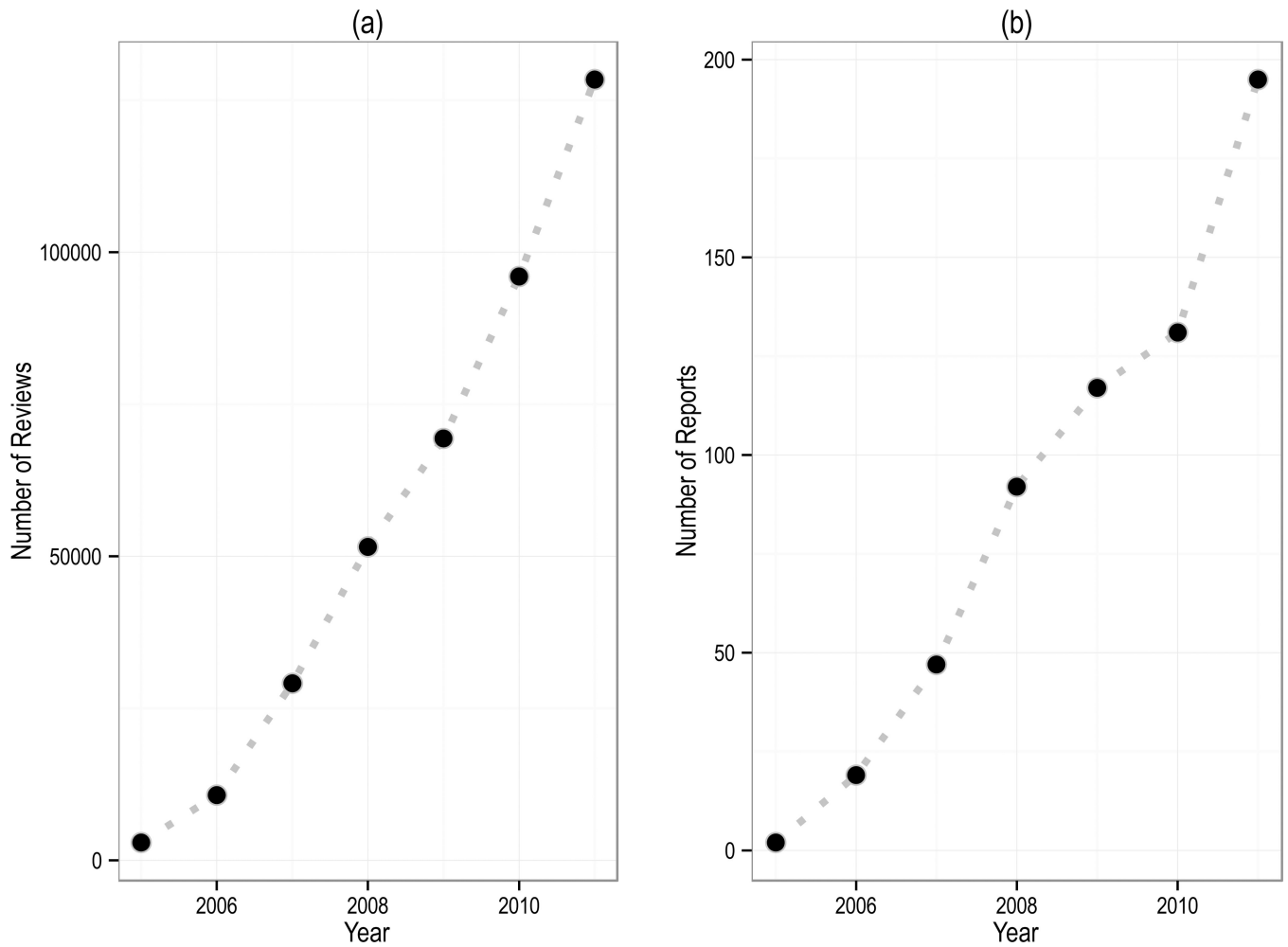


Figure 1. The volume of yearly reviews and reports of foodborne illness: (a) Yearly volume of foodservice business reviews and (b) yearly volume of foodborne illness reports. The increase in the number of reviews and illness reports could be due to the increase adoption of social media and related technologies for communication.

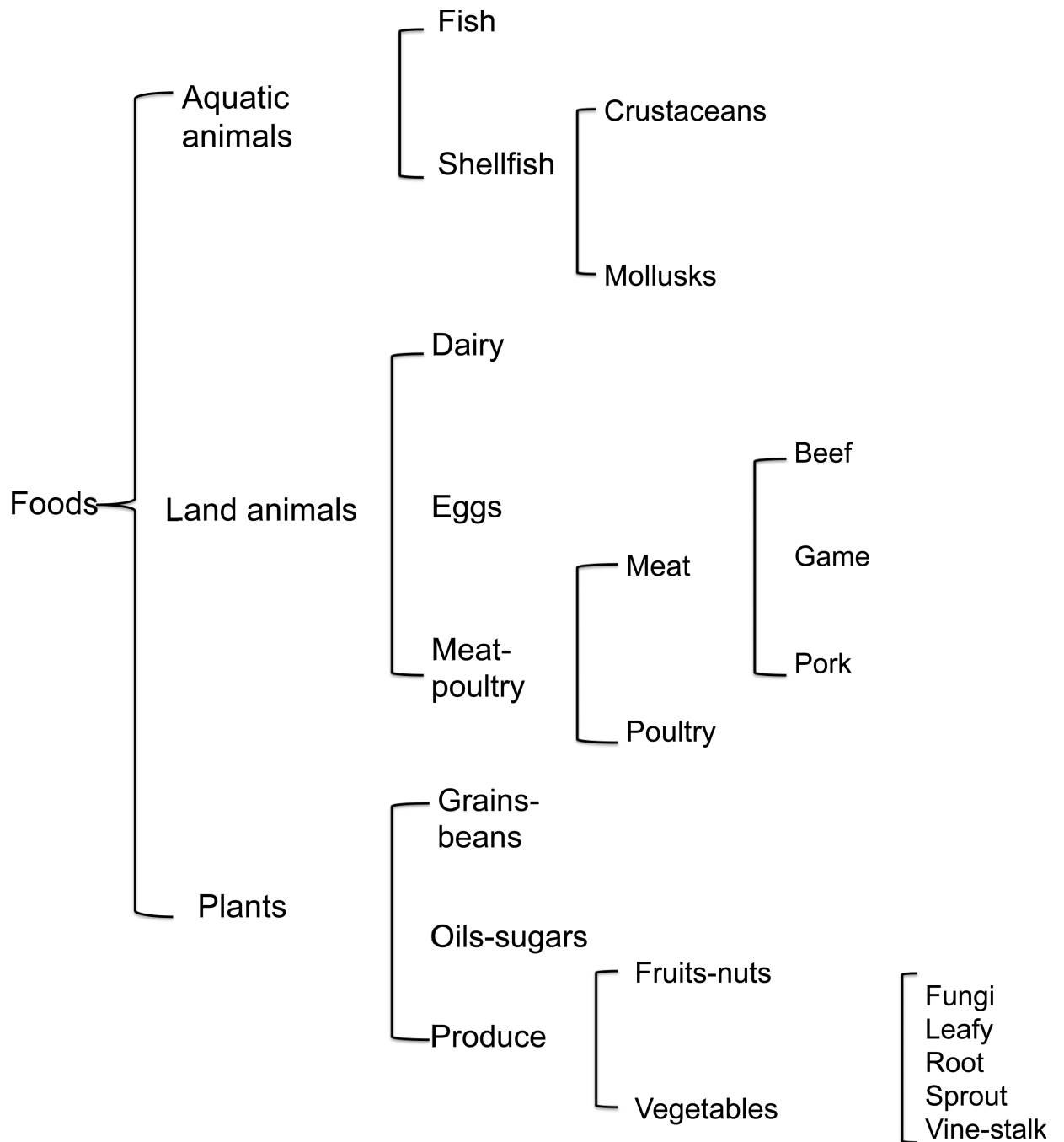
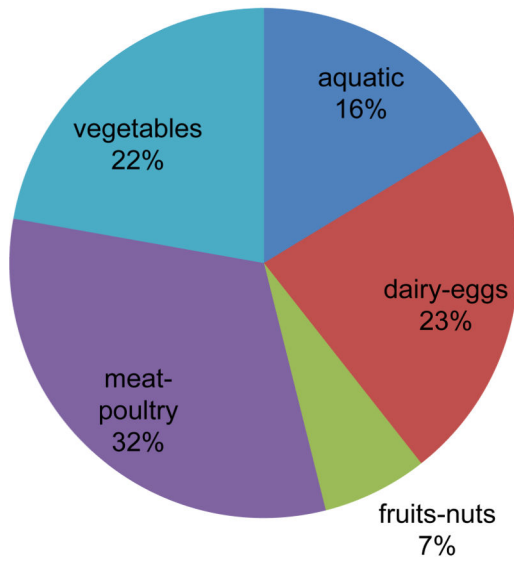


Figure 2. Hierarchy of food categories. Food categories are extracted from CDC publications on foodborne illness attribution.

(a) Yelp: 2005-2011



(b) CDC: 2005-2011

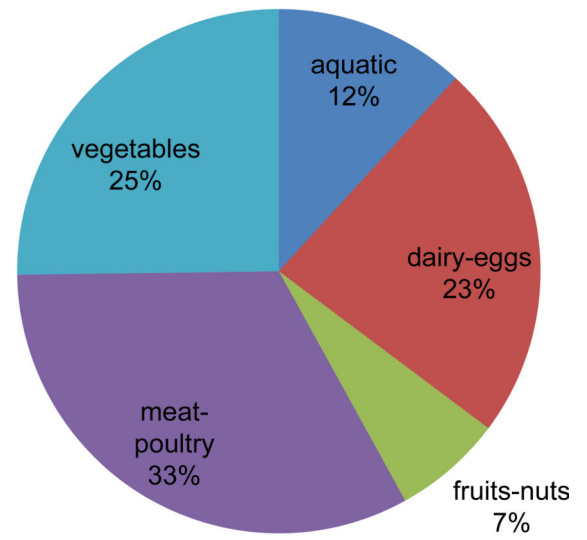


Figure 3. Categorization of foods implicated in Yelp reviews and CDC outbreak reports into five broad categories. Proportion of implicated foods in each of the five categories for Yelp (a) and CDC (b).

Table 1

Sample reports of alleged foodborne illness. Keywords are in bold.

<p>business_id=rb1ZR9xtCUgWjE19AU2y8w user_id=-1rqMSXzoQ7iYTRipDNhPA stars=1 date=2009-01-10</p> <p>I got HORRIBLE food poisoning from this place. If I could give it negative stars I would. And no, I'm not a lightweight: I eat Indian food all the time (and have even been to India) without getting sick. I know it was from the place because it was the only thing I had eaten that was out of the ordinary for the entire week. As it turns out, one of my coworkers had the same experience the same week from the same place, although unfortunately he only told me later and thus I was not able to avoid it. So, in summary, yikes! if you value your health, stay away!</p>
<p>business_id=279Aj_4Hd7EhoAZOip42g user_id=j0FOcXf6WQeVqlQVdAEt4w stars=1 date=2005-07-11</p> <p>I went here on a thursday during their free taco day...I dont know about you all but I dont find getting cold and hot flashes up my spine while puking at 4 in the morning very exciting. Me and my friend got f*cking food poisoning there! I didnt feel better until 5pm the next day. Tch...I am still mad at that...Damn taco meat must've been rat meat. I guess free food means free sh*t at their restaurant. I'll still go there for the two dollar beer special but the bartenders' attitude could be a little less b..chy, I dunno just a thought.</p>
<p>business_id=x52nVXRLWAwf3Rw76jcKMg user_id=MGL6GNXBjchbHx2D70MFbg stars=1 date=2010-01-02</p> <p>Epic Fail. Yesterday, I looked at the reviews and decided to post a four-star review, as I headed over to Zorba's to meet a few out-of-town friends." Why such a bad rap?" I thought -- and figured I'd help boost the reviews of this place that I'd been to twice before, and enjoyed. Well, I went there yesterday for lunch. Today, I woke up deathly ill, and proceeded to kick off 2010 by vomiting. Nice. I'm still sick but my family is taking care of me. The three of us had different items -- not sure what took us all down -- but we suspect Zorba's as we all went our separate ways and are all deathly ill today. Now I will add that I'm sure they run a good business and are decent people. But food poisoning is the one thing that cannot happen when you run a restaurant. I will touch base with them to see if they will do anything for us.</p>

Table 2

Percentage of implicated foods within each category by year.

Year	2006		2007		2008		2009		2010		2011	
	CD	Yelp	CD	Yelp	CD	Yelp	CD	Yelp	CD	Yelp	CD	Yelp
Beef	8.22 %	9.09 %	10.5 8%	2.70 %	11.8 1%	9.01 %	13.2 4%	3.17 %	6.45 %	6.67 %	8.29 %	8.24 %
Crustaceans	2.51 %	0.00 %	1.11 %	0.00 %	0.37 %	2.70 %	0.00 %	2.12 %	0.40 %	3.08 %	0.92 %	0.75 %
Dairy	13.2 4%	9.09 %	13.0 9%	17.5 7%	12.5 5%	11.7 1%	13.7 3%	15.8 7%	12.9 0%	11.2 8%	13.3 6%	8.99 %
Eggs	5.48 %	0.00 %	5.57 %	5.41 %	7.38 %	3.60 %	10.2 9%	4.23 %	9.68 %	3.08 %	4.15 %	5.24 %
Fish	4.79 %	0.00 %	4.74 %	2.70 %	5.54 %	6.31 %	3.92 %	2.12 %	5.65 %	3.08 %	3.69 %	4.49 %
Fruits-nuts	3.88 %	0.00 %	5.57 %	5.41 %	5.17 %	5.41 %	6.37 %	5.29 %	3.23 %	5.64 %	5.99 %	3.75 %
Fungi	0.91 %	0.00 %	0.84 %	2.70 %	0.37 %	0.00 %	0.00 %	1.06 %	0.40 %	1.54 %	0.92 %	0.75 %
Game	0.00 %	0.00 %	0.00 %	0.00 %	0.00 %	0.00 %	0.00 %	0.00 %	0.40 %	0.00 %	0.92 %	0.00 %
Grains-beans	17.1 2%	36.3 6%	18.9 4%	28.3 8%	18.8 2%	32.4 3%	18.6 3%	28.0 4%	20.9 7%	29.2 3%	19.3 5%	29.9 6%
Leafy	10.2 7%	9.09 %	5.29 %	1.35 %	10.3 3%	4.50 %	5.39 %	6.88 %	11.2 9%	3.59 %	8.29 %	5.62 %
Meat*	2.05 %	9.09 %	0.84 %	1.35 %	0.00 %	1.80 %	0.49 %	3.70 %	0.40 %	3.08 %	0.00 %	5.24 %
Mollusks	2.74 %	0.00 %	0.84 %	0.00 %	0.74 %	2.70 %	0.98 %	1.59 %	2.42 %	1.03 %	5.53 %	2.62 %
Pork	3.20 %	0.00 %	4.46 %	2.70 %	4.06 %	4.50 %	4.41 %	1.59 %	3.23 %	4.10 %	6.45 %	2.25 %
Poultry	9.59 %	27.2 7%	12.8 1%	10.8 1%	9.59 %	8.11 %	9.31 %	10.0 5%	9.27 %	9.74 %	10.1 4%	8.61 %
Root	4.34 %	0.00 %	3.34 %	5.41 %	0.37 %	3.60 %	1.47 %	3.70 %	1.21 %	5.13 %	0.92 %	5.24 %
Shellfish*	0.23 %	0.00 %	0.00 %	0.00 %	0.00 %	0.00 %	0.00 %	0.00 %	0.00 %	0.00 %	0.00 %	0.00 %
Sprout	0.00 %	0.00 %	0.84 %	0.00 %	0.74 %	0.00 %	1.47 %	0.00 %	1.21 %	0.00 %	1.38 %	0.75 %
Vegetables*	0.46 %	0.00 %	1.39 %	0.00 %	0.37 %	0.00 %	0.49 %	1.06 %	0.00 %	1.54 %	0.00 %	0.37 %
Vine-stalk	10.9 6%	0.00 %	9.75 %	13.5 1%	11.8 1%	3.60 %	9.80 %	9.52 %	10.8 9%	8.21 %	9.68 %	7.12 %

(*) indicates all foods in this category other than those specific categories listed in Figure 2.