ORIGINAL ARTICLE

A spatial, temporal, and molecular epidemiological study of hospitalized patients infected with community-acquired or healthcare-associated *Clostridium difficile* in the Niagara Region, Ontario, Canada between September 2011 and December 2013

Maryam Salaripour, MSc, MPH, PhD;[1] Jennie Johnstone, MD, PhD, FRCPC;[2,3,5] George Broukhanski, PhD, MSc;[2,6] Michael Gardam, MSc, MD, CM, MSc, FRCPC[1,4,5]

[1]*Department of Health Policy and Management, York University, Toronto, ON, Canada*
[2]*Public Health Ontario, Toronto, ON, Canada*
[3]*St. Joseph's Health Centre, Toronto, ON, Canada*
[4]*Humber River Hospital, Toronto, ON, Canada*
[5]*Department of Medicine, University of Toronto, ON, Canada*
[6]*Department of Laboratory Medicine and Pathobiology, University of Toronto, ON, Canada*

Corresponding author:
Maryam Salaripour, MSC, MPH, PhD
School of Health Policy and Management
York University
Room 409 HNES Building
4700 Keele St.
Toronto, ON M3J 1P3
Canada
Email: msalaripour@hotmail.com; msalarip@yorku.ca

Alternate corresponding author:
Dr. Michael Gardam, MSc, MD, CM, MSc, FRCPC
Chief of Staff, Humber River Hospital
1235 Wilson Ave.
Toronto, ON M3M 0B2
Canada
Email: mgardam@hrh.ca

ABSTRACT

Objectives: To investigate and compare the incidence, geographical distribution, temporal patterns, and genetic relatedness of hospitalized patients with community-acquired *Clostridium difficile* infections (CA-CDI) and healthcare-associated *C. difficile* infections (HA-CDI) in the Niagara Region, Ontario over the time of a large, multi-hospital outbreak.

Methods: We conducted a retrospective case series study of the consecutive hospitalized confirmed CDI cases between September 2011 and December 2013 using SaTScan statistics and Statistical Process Control.

Results: Using provincial guidelines on classification of *C. difficile* cases, we estimated that, of the 629 CDI cases, 318 were CA-CDI and 311 were HA-CDI. The rate per 1,000 patient days for the entire study period for the hospitalized CA-CDIs was 3.9 CDIs/1,000 patient days and 3.8 CDIs/1,000 patient days for HA-CDIs. We identified spatial clusters for CA-CDIs using the first three digits of the patients' home postal codes. A temporal cluster of HA-CDI was identified after a period of time when a high number of CA-CDI cases were hospitalized. Molecular typing was done on 6% (40/629) of patients that met study definition; 13 were CA-CDI and 27 were HA-CDI. The majority (44.4%) of the NAP1 strains (12 of the 27 tested) were seen in patients with HA-CDI. Various unrelated strains were also identified.

Conclusions: Geographical clustering, temporal features, and genotypic features of CDI cases appear to be unique to CDI cases in the community, compared to those in hospital. Nonetheless, understanding the potentially bi-directional transmission pathways between hospitalized CA-CDI incidence and HA-CDI manifestation, as well as the community drivers of CA-CDIs, can inform clinical and public health patient safety and prevention policies.

KEYWORDS

Community and healthcare-associated *Clostridium difficile*; spatial; temporal; clustering

Acknowledgements: Special thanks to Infection Prevention and Control and Decision Support staff in the Niagara Health Service.
Conflicts of interest: None.
Funding: None.
Ethics approval: The protocol for this study was approved by York University's Office of Research Ethics and Niagara Health Service's Research Ethics Board.

INTRODUCTION

Clostridium difficile (*C. difficile*) has emerged as a significant source of infectious diarrhea beyond hospital settings, resulting in community-acquired *C. difficile* infections (CA-CDI) [1, 2]. The community reservoirs of CA-CDI remain unclear and many of those infected with CA-CDI do not have the conventionally established risk factors for healthcare-associated *C. difficile* infections (HA-CDI) [1, 3]. CA-CDI has been linked to various environmental sources such as floodwater, rivers, lakes, marine sediments, food, farm animals, and household pets [1, 4, 5]. Yet unlike other gastrointestinal infections, increases in CA-CDI rates in the northern hemisphere have not been associated with the warmer summer months [6-8]. Conversely, research to date has pointed to the increased use of antimicrobials in winter and spring months as a contributing risk factor in both CA-CDIs and HA-CDIs [9, 7].

Researchers also debate whether asymptomatic, previously hospitalized patients can be a source of transmission in the community and/or whether hospitalized CA-CDI plays a role in spreading the spores in hospital settings [10, 11]. In a bid to answer these questions, we investigated and compared the geographical distribution, temporal patterns, and genetic relatedness of CA-CDI and HA-CDI cases admitted to the Niagara Health System (NHS) between the summer of 2011 and the winter of 2013, during which period a series of *C. difficile* outbreaks occurred in the region's hospitals.

METHODS

Study design, study period, and setting

The design featured a retrospective case-series study of consecutive patients with confirmed CDI infections hospitalized in NHS hospitals between September 2011 and December 2013. NHS hospitals are the service providers for the Niagara Region in Ontario; they offer a wide range of programs and services to a catchment area spanning 12 municipalities with a population of approximately 430,000 [12].

Case definition, identification, data source, and privacy

Case definition and eligibility criteria followed the provincial guidelines for CDI prevention in healthcare settings [13], as reflected in NHS infection prevention and control (IPAC) policies (see Appendix A).

Cases of CDI were identified after laboratory testing of stool samples from symptomatic patients. Daily surveillance by IPAC service personnel at NHS sites confirmed the laboratory testing results. Confirmed cases were then approved and finalized in consultation with an external infectious diseases and infection control physician.

Data for this study were aggregated in a central database. Data came from each of the NHS hospitals' IPAC offices, administrative databases, and medical records. For more accurate data collection, expert personnel in the NHS Decision Support department conducted a retrospective query in its databases. Where data were missing, an electronic record review and a paper chart review were conducted using name, date of admission, and site-specific medical records numbered to match the records. A de-identified data set was used for final analysis.

C. difficile testing and strain typing methods

Between September 2011 and April 2012, all CDI samples were sent to an academic hospital laboratory that used DNA amplification technique to identify toxin-producing CDI strains. The BD GeneOhm™ Cdiff Assay had a sensitivity of 93.8% and a specificity of 95.5% [14]. From April 2012 to December 2013, NHS sent the CDI samples to an external commercial laboratory that used a Nucleic Acid Amplification Test (NAAT), the BD MAX™ Cdiff Assay, with a sensitivity of 96.3% and a specificity of 92.4% [15]. The provincial reference laboratory performed strain typing of the *C. difficile* isolates using a pulsed-field gel electrophoresis (PFGE) technique, a standard National Medical Laboratory procedure.

Statistical analysis

We stratified the CDI cases using CA-CDI and HA-CDI incidences. Data included CDI discreet count values, month and year of laboratory testing, the first three digits of the eligible CDI patients' postal codes or forward sortation area (FSA), and total patient days for all NHS sites per month for rate calculation. Rate per 100,000 population was calculated for CA-CDIs, and rate per 1,000 patient days was computed for HA-CDIs. When needed, we used information on the Niagara Region's population from Statistics Canada's 2011 census for data analysis. Monthly incidence measures were calculated and out-of-control ranges searched using Statistical Process Control (SPC).

Spatial Cluster Analysis

Complementing geographical distribution maps with spatial randomness statistical tests indicate whether the clustering is an act of chance or the result of an underlying risk factor. We performed a purely spatial and spatio-temporal scan of the CA-CDI and HA-CDI cases to test for the presence of patterns in their approximate geographical origin and conducted a space-time permutation study to identify clusters independent of time and location. The application of spatial Scan Statistics allows researchers to measure the significance and the location of a general or focused cluster [16] that subsequently leads to clues about the disease under investigation. Spatial scan statistics employ a likelihood ratio test to assess clusters of various sizes and adjust for multiple testing [17]. The Monte Carlo simulation of 999 randomizations of the data set ranks the likelihood of the cluster's significance [18]. Focused clusters are detected based on multiple circular (or other shaped) windows of variable sizes, scanning the given geographical area for the variable of interest. The null hypothesis of equivalent risk inside and outside the circular scan windows is rejected when the number of cases inside the cluster zone is more than the expected number of cases, independent of the specific geographical locations and administrative boundaries.

1. Purely spatial Scan Statistics for investigation of non-random clusters

Using a circular scan window centred on each possible point throughout the study area, this one-dimensional spatial Scan Statistic process compares the disease risk observed inside the window (cluster) with the risk outside the window (cluster). The most likely cluster has the highest likelihood ratio, with *p* values of 0.005 or less.

2. Spatio-temporal Scan Statistics for investigation of non-random clusters

The space-time Scan Statistics identify clusters throughout the study region by scanning for cases using a cylindrical window, where the base of the cylinder centres on one of the multiple centroids within the study area. The cylinder's height defines the time interval as a whole for the entire study period. The cylindrical window then scans the geographic base while changing the radius of the base as well as scanning for possible time intervals (changing the height of the cylinder).

3. Space-time permutation Scan Statistics for investigation of independent clusters

This model identifies the increased risk of a disease or differences in geographical distribution at different times by adjusting for time and space. Therefore, the number of observed cases in a cluster is compared to the expected number of cases if all cases were independent of each other in terms of infection time and spatial locations. The ability to adjust for purely temporal clusters of this type of scan means that it can highlight the locally initiated clusters.

For computation purposes, a Poisson distribution model was used while operating the SaTScan software. The first three digits of individuals' postal codes were used to identify the locations or smaller geographical units within the overall study area. The time precision was set by the day. Temporal and geographical checks were in place to ensure that all cases, controls, and populations were within the study's specified temporal period and geographical area. The maximum temporal cluster size was set for 50% of the study time and the maximum spatial cluster size was set for 50% of the study's at-risk population.

Temporal Cluster Analysis

SPC charts to investigate out-of-control abnormalities and outbreaks

The SPC approach was used to provide information on unusual variations and exceptional changes in CDI infection rates between months and seasons [19]. Rare events of disease clustering in a given time period are best explained by the Poisson process [20]. Therefore, this analysis used *u* control charts for discrete data (numerator) with a varying size of monthly patient days (denominator) to monitor the total number of incidents per month [19, 21] . Although in an industrial environment the use of process control charts with ±3-sigma control limit has been recommended (99.73% of all plot points in a normal distribution and stable process), use of a 3-sigma control limit has been questionable for healthcare [21]. Therefore, for epidemiological investigation of infections, more sensitive and less specific standards should be applied to increase the power and confidence of the "out of statistical control" state of CDI [20]. For this study, the control limits were set at ±2-sigma covering 95% of the plotted points; smaller variations in data could be identified, which, in practice, are signals for thorough epidemiological investigation [21]. Choosing a tighter control limit increases the rate of false positives or out-of-control points (type I error) to 5% for each plotted value (compared to 0.27%); this can also be clarified by epidemiological investigation.

Temporal Scan Statistics for investigation of non-random clusters

Scan Statistics identifies and evaluates clusters of cases in a purely spatial, purely temporal, or space-time setting [22]. A Bernoulli distribution is a 0/1 case-control type of binary data. To evaluate the temporal pattern of the CDIs and investigate non-random clustering, we used a purely temporal statistics test. HA-CDI were considered cases and CA-CDI were considered controls. Scan Statistics used multiple different window sizes to gradually scan across time and/or space and document the number of observed and expected observations inside the windows. The risk inside the clusters compared to outside the clusters, measuring for irregularity of the potential cluster, was based on a likelihood ratio [23]. The cluster that yielded the most extreme ratio was least likely to be by chance [23].

TABLE 1: Rate of hospitalized CA-CDI and HA-CDI for NHS hospitals between September 2011 and December 2013.

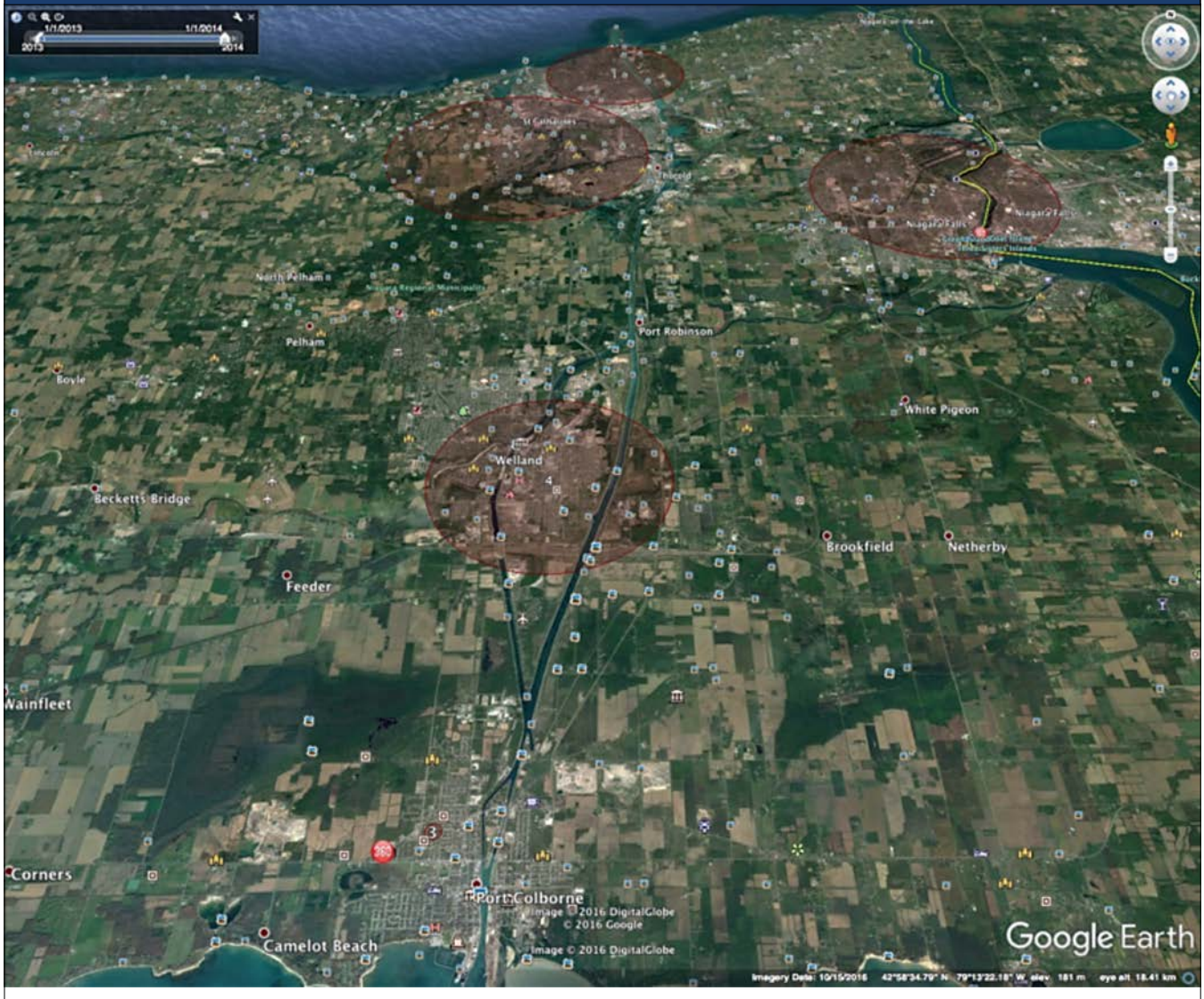
	2011	2012	2013	Overall rate for study period	Reported rates in other studies
CA-CDI (rate/100,000 population)	14.84	33.22	25.5		11.6 [36]
HA-CDI (rate/1,000 patient days)	3.7	3.36	4.34	3.83	0.3 in 2011 [29] 0.35 in 2012 0.33 in 2013

Return to TABLE OF CONTENTS

A purely temporal retrospective multivariate Scan Statistics was conducted, scanning for clusters with high rates using the Bernoulli model. The minimum temporal precision was set at one month and the maximum temporal cluster size was set at 50% of the study period. A maximum temporal cluster size limits the maximum size of the population at risk within the cluster to no more than 50% of the population at risk in the study [24, 16].

Random replication of the data set using computer simulation is a feature of Scan Statistics that adds to the power of the test. The number of replications under the null hypothesis for the standard Monte Carlo test was set at 999 to ensure statistical power for the Scan Statistic and the *p*-value calculation [22]. Under this setting, a high likelihood ratio rejects the null hypothesis and favours the clustering inside the scanning window(s) [22]. In this step of the temporal study, the null hypothesis assumed that the temporal clusters of hospitalized CA-CDI and HA-CDI occurred at the same time. The alternative hypothesis suggested the presence of clusters in hospitalized CA-CDI that did not show up at the same time as those in HA-CDI.

FIGURE 1A



Purely spatial scan statistics of the cases of CA-CDI in the Niagara Region based on their FSA or the first three digits of the postal codes between September 1, 2011 and December 31, 2013. Number of significant* clusters (in red) with high rates identified using discreet poison model ($p<0.05$). Five FSAs of the significant clusters:
L3K: $p<0.000$; N = 26; log likelihood ratio: 8.879242
L2G: $p<0.000$; N = 30; log likelihood ratio: 7.155066
L2E: $p<0.002$; N = 25; log likelihood ratio: 6.371665
L2N: $p<0.021$; N = 31; log likelihood ratio: 4.754386
L3B: $p<0.025$; N = 24; log likelihood ratio: 4.477314

Legends:
360 360 cities: Location of the panoramic cameras
small blue squares: Photos of landmark places on Google Earth

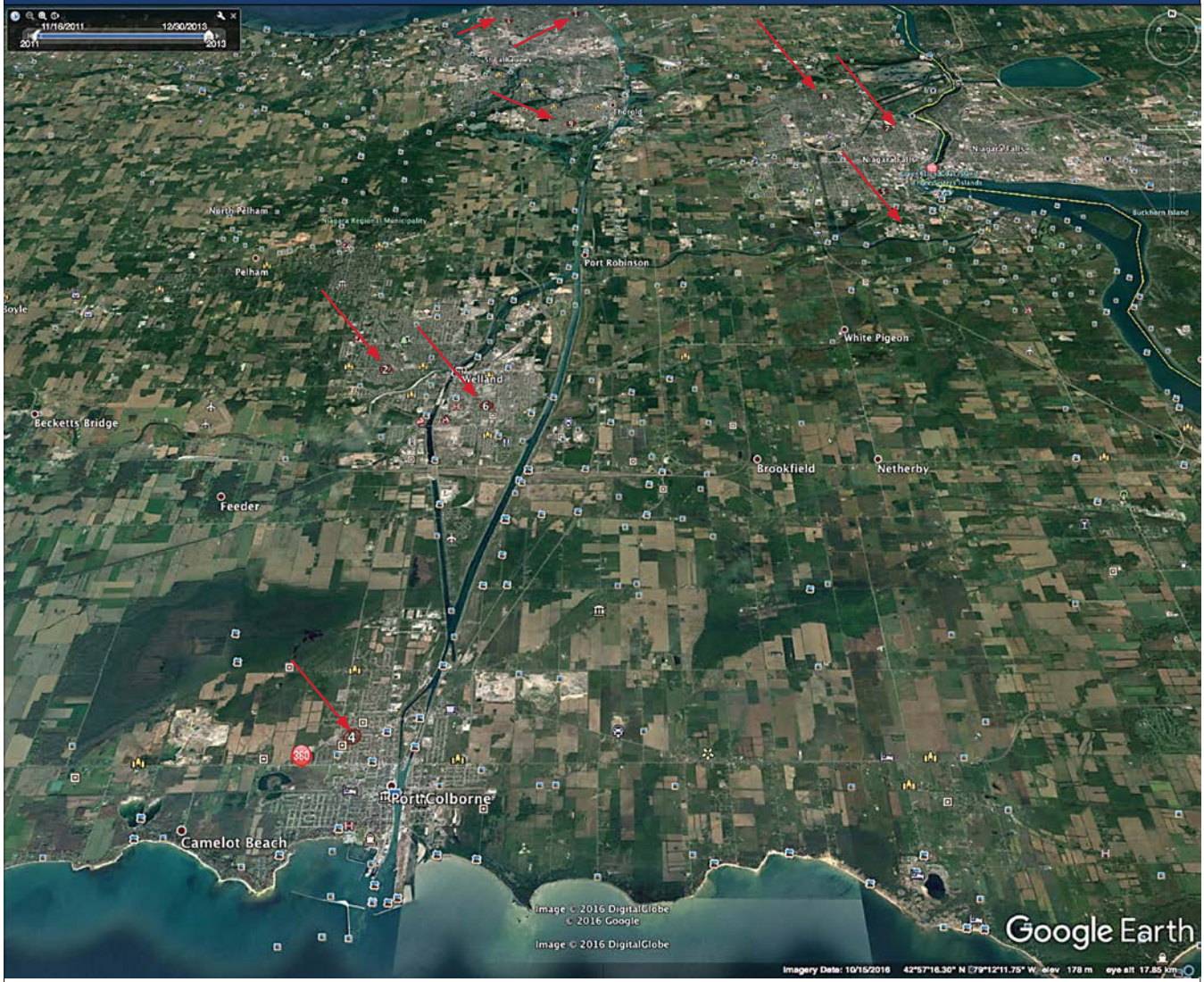
*A cluster is statistically significant when its log likelihood ratio is greater than the critical value, which is, for significance level Standard Monte Carlo Critical Values, 0.001: 7.924854; 0.01: 5.816313; and 0.05: 3.788838.

Test of seasonality

Using the additive seasonal cases, a seasonal mean of the absolute cases was calculated for each season to allow us to test for seasonality in our relatively small data sets. To find out whether seasonal properties have a role in the increase of CDI in certain periods, the time series data for both groups, CA-CDI and HA-CDI, were adjusted for a seasonal component [25]. Given the small number of seasonal points, an analytical approach was used rather than a graphical depiction of the seasonal influences, which is more common but mainly used for longer study periods. The additive seasonal indexes were calculated by subtracting the grand mean from each seasonal average. Subtracting each seasonal index from the associated seasonal measurement provided the seasonal adjusted values for each season [25].

Data were combined and stored in SPSS software, Version 21.0 (IBM Corp., Armonk, NY) and Microsoft® Excel for Mac, Version 15.27(161010). To determine the spatial and temporal Scan Statistics, we used SaTScan version 9.4.4 64-bit [26, 22]. Information on geocodes for FSAs was accessed from GPSVisualizer [27].

FIGURE 1B



Retrospective spatio-temporal scan statistics of the cases of CA-CDI in the Niagara Region based on their FSA or the first three digits of the postal codes between September 1, 2011 and December 31, 2013. Number of significant* clusters (in red) with high rates identified using discreet poison model ($p<0.05$), identified with an arrow. Nine FSAs of the significant* clusters were identified:

L2M: $p<0.000$; N = 16; log likelihood ratio: 26.189012; L3C: $p<0.000$; N = 14; log likelihood ratio: 25.998135
L2N: $p<0.000$; N = 12; log likelihood ratio: 21.666917; L3K: $p<0.000$; N = 20; log likelihood ratio: 20.866021
L2G: $p<0.000$; N = 0; log likelihood ratio: 19.762781; L3B: $p<0.000$; N = 9; log likelihood ratio: 17.984835
L2E: $p<0.000$; N = 15; log likelihood ratio: 17.067404, L2J: $p<0.000$; N = 10; log likelihood ratio: 16.349787
L2V: $p<0.004$; N = 8; log likelihood ratio: 13.471642

Legends:
360 cities: Location of the panoramic cameras
small blue squares: Photos of landmark places on Google Earth

*A cluster is statistically significant when its log likelihood ratio is greater than the critical value, which is, for significance level Standard Monte Carlo Critical Values, 0.001: 14.909737; 0.01:12.850207; and 0.05: 11.177605.

Return to TABLE OF CONTENTS

We categorized the results of molecular testing for each category as discreet counts and as proportions of the total specimens tested.

Ethics statement

The protocol for this study was approved by York University's Office of Research Ethics and Niagara Health Service's Research Ethics Board. This study entirely consisted of secondary data analysis of de-identified quality improvement patient data; therefore, the requirement for informed consent was waived.

Results

A total of 1,051 CDI cases were identified through laboratory detection of toxins produced by *C. difficile* strains, 629 of which met the eligibility criteria; 318 (50.1%) were CA-CDI and 311 (49.4%) were HA-CDI.

Table 1 lists the rate per 1,000 patient days for each study year for the HA-CDI category and the rate per 100,000 population for the CA-CDI category.

Spatial Scan Statistics

Figures 1A, 1B, and 1C provide the Scan Statistics of the purely spatial, spatio-temporal, and time-space permutations, respectively, of the hospitalized CA-CDI cases in the Niagara Region. Cluster ($p<0.005$) identification was based on their specimen collection date and their residential FSA information. The identified clusters have different geocodes, and the radii of the circular windows were set for 1 km for each cluster.

FIGURE 1C

Retrospective space-time permutation scan statistics of the cases of CA-CDI in the Niagara Region based on the FSA or the first three digits of the postal codes between September 1, 2011 and December 31, 2013. Number of significant* clusters ($p<0.05$). Three FSAs of the significant* clusters:
Location IDs included (L3K, L3B, L3C); $p<0.000$; N = 51, test statistic: 29.779385
Location IDs included (L2E, L2J, L2H, L2G); $p<0.000$; N = 45, test statistic: 19.770718
Location IDs included (L7T, L9A, L3M, L0R, L2R, L2N, L2S, L2M); $p<0.000$; N = 45, test statistic: 19.397479
*A cluster is statistically significant when its test statistic is greater than the critical value, which is,
for significance level Standard Monte Carlo Critical Values, 0.001: 11.580870; 0.01: 9.198556; and 0.05: 8.144111.

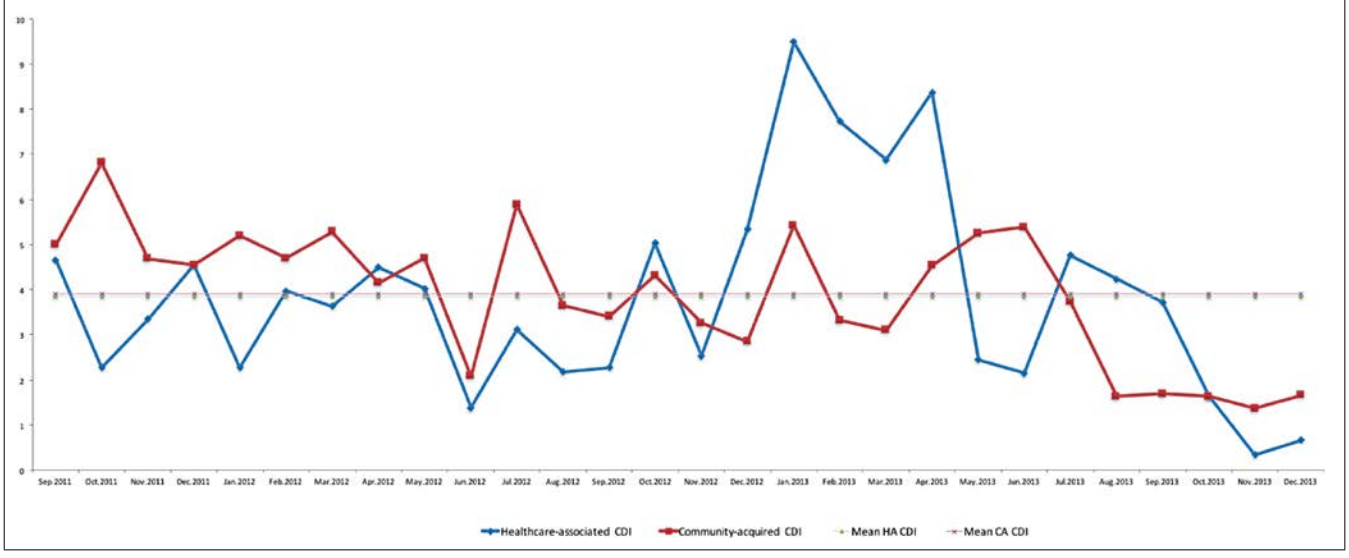
Legends:
360 360 cities: Location of the panoramic cameras
small blue squares: Photos of landmark places on Google Earth

CA-CDI Scan Statistics identified five very localized, purely spatial clusters ($p<0.05$), with one FSA attributing to each cluster (Figure 1A). The clustering of CA-CDI cases in the Niagara Region was predominantly positioned in urban zones (Figure 1B). Upon further exploration and plotting of public dwellings and communal residences (such as nursing homes, shelters, schools, or group homes), we noticed multiple assisted-living supportive housing demarcations within the perimeters of the spatial clusters of CA-CDI. Figure 1C signals the CDI cluster areas based on the number of observed to expected cases.

Time series analyses and SPC charts

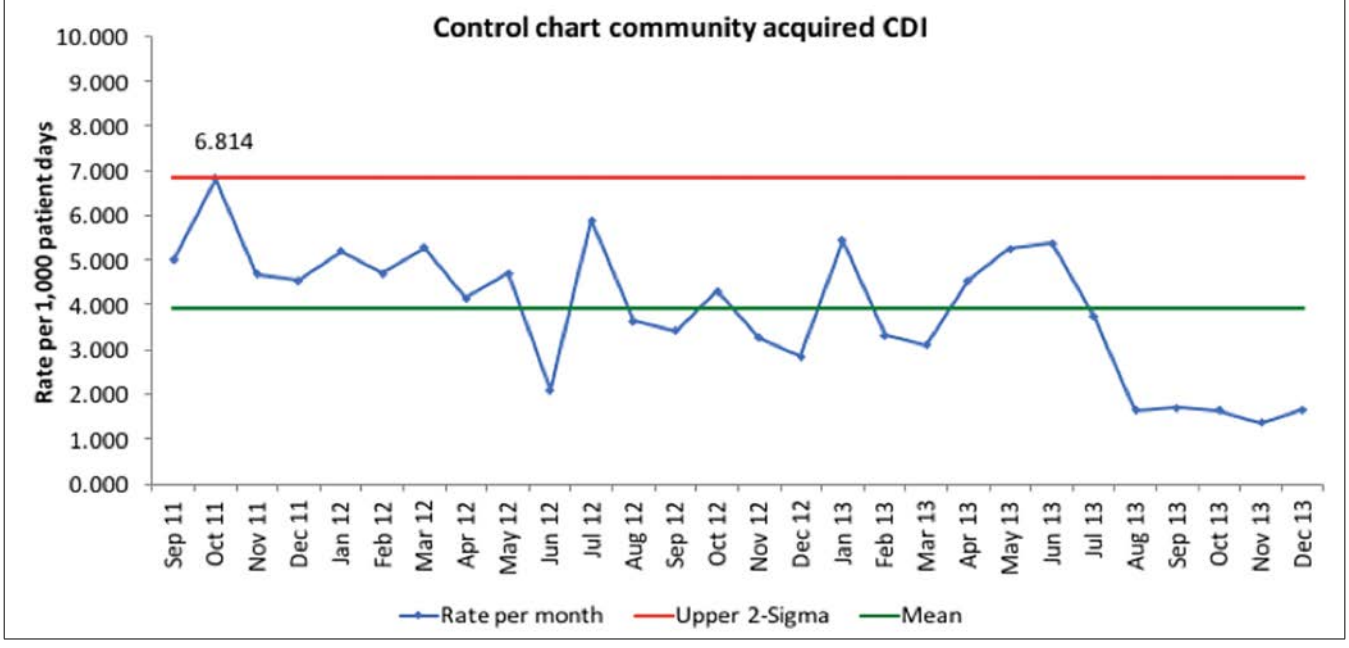
Figures 2A, 2B, and 2C explore the time series pattern of CDIs in NHS hospitals. When the control limit is set at ±2 sigma, the control chart for CA-CDIs indicates many months of higher-than-average CA-CDI rates with no out-of-control range. The control charts show an out-of-control period for HA-CDIs starting in January 2013 and lasting until April 2013. To confirm this result and to understand whether the increase in CA-CDI and HA-CDI cases co-occurred, we turned to purely temporal analysis.

FIGURE 2A



Temporal visualization of CDI cases in the Niagara Region. Comparison of time series trends of CA-CDI and HA-CDI patients hospitalized in NHS hospitals between September 2011 and December 2013.

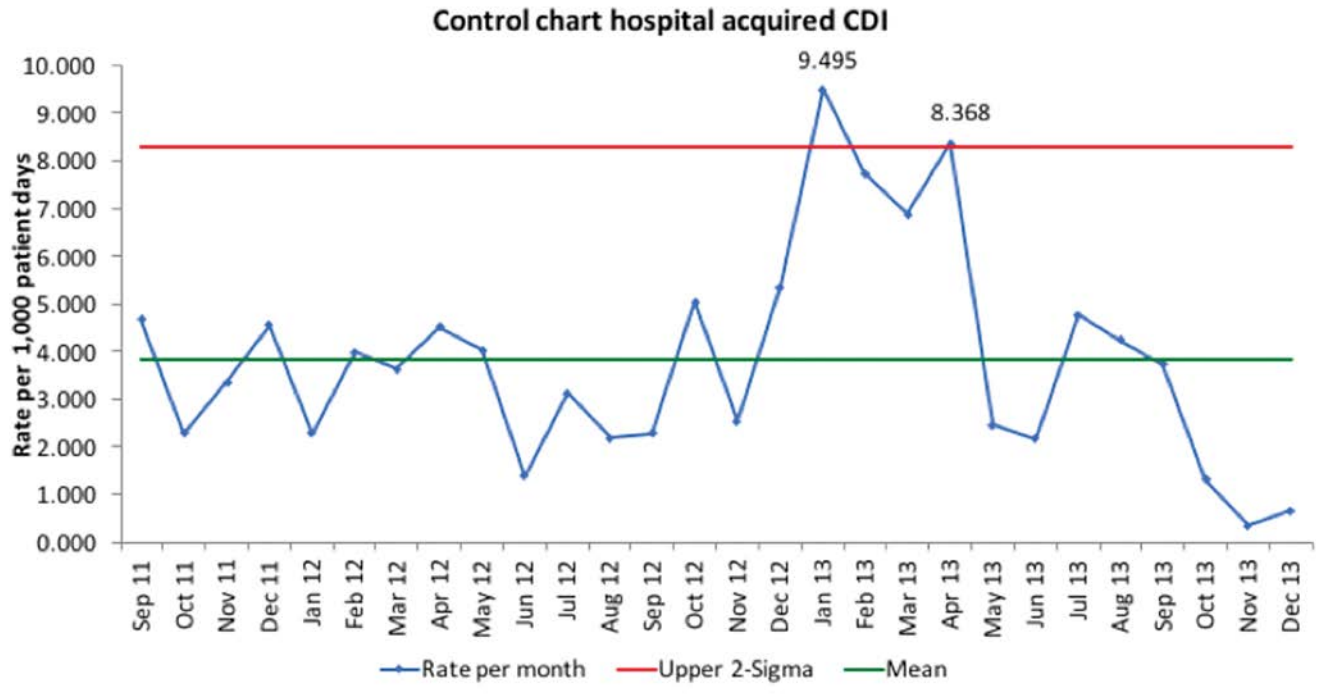
FIGURE 2B



SPC display of hospitalized CA-CDI rates per 1,000 admissions hospitalized in NHS hospitals between September 2011 and December 2013.

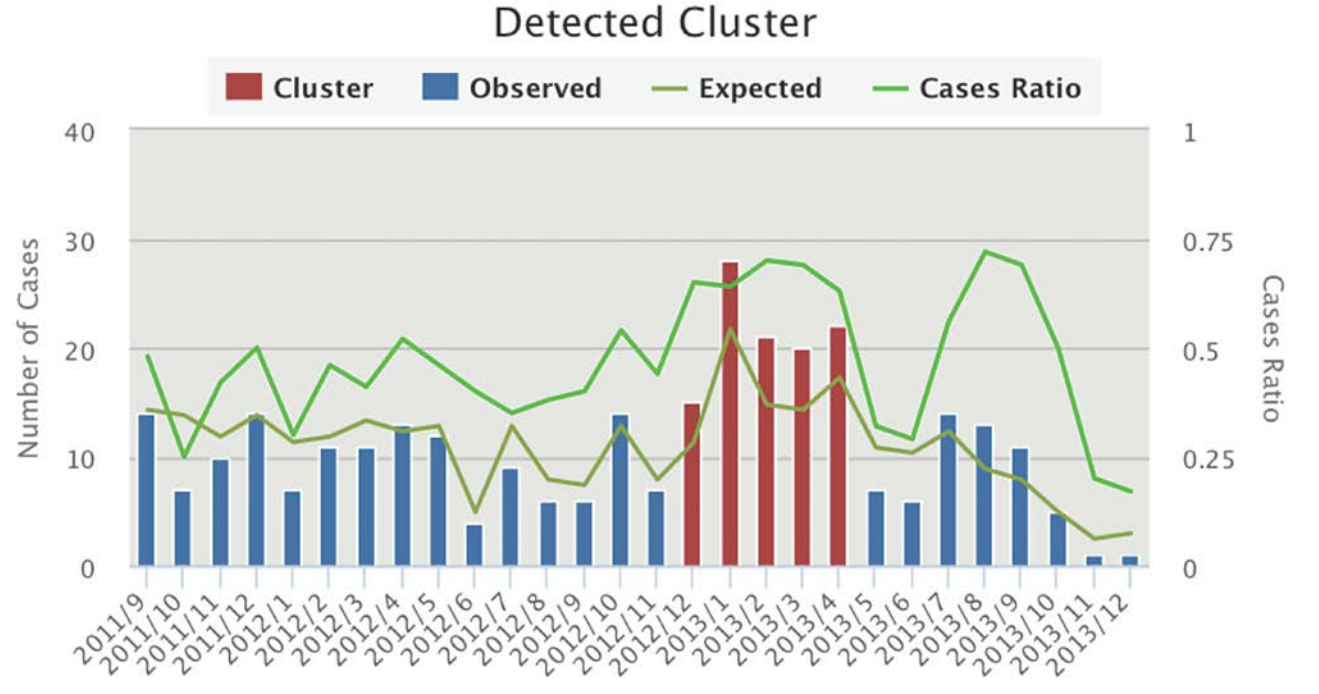
Return to TABLE OF CONTENTS

FIGURE 2C



SPC display of hospitalized HA-CDI rates per 1,000 admissions hospitalized in NHS hospitals between September 2011 and December 2013.

FIGURE 3



Purely temporal analysis scanning for clusters with high rates. A retrospective study of CDI cases in NHS hospitals between September 2011 and December 2013 using the Bernoulli model, SaTScan v9.4.4. Information on the detected temporal cluster:
Time frame: 2012/12/01 to 2013/4/30; Log likelihood ratio: 12.027272; Monte Carlo rank: 1/1,000; *P*-value: 0.001
A cluster is statistically significant when its log likelihood ratio is greater than the critical value, which is, for significance level:
Gumbel critical values: 0.00001: 13.971596 and 0.0001: 11.659499; Standard Monte Carlo critical values: 0.001: 8.060346; 0.01: 7.215835; and 0.05: 5.575158

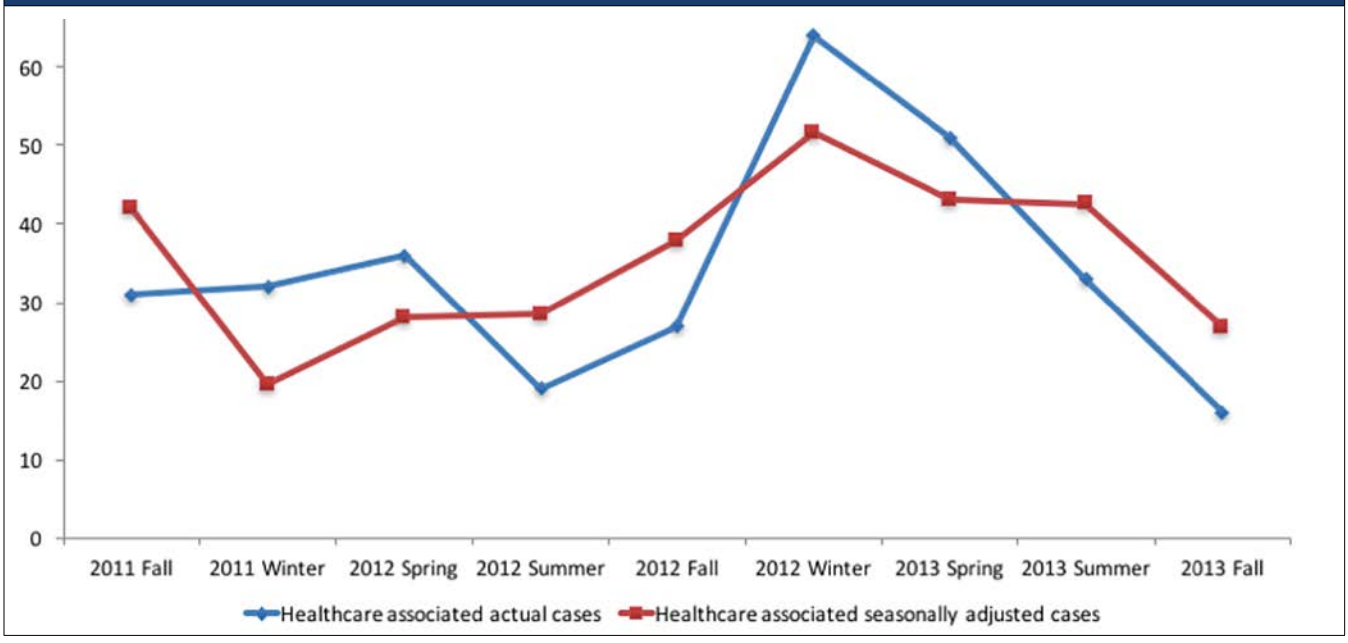
Temporal scan statistics

Figure 3 illustrates a cluster of HA-CDI that was identified between December 2012 and April 2013, following a period of high CA-CDI hospitalization. Identification of a cluster rejects our null hypothesis that the cases in hospitals and the community happened at the same time. Instead, cases acquired in the community occurred at a different time than those acquired in hospitals during the period of the cluster.

Test of seasonality

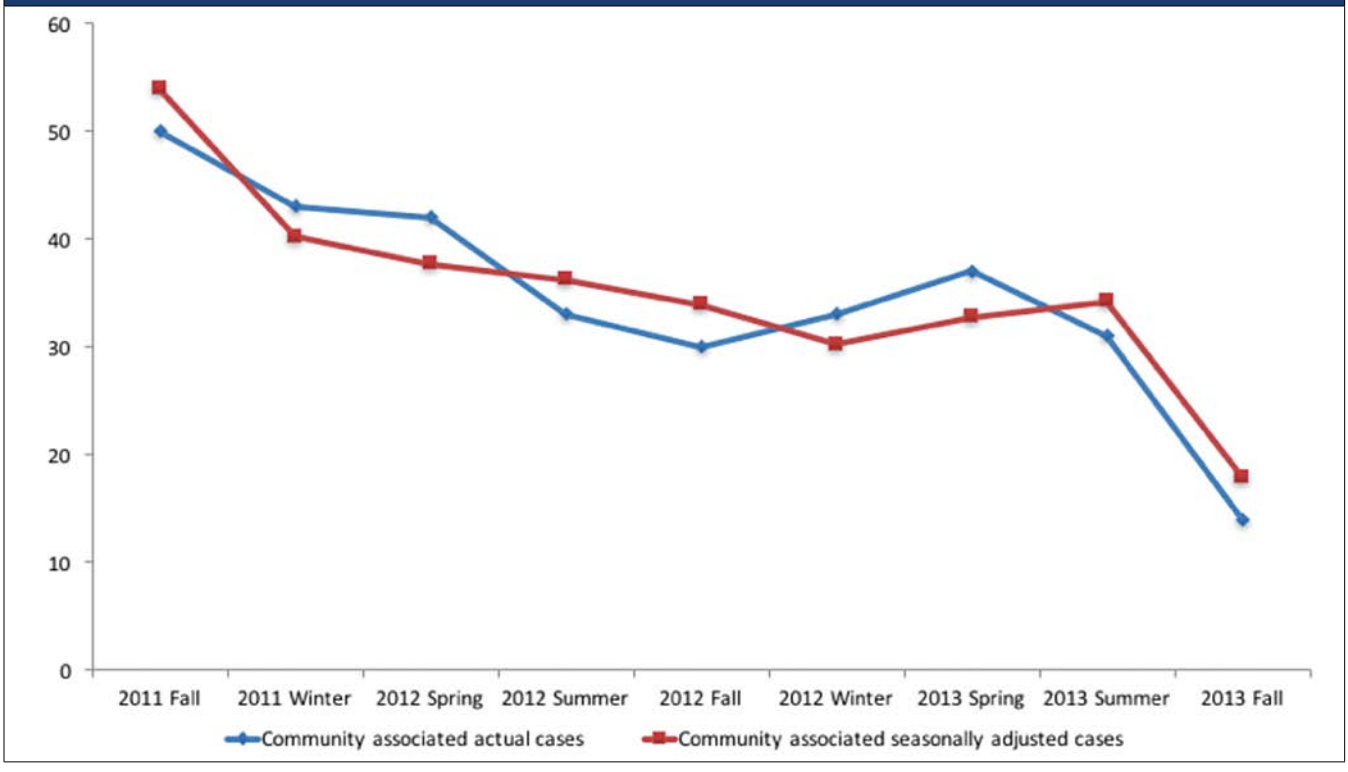
The crude grand seasonal mean for the study period for all seasons was 35 for CA-CDIs and 36 for HA-CDIs. To better

FIGURE 4A



Crude and seasonality adjusted values for CA-CDI cases in the Niagara Region between September 2011 and December 2013.

FIGURE 4B



Crude and seasonality adjusted values for HA-CDI cases in the Niagara Region between September 2011 and December 2013.

Return to TABLE OF CONTENTS

understand the effect of the seasons as an influencing factor on the prevalence of CA- and HA-CDIs, we calculated the additive seasonal indexes and numerically plotted the computed seasonal effects for all seasons in the study period. Graphical evaluation of the crude and seasonally adjusted cases for CA- and HA-CDIs indicated a lower seasonal influence in the former than the latter (see Figures 4A and 4B).

C. difficile strain typing

Overall, 6% (40/629) of the study cases were tested for molecular typing. 4% (13/318) of CA-CDI specimens were tested and PFGE identified various strains, including: 2/13 (15%) NAP1 strains; the rest (85%) comprised other unrelated strains (A, B, C, D, I, M, N). 9% (27/311) of the HA-CDIs were tested for strain identification: 12/27 cases (44%) were NAP1strain, 2/27 (7%) were non-NAP1, and the rest (48%) were other unrelated strains (A, B, D, L, M, N, O, T, V).

DISCUSSION

In our case series study, we found differences between the temporal patterns of the hospitalized CA-CDI and HA-CDI cases and a unique pattern of spatial distribution for CA-CDIs. Our study did not reveal a seasonality pattern for the CA-CDI cases and we discovered that cases of CA-CDI and HA-CDI were temporally independent. Although our study was conducted only on hospitalized patients with CA-CDIs, the overall incidence was notably higher for the Niagara Region (14.84 in 2011; 33.22 in 2012; and 25.5 in 2013) than for studies done in the UK in 2004 (22.0 per 100,000 population), Connecticut in 2006 (6.9 per 100,000 population), and Philadelphia in 2005 (7.6 per 100,000 population) [28]. Similarly, given the fact that it experienced many outbreaks during the study period, the Niagara Region's HA-CDI rates were markedly higher (3.83/1,000 patient days) than the average rates/1,000 patient days for the entire province of Ontario, which were 0.30 and 0.33 in 2011-12 and 2012-13, respectively [29].

Spatial clusters of CA-CDI in our study were indicative of substantial accumulation of community cases that were admitted to the NHS hospitals from urban zones. This is in contrast to recent studies, which suggest a positive association between environmental elements such as flooding [5], rainfall, exposure to agricultural structures (exposure to soil, livestock, or raw animal products), bathing in potentially contaminated watercourses, and an increased risk of CA-CDI [30-32]. On the other hand, the proximity to communal dwellings such as nursing homes has been recognized as a contributing factor to increased risk of CDI in the community. This may therefore support the possibility of CA-CDI cases originating from shared community residences such as assisted-living supportive housings. In Ontario, the *Long-Term Care Homes Act* (S.O.2007, c.8.) [33] and the *Retirement Homes Act* (S.O.2010, c.11) [34] specify the need for infection prevention and control training and practices in these settings, but the legislation does not pertain to other fast-growing communal dwellings, such as assisted living or supportive housing.

The seasonal associations found in other CDI studies were not evident in our NHS study. Some of the studies that established a seasonal pattern suggest that the increase of CA-CDI in winter months can be attributed to the rise in antimicrobial prescribing practices during the influenza season [7, 35, 36]. However, reports of hospitalized and community-based CDI in the southern hemisphere did not substantiate the previous claim and pointed to the increased incidence in summer months, where they assumed a role for imported fresh produce for this pattern [6]. In our study, lack of a seasonal pattern may be explained by the presence of *C. difficile* in the community through other reported sources such as retail meat, farm services, soil, pets, and domestic animals [37-42].

The purely temporal study of CA-CDI and HA-CDI cases established a hospital-associated cluster spanning from December 2012 to April 2013, where a rise in CA-CDI cases predated the HA-CDI's temporal cluster (see Figure 3). One hypothesis could be that the asymptomatic carriage of HA-CDIs after discharge from NHS hospitals in the weeks or months preceding our study period might have contributed to an increase of CA-CDI patients in the community and their return to the hospitals to receive care. Another possibility is that the admission of non-suspected CA-CDI cases due to a lack of established risk factors upon admission to hospitals might have prompted HA-CDI outbreaks. Despite the moderate homogeneity between the HA-CDI strains that could point to a nosocomial transmission (12 of 27 were NAP1), more than half of the HA-CDI outbreak strains did not show a molecular relatedness. This may be explained by the introduction of multiple unrelated strains through direct or indirect contact with the CA-CDI patients admitted to NHS hospitals.

Prospective CA-CDI surveillance, added to strain typing programs inclusive of CA-CDI and HA-CDI, can identify the transmission pathways and the unique risk factors associated with CA-CDI. Added to the traditional surveillance methods used in hospitals, community surveillance of CA-CDI can inform the discourse of this infection's unique risk factors. In addition, research informed by geographical homogeneity can provide better understanding of the causal factors attributed to the infection's community clustering.

Our study was limited to hospitalized CA-CDI cases; we had no knowledge of the CA-CDI patients who did not need hospitalization. This limited the generalizability of our findings. Moreover, because we lacked access to the full postal code information, we could not document the precise location of the CA-CDIs. This reduced our ability to pinpoint the location of potential public sources of infection in the community. Furthermore, due to the short study period, the power of the seasonality analysis was limited and the identified patterns (or lack thereof) could have been influenced by multiple outbreaks during our study period. Our strain typing assessment was limited to those tested as a result of an outbreak investigation and composed a small proportion of all CDIs. The risk of misclassification of CDI cases in this study was reduced by using a comprehensive surveillance database, which was based on a case definition, case confirmation, and expert consultation.

Epidemiological evaluation by means of administrative and quality improvement databases allowed for a large-scale data set and reduced the risk of recall bias.

The temporal independence of the CA- and HA-CDI cases, the higher-than-expected number of hospitalized CA-CDI cases, and the multiple reported HA-CDI outbreaks with a large proportion of unrelated molecular patterns all point to a possible association between the appearance of hospitalized CA-CDI cases and hospital outbreaks. Other studies have hypothesised a positive correlation between increased HA-CDI rates in hospitals and community prevalence and have suggested that hospital cases could be a driver of CDI in the community [10, 43, 44]. Some studies postulated a community reservoir as a potential attributing base for this infection into the hospitals [5, 39, 40]. Novel research programs that combine hospital and community findings can detect the direction of CA-CDI transmission. A better understanding of the epidemiology and the community drivers of CA-CDI will guide hospital and community patient safety policies, inform public health programs, and improve quality of health at a population level.

APPENDIX A: NHS' definitions of CDI, HA-CDI, and CA-CDI used between September 2011 and December 2013 for surveillance and case identification.

NHS Definition of CDI

• Diarrhea
WITH
• Laboratory confirmation of *C. difficile* (e.g., by positive toxin A/B assay or PCR);
OR
• Visualization of pseudomembranes on sigmoidoscopy or colonoscopy;
OR
• Histological/pathological diagnosis of pseudomembranous colitis.

Definition of HA-CDI	Definition of CA-CDI
An HA-CDI case is defined as a patient who has not had CDI in the past eight weeks but meets one of the following criteria: • They do not present with CDI upon admission but show onset of symptoms >72 hours after admission. • The infection was present at time of admission but was related to a previous admission to the same facility within the last four weeks.	A CA-CDI case matches the case definition for CDI and does *not* match the HA-CDI definitions. In other words: • CDI symptoms *were* present upon admission or symptom onset *was less* than 72 hours after admission. • No exposure to any healthcare facility occurred within the last four weeks, or the source of infection cannot be determined and the patient has not had HA-CDI in the last eight weeks.

REFERENCES

1. Gupta, A., & Khanna, S. (2014). Community-acquired *Clostridium difficile* infection: An increasing public health threat. *Infection and Drug Resistance, 7*, 63-72.
2. Khanna, S., & Pardi, D. S. (2010). The growing incidence and severity of *Clostridium difficile* infection in inpatient and outpatient settings. *Expert Review of Gastroenterology & Hepatology, 4*(4), 409-416.
3. Khanna, S., Pardi, D. S., Aronson, S. L., Kammer, P. P., Orenstein, R., St. Sauver, J. L., Harmsen, W. S., & Zinsmeister, A. R. (2012). The epidemiology of community-acquired *Clostridium difficile* infection: A population-based study. *American Journal of Gastroenterology, 107*(1), 89-95.
4. Hoover, D. G., & Rodriguez-Palacios, A. (2013). Transmission of *Clostridium difficile* in foods. *Infectious Disease Clinics of North America, 27*(3), 675-685.
5. Lin, C. J., Wade, T. J., & Hilborn, E. D. (2015). Flooding and *Clostridium difficile* infection: A case-crossover analysis. *International Journal of Environmental Research and Public Health, 12*(6), 6948-6964.
6. Furuya-Kanamori, L., McKenzie, S. J., Yakob, L., Clark, J., Paterson, D. L., Riley, T. V., & Clements, A. C. (2015). *Clostridium difficile* infection seasonality: Patterns across hemispheres and continents – A systematic review. *PLoS ONE, 10*(3), e0120730.
7. Brown, K. A., Daneman, N., Arora, P., Moineddin, R., & Fisman, D. N. (2013). The co-seasonality of pneumonia and influenza with *Clostridium difficile* infection in the United States, 1993-2008. *American Journal of Epidemiology, 178*(1), 118-125.
8. Burckhardt, F., Friedrich, A., Beier, D., & Eckmanns, T. (2008). *Clostridium difficile* surveillance trends, Saxony, Germany. *Emerging Infectious Diseases, 14*(4), 691-692.
9. Furuya-Kanamori, L., Stone, J. C., Clark, J., McKenzie, S. J., Yakob, L., Paterson, D. L., Riley, T. V., Doi, S. A., & Clements, A. C. (2015). Comorbidities, exposure to medications, and the risk of community-acquired *Clostridium difficile* infection: A systematic review and meta-analysis. *Infection Control & Hospital Epidemiology, 36*(2), 132-141.
10. Walker, A. S., Eyre, D. W., Wyllie, D. H., Dingle, K. E., Harding, R. M., O'Connor, L., Griffiths, D., Vaughan, A., Finney, J., Wilcox, M. H., Crook, D. W., & Peto, T. E. (2012). Characterisation of *Clostridium difficile* hospital ward-based transmission using extensive epidemiological data and molecular typing. *PLoS Med, 9*(2), e1001172.
11. Sethi, A. K., Al-Nassir, W. N., Nerandzic, M. M., Bobulsky, G. S., & Donskey, C. J. (2015). Persistence of skin contamination and environmental shedding of *Clostridium difficile* during and after treatment of *C. difficile* infection. *Infection Control & Hospital Epidemiology, 31*(1), 21-27.
12. Statistics Canada. *Focus on Geography Series, 2016 Census.* Retrieved from https://www12.statcan.gc.ca/census-recensement/2016/as-sa/fogs-spg/Index-eng.cfm
13. Provincial Infectious Diseases Advisory Committee. (2011). *Best practices for infection prevention and control programs in Ontario in all health care settings*. Retrieved from https://www.publichealthontario.ca/en/eRepository/BP_IPAC_Ontario_HCSettings_2012.pdf
14. Becton Dickinson. (2008). BD GeneOhm™ Cdiff Assay – For the rapid and accurate detection of *C. difficile* infections. 2010. Contract No.: 2.

Return to TABLE OF CONTENTS

Retrieved from https://www.bd.com/resource.aspx?IDX=17953
15. Dalpke, A. H., Hofko, M., Zorn, M., & Zimmermann, S. (2013). Evaluation of the fully automated BD MAX Cdiff and Xpert *C. difficile* assays for direct detection of *Clostridium difficile* in stool specimens. *Journal of Clinical Microbiology, 51*(6), 1906-1908.
16. Kulldorff, M., & Nagarwalla, N. (1995). Spatial disease clusters: Detecton and inference. *Statistics in Medicine, 14*(8), 799-810.
17. Heffernan, R., Mostashari, F., Das, D., Karpati, A., Kulldorff, M., & Weiss, D. (2004). Syndromic surveillance in public health practice, New York City. *Emerging Infectious Diseases, 10*(5), 858-864.
18. Kulldorff, M. (1997). A spatial scan statistic. *Communications in Statistics – Theory and Methods, 26*(6), 1481-1496.
19. Benneyan, J. (1998). Statistical quality control methods in infection control and hospital epidemiology, Part I: Introduction and basic theory. *Infection Control & Hospital Epidemiology, 19*(3), 194-214.
20. Benneyan, J. (1998). Statistical quality control methods in infection control and hospital epidemiology, Part II: Chart use, statistical properties, and research issues. *Infection Control & Hospital Epidemiology, 19*(4), 265-283.
21. Sellick, J. J. (1993). The use of statistical process control charts in hospital epidemiology. *Infection Control and Hospital Epidemiology, 14*(11), 649-656.
22. Kulldorff, M. (2005). *SaTScan software for the spatial, temporal and space-time scan statistics, 2005*. Retrieved from https://www.satscan.org/
23. Lawson, A. B., Banerjee, S. N., Haning, R. P., & Yugarte, M. D. (2016). *Handbook of spatial epidemiology*. Boca Raton, FL: Taylor & Francis Group.
24. Elliott, P., & Wartenberg, D. (2004). Spatial epidemiology: Current approaches and future challenges. *Environmental Health Perspectives, 112*(9), 998-1006. doi: 10.1289/ehp.6735
25. Carlberg, C. (2013). *Predictive analytics: Microsoft Excel*. Indianapolis, IN: QUE.
26. Kulldorff, M., & Information Management Services. (2009). SaTScan™ v8.0: Software for the spatial and space-time scan statistics. Retrieved from http://www.satscan.org
27. Schneider, A. (2018). GPS Visualizer: Do-it-yourself mapping. Retrieved from http://www.gpsvisualizer.com/
28. Centers for Disease Control and Prevention. (2008). Surveillance for community-associated *Clostridium difficile* – Connecticut, 2006. *Morbidity and Mortality Weekly Report, 57*, 340-343.
29. Health Quality Ontario. (2014). *Measuring up: A yearly report on how Ontario's health system is performing*. Retrieved from http://www.hqontario.ca/portals/0/Documents/pr/measuring-up-yearly-report-en.pdf
30. Furuya-Kanamori, L., Robson, J., Soares Magalhães, R. J., Yakob, L., McKenzie, S. J., Paterson, D. L., Riley, T. V., & Clements, A. C. (2014). A population-based spatio-temporal analysis of *Clostridium difficile* infection in Queensland, Australia over a 10-year period. *Journal of Infection, 69*(5), 447-455.
31. Kistemann, T., Zimmer, S., Vagsholm, I., & Andersson, Y. (2004). GIS-supported investigation of human EHEC and cattle VTEC O157 infections in Sweden: Geographical distribution, spatial variation and possible risk factors. *Epidemiology & Infection, 132*, 495-505.
32. Anderson, D. J., Rojas, L. F., Watson, S., Knelson, L. P., Pruitt, S., Lewis, S. S., Moehring, R. W, Sickbert Bennett, E. E., Weber, D. J., Chen, L. F., & Sexton, D. J. (2017). Identification of novel risk factors for community-acquired *Clostridium difficile* infection using spatial statistics and geographic information system analyses. *PLoS ONE, 12*(5), e0176285.
33. *Long-Term Care Homes Act*. S.O.2007, c.8. Retrieved from https://www.ontario.ca/laws/statute/07l08
34. *Retirement Homes Act*. S.O.2010, c.11. Retrieved from https://www.ontario.ca/laws/statute/10r11
35. Gilca, R., Fortin, E., Frenette, C., Longtin, Y., & Gourdeau, M. (2012). Seasonal variations in *Clostridium difficile* infections are associated with influenza and respiratory syncytial virus activity independently of antibiotic prescriptions: A time series analysis in Québec, Canada. *Antimicrobial Agents & Chemotherapy, 56*(2), 639-646.
36. Kuntz, J. L., Chrischilles, E. A., Pendergast, J. F., Herwaldt, L. A., & Polgreen, P. M. (2011). Incidence of and risk factors for community-associated *Clostridium difficile* infection: A nested case-control study. *BMC Infectious Diseases, 11*, 194.
37. Borriello, S. P., Honour, P., Turner, T., & Barclay. F. (1983). Household pets as a potential reservoir for *Clostridium difficile* infection. *Journal of Clinical Pathology, 36*(1), 84-87.
38. Baumgardner, D. J. (2012). Soil-related bacterial and fungal infections. *Journal of the American Board of Family Medicine, 25*(5), 734-744.
39. Rodriguez-Palacios, A., Staempfli, H. R., Duffield, T., & Weese, J. S. (2007). *Clostridium difficile* in retail ground meat, Canada. *Emerging Infectious Diseases, 13*(3), 485-487.
40. Weese, J. S., Avery, B. P., Rousseau, J., & Reid-Smith, R. J. (2009). Detection and enumeration of *Clostridium difficile* spores in retail beef and pork. *Applied and Environmental Microbiology, 75*(15), 5009-5011.
41. Stone, N. E., Sidak-Loftis, L. C., Sahl, J. W., Vazquez, A. J., Wiggins, K. B., Gillece, J. D., Hicks, N. D., Busch, J. D., Keim, P., & Wagner, D. M. (2016). More than 50% of *Clostridium difficile* isolates from pet dogs in Flagstaff, USA, carry toxigenic genotypes. *PLoS ONE, 11*(10), e0164504.
42. Anderson, D. J., Podgorny, K., Berríos-Torres, S. I., Bratzler, D. W., Dellinger, E. P., Greene, L., Nyquist, A., Saiman, L., Yokoe, D. S., Maragakis, L. L., & Kaye, K. S. (2014). Strategies to prevent surgical site infections in acute care hospitals: 2014 update. *Infection Control and Hospital Epidemiology, 35*(6), 5-27.
43. Furuya-Kanamori, L., Yakob, L., Riley, T. V., Paterson, D. L., Baker, P., McKenzie, S. J, Robson, J., & Clements, A. C. A. (2016). Community-acquired *Clostridium difficile* infection, Queensland, Australia. *Emerging Infectious Diseases, 22*(9), 1659-1661.
44. Yakob, L., Riley, T. V., Paterson, D. L., Marquess, J., & Clements, A. C. (2014). Assessing control bundles for *Clostridium difficile*: A review and mathematical model. *Emerging Microbes & Infections, 3*(6), e43.