

Original Paper


Cite this article: Smith CR, Kershaw T, Johnson K, Meghnath K (2019). An outbreak of hepatitis A in Canada: The use of a control bank to conduct a case-control study. *Epidemiology and Infection* **147**, e300, 1–8. <https://doi.org/10.1017/S0950268819001870>

Received: 11 July 2019
Revised: 24 September 2019
Accepted: 13 October 2019

Key words:
Hepatitis A; outbreaks; case-control study

Author for correspondence: C. R. Smith,
E-mail: courtneyr.smith@canada.ca

An outbreak of hepatitis A in Canada: The use of a control bank to conduct a case-control study

C. R. Smith¹ , T. Kershaw¹, K. Johnson² and K. Meghnath³

¹Outbreak Management Division, Public Health Agency of Canada, Guelph, Ontario, Canada; ²Enteric, Zoonotic and Vectorborne Disease Unit, Public Health Ontario, Toronto, Ontario, Canada and ³Enteric, Zoonotic, Vectorborne and Environmental Infections Program, BC Centre for Disease Control, Vancouver, British Columbia, Canada

Abstract

An outbreak of 18 cases of hepatitis A virus infection across five Canadian provinces was investigated. Case onsets occurred between October 2017 and May 2018. A retrospective matched case-control study was conducted to identify the likely source of the outbreak. Three matched controls were recruited for each case using a previously established control bank, supplemented by landline and cell phone call lists. Univariate and multivariate matched analyses were conducted to identify a potential outbreak source. Seventy-two per cent of controls were recruited through the control bank, and required on average 25.5 calls per recruited control; 20% of controls were recruited through a landline sample and 8% of controls were recruited through a cell phone sample, requiring an average of 847.3 and 331.7 calls per recruited control, respectively. Results of the analysis pointed to shrimp/prawns (odds ratio (OR) 15.75, $p = 0.01$) and blackberries (OR 7.21, $p = 0.02$) as foods of interest, however, an outbreak source could not be confirmed. The control bank proved to be a more efficient method for control recruitment than random call lists. Expanding the control bank size and using alternative methods, such as online surveys, may prove beneficial for increasing the timeliness of a case-control study during an outbreak investigation.

Introduction

Hepatitis A is a disease caused by the hepatitis A virus (HAV). HAV can be transmitted from person to person or through the consumption of contaminated food products [1, 2]. HAV is prone to foodborne outbreaks as it is able to withstand most processes that are commonly used to control bacterial pathogens in food, such as mild pasteurisation and exposure to high temperatures [3]. HAV is also resistant to desiccation, and can remain infectious for several months on frozen foods [3]. Foodborne outbreaks of HAV have been associated most commonly with fresh and frozen produce, such as semi-dried tomatoes [4, 5, 6], frozen berries [7, 8] and pomegranate arils [9]. Outbreaks of HAV have also been associated with various shellfish such as raw scallops [10] and oysters [11, 12]. Hepatitis A is a nationally notifiable disease in Canada, and is reported at a rate of 0.68 per 100 000 persons [13].

Many investigations into foodborne outbreaks include case-control studies as a means to identify a suspect source. Recruitment strategies for controls in these studies vary, but can include random digit dialling, population registries, nearest-neighbour or friend recruitment and hospital or clinic recruitment [14, 15]. In a research setting, the critical focus of a case-control study design is often the development of a scientifically sound and cost-effective method for recruiting controls [14]. Alternatively, in an outbreak context, the critical focus is often the most time-sensitive means of recruitment, as public health action often depends on quick results. This is particularly the case for outbreaks of HAV, where post exposure prophylaxis is possible within the first 14 days after exposure [16].

Although the use of random digit dialling or listed phone numbers are common for case-control studies, the number of calls to recruit a single control is often high due to non-answers, invalid numbers and a lack of desire to participate. For example, a cryptosporidium investigation by the United States Centers for Disease Control and Prevention required approximately 11 400 calls to complete 151 cases and 302 control interviews [17]. Response rates for random digit dialling methods have seen a dramatic decline in recent years, due in part to the popularity of caller identification [15, 18, 19]. The use of a pre-established control bank of consenting participants could result in a higher response rate than that of random-digit dialling or calling listed numbers within the general population, potentially reducing time and cost to complete a case-control study [20]. A further limitation for random digit dialling methods is the lack of ability to make targeted calls, adding an additional challenge to recruiting controls in the context of a matched study. A control bank, complete with demographic

© Public Health Agency of Canada 2019. This is an Open Access article, distributed under the terms of the Creative Commons Attribution-NonCommercial licence (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

information for each individual, is one solution to this issue. Several case-control studies have used the control bank approach to recruit controls, including studies in Australia for outbreaks of *salmonella*, *campylobacter* and *Escherichia coli* [21, 22, 23], and a Canadian study examining an outbreak of *cyclospora* [20].

In July 2018, 18 cases of laboratory-confirmed HAV infection with genotype 1a were identified across five geographically dispersed provinces in Canada. All cases had one of two genetically related ribonucleic acid (RNA) fingerprints. Although genotype 1a is common in Canada, the RNA nucleotide sequences specific to this outbreak had not been previously associated with any other outbreak. Onset dates ranged between October 2017 and May 2018. Given the multi-jurisdictional nature of the outbreak, case-to-case transmission was considered unlikely. High risk exposures associated with HAV transmission, such as intravenous drug use and homelessness, were not reported among cases. A national outbreak investigation was initiated to identify the likely source of illnesses and determine appropriate public health action. Initial frequency-based analysis of exposure data did not identify a suspect source. A case-control study was conducted to assess whether the frequency of exposures to various food items was significantly different among cases and controls. A control bank, established via a previous study [24], was chosen as the primary method to recruit controls. The objective of this paper is to describe the results of the case-control study, and examine the advantages and disadvantages of using a control bank to facilitate the recruitment of controls in an outbreak investigation.

Methods

Study design

The study was a retrospective matched case-control design. Three controls were recruited for each case and matched based on the first digit of Forward Sortation Area (FSA), age group (0–19, 20–49, 50–69, 70+) and sex (male, female). The FSA represents a geographical unit in Canada. The first digit of the FSA is a letter that identifies the province or territory. For two provinces, this first digit further specifies a particular part of the province. Cases in this study were distributed across five provinces, suggesting that geography did not appear to be closely associated with exposure. However, geography was included as a matching criteria as dietary habits differ significantly across Canada [24], and the investigators wanted to ensure that the control group would have the same geographic diversity as the cases. Given that the control bank consisted of a higher proportion of women than men, and a higher proportion of older adults than younger adults, additional matching based on age and sex ensured that the controls selected would not be biased based on the sampling frame.

Selection of cases

A confirmed case was defined as a resident of or visitor to Canada with laboratory-confirmed HAV infection with genotype 1A and one of two genetically related outbreak RNA fingerprints; an onset date on or after 1 October 2017; and no close contact with a confirmed case 15 to 50 days prior to illness onset. A secondary case was defined as a resident of or visitor to Canada with laboratory-confirmed HAV infection, close contact with a confirmed case 15 to 50 days prior to illness onset and symptom onset at least 15 days after the laboratory confirmed case. Cases were included in the study if they met the confirmed case

definition, had a completed questionnaire, and did not report travel outside of Canada during their exposure period.

Genotyping and RNA fingerprinting of clinical isolates was completed at the National Microbiology Laboratory to compare isolates. A 373 nucleotide fragment in the VP1-2A region was amplified to determine genotype and RNA fingerprint. Isolates were considered to have the same RNA fingerprint when all 373 nucleotides were identical.

Selection of controls

Healthy community controls were recruited by a hired third party contractor using a control bank established as part of a separate Canadian study [24]. The control bank was established in 2014 and 2015 and contained contact and demographic information for Canadians that provided consent to have their information collected and stored for future enteric disease outbreak investigations. As of July 2018, when the case control study was initiated, the control bank included 2113 Canadians from all provinces and territories, of which 122 were age, gender and geographic matches to the cases in the current study. When recruitment from the control bank was exhausted, controls were contacted through the use of random landline and cell phone numbers that were obtained from a sampling company. Individuals were excluded from being controls if they had a previous HAV infection and/or ever had symptoms consistent with HAV infection (i.e. jaundice, dark urine or pale/clay-coloured bowel movements), if they were previously vaccinated against HAV, if they travelled outside of Canada in the 15–50 days prior to interview, and if they travelled outside of their province for >14 days total in the 15–50 days prior to interview.

Exposures

Food items were selected for inclusion in the case-control study based on a review of case exposure information and an examination of potential sources of HAV infection from a literature review. The food items included in the study were: ham deli meat, shrimp/prawns, yogurt, fruit smoothies, dried fruits (such as raisins, figs, dates, cranberries and/or apricots), avocado, semi-dried tomatoes, green onion, bagged or pre-washed lettuce in a salad mix, other bagged salad such as broccoli slaw or coleslaw, spinach, strawberries (fresh or frozen), raspberries (fresh or frozen), blackberries (fresh or frozen), blueberries (fresh or frozen) and other berries (fresh or frozen).

Data collection

Exposure information for cases was collected through initial interview by local public health officials using routine provincial HAV questionnaires, as per routine practice. After a national outbreak investigation was initiated, select cases were re-interviewed by the Public Health Agency of Canada (PHAC) to gather additional detail on food items and to ask about exposures not included in the routine interview. Cases were selected for re-interview based on the level of detail captured for exposures in the initial interview. Controls were interviewed by a hired third party contractor using a control questionnaire script. The exposure period was 15–50 days prior to symptom onset for cases, and 15–50 days prior to interview date for controls.

Analysis

To examine the usefulness of the control bank, the average number of calls required per recruited control was calculated for each recruitment method by gender and age group. The response rate was also calculated for each recruitment method by dividing the number of recruited controls by the total number of individuals contacted. The proportion of potential controls excluded as a result of non-response, not meeting matching criteria, not meeting inclusion criteria, refusals and numbers out of service was also calculated for each recruitment method.

To investigate a potential suspect source for the outbreak, a matched analysis of cases and controls was conducted. McNemar's odds ratios (ORs) for matched pairs were calculated to provide an estimate of risk associated with exposure and contraction of HAV. Exposures identified in univariate analysis with p -value < 0.2 and with $OR > 1$ were included in a multivariate analysis. Models were made using backward stepwise selection, with variables remaining in the model if they changed the significant coefficients by more than 20%. All analyses were conducted in Stata at an α level of 0.05.

Results

Of the 18 cases in this outbreak investigation, 12 met the inclusion criteria for the study. Six cases were excluded because exposure information was unavailable ($n = 3$), because they travelled outside of Canada during their exposure period ($n = 2$), or because they were a secondary case ($n = 1$). The demographic characteristics of the cases are outlined in Table 1.

Three controls were matched to each case, for a total of 36 controls. Twenty-six (26/36; 72.2%) of the controls were obtained through the control bank. A total of 662 calls were required to complete interviews for the 26 controls recruited through this method, for an average of 25.5 calls per control (Table 2). The response rate for controls recruited from the control bank was 21%. Females between the ages of 20–49 were the most difficult to recruit using this method, requiring an average of 52.3 calls per control. Seven (7/36; 19.5%) of the controls were obtained through a random dialling from a listed landline sample. A total of 5193 calls were required to complete interviews for the seven controls recruited through this method, for an average of 847.3 calls per control. The response rate for the random listed landline sample was 0.3%. For this recruitment method, males between the ages of 20–49 were the most difficult to recruit, requiring an average of 2347 calls per control. Lastly, three (3/26; 8.3%) of the controls were obtained through a random cell phone sample. A total of 995 calls were required to complete interviews for the three controls recruited through this method, for an average of 331.7 calls per control. The response rate for the random cell phone sample was 0.3%. All three controls recruited with this method were males between the ages of 20–49. Overall, it took 4 weeks to complete interviewing of all 36 controls.

Reasons for exclusion varied by method of recruitment (Fig. 1). When using the control bank to recruit controls, the most common reasons for exclusion were the individual not answering the phone (46% of exclusions) or not meeting the inclusion criteria (29% of exclusions). For the landline sample, the most common reasons for exclusion were the individual not answering the phone (33% of exclusions) and the number not being in service/wrong number (27% of exclusions). For this

Table 1. Demographic characteristics of cases in the outbreak and cases included in the case-control study

Variable	All outbreak cases ($n = 18$)	Cases included in the case-control study ($n = 12$)
Age		
Mean	41	46
Median	40	48
Range	3–69	12–69
Sex		
Male	9/18 (50%)	6/12 (50%)
Female	9/18 (50%)	6/12 (50%)
Province		
1	4/18 (22.2%)	1/12 (8.3%)
2	2/18 (11.1%)	1/12 (8.3%)
3	2/18 (11.1%)	2/12 (16.7%)
4	4/18 (22.2%)	4/12 (33.3%)
5	6/18 (33.3%)	4/12 (33.3%)

method of recruitment, exclusions due to refusal or language barrier were also quite high (19% of exclusions). Lastly, for recruitment via the cell phone sample, the most common reasons for exclusion were the number not being in service/wrong number (51% of exclusions) and the individual not answering the phone (40% of exclusions).

Food frequencies for cases and controls are summarised in Table 3. Of the 12 cases, five had exposure information based on initial interview alone, while seven cases had exposure information based on initial interview and a re-interview conducted by PHAC. The most frequently reported exposures by cases were: ham deli meat (6/7; 86%), shrimp/prawns (8/10; 80%), yogurt (5/5; 100%), bagged salad (7/8; 88%), fresh strawberries (5/5; 100%), fresh raspberries (6/7; 86%), fresh blackberries (7/8; 88%), any blueberries (10/11; 91%) and fresh blueberries (9/10; 90%). The most frequently reported exposures by controls were ham deli meat (31/36; 86%), any strawberries (33/36; 92%), fresh strawberries (32/36; 89%) and any blueberries (29/36; 81%). Shrimp/prawns had 16 times greater odds of being consumed by cases compared to controls, which was statistically significant at the 0.05 α level ($OR 15.75$, $p = 0.01$; Table 3). Blackberries had seven times greater odds of being consumed by cases compared to controls, which was also statistically significant at the 0.05 α level ($OR 7.21$, $p = 0.02$). All other food exposures were either not significant at the 0.05 α level, or indicated a protective effect.

Exposures identified in univariate analyses with a p -value < 0.2 and with an $OR > 1$ were included in a multivariate, conditional logistic regression model. These variables included shrimp/prawns, other bagged salad, and blackberries (Table 3). Due to small cell sizes, the full model did not have sufficient data to produce results. Following the procedure of backward stepwise regression, subsequent models were run, removing variables one at a time, in order of least to most significant based on univariate analyses. Subsequent multivariate models also had insufficient data to produce results.

Table 2. Numbers of controls recruited and calls made for each recruitment method. Control bank recruitment was exhausted before proceeding to the landline and cell phone samples

	Control bank					Random listed landline sample					Random cell phone sample				
	Number of individuals called (i.e. available in bank)	Number of phone calls made	Number of controls recruited (% of category total)	Average number of individuals called per control	Average number of calls per control	Number of individuals called	Number of phone calls made	Number of controls recruited (% of category total)	Average number of individuals called per control	Average number of calls per control	Number of individuals called	Number of phone calls made	Number of controls recruited (% of category total)	Average number of individuals called per control	Average number of calls per control
Controls recruited (<i>n</i> = 36)															
0–19 females (<i>n</i> = 3)	10	104	3 (100)	3.3	34.6	–	–	–	–	–	–	–	–	–	–
0–19 males (N/A)	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
20–49 females (<i>n</i> = 6)	34	209	4 (67)	8.5	52.3	261	349	2 (33)	130.5	174.5	–	–	–	–	–
20–49 males (<i>n</i> = 9)	22	150	4 (44)	5.5	37.5	1781	4694	2 (22)	890.5	2347	993	995	3 (100)	331	331.7
50–69 females (<i>n</i> = 9)	29	145	7 (78)	4.1	20.7	75	147	2 (22)	37.5	73.5	–	–	–	–	–
50–69 males (<i>n</i> = 9)	27	54	8 (89)	3.4	6.8	3	3	1 (11)	3	1	–	–	–	–	–
Total	122	662	26	4.7	25.5	2120	5931	7	302.9	847.3	993	995	3	331	331.7

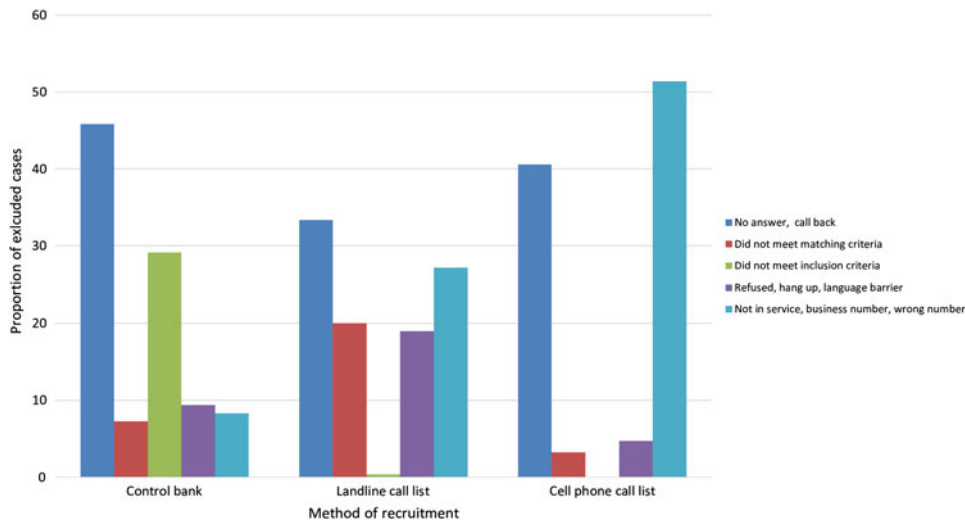


Fig. 1. Reasons for exclusion by method of recruitment.

Discussion

Use of a control bank

Recruitment of neighbourhood or friend controls is one common method of control recruitment in a fast-paced outbreak scenario. In the current outbreak, however, there were several limitations that prevented the use of this method. First, relying on neighbourhood controls assumes that all cases would be able to recruit a minimum of 1–3 controls. However, in our outbreak not all cases were able to be re-contacted after the conclusion of their initial interview; using this method would therefore result in several cases without suitable controls. Second, neighbours or friends may be more likely to share similar food habits as their corresponding case, leading to an underestimate of any effect [25]. With a small-sized outbreak, the underestimate of an effect could be detrimental to identifying a potential source. To address these limitations, recruitment of controls via the control bank was chosen as the primary recruitment method.

Overall, the use of the control bank for recruitment in this case-control study was successful. This method had a higher response rate and required the fewest number of calls per control compared to random dialling, likely as a result of several factors. The control bank consists of individuals who have previously agreed to participate in such studies, increasing the response rate when compared to calls to the general population. In this study, the response rate for the control bank was almost 70 times greater than that of the landline or cell phone samples. Similarly, in a recent case-control study related to an outbreak of cyclosporiasis, the response rate for the control bank was 60% compared to only 24% via random digit dialling [20].

The control bank also allows for targeted calls to individuals based on the matching variables of the study. For example, the control bank included information on area of residence for each individual, which was useful for geography-based matching. Given that area codes are not always representative of an individual's home address [15], the use of random digit dialling or call lists likely increases the number of calls needed to obtain a control if geography-based matching is required. The inclusion of age and gender information within the control bank also proved advantageous in comparison with landline or cell phone samples in the context of age and gender matching. Interestingly, the control bank group had the highest proportion of exclusions due to not meeting the inclusion criteria.

Compared to landline and cell phone sample recruitment, more of the control bank calls were able to successfully reach a matched individual, and therefore complete the inclusion criteria questions. In addition, compared to the general population (i.e. individuals contacted via the landline or cell phone call lists), it is possible that those individuals who agreed to be included in the control bank are also more likely to be vaccinated for HAV, perhaps a variation of the 'healthy volunteer effect' [26].

Although recruitment via the control bank was the most efficient method for control recruitment, the decision to include three separate matching variables extended the time needed to find eligible controls through all recruitment methods. In this study, it took 4 weeks to recruit 36 controls. Information on the length of time required to recruit controls in other similar case-control studies is limited. Given that case-control studies are often used to collect timely data to inform public health action, reducing the number of matching variables may be an important consideration to expedite the data collection process. A solution may be to control for certain variables in analysis when possible, rather than design.

It's also important to highlight that the initial set up of a control bank may be resource intensive. The current control bank was created through the completion of a separate study, taking place over a period of a year [24]. Although an existing control bank takes a minimal effort to maintain, its creation requires significant human resources to recruit controls, and may not be feasible for all jurisdictions. Over time, a control bank might become depleted if individuals move, change phone numbers or experience volunteer fatigue, resulting in a need to replenish the bank. In the current study, an expanded control bank would have likely expedited control recruitment. Having more individuals available in each geography, age and gender group may have avoided the need for landline and cell phone samples to supplement recruitment, and in turn hasten the recruitment process. For this reason, expanding the control bank would likely be of value for future case-control studies. Although the control bank was the most efficient method of recruitment in this study, it still required significant human resources to make the 662 calls necessary to obtain the 26 controls from the bank.

The finding that younger age groups were more difficult to recruit is likely indicative of a shift from a preference of phone contact to online contact, a transition that emulates the shift of the previous decade from landlines to cell phones [15]. In this sense future case-control studies may benefit from different avenues for control

Table 3. Food frequencies for cases ($n = 12$) and controls ($n = 36$) and matched ORs, confidence intervals and p -values for each food item

Food item	Exposed		Not exposed		% Exposed		OR	Lower confidence limit	Upper confidence limit	p value
	Cases	Controls	Cases	Controls	Cases	Controls				
Meat and seafood										
Ham deli meat	6	31	1	5	86	86	1.00	0.08	11.93	1.00
Shrimp or prawns	8	7	2	29	80	19	15.75	1.90	130.20	0.01
Vegetables										
Semi-dried tomatoes	0	4	7	32	0	11	x			x
Green onions	2	26	5	10	29	72	0.27	0.05	1.49	0.13
Bagged salad (lettuce in a salad mix)	7	27	1	9	88	75	1.00	0.20	4.95	1.00
Other bagged salad (e.g. broccoli slaw or coleslaw)	6	17	2	19	75	47	2.77	0.24	32.33	0.42
Spinach	6	22	2	14	75	61	x	x	x	x
Dairy										
Yogurt	5	28	0	8	100	78	x	x	x	x
Fruit										
Fruit smoothie	2	15	8	21	20	42	2.77	0.24	32.33	0.42
Dried fruits	7	23	4	13	64	64	x	x	x	x
Raisins	3	16	1	20	75	44	1.40	0.38	5.20	0.61
Figs	1	1	1	35	50	3	x	x	x	x
Dates	1	4	1	32	50	11	0.26	0.05	1.44	0.12
Dried cranberries	3	16	1	20	75	44	1.76	0.19	16.48	0.62
Dried apricots	1	2	1	34	50	6	3.12	0.56	17.23	0.19
Avocado	5	16	5	20	50	44	1.00	0.17	5.77	1.00
Strawberries	6	33	6	3	50	92	0.07	0.01	0.59	0.01
Fresh strawberries	5	32	0	4	100	89	x	x	x	x
Frozen strawberries	0	7	5	29	0	19	x	x	x	x
Raspberries	7	28	4	8	64	78	0.48	0.11	2.07	0.32
Fresh raspberries	6	26	1	10	86	72	4.07	0.34	48.23	0.27
Frozen raspberries	1	5	6	31	14	14	x	x	x	x
Blackberries	8	9	3	27	73	25	7.21	1.46	35.60	0.02
Fresh blackberries	7	7	1	29	88	19	x	x	x	x
Frozen blackberries	1	4	7	32	13	11	1.00	0.10	9.61	1.00
Blueberries	10	29	1	7	91	81	2.28	0.17	30.72	0.53
Fresh blueberries	9	27	1	9	90	75	2.81	0.30	26.19	0.36
Frozen blueberries	1	5	9	31	10	14	0.57	0.06	5.33	0.62
Other berries	1	5	8	31	11	14	0.57	0.06	5.33	0.62
Fresh other berries	0	2	1	32	0	6	x	x	x	x
Frozen other berries	0	0	1	34	0	0	x	x	x	x

OR, odds ratio; x, odds ratio could not be calculated.

recruitment, such as online surveys. Online surveys may result in expedited data collection, facilitate recruitment of demographic groups that are more difficult to reach by phone, and lower study costs. However, online surveys may serve to exclude other groups, such as seniors. In addition, online surveys could lead to unrepresentative samples, depending on where the survey is hosted and the population that visits a given website.

Given that this study was designed primarily to identify a suspect source of the outbreak, and not to conduct a comparison of recruitment methods, statistical analyses were not conducted on recruitment numbers. In addition to a very small sample size across all recruitment methods, some demographic subgroups were not represented in landline or cell phone recruitment, as the target number of controls was reached with the control

bank alone (e.g. 0–19 year old females). The observations made in this study serve to provide a crude comparison of various methods of recruitment for consideration by other public health professionals when conducting studies for outbreak investigations.

Results of the study

Findings of the case-control study univariate analysis pointed to shrimp/prawns and blackberries as food items of interest. Because there were no leftovers available from case homes for sampling, or product specificity (e.g. brand information) from cases that reported exposures to these foods, it was not possible to conduct traceback or an environmental investigation at the production level.

Although a food safety investigation was not possible, shrimp/prawns and blackberries both present plausible sources for this outbreak. When considering all 17 confirmed cases in this outbreak, shrimp/prawns was one of the most commonly reported exposures. The long shelf-life of frozen shrimp/prawns aligns well with the long duration of the outbreak (8 months), thereby increasing its plausibility as a suspect source. Outbreaks of HAV are commonly associated with shellfish, typically occurring as a result of contamination of harvest beds with human sewage [27, 28]. Shellfish implicated in HAV outbreaks have included raw oysters [13, 29, 30], clams [31, 32] and less commonly, shrimp [33].

Blackberries also produced a significant OR in the univariate analysis, with most cases reporting fresh blackberry exposure. The OR may be underestimated, as controls were interviewed exclusively in the summer, a time when berry consumption is typically higher [24]. One case reported visiting an area of Mexico known for blackberry production during their exposure period. In addition, general import patterns indicate that blackberries in Canada are imported from October to May, with a switch to domestic suppliers in the summer months (unpublished data). This import pattern aligns with the timing of case onsets in this outbreak, which also occurred from October to May. Berries have been identified as the source of numerous HAV outbreaks globally. For example, mixed frozen berries were implicated in outbreaks in Italy [34] and Ireland [35], frozen strawberries were implicated in a multinational Nordic outbreak [8] and fresh blueberries were identified as a suspect source in an outbreak in New Zealand [36].

The high ORs of blackberries and shrimp/prawns may also be an indication of a potential confounding variable, as these exposures are not commonly reported in the Canadian population [24]. Without a multivariate analysis, it is difficult to determine the independent effect of shrimp/prawns or blackberries on the outcome. Limited product specificity from cases and an absence of any product leftovers also prevented a food safety investigation, and therefore, the confirmation of a source.

Interestingly, strawberries appeared to have a protective effect, with an OR less than 1. This again might be a result of the seasonality of control interviewing. While cases had their exposure periods distributed throughout the year, all controls were interviewed in the month of August. This could have increased the frequency of strawberry consumption for controls compared to cases, thereby creating the appearance of a protective effect.

This case control study has several limitations. First, while case exposures were dispersed over an 8 month period, from October 2017 to May 2018, control exposures were collected only in the month of August. As mentioned previously, this difference may

affect the comparability of food frequencies between the two groups. Second, accuracy of recall may have differed between cases and controls, as cases were often interviewed and/or re-interviewed weeks or months after their symptom onset; alternatively, controls recalled exposures in the 15–50 days preceding the date of interview. There were only 12 cases included in this case-control study, with some cases having incomplete data on various exposures of interest. This limitation prevented the calculation of some matched ORs and resulted in reduced precision and large confidence intervals. In addition, small cell sizes prevented the completion of a multivariate conditional logistic regression model. Without a multivariate model, it is possible that confounding impacted results.

Conclusion

Findings of the study pointed to shrimp/prawns or blackberries as foods of interest in this outbreak of HAV in multiple provinces across Canada. Although the source could not be confirmed, the methodology of the study allowed for a useful comparison of recruitment strategies. Control banks continue to prove their utility over landline and cell phone samples, but other avenues of recruitment, such as online surveys, may be increasingly useful in the future.

Acknowledgements. The authors would like to acknowledge the contributions of Anton Andov, Eleni Galanis, Marsha Taylor, Victor Mah, Colette Gaulin, Caroline Duchesne, Bev Billard and Sarah Fleming. The authors would also like to thank local public health colleagues and public health laboratories in British Columbia, Alberta, Ontario, Quebec and Nova Scotia as well as the British Columbia Centre for Disease Control, Alberta Health, Alberta Health Services, Public Health Ontario, Ontario Ministry of Health and Long-Term Care, Ministère de la Santé et des Services sociaux du Québec and the Nova Scotia Department of Health and Wellness. Lastly, the authors would like to thank the Canadian Food Inspection Agency, Health Canada and the Public Health Agency of Canada for their support of this investigation.

Financial support. This work was supported by the Public Health Agency of Canada.

Conflict of interest. None.

References

1. Cuthbert JA (2001) Hepatitis A: old and new. *Clinical Microbiology Reviews* **14**, 38–58.
2. Acheson D and Fiore AE (2004) Hepatitis A transmitted by food. *Clinical Infectious Diseases* **38**, 705–715.
3. Cook N *et al.* (2018) Persistence of hepatitis A virus in fresh produce and production environments, and the effect of disinfection procedures: a review. *Food and Environmental Virology* **10**, 1.
4. Carvalho C *et al.* (2012) A possible outbreak of hepatitis A associated with semidried tomatoes, England, July–November 2011. *Eurosurveillance* **17**, 20083.
5. Donnan EJ *et al.* (2012) A multistate outbreak of hepatitis A associated with semidried tomatoes in Australia, 2009. *Clinical Infectious Diseases* **54**, 775–781.
6. Gallot C *et al.* (2011) Hepatitis A associated with semidried tomatoes, France, 2010. *Emerging Infectious Diseases* **17**, 566–567.
7. Gillesberg LS *et al.* (2013) Ongoing multi-strain food-borne hepatitis A outbreak with frozen berries as suspected vehicle: four Nordic countries affected, October 2012 to April 2013. *Eurosurveillance* **18**, 8–13.
8. Severi E *et al.* (2015) Large and prolonged food-borne multistate hepatitis A outbreak in Europe associated with consumption of frozen berries, 2013

- to 2014. *Eurosurveillance European Communicable Disease Bulletin* **20**, 21192.
9. **Collier MG *et al.*** (2014) Outbreak of hepatitis A in the USA associated with frozen pomegranate arils imported from Turkey: an epidemiological case study. *The Lancet Infectious Diseases* **14**, 976–981.
 10. **Viray MA *et al.*** (2018) Public health investigation and response to a hepatitis A outbreak from imported scallops consumed raw-Hawaii, 2016. *Epidemiology and Infection* **147**, 1–8. doi: <https://dx.doi.org/10.1017/S0950268818002844>.
 11. **Bialek SR *et al.*** (2007) Use of molecular epidemiology to confirm a multi-state outbreak of hepatitis A caused by consumption of oysters. *Clinical Infectious Diseases* **44**, 838.
 12. **Guillois-Becel Y *et al.*** (2009) An oyster-associated hepatitis A outbreak in France in 2007. *Eurosurveillance* **14**, 19144.
 13. **Government of Canada notifiable disease dataset extraction.** Available at <https://diseases.canada.ca/notifiable/extract-dataset> (Accessed 14 September 2019).
 14. **Bernstein L** (2006) Control recruitment in population-based case-control studies. *Epidemiology* **17**, 255.
 15. **Kalton G and Piesse A** (2007) Survey research methods in evaluation and case-control studies. *Statistics in Medicine* **26**, 1675.
 16. **Link-Gelles R, Hofmeister MG and Nelson NP** (2018) Use of hepatitis A vaccine for post-exposure prophylaxis in individuals over 40 years of age: a systematic review of published studies and recommendations for vaccine use. *Vaccine* **36**, 2745.
 17. **Fox LM *et al.*** (2005) Emergency survey methods in acute cryptosporidiosis outbreak. *Emerging Infectious Diseases* **11**, 729.
 18. **Kempf AM and Remington PL** (2007) New challenges for telephone survey research in the twenty-first century. *Annual Review of Public Health* **28**, 113.
 19. **Wang PP *et al.*** (2009) Validity of random-digit-dialing in recruiting controls in a case-control study. *American Journal of Health Behavior* **33**, 513.
 20. **Morton V *et al.*** (2019) Use of a case-control study and control bank to investigate an outbreak of locally acquired cyclosporiasis in Canada, 2016. *Canada Communicable Disease Report* **45**, 225.
 21. **McPherson M *et al.*** (2009) Serogroup-specific risk factors for Shiga toxin-producing *Escherichia coli* infection in Australia. *Clinical Infectious Diseases* **49**, 249.
 22. **Munnoch SA *et al.*** (2009) A multi-state outbreak of Salmonella Saintpaul in Australia associated with cantaloupe consumption. *Epidemiology and Infection* **137**, 367–374.
 23. **Stafford RJ *et al.*** (2007) A multi-centre prospective case-control study of campylobacter infection in persons aged 5 years and older in Australia. *Epidemiology and Infection* **135**, 978–988.
 24. **Canada** (2015) Infectious Disease Prevention and Control Branch. Foodbook Report.
 25. **Wacholder S *et al.*** (1992) Selection of controls in case-control studies: II. Types of controls. *American Journal of Epidemiology* **135**, 1029–1041.
 26. **Delgado-Rodriguez M and Llorca J** (2004) Bias. *Journal of Epidemiology and Community Health* **58**, 635.
 27. **Bellou M, Kokkinos P and Vantarakis A** (2013) Shellfish-borne viral outbreaks: a systematic review. *Food and Environmental Virology* **5**, 13–23.
 28. **Iwamoto M *et al.*** (2010) Epidemiology of seafood-associated infections in the United States. *Clinical Microbiology Reviews* **23**, 399.
 29. **Shieh YC *et al.*** (2007) Molecular confirmation of oysters as the vector for hepatitis A in a 2005 multistate outbreak. *Journal of Food Protection* **70**, 145–150.
 30. **Desenclos JC *et al.*** (1991) A multistate outbreak of hepatitis A caused by the consumption of raw oysters. *American Journal of Public Health* **81**, 1268.
 31. **Sánchez G *et al.*** (2002) Molecular characterization of hepatitis A virus isolates from a transcontinental shellfish-borne outbreak. *Journal of Clinical Microbiology* **40**, 4148–4155.
 32. **Tang YW *et al.*** (1991) A serologically confirmed, case control study, of a large outbreak of hepatitis A in China, associated with consumption of clams. *Epidemiology and Infection* **107**, 652.
 33. **Ashwell M *et al.*** (2004) Hepatitis A outbreak associated with a Mother's Day'yum cha'meal, Sydney, 1997. *New South Wales Public Health Bulletin* **15**, 6–9.
 34. **Rizzo C *et al.*** (2013) Ongoing outbreak of hepatitis A in Italy: preliminary report as of 31 May 2013. *Eurosurveillance* **18**, 20518.
 35. **Fitzgerald M *et al.*** (2014) Outbreak of hepatitis A infection associated with the consumption of frozen berries, Ireland, 2013 - linked to an international outbreak. *Eurosurveillance* **19**, 20942.
 36. **Calder L *et al.*** (2003) An outbreak of hepatitis A associated with consumption of raw blueberries. *Epidemiology and Infection* **131**, 745–751.