Comparison of Restaurant Inspection Report Results and its Corresponding Star Ratings on Yelp and Google Reviews

Elaine Kong[1], Helen Heacock[2]

1 Lead Author, B. Tech Student, School of Health Sciences, British Columbia Institute of Technology, 3700 Willingdon Ave, Burnaby, BC V5G 3H2

2 Supervisor, School of Health Sciences, British Columbia Institute of Technology, 3700 Willingdon Ave, Burnaby, BC V5G 3H2

ABSTRACT

Background: In the current culture of dining-out, there is a greater emphasis on the overall dining experience at restaurants and less of a concern regarding food safety. The public often relies on consumer-generated review websites, such as Yelp and Google Reviews, to decide on where to eat. Each restaurant is often rated out of 5-stars based on factors such as customer service and food quality. The public perceives a restaurant with a 1-star rating poorly, whereas a restaurant with a 5-star rating is seen as excellent. Moreover, the aspect of food safety is determined by Environmental Health Officers (EHOs) who conduct inspections and assign hazard ratings to restaurants, which describe them as a low, moderate, or high-risk food premises. These inspection report results can be disseminated to the public online or through a placard system by the health authority. Currently, in most cities, there is no linkage or display of inspection report results on consumer-generated review websites.

Methods: Secondary data was collected from publicly available online sources: Fraser Health's restaurant inspection reports and two consumer-generated restaurant review websites – Yelp and Google Reviews. The author analyzed 170 randomly selected restaurants from the three most populous cities under Fraser Health's jurisdiction (British Columbia, Canada): Surrey, Burnaby, and Abbotsford. Only independent restaurants and their routine inspection reports were considered in this study. The following data was obtained from each of the restaurant's available routine inspection reports: current hazard rating, the average hazard score, and total number of critical violations (CVs). These variables were then compared to the current star rating found on Yelp and Google Reviews.

Results: A total of six statistical analyses were conducted: two chi-square tests and four correlational analyses. When comparing the current hazard rating of the restaurant and their current star rating using chi-square tests, $p = 0.0855$ for Yelp and $p = 0.0739$ for Google Reviews. Furthermore, in all four correlational analyses, a negative linear relationship was observed, but only three resulted in statistically significant results. When comparing the average hazard score of the restaurant's routine inspections and their current star rating, $p = 0.0591$ for Yelp (power = 47.21%) and $p = 0.0000$ for Google Reviews (power = 99.97%). When comparing the restaurant's total CVs from routine inspections and their current star rating, $p = 0.0001$ for Yelp (power = 97.29%) and $p = 0.0000$ for Google Reviews (power = 100%).

Conclusions: The findings of this study demonstrated that prescribed food safety evaluations largely align with the customer perception of restaurants. Although three out of six statistical tests resulted in statistically significant results, overall, it appears that restaurants with a higher star rating have lower number of CVs and lower average hazard scores. Even though this ideal relationship was established, the importance of safe food handling practices and serving safe food to the public should not be overlooked. Consumer-generated restaurant review websites are an excellent avenue to promote food safety within the overall culture of dining-out at restaurants.

Keywords: food safety, inspection report results, star ratings, restaurants, review websites, Yelp, Google Reviews

INTRODUCTION

Consumer-generated restaurant reviews on websites, such as Yelp, are often a customer's first choice to obtain recommendations when deciding on where to eat. Although these sites provide information on the overall dining experience and food quality of the restaurant, the aspect of food safety is noticeably missing. Environmental Health Officers (EHOs) are public health professionals who conduct inspections of restaurants and evaluate their food safety. In British Columbia (BC), EHOs assign hazard ratings that describe restaurants as "Low, Moderate, or High" risk.

This study examined the existence of associations and correlations between a restaurant's inspection report results and its corresponding star ratings found on the following review websites: Yelp and Google Reviews. The findings of this study can assist the environmental health field with identifying potential knowledge gaps between prescribed food safety scores and customer perceptions of restaurants.

LITERATURE REVIEW

Purpose of Restaurant Inspections

According to the Public Health Agency of Canada (PHAC), an estimated four million Canadians are affected by foodborne illnesses each year with approximately 238 related deaths occurring annually (PHAC, 2016). Hence, foodborne illnesses can carry a heavy economic burden on the healthcare system and can adversely affect the lives of Canadians. Fortunately, the occurrences of foodborne illnesses can be prevented through food safety education and safe food handling practices.

EHOs play a key role in the prevention of foodborne illnesses. In order to protect public health, EHOs are responsible for identifying potential health hazards during inspections and taking the appropriate corrective actions to mitigate or eliminate the hazard. Under the Public Health Act [SBC 2008] c. 28, they are given the powers to conduct inspections of restaurants to ensure food is being safely prepared and served to the public.

According to Bryan (1988), certain food handling practices, such as improper cooling and advance preparation, can contribute to the occurrence of foodborne illnesses – these practices are recognized as critical violations (CVs) by EHOs. In BC, the overall violation score culminates into a hazard rating of "Low, Moderate, or High." A high hazard rating indicates that a restaurant contains a number of high-risk practices that could cause foodborne illness. Given that EHOs are conducting inspections onsite, they are in a crucial position to provide education to the operator and front-line workers on how to ensure that safe food is being served to their customers. Therefore, restaurant inspections provide an avenue to reinforce a culture in the food industry that prioritizes food safety, in addition to informing paying customers on which restaurants are successfully complying with the legislation.

Consumer-Generated Online Review Websites

Social media and internet-based platforms have drastically increased connectivity and communication amongst people around the world. Restaurants no longer depend on promoting their business through word-of-mouth between family and friends; instead, operators can monitor their progress through online review websites. These sites contain general information about the restaurants, such as their contact information, their menu, and hours of operation. Above all, these websites allow users to write reviews of their personal experiences with the restaurant, upload photos of the food, and leave a rating out of 5-stars. Therefore, the average rating displayed for the restaurant are subjected to personal preferences and biases.

Although consumer-generated review websites allow greater exposure for operators to promote their businesses, several studies have indicated that there are a few caveats to consider. Park, Kim, and Almanza (2016) point out that Yelp customers often use a 5-star rating scale to assess food quality, service quality, pricing, and overall restaurant ambience. Similarly, a 2016 study looking at whether online reviews could predict health inspection results observed that many customer reviews are focused on service, the temperature of the food, and the restaurant environment (Wong, Chinaei,

& Rudzicz, 2016). These studies demonstrate that customers perceive other external factors of a restaurant as more important than food safety.

Although there are algorithms to filter out disingenuous ratings, there is still a risk that consumer-generated websites permit the potential for fake restaurant reviews to be posted. Operators may attempt to boost their restaurant ratings with paid reviewers, or customers may purposely leave negative and controversial posts with the intention of harming the business' reputation (Kang et al., 2013). Even though consumer-generated review websites may be fraught with inconsistencies and biases, a majority of customers rely on Google and Yelp to gather useful information during their decision-making process, which may be influenced by their social or economic choices (Bloem, 2018). Therefore, review websites can give a substantial amount of power to customers in crafting the restaurant's reputation and how the public may perceive it.

Inspection Report Disclosure Systems

Inspection report results can be disclosed in several ways. In some provinces, such as BC, the public has access to inspection reports online through the health authority websites. They can obtain information on the restaurant's hazard rating, number of CVs and non-CVs, and the frequency of inspections. Other cities, such as Toronto, New York, and Los Angeles, use a placard system that physically displays the inspection report results at the place of business (Thompson, de Burger, & Kadri, 2005; Wong et al., 2015). According to the World Cities Culture Forum (2019), the former cities have 272.5, 309.6, and 290.8 restaurants per 100 000 population, respectively. This suggests that a placard system may assist larger cities with a high number of restaurants per capita disseminate information about food safety in an efficient way. However, there is limited scientific evidence that has successfully determined the most effective disclosure method (Lee, 2013). This is difficult to establish, because these systems are dependent on a variety of factors, such as the consistency of inspector expectations, the public's understanding of inspection reports, and the ease for the public to interpret the results.

The implementation of publicly accessible disclosure systems has often been associated with improvement in operator compliance and restaurant inspection scores. The health department in Moncton, New Brunswick conducted a follow-up evaluation into their online disclosure system, eight years after implementation – they observed an increased compliance amongst restaurants (Lewis, 2015). In addition, Toronto's DineSafe program was implemented in 2001 and involves a color-coded placard system of red, yellow, and green – where green means a pass for the restaurant, yellow means a conditional pass, and red indicates a closure (City of Toronto, 2019). A 2005 case study of the effectiveness of the DineSafe system found that between 2001 and 2003, compliance rates of restaurants in Toronto increased from 78% to 88% – that is, there was an increase in the number of restaurants with a green placard (Thompson, de Burger, & Kadri, 2005). This suggests that the physical disclosure of inspection results to the public at the place of business assists with influencing operator compliance.

Despite the various ways to disseminate the results of an inspection to the general public, ultimately the effectiveness of the disclosure system depends on the public's perception of what the results mean. For inspection results to be meaningful, the general public requires a basic knowledge and understanding on how to interpret them. Dundes and Rajapaksa (2001) conducted surveys to look at how college students perceive a restaurant's letter grade of C or score of 86%. As expected, the college students displayed a limited knowledge set of what the results mean; they often associated it with the cleanliness of the restaurant (Dundes & Rajapaksa, 2001). However, a letter-grade result appears to significantly affect how people perceive and rank restaurants. For instance, the authors found that 58% of college students would not eat at a restaurant with a C-letter grade (Dundes & Rajapaksa, 2001). Through customer perception surveys on the general public, another study found that 76% of people would be satisfied with eating at a restaurant with an A-letter grade (Wong et al., 2015). Although letter-grade disclosure systems are easy for the public to comprehend, Choi et al.

(2013) indicate that a narrative approach, where violation descriptions are available for the public to view, assists with enhancing the public's understanding of food safety. Overall, the reporting format of the results and type of disclosure system used can have a significant impact on how customers absorb and perceive the information given to them (Choi et al., 2013).

Relationship Between Inspection Report Results and Online Reviews

With the rising popularity of social media, there is also growing interest in the environmental public health field in using user-generated internet-based platforms as a way to monitor the health of society. In recent years, a few studies have begun looking at how social media relates to inspection reports (Kang et al., 2013; Park, Kim, & Almanza, 2016; Wong, Chinaei, & Rudzicz, 2016). Kang et al. (2013) conducted the first empirical study that textually analyzed Seattle restaurant reviews from Yelp and compared them to their corresponding hygiene inspection report results. After removal of deceptive reviews, the authors determined that observations from the public left on Yelp predicted the hygiene inspection penalty scores with 82% accuracy (Kang et al., 2013). Even more, the authors also found that average online ratings were negatively correlated with the inspection penalty scores (Kang et al., 2013). This means that restaurants with higher customer ratings had fewer penalties, which is to be expected. A similar study conducted on restaurants in the Kitchener-Waterloo region of Ontario expanded on the previous empirical study and found a positive predictive-relationship between reviews and inspection results with 90% accuracy (Wong, Chinaei, & Rudzicz, 2016).

Generally, review websites appear to be an indicator of restaurant hygiene, which suggests that public concerns revolve mainly on sanitary conditions instead of food handling practices. This is a concern since poor food handling practices are the major contributor to the occurrence of foodborne illnesses (Bryan, 1988). As noted, food safety is often a decision criterion that is left out on consumer-generated review websites. However, in 2013, Yelp partnered with select cities, such as New York and Los Angeles, to display health inspection report scores on their corresponding restaurant review page (Booth, 2014). Food safety is integrated with Yelp restaurant reviews by displaying the inspection scores and taking a narrative approach where the violations are described in detail (Booth, 2014). Due to the numerous factors involved in customer perception, evaluating the effectiveness of this method has proven to be difficult and complicated. A recent study observed the effect of digitized disclosure of inspection reports on Yelp in San Francisco and found insignificant changes in overall restaurant hygiene scores (Dai & Luca, 2018). However, the authors note that the design in which inspection results are displayed to the public play an important role in improving the effectiveness of the disclosure system (Dai & Luca, 2018).

Purpose of Study

The purpose of this research project was to determine associations and correlations between restaurant inspection report variables (ie. hazard rating, hazard rating score, CVs) and the corresponding star ratings on the following review websites: Yelp and Google Reviews. The main research question addressed was whether customer perception of restaurants align with prescribed food safety evaluations.

MATERIALS AND METHODS

Materials Used

This research study involved the collection of secondary data from various websites: Fraser Health (through HealthSpace), Yelp, and Google Reviews. All data sources were available online for public access. Microsoft Excel was used to record data and generate randomized numbers to select the samples used in this study (Microsoft Office 365, 2019). It was also used to run the descriptive statistics of the collected data. Likewise, Number Cruncher for Statistical Systems (NCSS) was the statistical software used to perform the inferential statistics; specifically, chi-square tests and correlation analyses were conducted on the data (NCSS Statistical Software, 2019).

Standard Method

The following cities were selected for this study based on the three most populous cities in Fraser Health: Surrey, Burnaby, and Abbotsford (Statistics Canada, 2019). The randomized selection of restaurants to sample was loosely adapted from a study conducted by Yip et al. (2015) on inspection violations in community care facilities in BC. Each full page of the Fraser Health inspection report website under "View: all" contained 30 facilities, while the last page varied in the number of facilities. The author counted the number of pages under each city and the number of restaurants on the last page at the time of the study. Out of these pages, Microsoft Excel was used to obtain randomized numbers to indicate which page would be selected for further processing. For this study, seven pages were selected for each city. In total, 21 pages were analyzed. Every facility on each page was examined to determine if it fulfilled the established inclusion and exclusion criteria. A total of 170 samples were selected.

The following data was collected from each sample and recorded in Microsoft Excel: date of data collection, city, restaurant name, current hazard rating, total number of routine inspections, total number of CVs and total hazard score from routine inspections. Next, each restaurant was looked up in Google Reviews and Yelp to obtain the current star rating and the total number of reviews. If the restaurant did not exist in either of the review websites, "N/A" was recorded and it would not be considered in the corresponding statistical analyses.

Inclusion and Exclusion Criteria

This study focused on restaurants that are independently owned and are not owned by a corporation, as in the case of chain restaurants (Tung & Heacock, 2018). Since there is no universally established definition to describe a chain restaurant in the literature, this study defined a chain restaurant "as a food premises with a minimum of 10 identical establishments across Canada and/or globally" (Tung & Heacock, 2018). For this research study, the following food premises were excluded: chain restaurants, convenience or grocery stores, mobile food premises (ex. food trucks, food

carts), concession stands, and food premises that indicated that it was located on an institution's property (ex. specified hospital site). To help with identifying whether the food premises fulfilled the inclusion criteria, the HealthSpace system lists each facility as a certain "facility type" (Fraser Health Authority, 2019). Samples that had a facility type as a "restaurant" were included, while all other descriptions were excluded. For each selected restaurant, only data from routine inspection reports were considered, while follow-up inspection reports were excluded.

Ethical Considerations

The secondary data collected in this study is available for the public, therefore there was minimal ethical concerns to take into consideration (Cseke et al., 2014). Unique identifiers were used to label the restaurants sampled to keep the food premises anonymous.

STATISTICAL ANALYSES AND RESULTS

Description of Data Collected

Table 1 lists the types of data that were collected and analyzed from Fraser Health restaurant inspection reports and the consumer-generated review websites.

Table 1. *List of nominal and numerical data used in study.*

Fraser Health Inspection Reports	
Nominal Data	– Current hazard rating – City
Numerical Data	– Hazard rating score – Number of routine inspections – Number of CVs
Review Websites – Yelp, Google Reviews	
Nominal Data	– Current star rating: "equal to or less than 3.5-stars" or "more than 3.5-stars"
Numerical Data	– Current star rating: 0 to 5-stars – Number of reviews posted

The proportion of restaurants from each city is summarized in Figure 1. Out of the 170 randomly collected samples, 45 restaurants were

from Surrey, 73 from Burnaby, and 52 from Abbotsford. In addition, 96% of restaurants had a current hazard rating of "Low" as displayed in Table 2.
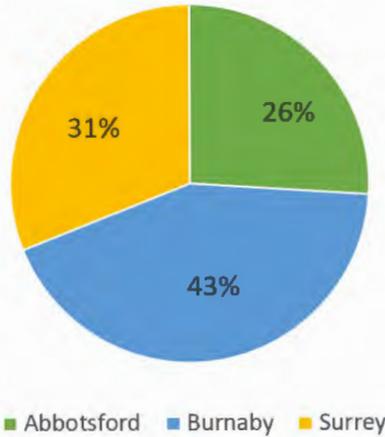

Figure 1. Percentage of restaurants sampled in each city out of a total of 170 restaurants.

Table 2. *Number of restaurants with each type of current hazard rating.*

Current Hazard Rating	Number of Restaurants	Percentage[a]
Low	163	95.9%
Moderate	6	3.5%
High	1	0.6%

[a] Out of a total of 170 restaurants.

Although 170 restaurants were analyzed, 18 samples did not have a Yelp page and one sample did not have a Google Reviews page. Therefore, out of the 170 collected samples, 152 restaurants had a Yelp page and 169 restaurants had a Google Reviews page. Yelp star ratings are separated by increments of 0.5-stars, while Google Reviews star ratings are separated by increments of 0.1-stars. Figure 2 indicates that 139 restaurants or 91% of restaurants with a Yelp page, had a current star rating of above 3.0-stars. In addition, 116 restaurants or 76% of restaurants with a Yelp page, had star ratings that were between 3.0-stars and 4.0-stars. In comparison, all the restaurants, except for two, had a current star rating above 3.0-stars in Google Reviews. Figure 3 demonstrates that 154 restaurants or 91% of restaurants with a Google

Reviews page, had a current star rating of above 3.7-stars. Overall, the star ratings were skewed towards a higher rating in Google Reviews as compared to Yelp.
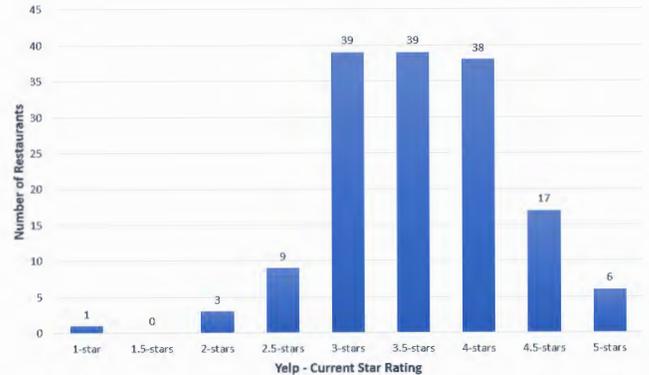

Figure 2. Number of restaurants with corresponding current star rating on Yelp.
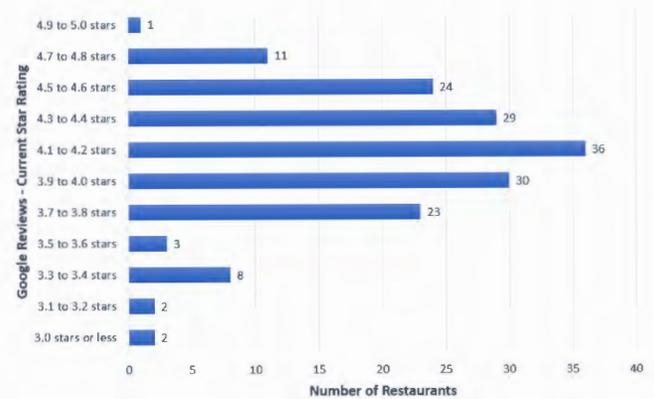

Figure 3. Number of restaurants with corresponding current star rating on Google Reviews.

Descriptive Statistics

Microsoft Excel was used to calculate the descriptive statistics of the numerical data collected in this study. Table 3 compares the current star ratings collected on Yelp and Google Reviews. In general, restaurants rated on Google Reviews had higher star ratings than Yelp as shown by the mean, median, and mode. Restaurants rated on Google Reviews had an average star rating of 4.1 stars and no restaurant had an average star rating below 2.8 stars. Furthermore, the comparison of the number of reviews of restaurants found on Yelp and Google Reviews is displayed in Table 4. In general, Google Reviews contained more

6

consumer-generated reviews than Yelp. This is demonstrated by the large difference between the mean number of reviews in Yelp and Google Reviews.

Table 3. *Descriptive statistics on current star ratings.*

Descriptive Statistics	Yelp	Google Reviews
Count	152	169
Minimum	1.0	2.8
Maximum	5.0	5.0
Mean	3.56	4.12
Median	3.5	4.1
Mode	3.0	4.2
Standard deviation	0.69	0.39

Table 4. *Descriptive statistics on total number of reviews.*

Descriptive Statistics	Yelp	Google Reviews
Count	152	169
Minimum	1	2
Maximum	334	1968
Mean	31.24	224.06
Median	17	156
Mode	2	196
Standard deviation	42.93	272.13

Inferential Statistics

NCSS was used to conduct chi-square tests and correlation analyses for each consumer-generated review website. For each statistical test, the null (H_0) and alternate (H_a) hypothesis were defined, and based on the obtained p-value, the appropriate conclusion was given. A total of six statistical tests were performed in this study.

Inferential Statistics – Chi-Square Test

To determine if an association exists between current hazard rating and current star rating of the restaurant on each consumer-generated review website, two chi-square tests were used. The results of the chi-square tests for Yelp and Google Reviews are shown in Table 5. Note that a beta or Type II error indicates that H_0 is true, but in reality, H_a is true and a difference between hazard rating and star rating truly exists

(Heacock, 2019). Increasing sample size can help minimize this type of error (Heacock, 2019).

Table 5. *Results of chi-square tests.*

Consumer-generated review website: Yelp
H_0: There is no association between the current hazard rating of the restaurant and current star rating on Yelp.
H_a: There is an association between the current hazard rating of the restaurant and current star rating on Yelp.
Result: p-value = 0.08548
Conclusion: Do not reject H_0 and conclude that there is no statistically significant association between the restaurant's current hazard rating and the current star rating on Yelp.
Note: Potential beta-error.
Consumer-generated review website: Google Reviews
H_0: There is no association between the current hazard rating of the restaurant and current star rating on Google Reviews.
H_a: There is an association between the current hazard rating of the restaurant and current star rating on Google Reviews.
Result: p-value = 0.07393
Conclusion: Do not reject H_0 and conclude that there is no statistically significant association between the restaurant's current hazard rating and the current star rating on Google Reviews.
Note: Potential beta-error.

Inferential Statistics – Correlation Analysis

For each consumer-generated review website, two numerical inspection report variables were analyzed: average hazard score and total CVs. The results of the four correlation analyses for Yelp and Google Reviews are

7

shown in Table 6. Note that a larger statistical power indicates a higher probability that the test results are valid.

Table 6. *Results of correlational analyses.*

Inspection report variable: Hazard Score
H_0: There is no linear relationship between average <u>hazard score</u> of routine inspections and current star rating value on <u>Yelp</u>. H_a: There is a linear relationship between average <u>hazard score</u> of routine inspections and current star rating value on <u>Yelp</u>. **Result:** p-value = 0.0591 **Conclusion:** Do not reject H_0 and conclude that there is no linear relationship between the average <u>hazard score</u> of the restaurant's routine inspections and the current star rating on <u>Yelp</u>. *Note: Potential beta-error. Power = 47.21%.*
H_0: There is no linear relationship between average <u>hazard score</u> of routine inspections and current star rating value on <u>Google Reviews</u>. H_a: There is a linear relationship between average <u>hazard score</u> of routine inspections and current star rating value on <u>Google Reviews</u>. **Result:** p-value = 0.0000 **Conclusion:** Reject H_0 and conclude that there is a linear relationship between the average <u>hazard score</u> of the restaurant's routine inspections and the current star rating on <u>Google Reviews</u>. *Note: Power = 99.97%.*
Inspection report variable: Critical Violations
H_0: There is no linear relationship between <u>total CVs</u> of routine inspections and current star rating value on <u>Yelp</u>.

H_a: There is a linear relationship between <u>total CVs</u> of routine inspections and current star rating value on <u>Yelp</u>. **Result:** p-value = 0.0001 **Conclusion:** Reject H_0 and conclude that there is a linear relationship between the <u>total CVs</u> of the restaurant's routine inspections and the current star rating on <u>Yelp</u>. *Note: Power = 97.29%.*
H_0: There is no linear relationship between <u>total CVs</u> of routine inspections and current star rating value on <u>Google Reviews.</u> H_a: There is a linear relationship between <u>total CVs</u> of routine inspections and current star rating value on <u>Google Reviews.</u> **Result:** p-value = 0.0000 **Conclusion:** Reject H_0 and conclude that there is a linear relationship between the <u>total CVs</u> of the restaurant's routine inspections and the current star rating on <u>Google Reviews</u>. *Note: Power = 100%.*

Three out of the four correlational analyses resulted in statistically significant findings. Figure 4 shows the linear relationship between the restaurant's average hazard score and current star rating on Google Reviews. The correlation coefficient is -0.3884, which indicates a negative and fairly strong relationship between the variables. This means that a higher star rating on Google Reviews is fairly correlated with a lower average hazard score from the restaurant's inspection reports.
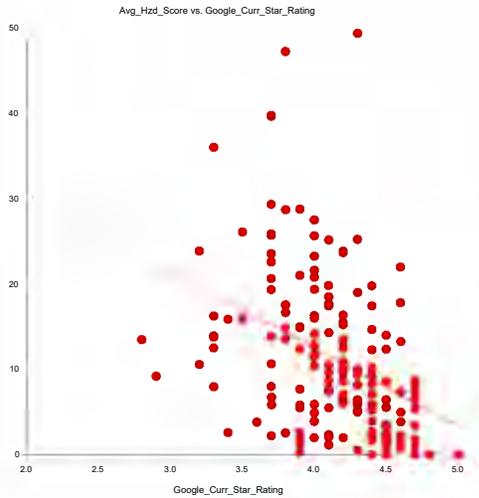
Figure 4. NCSS linear regression plot of restaurant's average hazard score and current star rating on Google Reviews.

Similarly, Figure 5 shows the relationship between the restaurant's total CVs in routine inspections and current star rating on Yelp. The correlation coefficient is -0.3042, which means a higher star rating on Yelp is fairly correlated with a lower number of total CVs from the restaurant's inspection reports.
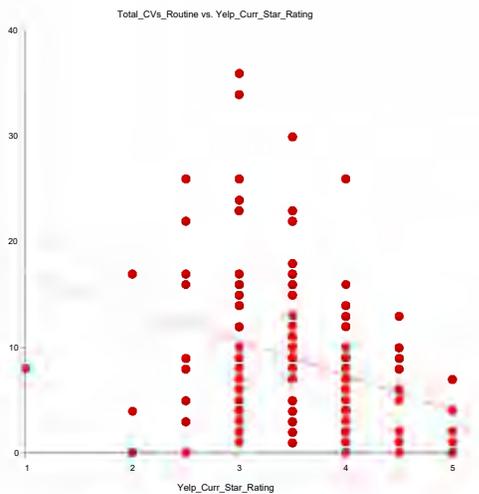

Figure 5. NCSS linear regression plot of restaurant's total CVs from routine inspection reports and current star rating on Yelp.

Another negative relationship is seen in Figure 6 between the restaurant's average hazard score and current star rating on Google Reviews. The correlation coefficient is -0.4517, which

indicates that a higher star rating on Google Reviews is fairly correlated with a lower number of total CVs from the restaurant's inspection reports.
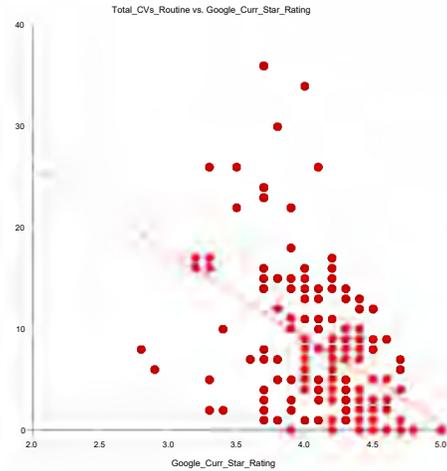

Figure 6. NCSS linear regression plot of restaurant's total CVs from routine inspection reports and current star rating on Google Reviews.

DISCUSSION

This study compared different restaurant inspection report variables (ie. hazard rating, hazard score, CVs) to the current star rating on Yelp and Google Reviews, which resulted in a mix of statistically significant and non-statistically significant findings.

The restaurant's total number of CVs in routine inspections and its current star rating on Yelp and Google Reviews were negatively correlated and statistically significant. Restaurants with few CVs had high star ratings on both review websites that were analyzed. Since CVs are an indication of practices that are likely to contribute to foodborne illnesses, then it appears that the customer perception of restaurants based on star ratings align with the food safety aspect of the restaurants conducted by inspectors. That is, this is the ideal relationship we would expect where restaurants that are rated highly have fewer CVs. However, it is important to note that the average star ratings on Yelp and Google Reviews are skewed towards a high rating – above 3-stars. Review

websites are subjected to external factors that may influence the displayed star rating; for instance, businesses may encourage people to post positive reviews or get the website to remove overly negative reviews, thereby affecting the overall average rating of the restaurant. In addition, most restaurant operators are aware of the expectations EHOs have when they conduct inspections and it is in the operator's best interest to achieve fewer CVs to maintain a low hazard rating.

Interestingly, when comparing the restaurant's average hazard score and its current star rating, there was a statistically significant negative relationship in Google Reviews only – there was a close to statistically significant finding for Yelp (p = 0.0591). This means that restaurants with lower average hazard scores tend to have higher star ratings. Similar to the comparison with CVs, it appears as though there is an ideal relationship where customer perception of restaurants aligns with the aspect of food safety. The discrepancy in the statistically significant findings between Google Reviews and Yelp may be due to the large difference in the total number of reviews between each website. On average, restaurants in Google Reviews had 224 ratings as compared to restaurants in Yelp which had an average of 31 ratings. That is, each restaurant star rating in Google Reviews was the average of a larger sample size when compared to Yelp.

The issue of sample size is clear in the statistical tests that were conducted in this study, because three out of six tests had potential beta-errors, which resulted in non-statistically significant findings. It occurred in the two chi-square tests, and the correlation analysis of the hazard score of routine inspections and the current star rating value on Yelp.

Notably, chi-square tests are sensitive to small frequencies and NCSS denotes a warning if at least one cell contains an expected value of less than five. This study identified 163 out of 170 restaurants with a current hazard rating of "Low", while six restaurants had a "Moderate" and one restaurant had a "High" hazard rating. As a result, the chi-square tests had several cells in the expected (and observed) counts table with a value of less than five. In this way, the results of the chi square tests are not robust which is

demonstrated by the non-statistically significant findings. Furthermore, the strong skew towards "Low" hazard ratings is influenced by the operator and the health authority. Firstly, it is in the operator's best interest to maintain a hazard rating of "Low" on the most current inspection, because this is the rating that gets displayed to the public as the "current hazard rating" on HealthSpace. Secondly, if a restaurant is evaluated as a "Moderate" or "High" hazard rating, it is in the health authority's best interest to promptly manage the facility to ensure that the food is safe for the public to consume, thereby converting the restaurant back to a "Low" hazard rating.

Overall, the findings of this study are consistent with the few studies in the scientific literature that have compared restaurant inspection reports to consumer-generated websites, specifically Yelp (Kang et al., 2013; Park, Kim & Almanza, 2016). Kang et al. (2013) determined that there was a negative correlation between inspection penalty scores and the restaurant's average rating on Yelp. Similarly, a 2016 study found a statistically significantly negative relationship between violation scores and Yelp ratings in analysis of chain restaurants (Park, Kim & Almanza, 2016). Besides using Yelp, the present study contributes to the literature by analyzing restaurant inspection reports to another consumer-generated review website, Google Reviews.

At the time of this study, the 2019-2020 COVID-19 pandemic prompted the BC provincial government to suspend dine-in operations at restaurants and limit them to take-out or delivery services only (Government of BC, 2020). The aftermath of this pandemic will likely result in a massive shift in dining culture, especially in terms of how customers perceive restaurants and food in general. The public may be less concerned about the restaurant ambience and service quality, rather they will be more cognizant of the due diligence of food premises in maintaining sanitary and safe practices. Even more, the star ratings and comments seen on restaurant review websites, like Yelp, may no longer be enough to satisfy a potential customer to decide on where to eat. Therefore, the aspect of food safety as supported by restaurant inspection reports may soon become a crucial

and decisive factor in a customer's decision-making process on where to eat.

LIMITATIONS

The primary limitations of this study that affect the reliability and validity of the results was the internal inconsistency of the secondary data sources that were used.

Issues with HealthSpace Inspection Reports

HealthSpace is the software used by Fraser Health to post inspection reports online (Fraser Health Authority, 2019). For each report, it is labelled as either a routine inspection or a follow-up inspection. When data was being collected, the author noticed that several of the reports were either mislabelled or limited by the software to be categorized as either routine or follow-up. This study only analyzed inspection reports labelled as "routine" on HealthSpace, because routine inspections involve a complete, thorough assessment of the restaurant and food handling practices. Follow-up inspections occur under the discretion of the EHOs; it involves inspection of a specific issue. In this way, there will likely be fewer CVs and non-CVs in a follow-up inspection as compared to a routine inspection. Therefore, reports that were mislabelled do not accurately reflect the study's intention to review routine inspection reports only to evaluate the restaurant's overall food safety practices. Furthermore, the online database does not contain a complete history of the restaurant's inspections from when the restaurant first opened. It appears that when there is new ownership of the food premises, all the previous inspection reports are removed from public viewing. In this way, the data collected on each restaurant was not a historical representation of the routine inspections. Therefore, it does not make it an equivalent comparison to the restaurant's current star rating which is a historical average of consumer ratings. As such, the average hazard score and total number of CVs of routine inspections in this study is not an accurate representation of the restaurant itself due to the issues described above.

Another issue with the public HealthSpace database was that inspection reports from 2016 and onwards contained the detailed inspection information, which included the CVs, hazard score, and comments. Inspection reports prior to 2016 were presented in a different format that provided less information. The major issue with this brief format was that there would be an assigned hazard rating, but no corresponding hazard score. Therefore, the missing study variables in some of the routine inspection reports affected the accurate collection of data for the statistical tests conducted in this research project.

Moreover, there were several suspected computer errors in some of the inspection reports on HealthSpace. Some reports did not have matching hazard scores and hazard ratings. For example, a report would contain CVs and non-CVs, yet the hazard score was displayed as zero. In addition, each hazard rating was categorized as "Low," "Moderate," or "High" with corresponding hazard scores "0 to 15," "16 to 29," and "30 and higher," respectively. Several reports displayed hazard ratings that had hazard scores that did not fall in the described categories. Again, these suspected computer errors affected the accurate portrayal of the food safety scores from each restaurant that was analyzed in this study.

Besides the technical issues with HealthSpace, inspection results are also limited by human intervention. Since inspection reports on HealthSpace evaluate restaurants based on a scoring system, there can be a potential inspector bias. That is, violation scores may be more lenient or stringent depending on the area's designated EHO. The policies and management expectations may also vary between geographical regions, which may also influence hazard ratings and scores depending on the city the restaurant is in.

Issues with Yelp and Google Reviews

Consumer-generated review websites are inherently subjective sources. The posted reviews and ratings are dependent on personal preferences and individual experiences with the restaurant. Notably, online reviews can be manipulated where businesses may encourage people to post positive reviews. Even more, businesses may use financial means to have positive reviews posted or have negative reviews

removed from these websites. As such, the current restaurant star ratings on Yelp and Google Reviews in this study is likely skewed towards higher ratings, closer to 5-stars instead of 1-star.

Time and Resource Constraints

Since this study was a student research project, the author was limited by time and resources. The study had to comply by a strictly defined timeline. As such the author could only collect data on 170 restaurants from three cities in Fraser Health. This led to three out of six statistical tests that were deemed not statistically significant with potential beta or Type II errors due to the small sample size. Furthermore, the data was manually collected from publicly accessible secondary sources. Therefore, the study is also subjected to human errors in accurately recording information for each sample.

Future Recommendations

The three cities in this study were chosen based on the largest population sizes in Fraser Health. In this way, the findings of this study can be generalized to urban areas that have online inspection report disclosures using HealthSpace. Sampling restaurants from a larger number of cities in Fraser Health or another health authority that uses HealthSpace would help increase the external validity of the study. In the future, the author recommends getting access to the complete HealthSpace database of restaurant inspection reports from Fraser Health to improve the internal validity of the study. If the data were available on an excel spreadsheet, specific variables could be extracted in order to compare it to the star ratings on Yelp and Google Reviews. This would help increase sample size of the analysis and reduce human errors that may occur when collecting and recording data by hand.

Since a large majority of restaurants had current hazard ratings of "Low" due to external biases, it resulted in unstable chi-square test results where many cells contained low observed and expected values. Due to the lack of robustness and limitations of chi-square tests in the context of this study, it is recommended to focus on correlation analyses or other statistical tests that do not depend on nominal data like current hazard rating.

KNOWLEDGE TRANSLATION

The findings from this study could assist in evaluating the degree of disconnect between customer perception and evidence-based restaurant inspections. Consumers appear to place more emphasis on online review websites to look for restaurant recommendations, rather than the food safety aspect of the restaurant through inspection reports. A collaboration between these two sources of information and perspectives may be considered. In a few American states, the health inspection scores of restaurants are integrated into Yelp websites.

Moreover, future programs may be developed to educate the general public on the importance of inspection reports and how to interpret them. A module on how to understand inspection reports could be integrated into a FOODSAFE program. Even more, this study could also prompt health authorities to re-evaluate the effectiveness of their current way to disseminate hazard ratings and inspection results to the public. A multifaceted approach to disclosing inspection results may be considered. Instead of having inspection reports available online, a visual component can also be displayed at the place of business. This can be in the form of a placard system, similar to DineSafe in Toronto, or it can be in a form that notifies customers that the food premises is inspected by a specified health authority along with a means to redirect them to the inspection report online (ex. website link, QR barcode).

FUTURE RESEARCH

To further evaluate the relationship between consumer-generated review websites and prescribed food safety scores, future student research projects may include:

- Expand the current study to include additional cities in Fraser Health (while increasing the sample size)
- Conduct a similar study using online inspection reports from another health authority in BC

12

- Conduct a similar study using inspection report results that are displayed in a placard system (ex. DineSafe Toronto)
- Conduct the same study next year (after the 2019-2020 COVID-19 pandemic) and compare the results to the current study

CONCLUSION

The overall dining experience at a restaurant is often placed at higher importance than food safety in the eyes of the consumer. In this way, operators may tend to focus on perfecting food quality and service to promote the business. However, maintaining safe food handling practices should not be overlooked. EHOs play a vital role in preventing CVs that could lead to foodborne illnesses in order to protect the health of the public and reduce the financial burden on the healthcare system. The findings of this study determined that the consumer perception of restaurants based on online star ratings agree with its corresponding food safety evaluations through inspection report results. That is, higher star ratings of restaurants correlated with lower average hazard scores and lower number of CVs. Although an ideal relationship was established between these two contrasting perspectives, this study further prompts current discussions on how to best disclose and relay food safety information to the public.

ACKNOWLEDGEMENTS

COMPETING INTERESTS

The authors declare that they have no competing interests.

REFERENCES

Bloem, C. (2018). Which Online Review Site Do More People Trust? (And How it Can Help You Improve Your Products and Services). Retrieved October 17, 2019, from https://www.in c.com/craig-bloem/want-to-grow-your -small-business-use-online-review-sites-right-way.html

Booth, D. (2014). Yelp partners with health departments to improve food safety. *Journal of Environmental Health*, *76*(8), 52–55.

Bryan, F. L. (1988). Risks of Practices, Procedures and Processes that Lead to Outbreaks of Foodborne Diseases. *Journal of Food Protection, 51*(8), 663-673. doi:10.4315/0362-028x-51.8.663

Choi, J., Miao, L., Almanza, B., & Nelson, C. D. (2013). Consumers' Responses to Restaurant Inspection Reports: The Effects of Information Source and Message Style. *Journal of Foodservice Business Research*, *16*(3), 255–275. https://doi.org/10.1080/15378020.2013.81 0536

City of Toronto. (2019). About DineSafe. Retrieved November 17, 2019, from https://www.toronto.ca/community-people/health-wellness-care/health-programs-advice/food-safety/dinesafe/about-dinesafe/

Cseke, P., Heacock, H., Sidhu, B., Mcintyre, L., & Wilcott, L. (2014). Inspecting inspection reports, does the type of restaurant change the risk?. *BCIT Environmental Health Journal*. Retrieved from https://circuit.bcit.ca/repository/islandora/object/repository%3A20

Dai, D., & Luca, M. (2018). Digitizing Disclosure: The Case of Restaurant Hygiene Scores. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3 131900

Dundes, L., & Rajapaksa, S. (2001). Scores and grades: A sampling of how college students and food safety professionals interpret restaurant inspection results. *Journal of Environmental Health*, *64*(5), 14–19.

Fraser Health Authority. (2019). Food Premises Facilities. Retrieved from https://www. healthspace.ca/Clients/FHA/FHA_Website.nsf/Food?OpenPage&

Government of BC. (2020). *Owners and operators of nightclubs and food and drink service – March 2020.* Retrieved April 3, 2020 from https://www2.gov. bc.ca/assets/gov/health/about-bc-s-health-care-system/office-of-the-provincial-health-officer/reports-publications/covid-19-pho-order-nightclubs-food-drink-services.pdf

Heacock, H. (2019). *Module 4: Introduction to Inferential Statistics* [PowerPoint presentation]. Retrieved from D2L.

Kang, J. S., Kuznetsova, P., Luca, M., & Choi, Y. (2013). Where not to eat? Improving public policy by predicting hygiene inspections using online reviews. *EMNLP 2013 - 2013 Conference on Empirical Methods in Natural Language Processing, Proceedings of the Conference*, (October), 1443–1448.

Lee, B. (2013). Food Safety Interventions. *National Collaborating Centre for Environmental Health*, (September), 1–53.

Lewis, E. (2015). Public inspection reports improving compliance. *The Times – Transcript.* Retrieved October 5, 2019, from https://ezw.lib.bcit.ca/login?url=https://search-proquest-com.ezw.lib.bcit.ca/docview/1645005126?accountid=26389

Microsoft Office 365. (2019). Microsoft Excel [Computer sofware]. Retrieved from https://products.office.com/en-ca/excel

NCSS Statistical Software. (2019). NCSS 2019 Data Analysis [Computer software]. Retrieved from https://www.ncss.com/

Park, H., Kim, J., & Almanza, B. (2016). Yelp versus inspection reports: Is quality correlated with sanitation in retail food facilities? *Journal of Environmental Health*, *78*(10), 8–12.

Public Health Act, SBC 2008, c. 28. Retrieved from http://www.bclaws.ca/civix/document/id/complete/statreg/08028_01

Public Health Agency of Canada. (2016). Yearly food-borne illness estimates for Canada. Retrieved October 6, 2019, from https://www.canada.ca/en/public-health/services/food-borne-illness-canada/yearly-food-borne-illness-estimates-canada.html#ov

Statistics Canada. (2019). *Population and dwelling count highlights tables, 2016 census: Population and dwelling counts, for Canada, provinces and territories, and census subdivisions (municipalities), 2016 and 2011 censuses – 100% data.* Retrieved November 16, 2019 from https://www 12.statcan.gc.ca/census-recensement /2016/dp-pd/hlt-fst/pd-pl/Table.cfm? Lang=Eng&T=302&SR=1&S=3&O=D&RPP=9999&PR=59&CMA=0#tPopDwell

Thompson, S., de Burger, R., & Kadri, O. (2005). The Toronto food inspection and disclosure system: A case study. *British Food Journal*, *107*(3), 140–149. https://doi.org/10.1108/00070700510586461

Tung, J. & Heacock, H. (2018). Critical food safety violations in Surrey: relationship to community median household income and restaurant type. *BCIT Environmental Health Journal*. Retrieved from https://circuit.bcit.ca/repository/islandora/object/repository%3A820

Wong, M. R., McKelvey, W., Ito, K., Schiff, C., Jacobson, J. B., & Kass, D. (2015). Impact of a letter-grade program on restaurant sanitary conditions and diner behavior in New York City. *American Journal of Public Health*, *105*(3), e81–e87. https://doi.org/10.2 105/AJPH. 2014.302404

Wong, S., Chinaei, H., & Rudzicz, F. (2016). *Predicting health inspection results from online restaurant reviews*. Retrieved from http://arxiv.org/abs/1603.05673

World Cities Culture Forum. (2019). Number of restaurants per 100,000 population. Retrieved November 17, 2019, from http://www.worldcitiescultureforum.com/data/number-of-restaurants-per-100.000-population.

Yip, T., Heacock, H., Sidhu, B., & Mcintyre, L. (2015). Inspection violations and community care facilities in British

Columbia, Canada. *BCIT Environmental Health Journal*. Retrieved from https://circuit.bcit.ca/repository/islandora/object/repository%3A59

REFERENCE TO COMPLETE PAPER

Kong, E. (2020). Comparison of restaurant inspection report results and its corresponding star ratings on Yelp and Google Reviews. *BCIT Environmental Health Journal.*