Determining Risk Factors of a Non-Point Source Outbreak of Campylobacter Cases Using Case-Case and Case-Control Studies

Kristen Pogreba-Brown1, Kacey Ernst1, Lisa Woodson1, Robin B Harris1

1University of Arizona, Epidemiology and Biostatistics, United States

*Corresponding author: Kristen Pogreba-Brown, Assistant Professor, University of Arizona, Epidemiology and Biostatistics, 1295 N. Martin, PO Box 245211, Tucson, AZ 85724, United States, mobile: 5202753945; E-mail: kpogreba@email.arizona.edu

Received date: Sep 19, 2015; Accepted date: Oct 23, 2015; Published date: Oct 28, 2015

Abstract

**Background:** Investigating foodborne outbreaks is a resource and time-intensive process using traditional case-control methodology. The use of case-case studies in outbreak investigations is not well studied, although they require fewer resources to conduct and limit selection and recall bias. In this study we investigated a cluster of Campylobacter infections using almost simultaneous case-control and case-case studies to compare results from the two methodologies.

**Methods:** In 2011 a significant increase in Campylobacter cases was detected in Pima County, AZ through routine surveillance. To determine potential sources of the outbreak we conducted two studies. The case-control study used randomly selected non-ill controls. The case-case study used historical surveillance data. Logistic regression analysis was used to determine risk factors for infection.

**Results:** Statistically significant risk factors associated with disease differed by design with travel (OR=4.1), Hispanic (OR=4.5), and youth (OR=3.6) in the case-control and untreated water (OR=5.4) and fresh eggs (OR=2.5) in the case-case. Effect modification by travel was found for untreated water (OR=14.0 for travelers vs. OR=1.5 for non-travelers) and eggs (OR=11.5 for travelers vs. OR=1.5 for non-travelers).

**Conclusions:** Travel history, a commonly reported risk factor, is a distal part of the exposure pathway. These studies exposed the more proximal cause to be largely attributed to travelers who had exposure to untreated water and fresh eggs. Case-case methods were found to be useful in outbreak investigations of a foodborne illness. This outbreak is also an example where a student response team response with a local public health department.

Keywords: Case-case studies; Case-control studies; Foodborne outbreaks; Campylobacter.

Introduction

Case-control studies are the primary design used in foodborne outbreak investigations. Yet the resources required for conducting frequent case-control studies are often beyond the capacity of local health departments. Developing rapid, minimal resource methods to identify factors associated with an outbreak could lead to more consistent and timely response. Case-case studies are a potential alternative for some outbreak investigations [1-5]. Case-case designs have been utilized in studies of other diseases such as cancer [6] and more recently been used for enteric diseases [4,7-9]. In this study, historical cases from surveillance data were compared to outbreak cases to identify the possible source contributing to an increase in cases.

Campylobacter is the leading cause of bacterial gastroenteritis globally [10]. In the United States 45,000 laboratory confirmed cases are reported each year [11], but are estimated to cause 2.4 million cases annually [12]. Most of these cases are sporadic in nature [13-19], however, for outbreaks that do occur, most are linked to unpasteurized dairy [20, 21] and untreated water [22]. While these studies help to determine the associated risk factors for the disease, they are costly and time consuming. This often makes them prohibitive to conduct for health departments outside of large, point source outbreaks. In Arizona, the annual rate of reported campylobacteriosis has been on average 1.23 times higher than the national average over the past ten years [23,24] with an average of 900 laboratory confirmed cases reported each year [25].

In mid-September of 2011, the Pima County Health Department (PCHD) in southern Arizona noted a sharp increase in the number of Campylobacter spp. infections reported through their routine surveillance system. Collaboration between PCHD and the University of Arizona’s student outbreak response team (SAFER) [26] sought to identify the possible source for this outbreak and to evaluate the efficiency of two parallel study designs; a case-case study and a case-control study.

Methods

Aberration Detection: A standard aberration detection algorithm was used to calculate the running mean of historical cases for the time period 2006-2010 [27-29]. Current cases were grouped in 4 week intervals to be compared to five years of historical data for the same 4 week intervals. These weeks represent the Center for Disease Control's Morbidity and Mortality Weekly Report (MMWR) reporting week (1-52) and not necessarily the onset date of the case. Average yearly
case counts were determined for successive time intervals (the current time period, the 4 weeks before and the 4 weeks after) [27,30]. Outbreaks are suspected when the number of current cases is above two standard deviations of the running mean [30].

Case-Control Study

Case definition and interviews: ‘Outbreak’ cases were defined as laboratory confirmed Campylobacter spp. cases identified through the county health department’s routine surveillance system from July to September 2011 shown in Figure 1. Pulse Field Gel Electrophoresis (PFGE) was not being conducted on Campylobacter cases in Arizona at the time, so it was not possible to identify genetically linked cases. Outbreak cases were interviewed by county epidemiologists using a standard investigation form. De-identified data were provided to the University team on demographics symptoms, travel, and exposure history to common risk factors for Campylobacter infection [13,15-17,31,32].

Figure 1: Reported Campylobacter Cases from July-September for 2011 and 2007-2010.

Control definition, recruitment, and interviews: Controls were identified through random digit dial by generating phone numbers in Excel (Microsoft Corp. Seattle, WA) using county pre-fixes. Ineligible households were those where residents spoke a language other than English, or no one met the needed age and gender criteria. Controls were matched to cases by gender and age group to be consistent with other published studies [13,15,19,33]. The goal was a 2:1 frequency match for controls to cases. Controls were excluded if they reported a diagnosed infection due to Campylobacter in the last month. For minors, the available guardian was either interviewed or consented to interview the child.

The control questionnaire mirrored the case interview for most potential risk factors. As there was a three month delay between case and control interviews due to lags in surveillance and IRB approval for University participation, exposures that were potentially less subject to recall bias (i.e. untreated water exposures and social gathering attendance) were asked for the period of time of the identified outbreak (July-September). Food consumption of specific items for the two weeks prior to the interview was asked to more closely match the recall period required of cases. Surveys were administered over the telephone by members of the University’s student outbreak response team using the survey program Illume (DatStat Corp, Seattle, WA).

Case-Case Study

Cases: The same cases were used for the case-case study as the case-control study.

Historical Cases: Historical comparison cases were identified from surveillance data on campylobacteriosis cases reported to PCHD from January 2007 to September 2010. Of the 681 potential cases from this time period, 105 without risk factor data were excluded. Comparison cases were restricted to those cases occurring during the same season as the outbreak cases: July – August of 2007 – 2010 (n = 186). The same questionnaire was used for both the historical and outbreak cases as shown in Figure 1.

Statistical Analyses

While the spike in cases was noted in August (weeks 33-36), cases with onset dates from July – September were included as the outbreak cases to account for any earlier or later exposures or reporting. When the months of July (N=10), August (N=26) and September (N=12) were combined, a total of 48 suspect outbreak cases were included in the analyses.

Counts and frequencies were determined for demographics, symptoms and risk factors for controls and both sets of cases. Fisher’s exact tests and Student’s t-tests were performed to determine differences in demographic characteristics between the outbreak cases, phone controls and historical cases. Logistic regression analysis was used to estimate odds ratios with 95% confidence intervals, although interpretation of odds ratios differed by study design. Multivariate models were used to evaluate ethnicity and travel history as either potential confounders or effect modifiers. All analyses were conducted in Stata version 11.0 (Stata Corp, College Station, TX).

Results

Aberration detection

Figure 2 shows the running means by week for 2011 cases compared to the 2007-2010 historical cases. These cases were largely sporadic in nature with no reported outbreaks in the county during this time period. The reported cases in weeks 33-36 (N=27) of 2011 were 2
standard deviations above what was expected given historical data and were almost double the historical average of 15 cases during the same time period indicating a potential outbreak of Campylobacter infection had occurred.

Figure 2: 2011 Laboratory confirmed Campylobacter cases in Pima County, AZ and previous 5 year running mean. Note: Note: July Cases – Weeks 25-36; August Cases – Weeks 32-41; September Cases – Weeks 36-39

Case-Control and Case-Case Studies

Response Rate: During the 22 days of data collection, 2,016 randomly selected phone numbers were called; 624 (31%) numbers were never answered, 784 (39%) were disconnected and 156 (7.7%) were businesses.

Of the 452 households contacted, 31 either spoke only Spanish (N=17) or did not meet the age and gender criteria for the frequency match (N=14). Sixty-eight (N=68) completed interviews (15% response rate from eligible households).

On average, these interviews required 2.8 hours of student time for each successful interview, though each interview took approximately 20-25 minutes to complete.

Demographics of Study Populations: Table 1 compares demographic characteristics of the outbreak cases (N=48), historical cases (N=186) and phone controls (N=68).

When comparing the outbreak cases to the phone controls, there were no statistically significant differences by gender or age, indicating the frequency matching was successful.

However, 44% of cases were Hispanic compared to only 28% of phone controls (p =.002). The percentage of Hispanic residents in the county as a whole was 34.6% [34].

Those under the age of 20 made up 35% of the outbreak cases, with the majority between 10-19 years old. There were more boys reported in the outbreak cases (60% of the 1-9 year olds were boys and 64% of the 10-19 year olds) (data not shown).

For the phone controls, women accounted for a higher proportion than outbreak cases (56% versus 65% - no statistical difference).

Table 1: Demographic characteristics of Campylobacter outbreak cases, historical cases, and controls

Spatial Distribution: Outbreak cases were found to occur disproportionately in three zip codes that accounted for 46% of all cases in July-September 2011.

These zip codes were geographically adjacent and include 15.2% [34] of the total population for the county.

During these three months, 20% of cases came from zip code 1, 15% from zip code 2 and 11% from zip code 3, compared to the historical cases residing in zip code 1 (12%), zip code 2 (6.3%) and zip code 3 (2.3%).

None of the phone controls resided within these three zip codes. Within these zip codes, Hispanics made up the majority of cases, however they also accounted for the majority of residents according to the 2010 U.S. Census [34].

Analyses of common risk factors were completed for these three areas alone with no statistically significant differences identified compared to all other outbreak cases (data not shown).
Symptomology: Symptoms were compared for the outbreak cases and the respective historical cases (Table 2). Most symptoms were reported consistently between the two sets of cases. However, hospitalizations were reported more commonly among cases from July -September 2011 as compared to cases for those same months in the previous four years (25% vs. 15%, p=0.1 using Fisher’s exact test).

### Risk Factor Comparisons

Case-Control Analysis: For consumption of various food items and lifestyle behaviors, cases were more likely to report consuming unpasteurized dairy (15% vs. 10%), and have a travel history (29% vs. 10%), particularly outside of the U.S. (25% vs. 10%) than controls (Table 3). The statistically significant crude risk factors were any travel history in the last 30 days (OR=4.07; 95% CI 1.49-11.13), travel to Mexico (OR=3.24; 95% CI 1.49-11.13), hotel anywhere outside of the U.S., including Mexico (OR=3.88; 95% CI 1.33-11.28) and being Hispanic (OR=4.53; 95% CI 1.79-11.5). Eating poultry was found to be ‘protective’ but not statistically significant and no distinction was made between consumption in the home or away from home in the questionnaire. Exposure to untreated water had no association for all three months combined, but when exposure for controls was limited to only the month of August, the odds ratio increased from 1.01 to 1.87 although neither was statistically significant. Fresh eggs could not be analyzed because the question was asked differently for the cases and controls. When the models were adjusted for age and gender, the odds ratios for each decreased with travel history in the last 30 days (OR=3.3; 95% CI 1.06-9.56), travel to Mexico (OR=2.6; 95% CI .8-8.7), and being Hispanic (OR=3.6; 95% CI 1.2-10.5) remaining significant.

### Table 2: Comparison of symptoms between outbreak cases and historical cases

<table>
<thead>
<tr>
<th>Symptom</th>
<th>Outbreak Cases N=48</th>
<th>Phone Controls N=68</th>
<th>July-Sept 2007-2010 N=186</th>
<th>Outbreak Cases vs. Null</th>
<th>Outbreak Cases vs. Phone Controls</th>
<th>Outbreak Cases vs. Phone Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N (%)</td>
<td>N (%)</td>
<td>Total (%)</td>
<td>OR (CI) [p-value]</td>
<td>OR (CI) [p-value]</td>
<td>OR (CI) [p-value]</td>
</tr>
<tr>
<td>Diarrhea</td>
<td>48 (92%)</td>
<td>15 (22%)</td>
<td>2.75 (11) (5.9%)</td>
<td>3.4 (1.26-9.3) [.016]</td>
<td>1.01 (.38-2.67) [.99]</td>
<td>.77 (.27-2.15) [.62]</td>
</tr>
<tr>
<td>Bloody stools</td>
<td>44 (92%)</td>
<td>7 (10%)</td>
<td>2.65 (26) (14%)</td>
<td>1.01 (1.40-2.56) [.98]</td>
<td>1.73 (.55-5.42) [.34]</td>
<td>1.79 (.56-5.76) [.33]</td>
</tr>
<tr>
<td>Nausea</td>
<td>30 (63%)</td>
<td>15 (22%)</td>
<td>7.75 (31) (17%)</td>
<td>2.5 (1.14-5.44) [.02]</td>
<td>------ (.14 p=.000)</td>
<td>------ (.15 p=.000)</td>
</tr>
<tr>
<td>Vomiting</td>
<td>18 (38%)</td>
<td>15 (22%)</td>
<td>4.5 (18) (9.7%)</td>
<td>.94 (.30-3.02) [.93]</td>
<td>.62 (.18-2.14) [.45]</td>
<td>.69 (.19-2.46) [.57]</td>
</tr>
<tr>
<td>Fever</td>
<td>29 (61%)</td>
<td>7 (10%)</td>
<td>28.8 (115) (62%)</td>
<td>.70 (.30-1.66) [.42]</td>
<td>.38 (.13-1.09) [.07]</td>
<td>.38 (.13-1.09) [.07]</td>
</tr>
<tr>
<td>Headache</td>
<td>18 (36%)</td>
<td>11 (16%)</td>
<td>10.5 (42) (23%)</td>
<td>1.25 (55-2.83) [.59]</td>
<td>.62 (.26-1.42) [.28]</td>
<td>.56 (.23-1.42) [.23]</td>
</tr>
<tr>
<td>Muscle Aches</td>
<td>26 (54%)</td>
<td>15 (22%)</td>
<td>28.8 (115) (62%)</td>
<td>.70 (.30-1.66) [.42]</td>
<td>.38 (.13-1.09) [.07]</td>
<td>.38 (.13-1.09) [.07]</td>
</tr>
<tr>
<td>Fatigue</td>
<td>27 (56%)</td>
<td>14 (21%)</td>
<td>4.5 (18) (9.7%)</td>
<td>1.25 (55-2.83) [.59]</td>
<td>.62 (.26-1.42) [.28]</td>
<td>.56 (.23-1.42) [.23]</td>
</tr>
<tr>
<td>Hospitalized</td>
<td>12 (25%)</td>
<td>7 (10%)</td>
<td>10.5 (42) (23%)</td>
<td>1.25 (55-2.83) [.59]</td>
<td>.62 (.26-1.42) [.28]</td>
<td>.56 (.23-1.42) [.23]</td>
</tr>
</tbody>
</table>

1No statistically significant differences between any symptoms were found using Fisher’s exact test.

*No statistically significant differences between any symptoms were found using Fisher’s exact test.
Hispanic

<table>
<thead>
<tr>
<th></th>
<th>Case-Case Travelers</th>
<th>Case-Case Non-Travelers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(OR (CI) [p-value])</td>
<td>(OR (CI) [p-value])</td>
</tr>
<tr>
<td>Exposure to untreated water</td>
<td>14.0 (3.2-61.1) [.00]*</td>
<td>Null (0 cases w/ exposure)</td>
</tr>
<tr>
<td>Raw dairy consumption</td>
<td>.96 (2.3-3.9) [.96]</td>
<td>1.34 (33-5.5) [.68]</td>
</tr>
<tr>
<td>Eggs</td>
<td>11.2 (2.0-62.8) [.01]**</td>
<td>1.5 (57-3.9) [.41]</td>
</tr>
<tr>
<td>Runny or raw eggs</td>
<td>2.5 (2-31) [.48]</td>
<td>.74 (19-2.8) [.66]</td>
</tr>
<tr>
<td>Attend a gathering</td>
<td>2.2 (.63-7.7) [.22]</td>
<td>.86 (26-2.8) [.8]</td>
</tr>
<tr>
<td>Poultry</td>
<td>.6 (.13-2.7) [.51]</td>
<td>.75 (26-2.1) [.6]</td>
</tr>
<tr>
<td>Hispanic</td>
<td>4.58 (.5-42.3) [.18]</td>
<td>1.1 (41-3.0) [.8]</td>
</tr>
<tr>
<td>Contact w/ someone with diarrhea</td>
<td>.93 (.39-2.2) [.88]</td>
<td>.73 (25-2.1) [.56]</td>
</tr>
<tr>
<td>Contact with ill animal</td>
<td>undefined</td>
<td>.29 (.03-2.4) [.25]</td>
</tr>
<tr>
<td>Youth</td>
<td>2.33 (.7-7.6) [.16]</td>
<td>.72 (28-1.84) [.49]</td>
</tr>
<tr>
<td>Hospitalized</td>
<td>.88 (.09-8.3) [.9]</td>
<td>1.8 (.74-4.6) [.19]</td>
</tr>
</tbody>
</table>

Table 4: Case-Case Stratification by Travel History.

Discussion

Conducting case-control and case-case studies within the same populations and general time frame allowed for a comparison of these methods and evaluation of the utility of the case-case design. This method is commonly used when PFGE data is available, but this data is not always available to delineate the ‘outbreak’ cases from other routine cases. This meant that not all ‘outbreak cases’ were in fact, tied to an outbreak and were part of the normal endemic cases seen in the county. However, this type of misclassification would bias the results towards the null because these cases would be more like the historical cases used as a comparison. Use of historical cases drawn from the same surveillance system as the outbreak cases reduces selection bias [35]. Recall bias is also minimized because all cases had symptoms and were interviewed within a similar time period following their laboratory confirmation. These studies can be conducted with existing case data, eliminating the staff time and effort needed to find and interview controls. While direct cost comparisons were not possible for
this study, we identified a significant amount of time was necessary to recruit each control relative to pulling historical case data. Furthermore, the use of a student response team to recruit and interview controls expanded the health department’s capacity to further investigate this outbreak.

Travel history had the strongest measure of effect in the case-control study. Foreign travel has been associated with Campylobacter infection in a number of case-control studies [13,15,17,36] and, while including all foreign travel had a stronger strength of association, given the proximity to the U.S. Mexico border, the association particularly with Mexico was a finding that has relevance to public health planning in this region. Interestingly, when this association was adjusted for Hispanic ethnicity, it was clear that being Hispanic was a confounder in relationship between travel history and development of disease but not an effect modifier (ethnicity did not change the measure of association when stratified by travel history).

One limitation was in the recruitment of phone controls. While the recruitment goal was a 2:1 match of controls to cases, this did not occur in all the age groups, particularly among boys (age 1-19). This group represented 33% of all cases reported from outbreak cases but only 13% of the phone controls. This difference may have led to a larger effect size being attributed to youth in the case-control analyses.

In retrospect, we would not choose to match on gender for children to help the recruitment efforts of controls. In addition, since almost 58% of the outbreak cases resided in three adjoining zip codes, it may have been a better strategy to match based on residence of the cases. However, based on prior experiences, matching based on location greatly increases recruitment time and resources. Finally, there were no bilingual interviewers available for control interviews, however only 17 (0.8%) of the 2,017 control calls required a Spanish speaking interviewer and given the response rate, would have only resulted in another 2-3 completed control interviews.

Unlike reported studies where the case-case and case-control studies resulted in similar findings [4,37,38], identified risk factors differed somewhat between study designs. This may be explained by slight differences in questionnaire design between the two studies. The case questionnaire asked about ‘unpasteurized eggs’ while controls were asked only about ‘eggs’ with results being a higher percentage of controls, than cases reported eating eggs. While this may have in fact been the case, it is possible that the cases were confused with the term ‘unpasteurized eggs’. For untreated water, cases were asked about exposures that took place in the 2 weeks prior to their illness which was during the summer months. Since control interviews were conducted in November, information on exposures during the months July-September was gathered, instead of October/November when recreational water exposures would be lower. This decision resulted in controls reporting for a 12 week period of time, rather than the 2 weeks for the cases, moving the association towards the null. Given the longer exposure period, it was not surprising that the proportion of controls that reported a water exposure was higher than that of cases. However, when the case-control analysis was limited to only exposures reported for the 4 weeks of August, the odds ratio for exposure to untreated water increased to 1.87 (95% CI 1.65-5.4).

Given that travel was such a significant risk factor in the case-control results, it was examined for interaction within the case-case study. When the case-case study was stratified by travelers and non-travelers, travel history was a clear effect modifier of certain risk factors. Exposure to untreated water and fresh egg consumption, in particular, was higher among travelers. The strong association between travel and campylobacteriosis in the case-control study was consistent with previous studies [13,15,31,36,38,39]. The lack of an increase in odds ratio in the case-case analyses was not surprising, given rates of travel among cases have been reported to be as high as 20% [11] to 42% [39]. Surprisingly, given the high odds of travel associated with disease, the specific exposures related to travel have not been explored in the literature. The case-case study identified the likely exposures that were being overwhelmed by prevalence of travel among the cases in the case-control study. This method could be very useful to further explore the true exposure leading to disease among travelers.

### Conclusion

Risk factors were identified by both studies and provided valuable information on exposures associated with Campylobacter infections. More importantly, the use of concurrent study methods helped to identify more relevant exposures related to travel, since travel in itself is not a true pathogen exposure. These results show that case-case studies can be used not only as an efficient precursor to a case-control study, but can provide additional information about the transmission of disease. Given the limited resources available in many state and local health departments, new and innovative ways of comparison study designs should be considered and investigated.

### References
