May 2017

Determining the Effects of Social Media Monitoring to Identify Potential Foodborne Illness in Southern Nevada

Lauren Diprete
University of Nevada, Las Vegas, laurendiprete@gmail.com

DETERMINING THE EFFECTS OF SOCIAL MEDIA MONITORING TO IDENTIFY POTENTIAL FOODBORNE ILLNESS IN SOUTHERN NEVADA

By

Lauren Kirstin DiPrete

Bachelor of Science – Kinesiology
University of Nevada, Las Vegas
2009

A thesis submitted in partial fulfillment
of the requirements for the

Master of Public Health

Department of Environmental and Occupational Health
School of Community Health Sciences
Division of Health Sciences
The Graduate College

University of Nevada, Las Vegas
May 2017

UNLV | GRADUATE COLLEGE

Thesis Approval

The Graduate College
The University of Nevada, Las Vegas

April 6, 2017

This thesis prepared by

Lauren DiPrete

entitled

Determining the Effects of Social Media Monitoring to Identify Potential Foodborne Illness in Southern Nevada

is approved in partial fulfillment of the requirements for the degree of

Master of Public Health
School of Community Health Sciences

Brian Labus, Ph.D.
Examination Committee Co-Chair

Mark Buttner, Ph.D.
Examination Committee Co-Chair

Patricia Cruz, Ph.D.
Examination Committee Member

Clark Kincaid, Ph.D.
Graduate College Faculty Representative

Kathryn Hausbeck Korgan, Ph.D.
Graduate College Interim Dean

Abstract

Foodborne illness, commonly referred to as food poisoning, affects an estimated 1 in 6 Americans every year, despite the fact that it is entirely preventable. Many cases of foodborne illness go unreported; however, better reporting leads to faster health department response and containment. Social media monitoring, using software to identify trends in social media posts, is a novel new tool that has been tested in a variety of public health fields with promising preliminary results. The Southern Nevada Health District (SNHD) has employed social media monitoring software to identify potential foodborne illness within Southern Nevada. The purpose of this study was to determine the extent to which this tactic was effective in identifying high risk facilities that could be the source of disease, and then characterizing those high risk facilities based on the Food and Drug Administration's (FDA) five foodborne illness risk factors. This study revealed that restaurants flagged by the software performed worse on routine inspections than matched controls, both before and after adjusting the scores to account for every observation of risky food handling. Secondly, the data showed that in all inspections, contamination was the most frequently observed foodborne illness risk factor out of compliance. These findings show that social media monitoring can be a useful tool to guide inspectors to restaurants that may have an active lapse in food safety. Additionally, the fact that contamination was most frequently observed in both groups of restaurants shows that there is a need to educate food handlers and managers on effective contamination prevention techniques.

Acknowledgements

It takes a village to complete a thesis. An enormous thank you is owed to my committee whose thoughtful edits and astute guidance improved this document in a multitude of ways. Specifically, thank you to my chair Dr. Brian Labus who endured panicked messages and distraught phone calls and who gave this accurate and eloquent feedback on a since deleted sentence: “I kind of hate this.” Thanks to my co-chair, Dr. Mark Buttner for instructing me not only in this capstone work but also in both my very first class and very last class of my MPH career. A massive amount of gratitude to my family for providing the mental and emotional support needed to keep at this degree slowly and steadily over the course of five years. Large thanks to my friends for enduring the neglect that a graduate program requires. A sweet thank you is owed to Luke, my patient mutt who sacrificed many a walk to see the completion of this document. Many thanks go to the Southern Nevada Health District for sharing the data used in this research. And finally thanks to Adam Sadilek, Henry Kautz, Eric Portman, and the entirety of the original research team who turned the original pilot study from an idea to reality.

Table of Contents

Introduction

Foodborne illness, commonly referred to as food poisoning, is a group of illnesses caused by a variety of pathogens that lead to a number of symptoms, the most frequent being vomiting, diarrhea, and fever (FDA, 2014). Foodborne illness is estimated to affect 48 million people and leads to 3,000 deaths every year in the United States (CDC, 2016a). The majority of foodborne illness outbreaks originate from a meal eaten at a restaurant (CDC, 2016b). Researchers are forced to rely on estimates because foodborne illness is greatly underreported due to its often self-limiting nature in otherwise healthy individuals. The nature of foodborne illness transmission and infection is understood to be entirely preventable. For this reason, the United States Department of Health and Human Services has named Food Safety as a goal in their Healthy People 2020 initiative (Office of Disease Prevention and Health Promotion, n.d.).

According to the United States Food and Drug Administration (FDA), there are five risk factors that, if not addressed with active managerial control and safe food safety practices, will increase the likelihood of transmitting foodborne illness in a restaurant setting (2009). Active managerial control is any operator-instituted system for protecting food safety throughout a restaurant. Successful systems often focus on food handler knowledge and competency as well as instituting and monitoring proper procedures. The five foodborne illness risk factors requiring active managerial control are: 1) foods from unsafe sources, 2) poor personal hygiene, 3) inadequate cooking temperatures, 4) improper holding time and temperature, and 5) contaminated equipment and cross contamination (FDA, 2009). Examples of these five foodborne illness risk factors are provided in Appendix A.

Research has shown that foodborne illness outbreaks in food establishments can be related to one or more contributing factors. These contributing factors are arranged into three groups: contamination, survival, and proliferation (CDC, 2015). Contamination contributing factors include the food entering the facility contaminated during harvest, or bare hand contact by an ill food worker. Survival contributing factors include insufficient temperature during a cooking kill step or insufficient freezing for parasite destruction. Finally, proliferation contributing factors include insufficient hot holding temperatures, or a long, unmonitored cooling process allowing the food to linger in the temperature danger zone of 41°F - 135°F long enough to allow microbiological growth. To prevent contributing factors, a restaurant must control the five foodborne illness risk factors shown in Appendix A. In response to the science showing that outbreaks can be prevented by controlling the foodborne illness risk factors, many health departments tasked with regulating food establishments have based their restaurant inspection program on these risk factors. The Southern Nevada Health District (SNHD) is one such health authority.

Southern Nevada Health District Food Safety Program

SNHD is large local health department providing public health services to 2.1 million residents, which equates to 72% of the state's population (SNHD, n.d., b), as well as 42 million visitors annually (Las Vegas Valley Convention and Visitors Authority, n.d.). To accomplish this task, SNHD is comprised of 4 departments: administration, community health, environmental health, and clinical services. Preventing foodborne illness requires a team of cross-disciplinary personnel working in three of the four agency departments. Routine food inspections are accomplished within the environmental health department, while foodborne illness outbreak investigations require a larger response involving participants from three departments. Case finding and outbreak characterization is accomplished by the epidemiology program within community health. The community health department also includes the public health laboratory, which handles clinical and environmental sample testing. The site evaluations and manager interviews are accomplished by food operations and special programs staff within the environmental health department. Finally, the public relations for outbreak response is handled by the public information office within the administration department (SNHD, n.d., b).

In 2015, there were 358 cases of illness caused by reportable foodborne pathogens in Southern Nevada (SNHD, n.d., c). This includes salmonellosis, listeriosis, and campylobacteriosis, to name a few. This does not, however, include nonreportable pathogens, such as norovirus, which is the most common cause of foodborne illness (CDC, 2015). Reportable illnesses are any of a list of diseases that require medical personnel and laboratories to alert the local health department upon diagnosis as outlined in Nevada Revised Statute (NRS) chapter 441A (NAC, 2012a). Failure to report can be punishable by up to a $1000 fine per incidence. The purpose of reporting certain illnesses to the health department is so that SNHD can centralize the information, monitor trends, and respond appropriately to reduce the spread of illness within the community.

SNHD is granted authority to regulate food establishments to ensure active managerial control over the foodborne illness risk factors and proper food safety practices under NRS chapter 446 (NAC, 2012b). According to this law, SNHD is required to perform unannounced inspections on every permitted food establishment in its jurisdiction at least once every calendar year and up to as many times as necessary for the health department to determine safe operation and food handling. These inspections are based on the FDA's five foodborne illness risk factors discussed above. SNHD's risk-based inspections are formulated to give the highest demerits to critical violations that are likely to lead to customer illness, such as food handlers neglecting to wash their hands in between handling raw food and ready to eat foods. A lesser amount of demerits are given to major violations which, in conjunction with other factors, have the potential to lead to illness, such as a hand sink being out of hand soap necessary for proper handwashing. Finally, good food management practices are noted as well; however, these do not carry any demerit value. For example, a floor soiled with old food debris would fall under this category. It is not likely to contribute to foodborne illness at the time of the inspection, but if not addressed could lead to a pest infestation that could lead to food contamination (SNHD, n.d., a).

Under the SNHD inspection protocol, the demerits for a violation will only be counted once, no matter how many examples of that violation are found during the inspection. For example, the following food safety issues all fall under violation #14 which is worth 3 demerits: no sanitizer concentration left in sanitizer buckets, using a dish machine that is malfunctioning and unable to clean dishes properly, using a slicer that is soiled with old food debris, and dishes stored as clean being visibly soiled. Therefore, if a restaurant had all of these issues, it would only earn 3 demerits on its inspection report, which is a passing A grade. In contrast, a restaurant that had no sanitizer left in its sanitizer buckets during active food preparation (violation #14, 3 demerits), is cooling their food improperly (violation #8, 5 demerits), and has food left out on the counter unmonitored in the temperature danger zone (violation #9, 5 demerits) would earn an inspection score of 13 demerits which would result in a B downgrade (SNHD, n.d., a). Therefore, under the current SNHD system, restaurants that make multiple similar mistakes score better than restaurants that make the same amount of varied mistakes, which may not be an accurate reflection of the food safety within the restaurant.

SNHD routine food inspection uses a demerit system that results in grades. Zero to ten demerits on the unannounced, routine inspection leads to an A grade, eleven to twenty demerits results in a B downgrade, twenty-one to forty demerits earns a C downgrade, and finally, forty-one demerits or more, or the presence of an imminent health hazard, will lead to an immediate closure of a food facility. Because every establishment is expected to operate at an A level at all times, all restaurants earning a downgrade or closure are granted up to three weeks to make corrections and request a scheduled re-inspection. On the re-inspection, food establishments must earn zero to ten demerits, as well as have no repeat critical or major violations, to return to an A status. If a restaurant has a repeated critical or major violation, or earns more than ten demerits on a re-inspection, instead of passing, it will be downgraded one additional letter grade with B graded restaurants now earning a C grade (SNHD, n.d., a).

Despite a stringent prevention program, foodborne illness complaints and outbreaks do occur. SNHD receives complaints of possible foodborne illness from multiple sources. Most complaints of foodborne illness come either from medical professionals as explained above or from the sick individuals themselves via phone calls or online reporting. Because not all cases of foodborne illness are reported to the health district, SNHD recently implemented a novel tool to identify possible cases of foodborne illness in the community: social media monitoring software (DiPrete, 2016).

Social Media Monitoring

Social media is any website or application that allows its members to interact within an online community by posting user-generated content (Merriam Webster, n.d.). This user-generated content can vary widely depending on the social media site. For example, Snapchat allows users to post 10 second videos that will expire after 24 hours. Instagram is a venue for members to share pictures and other images. Facebook allows users to post status updates, or even share articles and news headlines. Twitter allows users to post updates called tweets that are restricted to a maximum of only 140 characters.

Social media usage by Internet users has been on the rise since its inception, and as of 2015, 65% of adults regularly used one or more forms of social media (Perrin, 2015). A common misconception is that social media use is only for younger generations. While it is true that the majority of users are younger (90% of young adults are active social media users), 77% of Internet users age 30-49 are on social media, and the fastest growing segment of social media users are older adults (Perrin, 2015). In 2015, 35% of adults 65 and older were on social media, tripling the amount of users in 5 years (Perrin, 2015). Those users who do participate in social media tend to be quite active as well. In fact, 76% of users with an account are logging into Facebook every single day (Greenwood et al., 2016). The large amount of users and their often prolific activity results in an extensive amount of data on social media. On Twitter alone, for example, over 500 million tweets are posted every day (Stricker, 2014).

Social media monitoring uses machine learning to recognize keywords or phrases on social media platforms and pairs that with the user’s spatiotemporal information to gain useful information from a community. This innovative tactic has been used in many facets of public health to better understand disease and to better protect the health of communities. By harnessing data from social media, researchers have been able to detect useful information quickly.

For example, researchers Signorini, Segre, & Polgreen (2011) retroactively studied Twitter posts during the H1N1 pandemic of 2009 and observed that signs of influenza-like illness on Twitter not only matched the official outbreak curve of the pandemic, but Twitter had this information up to 2 weeks earlier than traditional reporting. To discover this, researchers obtained tweets originating in the United States during the H1N1 pandemic activity. Their software, which was primed to identify multiple keywords indicating influenza-like illness, used these tweets to predict the location of illness over time based solely on the language of the tweets and their geolocation. When comparing the information from Twitter to traditional reports from the Influenza Sentinel Provider Surveillance Network and the CDC Influenza Reporting Regions information, researchers discovered that estimates from Twitter accurately matched the observed spread of disease. In addition to providing accurate information, the influenza information from the social media monitoring software was able to deliver its results one to two weeks before traditional methods. This is because it is able to bypass lags observed from the time it takes to diagnose and report illness (Signorini et al., 2011). Despite these promising results, it is important to note that while this method provides speed, it loses the specificity seen in traditional reporting methods. The information gained from Twitter is simply syndrome-based and unconfirmed; there is no medical diagnosis to confirm that the user tweeting about illness is actually experiencing the flu.

There are multiple instances of using Internet activity and social media monitoring to detect foodborne illness. For example, New York City’s Department of Health and Mental Hygiene used Yelp reviews to supplement their foodborne illness investigation program.

Harrison et al. (2014) implemented a software that scanned Yelp reviews for signs of foodborne illness. Health department staff then reviewed the 893 Yelp reviews flagged by the software and determined that 56% were likely describing a foodborne illness event. Of those describing foodborne illness, only 3% were ever officially reported through traditional means to the health department for investigation. Health department staff then initiated foodborne illness investigations from the information in the flagged Yelp reviews. From these investigations, health department staff was able to identify and respond to three distinct foodborne illness outbreaks that were previously unknown. In this study, researchers noted the value that social media monitoring brought to their foodborne illness investigation program; however, the time required to sift through results to find meaningful information was listed as a limitation. Further research will be required to refine the system to be more competent in flagging meaningful Yelp reviews (Harrison et al., 2014).

Yelp may also have the ability to assist food inspectors in assessing risk within a restaurant. One study in New York City compared a restaurant's Yelp score with its average health inspection score of the last four years. Park et al. (2016) found a correlation between poor Yelp scores and poor health inspection scores among chain restaurants. However, the same correlation did not hold true in independent restaurants. Another study attempted to answer the question of whether poor Yelp reviews can predict poor sanitation of a restaurant, and found promising results. While the previous study looked at Yelp scores alone, Schomberg et al (2016) utilized a more refined system that included the Yelp score, various keywords indicative of foodborne illness or poor sanitation within the review itself, and the individual Yelp review's usefulness rating, which is determined by a proprietary Yelp algorithm. When incorporating all of these sources of information to the social media monitoring software, researchers found more promising results. In fact, 77% of restaurants flagged by this social media monitoring software in New York City were found to have critical food safety violations in their health inspection reports, resulting in a B grade or worse (Schomberg et al., 2016).

In addition to Yelp, Twitter is a popular source for data mining to identify possible foodborne illness. The Chicago Department of Public Health used a program called Foodborne Chicago to identify tweets originating within Chicago city limits that contained the keyword "food poisoning." The software then sent an automatic response to the user with a link for the online foodborne illness reporting form. Researchers found that after instituting this software, reports of foodborne illness went up and the subsequent inspections identified more lapses in food safety than random, routine inspections (Harris et al., 2014). This social media monitoring software was limited in the sense that it would only identify one phrase: food poisoning, so it is likely that it may have missed other users with foodborne illness who may have used different phrasing in their tweets.

Finally, researchers in New York found in a retrospective study that Twitter could be a valuable resource in identifying restaurants that are likely to have poor food safety practices. Sadilek et al. (2013) used a social media monitoring software called nEmesis that identified evidence of foodborne illness in the language of tweets and then incorporated geolocation tags to identify the restaurants with potential lapses in food safety. Then researchers obtained the health inspection reports and found that restaurants flagged by the software were more likely to have poor sanitation scores on their health inspection reports (Sadilek et al., 2013). The researchers of this study and SNHD teamed up to launch the software in Southern Nevada in 2015.

nEmesis Software

The software used in the study referenced above, nEmesis, was created by researchers at the University of Rochester. The goal was to mine social media for evidence of disease. Originally, the software was used to predict the spread of flu-like illness in New York City (Brennan et al., 2013). Then, the researchers moved on to using the software to locate potential foodborne illness in New York City in a retroactive study design (Sadilek et al., 2013). Most recently, the research team paired up with SNHD to test the software in a prospective study (Sadilek et al., 2016).

nEmesis uses the content in a user's tweet as well as their geolocation tags to identify and locate potential illness in the community. The software employs an advanced language model that uses a wide range of keywords allowing it to be able to interpret nuances and phrases, such as feeling "under the weather." The language model can also differentiate between feeling "so sick to my stomach" which is meaningful to researchers, from feeling "so sick of homework" which would be irrelevant (Sadilek et al., 2016). The software is also able to account for the incubation period of foodborne illness and to help more accurately identify which restaurant could have potentially been the source of illness.

More specifically, nEmesis works in the following way. A user will tweet, or post an update, from a restaurant. If the geolocation service setting on the phone is turned on, then nEmesis will be able to "snap" the individual to that restaurant they are currently in by using Google Place. Note that this original tweet does not have to be about foodborne illness symptoms or even food at all; nEmesis simply identifies that a particular person was at a certain restaurant at a given time. Then, nEmesis will follow the user for the next two weeks. If, during those next two weeks, the user tweets about foodborne illness symptoms, then the software will score that sick tweet and associate it with the original restaurant. Restaurants then earn sick scores based on the number and severity of sick tweets posted by users after visiting. If a user tweeted from multiple restaurants before posting a subsequent sick tweet, then that sick tweet would simply be associated with all the previous restaurants during the incubation period. This means that most restaurants in the valley have a low level of noise, and it would take multiple sick tweets from multiple users to elevate the sick score above a specific threshold for response (DiPrete, 2016).

nEmesis Social Media Monitoring at SNHD

To determine if the list of sick restaurants compiled by nEmesis was accurate in the sense that it was comprised of restaurants with uncontrolled foodborne illness risk factors that could lead to foodborne illness, SNHD conducted a three month matched case control study (Sadilek et al., 2016). In the study, the restaurants flagged by the software were matched with similar restaurants based on area of town (using predefined geographic divisions created by SNHD) and type of restaurant. Different parts of town and different types of restaurants each have their own unique challenges, so researchers held these parameters the same to make fair comparisons. For example, if nEmesis flagged a fast food restaurant on the West side of town was matched with a fast food restaurant on the West side of town as a control. Flagged restaurants in hotels or food courts were excluded due to the difficulty in accurately identifying the correct restaurant in such a highly concentrated area. Blinded inspectors were then instructed to perform routine inspections at both the flagged restaurant and the control restaurant. The same inspector conducted both inspections for each pair to eliminate bias stemming from potential differences between inspectors (DiPrete, 2016).

Over the course of three months, 72 flagged restaurants were inspected as well as 72 controls. The results showed that adaptive inspections based on the software were 64% more likely to result in a C downgrade and on average earned 50% more demerits than controls (Sadilek et al., 2016). Therefore, researchers concluded that the social media monitoring software was effective in identifying restaurants that had the potential to spread foodborne illness and would benefit from intervention from SNHD inspectors.

Research Questions and Hypotheses

The purpose of this study was to determine the extent to which the software was effective in identifying high risk facilities that could be the source of disease, and then characterizing those high risk facilities based on the Food and Drug Administration's five foodborne illness risk factors. In this study, the data resulting from 70 pairs of inspection reports from the Sadilek et al. (2016) matched case control study conducted from January through March of 2015 were examined in a novel way in a secondary data analysis. While the previous analysis determined that using social media monitoring in Southern Nevada has provided a useful tool to guide inspectors to restaurants that are more likely to have active lapses in food safety, there are two areas that research did not explore. First, the previous study by Sadilek et al. (2015) did not consider the violations that may have been observed multiple times in a restaurant, but due to the current SNHD inspection structure are only counted once in the report. Secondly, the previous study did not evaluate the role of the five foodborne illness risk factors. In this study, both of these areas were explored using data from the original Sadilek et al. (2016) study.

Research Question 1: Adjusted Score

Did the social media monitoring software identify restaurants with more food safety issues than the controls based on adjusted scores?

H1o: The social media monitoring software did not identify restaurants with more food safety issues than the controls based on adjusted scores.

H1a: The social media monitoring software did identify restaurants with more food safety issues than the controls based on adjusted scores.

Research Question 2: Foodborne Illness Risk Factor

Did the social media monitoring software identify one foodborne illness risk factor that was cited more often than the others among restaurants flagged to be at high risk for spreading foodborne illness?

H2o: The social media monitoring software identified no difference in the number of times each risk factor was cited among restaurants flagged to be at a high risk for spreading foodborne illness.

H2a: The social media monitoring software did identify one risk factor that was cited more often than the others among restaurants flagged to be at a high risk for spreading foodborne illness.

Methodology

Adjusted Score

In this research, an adjusted score was created that that took into account the total number of food safety issues cited. SNHD currently only counts the amount of violation categories that were observed and disregards the amount of issues observed within each category. Each pair of full inspection reports was obtained and the total number of citations was counted, whether they fell under one violation category or many, to reach an adjusted score. Then, the adjusted score of the flagged facilities was compared with the matched control facilities.

Foodborne Illness Risk Factor

The second portion of the study identified whether the adaptive inspections based on the social media monitoring software resulted in one foodborne illness risk factor being cited more often than the others. Each violation number on the inspection report is based on a foodborne illness risk factor. Appendix B shows each violation and the base foodborne illness risk factor it is associated with (SNHD, n.d., a). The SNHD inspection report often lists multiple items under one violation number. For the purpose of clarity, Appendix B lists each item separately which is why the table may show multiple entries for the same violation number.

In Appendix B, it is evident that some violation numbers do not match directly with one risk factor. For example, violation number one contains a broad rule of operating within the parameters of the health permit. This rule could be violated in a multitude of different ways and the exact violation observed will explain which risk factor the violation falls under. When this violation was cited in the inspection report, the exact observation as described on the inspection report was examined and the appropriate risk factor was assigned on a case-by-case basis. Another example is violation 21 where the SNHD inspection report requires the manager to be knowledgeable in food safety. This violation ultimately correlates to all 5 risk factors. In the event that the manager is unknowledgeable, it is likely that there will be other food safety violations observed during the inspection. When this violation was cited on a report, it was excluded from further analysis in the study, relying on subsequent food safety violations with exact examples and direct risk factor correlations to relay the severity of food safety lapses. Finally, there are multiple violation numbers on the inspection report that are simply legal requirements not relating to a risk factor. These were also excluded for the purpose of this study. Violation number 23, for example, refers to adherence to the Nevada Clean Indoor Air Act. A restaurant that allows smoking is against the law, and would be cited as such; however, it has no bearing on food safety or this study, so for that reason it was excluded.

By using this method, the inspection report observations were translated from SNHD violations to their root foodborne illness risk factors. The number of times each risk factor was observed in each inspection was totaled. Then, the number of times each risk factor was cited in the flagged facilities was compared with each time it was cited in the control facilities using a matched case control analysis.

Results

Adjusted Score: Descriptive Epidemiology

Data from 70 pairs of restaurants, 140 restaurants total, were examined. An increase in the median number of demerits was observed in both the control and flagged groups when transforming from original to adjusted scores, as illustrated in Figures 1 and 2. Medians were used as a more accurate measurement since the data were not normally distributed. The median of the control group shifted slightly from a score of 6 to a score of 7 after adjustment. The median of the flagged group shifted more prominently from a score of 9 to a score of 12.

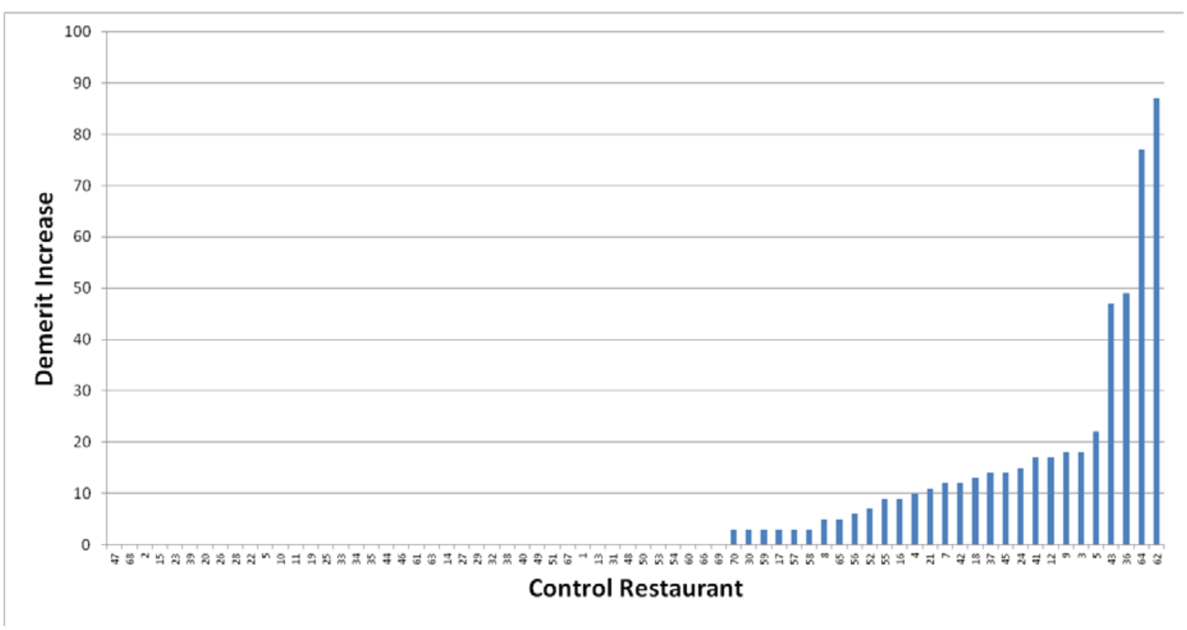

Figure 1. Difference in Scores after Adjustment for Control Restaurants

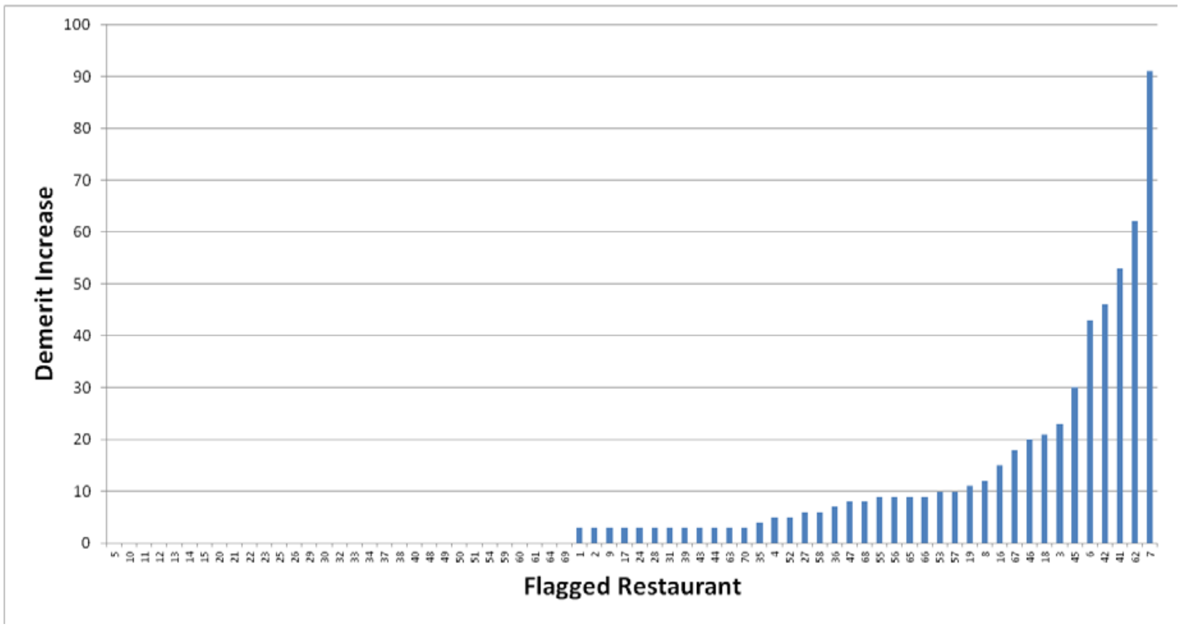

Figure 2. Difference in Scores after Adjustment for Flagged Restaurants

Adjusted Score: Statistical Analysis

Due to the matched case control study design of the original data, a pairwise analysis was necessary. A Related Samples Wilcoxon Signed Rank Test was utilized in order to account for the non-parametric distribution of data observed. This test allowed for the comparison of the median of differences between control adjusted scores and flagged adjusted scores.

It was determined that there were 38 positive differences, 21 negative differences, and 11 ties among the pairs. Figure 3 illustrates the difference in adjusted scores between the control restaurant and the flagged restaurant in each restaurant pair. With a significance level of 0.05, the p-value was 0.031, allowing for rejection of the null hypothesis. The social media monitoring software successfully identified restaurants with more food safety issues than control restaurants using an adjusted inspection score.

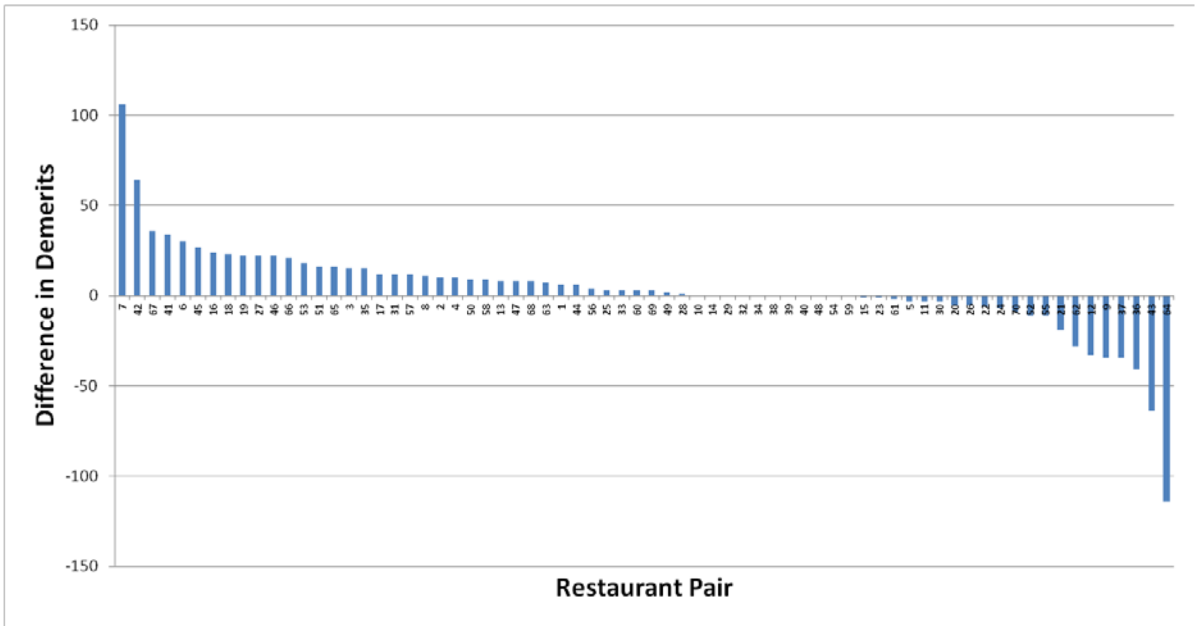

Figure 3. Difference Between Control Adjusted Score and Flagged Adjusted Score within Restaurant Pairs

Foodborne Illness Risk Factor: Descriptive Epidemiology

Two of the five risk factors were observed so infrequently in the inspections that they were excluded from analysis: Approved Source was observed three times out of 140 restaurants and Cook Temperatures was observed only one time. Figure 4 displays the total observations for each risk factor. Contamination was cited most frequently in both groups: 215 times in flagged restaurants and 194 times in control restaurants. Contamination concerns were also observed more times in an individual restaurant than the other risk factors (3.1 times per flagged restaurant and 2.8 times per control restaurant). Finally, contamination issues were observed in more restaurants than any other risk factor: in 55 flagged restaurants and in 55 control restaurants. Table 1 summarizes these observations.

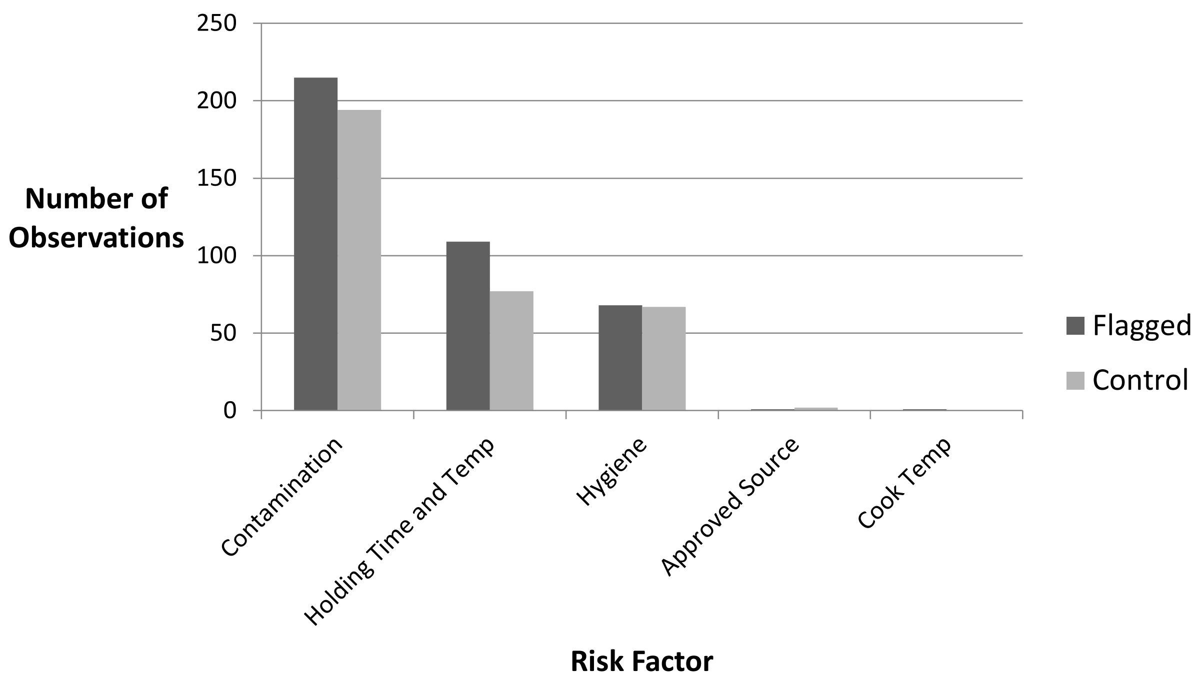

Figure 4. Total Number of Observations of Each Risk Factor

Table 1. Breakdown of Risk Factor Observations

Risk Factor	Flagged Group			Control Group		
	Total observations	Average times observed in a restaurant	Number of restaurants observed in	Total observations	Average times observed in a restaurant	Number of restaurants observed in
Hygiene	68	0.97	37	67	0.96	26
Holding Time and Temp	109	1.6	38	77	1.1	26
Contamination	215	3.1	55	194	2.8	55

Foodborne Illness Risk Factor: Statistical Analysis

For the same reasons outlined above, a Related Samples Wilcoxon Signed Rank Test was utilized. This test was performed individually for each risk factor in order to compare the median of the differences between the number of observations in the flagged restaurants and that of the control restaurants. For each risk factor, the difference in the number of times a risk factor was observed between flagged and control restaurants was not statistically significant so in each case the null hypothesis was retained. The p-values for each risk factor were as follows: Personal Hygiene was 0.755, Contamination was 0.359, and Holding Time and Temperature was 0.120.

Discussion

Adjusted Score

The first section of this study revealed that the software did, in fact, lead inspectors to restaurants with more food safety issues than controls based on adjusted scores. The original study using the nEmesis software by Sadilek et al. (2016) observed similar results when analyzing the original scores of control and flagged restaurants. The use of adjusted scores in this study showed a wider gap between control and flagged than what was observed previously. Specifically, adjusted scores showed a median score of 7 demerits for control and 12 for flagged, while original scores showed a median of 6 for control and 9 for flagged. Therefore, the gap between flagged and controls widened by nearly 50%, from 3 to 5 demerits, when using adjusted scores.

Adjusted scores were higher than original scores overall in both groups. The method of accounting for every food safety issue observed provides a more accurate reflection of a restaurant's control over food safety. In some instances, the difference in score was exceedingly large after adjusting: up to 91 demerits. In fact, using the adjusted score, 17 of the 140 restaurants would have been closed for exceeding 40 demerits, SNHD's threshold dictating an immediate closure of a restaurant. In actuality, only one restaurant was closed during the study period. It was also observed that as the original score increased, the gap between the original score and the adjusted score increased as well. This shows that the struggling restaurants may be struggling more than originally thought.

When looking at both groups of restaurants as displayed in Figures 1 and 2, the majority, 51%, had no change in their scores after adjustment. These restaurants typically had successful inspection results. Only 23% of restaurants had a significant change in score after adjustment of 10 demerits or more. This small group of restaurants was comprised of mostly poor inspection grades: 81% had an original score of a B-downgrade or worse. Therefore, the adjustment of scoring resulted in many passing restaurants continuing to pass and the struggling restaurants were left with a score that more accurately represents their food safety practices.

Restaurant inspection scoring is a complicated matter and there is no universal approach. While every jurisdiction bases their inspection program on the same FDA guidelines, their application of the Food Code and how to handle violations changes from place to place. There are many methods of scoring utilized throughout the country. San Francisco uses a point system where each restaurant starts with 100 points and each violation reduces the score, so high scores signify better food safety (San Francisco Department of Public Health, n.d.). In September of 2016 Washoe County, Nevada switched over from letter grade scoring similar to SNHD's to a stoplight pass/fail system with three inspection outcomes: green pass, yellow pass with conditions, or red closure (Washoe County Health District, n.d.; Kitchen, R., 2016). Due to the wide variety of strategies, it is common practice for health departments to periodically review their current protocols and consider ways to enhance their programs.

Foodborne Illness Risk Factor

Of the five foodborne illness risk factors, contamination was by far cited most often. This is perhaps in part because of the design of the inspection report. There are simply more ways for a contamination violation to occur so they are split up on the inspection report. For example, violation 11 covers contamination during storage, preparation, and service; violation 12 covers contamination by chemicals; violation 13 covers contamination by employees and customers;

and violation 14 covers contamination by equipment and surfaces. In short, there are more opportunities during an inspection for contamination to be cited than any other risk factor.

The fact that one risk factor was predominant in both the flagged and control groups shows that the flagged restaurants are not different from controls based on the types of risk factor issues observed. Instead, the difference between flagged restaurants and control restaurants lies in the amount of risk factor issues observed. This information means that the same education and intervention strategies designed for flagged restaurants can be beneficial for all restaurants since they are not inherently different in the types of violations observed.

Limitations

During the pilot test at SNHD, there was a delay between when the list of flagged facilities was received and when they were inspected. The delay was sometimes as long as 6 days between a restaurant being flagged by the software and inspected by SNHD staff. It is of course possible that conditions of the restaurant at the time of the potential exposure were drastically different than the conditions observed during the time of the inspection. For example, a sick worker could no longer be working, a broken refrigerator could have been fixed, or a vacationing manager could have returned to work. Additionally, in this system there is no confirmation of foodborne illness. Foodborne illness symptoms in a tweet can look similar to other diseases, such as Crohn’s disease, which is chronic and not contagious. Also, the software is accurate in its geolocation technology up to 50 feet. In the event of a small strip mall or a food court, for example, the accuracy of the restaurant selection suffers. Finally, while each set of restaurants was inspected by the same inspector, throughout the study multiple inspectors were involved. Despite required REHS (Registered Environmental Health/Sanitarian) certifications and ongoing training, there could be differences between how each inspector conducts an inspection.

Recommendations

The success of the nEmesis software in guiding inspectors toward struggling restaurants is yet another example of social media monitoring being used by public health professionals to protect their communities. This novel category of technology is a growing trend and it has the capability to become a valuable tool for health departments everywhere. In the battle against foodborne illness, this software can help alert inspectors to struggling restaurants faster in order to make corrective actions quicker and potentially prevent outbreaks. It is recommended that Health Departments consider incorporating similar software to help identify and respond to potential foodborne illness in all communities. Future research in this field should attempt to identify what effect, if any, utilizing this type of software has on the rates of confirmed foodborne illness in the community.

The use of adjusted scores in this study illustrated the true management of food safety in the restaurants. It is recommended that SNHD work with the food service industry, food service customers, and other stakeholders to update the scoring system in a way that incorporates every food safety issue observed during a restaurant inspection. In the event that the exact adjustment used in this study is implemented by SNHD, it is recommended that SNHD evaluates appropriate thresholds for downgrades. For example, with a new scoring system, perhaps 12 demerits or less might be a more reasonable cutoff for an A grade rather than the current 10 demerits.

In both the control and flagged groups, contamination was by far the most commonly observed risk factor out of compliance. SNHD currently has a Food Establishment Resource

Library website designed to help food workers and managers learn and enact good food safety practices. This online library offers fact sheets, food service logs, and sample procedures covering a variety of food safety topics in multiple languages. It is recommended that SNHD develop additional education resources specifically on the topic of contamination prevention to help restaurants manage this risk factor.

In response to identifying that contamination is the biggest food safety concern in restaurants, it is recommended that further research identifies novel and effective ways of preventing contamination. Prevention measures could come in the form of altering the built environment, improving management training and communications, or even the development of novel products such as anti-bacterial surfaces or improved sanitization equipment. A summary of recommendations is provided in Table 2.

Table 2. Recommendations by Target Group

	Target Group			
Recommendation	SNHD	All Health Departments	Food Service Community	Research Community
Incorporate social media monitoring technology to identify and respond to potential foodborne illness		X		
Identify the effect that this software has on the rates of illness within communities				X
Adjust scoring system to include all observations of food safety violations	X			
Develop and release training materials on preventing contamination violations	X	X		
Work to control the risk of contamination violations within the kitchen			X	
Develop systems and products to prevent contamination in the kitchen			X	X

Conclusion

The ultimate goal of this study was to identify additional ways to prevent foodborne illness within the community by using information from social media. Because foodborne illness is greatly underreported, novel methods are being tested to identify more cases sooner, to respond faster, and to curb the onset and size of outbreaks. In the first objective, the social media monitoring software was tested and shown to be effective in leading inspectors to restaurants with more food safety issues. This novel method of surveillance could become a new tool for health departments to identify restaurants in need of corrective action. With many health departments suffering from a lack of staffing and funding, this could be a way to allocate inspectors where and when they are needed most. In this study, an adjusted score was implemented which accounted for all food safety issues observed during an inspection. This adjusted method of scoring is recommended to provide the restaurant and public with a more accurate reflection of the food safety control in each restaurant. The second objective was to identify if one risk factor was more prominent in software-flagged restaurants. The data showed that contamination issues were observed most often in both groups. Due to these findings, it is recommended that education efforts for food workers be focused on preventing contamination. Further research should examine novel, effective ways of preventing contamination in the restaurant setting and to examine what effect a social media monitoring program has on the level of confirmed foodborne illness in a community.

Appendix A: The Five Foodborne Illness Risk Factors

Risk Factor Description	Risk Factor Example	Risk Factor Control Examples
Poor Personal Hygiene	-Improper handwashing and/or not washing hands when necessary -Bare hand contact with ready-to-eat foods -Food handlers working while ill with the following symptoms: 1. diarrhea 2. vomiting 3. sore throat with fever 4. jaundice 5. infected cuts, sores, or burns on hands or wrists	-Wash hands with warm, soapy water for 15-20 seconds and dry with a paper towel. -Wash hands when required: before putting on gloves, after removing gloves used for raw meat, when switching tasks, after touching one's face -Use utensils or proper gloves use when handling ready to eat foods -Institute an employee health policy to prevent food handlers from working while ill
Foods from Unsafe Sources	-Food prepared in an unpermitted location -Foods obtained from an unapproved source -Receipt of adulterated food	-Use only permitted, commercial grade kitchens -Source all foods from reputable, permitted suppliers -Inspect food upon receipt for wholesomeness
Improper Cooking Temperatures/Methods	- Inadequate cook temperature or method -Inadequate reheat temperature or method -Inadequate freezing temperature or amount of time to kill parasites in foods eaten raw	-Log cook temperatures -Ensure reheated foods reach a minimum of 165°F for 15 seconds -Obtain documentation from suppliers that fish are frozen to proper parameters to ensure parasite destruction.

Risk Factor Description	Risk Factor Example	Risk Factor Control Examples
Improper Holding Time and Temperature	-Improper hot and cold holding of potentially hazardous/time and temperature controlled for safety foods (PHF/TCS) -Improper cooling of PHF/TCSs -Lack of date/time marking for ready to eat PHF/TCSs -Improper use of time as a control	-Keep cold foods below 41°F and keep hot foods above 135°F -Cool foods from 135°F-70°F within 2 hours and from 70°F-41°F within the following 4 hours -Label foods with 6 day use by date -Ensure foods on time as a control are monitored and any foods left after 4 hours are discarded
Food Contamination	-Use of contaminated/improperly constructed equipment -Poor employee practices -Improper food storage or preparation -Exposure to chemicals	-Ensure proper warewash methods: soapy wash water above 110°F, rinse, sanitize, air dry -Educate and monitor food handlers -Store raw food below ready to eat food -Store chemicals below and away from food or food contact surfaces

Appendix B: Relationship between Violations and Risk Factors

Violation #	Violation Description	Risk Factor
1	Verifiable time as a control approved procedure when in use.	Improper Holding Time and Temperature
1	Operational plan, waiver, or variance approved and followed when required.	Varies
1	Operating within the parameters of the health permit.	Varies
2	Handwashing (as required, when required, proper glove use, no bare hand contact of ready to eat foods).	Poor Personal Hygiene
2	Foodhandler health restrictions as required.	Poor Personal Hygiene
3	Commercially manufactured food from approved source with required labels.	Foods from Unsafe Sources
3	Parasite destruction as required.	Improper Cooking Temperatures or Methods
3	Potentially hazardous foods/time temperature control for safety foods (PHF/TCS) received at proper temperature.	Improper Holding Time and Temperature
4	Hot and cold running water from approved source as required.	Poor personal hygiene
5	Imminently dangerous cross connection or backflow.	Food Contamination
5	Wastewater and sewage disposed into public sewer or approved facility.	Food Contamination
6	Food wholesome, not spoiled, contaminated or adulterated.	Food Contamination
7	PHF/TCSs cooked and reheated to proper temperatures.	Improper Cooking Temperatures or Methods
8	PHF/TCSs properly cooled.	Improper Holding Time and Temperature

Violation #	Violation Description	Risk Factor
9	PHF/TCSs at proper temperature during storage, display, service, transport, and holding.	Improper Holding Time and Temperature
10	Food and warewashing equipment approved, properly designed, constructed and installed.	Food Contamination
11	Food protected from potential contamination during storage and preparation.	Food Contamination
12	Food protected from potential contamination by chemicals.	Food Contamination
12	Toxic items properly labeled, stored, and used.	Food Contamination
13	Food protected from potential contamination by employees and consumers.	Food Contamination
14	Kitchenware and food contact surfaces of equipment properly washed, rinsed, sanitized, and air dried.	Food Contamination
14	Equipment for ware washing operated and maintained.	Food Contamination
14	Sanitizer solution provided and maintained as required.	Food Contamination
15	Handwashing facilities adequate in number, stocked, accessible, and limited to handwashing only.	Poor Personal Hygiene
16	Effective pest control measures.	Food Contamination
16	Animals restricted as required.	Food Contamination
17	Hot and cold holding equipment present, properly designed, maintained and operated.	Improper Holding Time and Temperature
18	Accurate thermometers (stem & hot/cold holding) provided and used.	Improper Holding Time and Temperature
19	PHF/TCSs properly thawed.	Improper Holding Time and Temperature
19	Fruits and vegetables washed prior to preparation or service.	Food Contamination
20	Single use items not reused or misused.	Food Contamination

Violation #	Violation Description	Risk Factor
21	Person in charge available and knowledgeable/management certification.	All 5
21	Food handler card as required.	Legal Requirement
21	Facility has an effective employee health policy.	Poor Personal Hygiene
22	Backflow prevention devices and methods in place and maintained.	Food contamination
23	Grade card and required signs posted conspicuously.	Legal Requirement
23	Consumer advisory as required.	Legal Requirement
23	Records/logs maintained and available when required.	Legal Requirement
23	Nevada Clean Indoor Air Act compliant.	Legal Requirement
23	PHFs labeled and dated as required.	Improper holding time and temperature
23	Food sold for offsite consumption labeled properly.	Foods from unapproved source

References

1. Brennan, S., Sadilek, A., Kautz, H. (2013). Towards Understanding Global Spread of Disease from Everyday Interpersonal Interactions. Twenty-Third International Conference on Artificial Intelligence.

2. Centers for Disease Control and Prevention (2015). National Environmental Assessment Reporting System 2014 Summary Report. Retrieved from: https://www.cdc.gov/nceh/ehs/nears/docs/2014-summary-report.pdf.

3. Centers for Disease Control and Prevention (2016a). Burden of Foodborne Illness: Overview. Retrieved from: https://www.cdc.gov/foodborneburden/estimates-overview.html.

4. Centers for Disease Control and Prevention (2016b). Surveillance for Foodborne Disease Outbreaks, United States, 2014, Annual Report. Atlanta, Georgia: US Department of Health and Human Services, CDC.

5. DiPrete, L. (2016). Using Social Media to Predict Foodborne Illness and Drive Inspections. National Environmental Health Association Annual Educational Conference. San Antonio, Texas.

6. Food and Drug Administration (2009). FDA Report on the Occurrence of Foodborne Illness Risk Factors in Selected Institutional Foodservice, Restaurant, and Retail Food Store Facility Types. Retrieved from: http://www.fda.gov/downloads/Food/GuidanceRegulation/RetailFoodProtection/FoodborneIllnessRiskFactorReduction/UCM224682.pdf.

7. Food and Drug Administration (2016). Foodborne Illnesses: What you need to know. Retrieved from: http://www.fda.gov/Food/ResourcesForYou/Consumers/ucm103263.htm.

8. Greenwood, S., Perrin, A., & Duggan, M. (2016). Social Media Update 2016. Pew Research Center. Retrieved from: http://www.pewinternet.org/2016/11/11/social-media-update-2016/.

9. Harris, J., Mansour, R., Choucair, B., et al. (2014). Health Department Use of Social Media to identify Foodborne Illness. MMWR. 63(32); 681-685. Retrieved from: https://www.cdc.gov/mmwr/preview/mmwrhtml/mm6332a1.htm.

10. Harrison, C., Jorder, M., Stern, H. et al. (2014). Using Online Reviews by Restaurant Patrons to Identify Unreported Foodborne Illness. MMWR 2014; 63. Retrieved from: https://www.cdc.gov/mmwr/pdf/wk/mm6320.pdf.

11. Kitchen, Rebecca (2016). New Rating System Set for Washoe County Restaurants. KOLO8 News Now. Retrieved from: http://www.kolotv.com/content/news/New-rating-system-for-Washoe-County-restaurants-391046241.html.

12. Las Vegas Valley Convention and Visitors Authority (n.d). Stats and Facts. Retrieved from: http://www.lvcva.com/stats-and-facts/.

13. Nevada Administrative Code (2012a). Chapter 441A-Infectious Diseases; Toxic Agents. Retrieved from: http://www.leg.state.nv.us/NAC/NAC-441A.html.

14. Nevada Administrative Code (2012b). Chapter 446- Food Establishments. Retrieved from: http://www.leg.state.nv.us/NRS/NRS-446.html.

15. Office of Disease Prevention and Health Promotion (n.d.). Food Safety. Retrieved from: https://www.healthypeople.gov/2020/topics-objectives/topic/food-safety.

16. Park, H., Kim, J., Almanza, B. (2016). Yelp Versus Inspection Reports: Is Quality Correlated With Sanitation in Retail Food Facilities?. NEHA Journal of Environmental Health. Retrieved from: http://neha.org/sites/default/files/flipping_book/june-2016-jeh/#8/z.

17. Perrin, A. (2015). Social Media Usage: 2005-20015. Pew Research Center. Retrieved from: http://www.pewinternet.org/2015/10/08/social-networking-usage-2005-2015/#.

18. Sadilek, A., Brennan, S., Kautz, H., Silenzio, V. (2013). nEmesis: Which Restaurants Should You Avoid Today?. First AAAI Conference on Human Computation and Crowdsourcing.

19. Sadilek, A., Kautz, H., DiPrete, L., Labus, B., et al, (2016). Deploying nEmesis: Preventing Foodborne Illness by Data Mining Social Media. Twenty-Eighth Annual Conference on Innovative Applications of Artificial Intelligence.

20. San Francisco Department of Public Health (n.d.). Food Safety Program: Restaurant Safety Scores. Retrieved from https://www.sfdph.org/dph/EH/Food/Score/default.asp.

21. Schomberg JP, Haimson OL, Hayes GR, Anton-Culver H (2016) Supplementing Public Health Inspection via Social Media. PLoS ONE 11(3):e0152117. doi:10.1371/journal.pone.0152117.

22. Signorini A, Segre AM, Polgreen PM (2011). The Use of Twitter to Track Levels of Disease Activity and Public Concern in the U.S. during the Influenza A H1N1 Pandemic. PLoS ONE 6(5): e19467. doi:10.1371/journal.pone.0019467.

23. Social media. (n.d.). Retrieved January 9, 2017, from https://www.merriam-webster.com/dictionary/social media

24. Southern Nevada Health District (n.d., a). Food Establishment Resource Library. Retrieved from: http://www.southernnevadahealthdistrict.org/ferl/inspection-violations.php.

25. Southern Nevada Health District (n.d., b). General Information. Retrieved from: http://www.southernnevadahealthdistrict.org/general-information.php.

26. Southern Nevada Health District (n.d., c). Statistics, Surveillance, & Reports. Retrieved from: https://southernnevadahealthdistrict.org/stats-reports/disease-statistics.php.

27. Stricker, G. (2014). The 2014 #YearOnTwitter. Twitter Blog. Retrieved from: https://blog.twitter.com/2014/the-2014-yearontwitter.

28. United States Food and Drug Administration (2014). BBB – Onset, duration, and symptoms of foodborne illness – Epidemiology summary table. Retrieved from http://www.fda.gov/Food/FoodborneIllnessContaminants/CausesOfIllnessBadBugBook/ucm071342.htm.

29. Washoe County Health District (n.d.). Food Safety News. Retrieved from: https://www.washoecounty.us/health/programs-and-services/food-protection-services/food-safety-news.php.

30. White RW, Tatonetti NP, Shah NH, et al (2013). Web-scale pharmacovigilance: listening to signals from the crowd. J Am Med Inform Assoc 2013; 20:404–408. Retrieved from: http://jamia.oxfordjournals.org/content/jaminfo/20/3/404.full.pdf.

Curriculum Vitae

Lauren DiPrete

Contact
laurendiprete@gmail.com
(702) 759-1504

Summary
Lauren DiPrete is passionate public health professional devoted to education and disease prevention. She began her work at the Southern Nevada Health District in food operations before moving onto grant coordination and research. Her current goals are investigating the potential benefits of using social media monitoring software in disease surveillance and response. She earned her B.S. in Kinesiology from UNLV and is currently pursuing her MPH.

Work
Southern Nevada Health District
Environmental Health Senior, 2015-Current
Responsible for coordinating the EHS-net Cooperative Agreement from the CDC focusing on social media monitoring research for the detection and response to foodborne illness

Environmental Health Specialist, 2011-2015
Responsible for conducting restaurant inspections, ensuring the food safety of large special events, and mentoring struggling restaurants to passing levels

Publications
Adam Sadilek, Henry Kautz, Lauren DiPrete, Brian Labus, Eric Portman, Jack Teitel, and Vincent Silenzio (2016). *Deploying nEmesis: Preventing Foodborne Illness by Data Mining Social Media*. Twenty-Eight Annual Conference on Innovative Applications of Artificial Intelligence.
-IAAI Notable paper award winner

Speaking Engagements
Food and Drug Administration's Retail Food Pacific Regional Conference
"Social Media Monitoring to Guide Inspections"
Reno, NV
September 2016
Audience of 300 federal regulatory, local regulatory, and industry professionals

National Environmental Health Association Annual Educational Conference,
"Social Media Monitoring to Guide Inspections"
San Antonio, TX
June 2016
Audience of 500 environmental health professionals

Three Square Partner Educational Meeting
“Food Safety in the Nonprofit Sector”
Las Vegas, NV
April 2015 and June 2015
Audience of 200 volunteer food handlers

Education
University of Nevada, Las Vegas
Bachelor in Science, Kinesiology
May 2009