The Contribution of Reinfections to Chlamydia Resurgence, Sexual Networks, and Spatial Clustering in Brant County, Ontario

Jenny Pereira Santos

A thesis submitted to the

Faculty of Graduate and Postdoctoral Studies

in partial fulfillment of the requirements for the

Degree of Master of Science: Epidemiology

School of Epidemiology, Public Health And Preventative Medicine

Faculty of Medicine

University of Ottawa

UNIVERSITY OF OTTAWA

FACULTY OF GRADUATE STUDIES

I declare sole authorship of the following materials with the co-authorship of Dr. Ann Jolly. I conducted the literature review and statistical analyses and interpreted the findings presented within, with review by Dr. Jolly. I compiled and wrote all of the following series of manuscripts in their entirety.

Acknowledgements

I would like to acknowledge the contributions by Rene Duplain and Sarah Simpkin from the University of Ottawa Library, geography services. They assisted me with the geographic identifying information and population data to allow for my sexual network and spatial analysis.

Abstract

Recent findings by Public Health Ontario (PHO) state that there were approximately 36, 346 confirmed cases of chlamydia in Ontario as of 2011. This represents an incidence rate that increased by 54% since 2006 rising from 177 to 272 per 100,000 in 2011. National rates only increased by 38% (210 to 290 per 100,000), meaning that Brant County rates surpassed both, increasing by over 100% (150 to 395 per 100,000).

The main objective of this series of manuscripts is to develop a clear profile of re-infected individuals in comparison to non-repeaters, while considering co-infections wit gonorrhea. The secondary objective was to determine the sexual network as well as spatial distribution patterns of cases in Brant County.

The study period is from January 1st, 2006 until December 31st, 2015, Data were extracted from the integrated public health information system (iPHIS). Basic descriptive statistics will be performed followed by a Cox-regression analysis in order to compare individuals who are repeaters with those who are not repeaters. Within the study period, there were 2,829 cases of chlamydia and 328 were reinfections. We identified twelve hotspots with high chlamydia infection rates of which, 58 per cent occurred within the previously identified core group, in the urban core of Brant County.

Table of Contents

Introduction

Brant County is located in central Ontario and had a population of approximately 125,099 in 2006, increasing to 129,288 in 2011. In 2006, the incidence rate for chlamydia at the BCHU was 150 per 100,000, which rose steadily to reach approximately 395 per 100,000 by 2011. Overall, this is an increase of over 100% in incidence rate.[1] Within that same time period, the incidence rate in Ontario only rose from 177 per 100,000 to 272 per 100,000, representing a 54% increase.[2] Across Canada, the incidence rate increased by approximately 38% between 2006 and 2011.[2] The incidence rate in Brant County is significantly higher than what we see provincially and nationally. This makes for a cause of concern in the region.

Within the Brant County Health Unit (BCHU), the majority of chlamydia infections are reported within the 20-24 year old age group for both males and females, with the next age group being the 15-19 year old females and the 25-29 year old males.[1] These age groups are also the age groups that account for the most cases of chlamydia across Ontario, with 65% of the total cases for 2011 being within the 20-24 year old age group (23,482 of 36,346). Although both men and women are just as likely to contract the infection, incidence rates in females are higher than those in males; due in part to a larger portion of females being tested, and males have a higher positivity rate than females.[3-6]

Aside from young age and sex, there are also various other factors that increase the risk of chlamydia infection which include: poor contraceptive use, working in the sex trade, having multiple sex partners in a short period of time or a recent new partner, as well as being involved in drug injection activities.[7] Additionally, it has been understood that individuals who have become infected with their first chlamydia infection and receive treatment (or their body

spontaneously clears the infection), are also at high risk for becoming reinfected within the first six months to two years post-treatment.[5]

As the mode of sexually transmitted infection (STI) transmission is often not homogeneous, reinfections and co-infections with gonorrhea are increasingly important, as they are among the main approaches (in addition to mapping high prevalence areas often referred to "core groups" using spatial methods) to identify high-risk groups for chlamydia infection within a given population.[5,8] As a high prevalence (core group) population was recently pinpointed within the downtown core of Brantford, it is of particular interest to evaluate the extent of reinfections with said core group and across Brant County. Doing so may shed light on the potential contribution reinfections made to the formation of a core group and even the specific characteristics of the individuals within it.

STI interventions and control programs currently advocate for higher screening rates, as they are extremely low, especially among individuals who present with the highest risk for infection and re-infection. However, there is evidence supporting the fact that improved partner notification (contact tracing) methods can contribute just as much to reducing chlamydia rates as increased screening rates.[9,10]

These data highlight the severity of the issue regarding chlamydia in Brant County and Canada. The present series of articles aims to expand on Canadian STI research as well as assisting with the implementation of intervention programs by examining the questions:

- What is the contribution of reinfections to chlamydia resurgence in Brant County, Ontario?
- Are there any notable sexual networks within Brant County?

- Are there significant clusters of infection in Brant County that account for the most chlamydia cases?
- Does this coincide with previous research on core groups in Brant?

The Ottawa Health Science Network Research Ethics Board granted ethics approval for the present study.

Chlamydia Resurgence and Reinfection in Canada: A Review of the Literature

Jenny P. Santos, BPH, MSc; Ann Jolly, PhD

Introduction to Chlamydia Trachomatis

Chlamydia is a sexually transmitted infection (STI) caused by the bacterium *Chlamydia trachomatis*.[6] Although both men and women are just as likely to contract the infection, incidence rates in females are higher than those in males − due in part to a larger portion of females being tested. However, males present with higher positivity rates than females; supporting the theory that males are the "hidden reservoir" for infection and are under diagnosed.[11] In contrast, the rate of recovery or duration of infection is not necessarily the same between the two sexes, as detecting the duration depends on whether the infection is symptomatic or not and requires researchers to withhold treatment, which is unethical. Males and females have different rates of asymptomatic infections, 50% and 70% respectively.[12]

In accordance with the Public Health Standards, the protocol in Ontario for chlamydia control, as well as other STIs, focuses mainly on the screening, re-testing of cases, and partner notification/contact tracing.[13] When cases are identified, it is required that they be immediately reported to the Ministry of Health so proper registration of the case can be ensured. Following this, cases need to be treated appropriately, followed up and re-tested after 6 months.[14] Partner notification procedures aim to provide examination and treatment to partners of index cases to identify new cases and to reduce the likelihood of reinfection of the original index case.[15] In addition to this secondary prevention method, screening of high-risk persons is the most common

primary prevention meaning preventing the individual from ever getting infected − both of which are necessary for effective management and control of chlamydia. [15, 16]

It is important to keep in mind that screening was initially only recommended for women, a decision not based on scientific data, as randomized clinical trials and cohort studies had not yet been carried out to look at the relationship between screening and chlamydia rates. [14] Carrying out such research in following years led to the observation that men were in fact a hidden reservoir for infection and therefore, were later included in recommendations.[11] Therefore, in order to develop evidence-based, improved and cost-effective screening and treatment recommendations, the evaluation of the costs and benefits of screening and therapy is necessary. This is recognized as the second highest priority in the national STI research priorities, with the first being antimicrobial resistance, which is another critical issues with treatment of chlamydia. [14, 17]

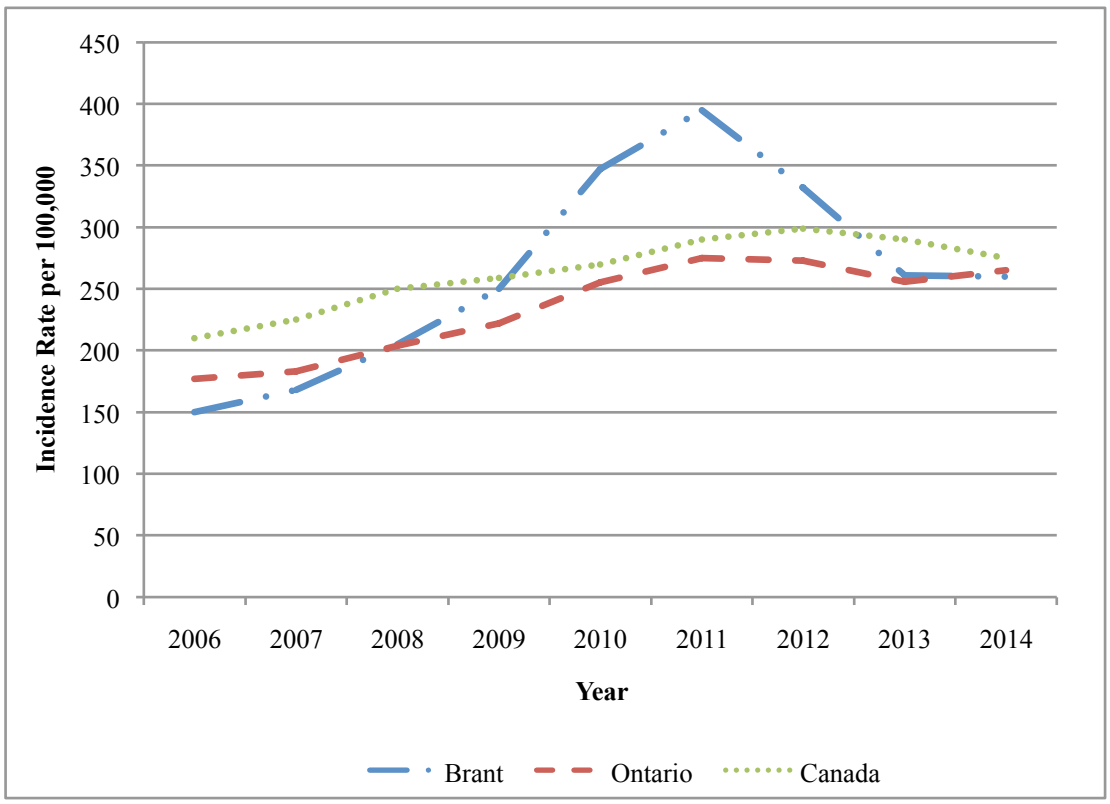
Rise of Chlamydia in Canada

The burden of chlamydia has become greater as incidence rates have steadily risen across Canada since 1997.[6] Chlamydia is the most commonly reported bacterial STI in Canada and its asymptomatic nature is especially troublesome when considering reinfection rates.[6] This increase is also an issue in many other countries across the globe, such as the United States of America (USA), Australia, and across Europe. [7] The focus of the present study will be Brant County, as they are currently experiencing some of the highest rates of chlamydia in Ontario aside from Toronto and some of the Northern Regions such as Porcupine, Northwestern and Thunder Bay. [2]

In Brant County specifically, patterns of infection have followed the global trend, with the most striking increases in incidence rate occurring from 2006 onward. In 2006, the incidence

rate for chlamydia in Brant County was 152 per 100,000, which rose to approximately 395 per 100,000 by 2011. These rates represent about a 160% increase in the overall incidence rate. Within that same time period, the incidence rate in Ontario also rose from 177 per 100,000 to 272 per 100,000, representing approximately a 54% increase in incidence rate.[2] The incidence rate in Brant County surpassed both Canadian and Ontario rates, as national rates only increased by approximately 38% within the same period, with 65,000 cases reported in 2006 and 94,690 cases in 2009 (*Figure 1*).[2] Even with the decrease after 2011, to approximately 260 cases per 100,000, there was little change between 2012 and 2014. This is concerning for Brant County as there was a drop in national rates during that time frame (299 per 100,000 to 275 per 100,000).

Figure 1: Chlamydia incidence rate in Brant County, Ontario, and Canada, 2006-2014



Possible Explanations Behind the Rise

There are speculations as to what may be facilitating the steady rise of chlamydia infections. However, there is no sole cause, as there are various factors that contribute both directly and indirectly. For example, the rise in chlamydia cases is partially due to increased rates of screening as there are more individuals getting tested, leading to more positive identifications.[6] The change in national screening recommendations highlighting men as hidden reservoirs also suggests the importance of contact tracing, increasing its' acceptance over time. [14] This relationship with contact tracing became more evident when incidence rates continued to rise even after infected women were getting screened and treated because they would become reinfected. [18, 19] Additionally, emerging evidence suggests that due to males being less likely to seek screening, they exhibit higher positivity rates (and chronic asymptomatic infections) in comparison to females, especially in general practice, increasing the need for males to also be screened. [11]

Another explanation for the rise in chlamydia cases involves the changes in the laboratory technologies themselves as well as our diagnostic capabilities. In terms of laboratory technologies, methods are more sensitive due to the introduction of polymerase chain reaction (PCR) technologies in 1997 (which is a form of the common nucleic acid amplification tests or NAATs) in order to identify *C. trachomatis*. [7] Diagnostic capabilities have also changed, as modern testing allows for swab samples during routine check-ups to be just as effective as wells as more acceptable as the use of urine samples. [7, 20] Aside from these changes in screening recommendations and laboratory technologies, early treatment is often associated with the interference of immunity to infection at the population-level (as natural clearance of an infection can take months to occur), increasing the likelihood for infection and reinfection. [21, 22]

This theory is referred to as the "arrested immunity hypothesis" and has been observed in various countries.[22] This hypothesis becomes evident when looking at incidence and sequelae rates in Canada before and after initial control programs were put in place during the early 1990s. Although at the time, both incidence and sequelae rates did decline significantly, the incidence rate began to soar again by the late 1990s and early 2000s. This implies a sharp increase in the prevalence of chlamydia (via reinfection rates), while sequelae rates continued to decrease, due to the evolutionary response of *C. trachomatis* to avoid detection by the screening tests.[21, 22] These facts raise some questions as to what types of control programs, such as ones aimed at decreasing incidence or ones aimed at decreasing prevalence, would be the most effective and whom they should be aimed towards to reach their full potential. From the evidence presented here, it seems that programs that aim at decreasing the true prevalence, such as focusing on preventing reinfections, would best decrease overall chlamydia infections. This is a theory that would be worth evaluating.

Thirdly, people may be more comfortable with the human immunodeficiency virus (HIV) risk as treatments have ensured that HIV is no longer associated with certain death (or immediate health difficulties). This in turn may have resulted in increased risky sexual behaviors favourable to chlamydia, especially in the high-risk young age groups.[5, 14] The relationship with HIV is bidirectional, as those with chlamydia are at approximately 3- to 6- fold higher risk of a concurrent infection with HIV.[7, 20, 23] Therefore, the relationship with HIV increases the impact of the various long-term health complications associated with a *C. trachomatis* infection including: pelvic inflammatory disease (PID), ectopic pregnancy, and infertility in women as well as urethritis, epididymitis and orchitis in men.[15, 7]

Finally, as previously stated, contact tracing is currently a main component of chlamydia management activities as there is increasing evidence associating contact tracing, especially of the current or most recent partner, with reduced reinfection and prevalence rates of chlamydia. [10,15, 24-26] Although these procedures are effective, it is estimated that they are not reaching all partners due to embarrassment in confronting partners, anonymous partners, not knowing the contact information for the partner, or the health professional either not being able to contact the individual or the patients' /contacts' refusal to be tested or give information. [15, 27] Therefore, poor uptake of contact tracing procedures may also account for a portion of the increase in chlamydia cases.

Role of Reinfections

As expected when rates of chlamydia began increasing, rates of reinfection followed − reaching 6-10% across Canada. [5, 19, 21] Reinfection rates are of particular interest as they are especially a concern for those who have the highest risk for infection − the young female population (aged 20-24 years). The young female population can account for as much as 30% of the overall reinfection rate in the given population. [19] Additionally, reinfection is also commonly associated with co-infection with other STIs, especially gonorrhea, as well as increased likelihood of sequelae. [5, 16, 19] Therefore, the cost and risk associated with reinfections is higher than those with single infections, and will increase as the number of positive tests increases.

Again, due to the asymptomatic nature of the infection, screening uptake can be exceptionally poor, contributing to higher rates of reinfection. [28] The latter is problematic as low screening rates increase the risk for transmission within the given population as well the long-term complications referred to above. [16] Men often have lower screening rates compared to

women, as they are less likely than women to be screened or tested regardless of whether their partners have recently been screened and treated.[12] This is where contact tracing plays a key role such that if it is not carried out to completion (named individuals are contacted, tested, treated, and followed-up), then treated women who return to previous relationships are likely to become reinfected via their male partners; who are unlikely to seek routine screening. [18] Additionally, identifying and locating sex partners of cases makes it possible to get an accurate illustration of the sexual networks at play in a given population–necessary to truly understand the pattern of reinfections and transmission. [29, 30] Therefore, without proper emphasis on complete contact tracing, public health professionals are unable to secure successful chlamydia control in the given population.

In an attempt to address rising reinfection rates, amendments on the national STI recommendations were made.[5] As of 2008, the Ontario Ministry of Health and Long-Term Care included the recommendation for the re-testing of individuals who received care for chlamydia and testing of males .[13] This reduces reinfection rates in women and it may indicate male partners that would benefit from screening and treatment. [31] In addition to contact tracing, another reason for this update was compelling evidence regarding the natural history of untreated chlamydia, which supported treatment being done promptly in order for patients to be re-screened and to improve contact tracing–decreasing the rate of transmission. [32]

Yorke, Hethcote, & Nold (1978) first defined core groups of STI transmission–comprised of individuals at a higher risk of infection than the individuals in surrounding areas. Core groups highlight high-risk individuals who are likely directly and indirectly responsible for cases outside their core group, causing the infection to remain endemic. [33] Therefore, members of core groups are often also those who are at high-risk for reinfection and forming larger networks of

transmission.[29] Constructing sexual networks when studying STIs is important, as individuals within different networks often differ in terms of sociological and behavioural factors. [29, 34] Although substantial evidence exists supporting the role played of sociological factors (such as the social determinants of health) in other aspects of sexual health, evidence is less consistent in regards to chlamydia, especially in Canadian context. [35] This network perspective not only yields important disease-control implications but also illustrates that the pattern of partners within core groups and complex networks has implications for transmission beyond what the number of sexual contacts or total links tells us. [36, 37]

Direction of Future Research

Researching chlamydia reinfection epidemiology will provide a better indication of those at the highest risk for infection and reinfection–aiding in the development of more effective prevention strategies. Reinfections contribute significantly to a higher incidence and prevalence of chlamydia as well as a higher risk for reproductive complications, especially in women. [7] Brant County would particularly benefit from determining the role of reinfections, as both incidence and prevalence rates are high. It is imperative to determine whether these new cases develop as a result of treatment failure or are actual reinfection cases as a result of incomplete contact tracing. Whether cases result from one or the other affects the management recommendations. [38]

Given the situation in Brant County, the relationship between contact tracing and reinfection rates to construct sexual networks would be beneficial to explore. There is no research on contact tracing in this population, as gathering the appropriate information on partners of cases is only beginning to become more of a priority for health care workers. This is

not an issue specific to Brant County; most populations across Canada and globally struggle with contact tracing utilization.[15] There is evidence supporting the ability to increase contact tracing uptake via resources including the utilization of the Internet or other software, which allows for professional use as well as chlamydia case and contact use.[39] Further studying these resources will provide the basis to implement them in areas, such as Brant County, where contact tracing is especially poor and/or where infection rates are high. This is attractive, as methods that strengthen contact tracing procedures have been found to contribute to as much of chlamydia control as methods used to strengthen screening procedures. [9, 26]

Once an understanding of the current contact-tracing situation in Brant County is developed, it will allow for an investigation of whether adding a mandatory "partner delivered therapy" (or expedited partner therapy [EPT]) component is effective in this population. Currently, EPT is not widely used in this population, as most sexual health services are delivered in a nurse led clinic, making it difficult to contact named partners of cases to ensure no allergy to treatment. There is research supporting the fact that patient delivered therapy as the main component of partner notification procedures leads to better health outcomes for chlamydia, therefore, this may warrant further investigation.[24, 40, 41] However, this is something that would only be beneficial if reinfection cases are determined to be true reinfections rather than treatment failure.

Although the present study will be focusing on the contribution of reinfections to the resurgence of chlamydia and the sexual networks involved, it is important to keep in mind that a number of public health approaches to control chlamydia need to be employed simultaneously in order to be effective in reducing the burden of disease. Shaw et. al. (2011) states that surveillance activities, in all avenues (primary prevention, diagnosis/management activities, provision of

opportunistic testing to at risk populations and provision of systemic population screening) are necessary to measure the effectiveness of chlamydia management policies and programs. If any one of these avenues are lacking, long-term control of the infection may not be possible. Potentially increasing the availability of walk-in clinics may decrease the delay between testing and treatment (and ultimately, infection and reinfection), as their provision has declined as of 1990. [42] This is particularly relevant in the case of Brant County, as there are preparations currently underway to implement walk-in services in the area of the county that was identified as being at the highest risk (the core group) of infection in downtown Brantford.

The present study aims at being one of the first to look at the relationship between screening factors, risks and incidence of reinfection, sexual networks constructed from case and contact data, and the rise in chlamydial infections. As research in STIs, especially in this area, is lacking in Canada, there is an immediate need to strengthen the research base in order to get appropriate interventions in place.

The Contribution of Reinfections to Chlamydia Resurgence in Brant County, Ontario

Jenny P. Santos, BPH, MSc; Ann Jolly, PhD

Abstract

Introduction: Chlamydial reinfection (CR) not only places additional demand on the health system but also increases the risk of long-term health complications. As of 2006, the incidence rate for chlamydia in Brant County was 152 per 100,000, which rose to a striking 395 per 100,000 by 2011. This study aims to explore the differences in demographic indicators, risk factors, reasons for testing, treatment, and contact tracing indicators between individuals who have experienced a CR and individuals who have experienced only a single infection.

Methods: A retrospective cohort was developed using notifiable disease data, extracted from the integrated public health system (iPHIS), on laboratory confirmed chlamydia cases in Brant County between January 1st 2006 and December 31st 2015. Public health nurses at the Brant County Health Unit collect data with a standardized case management sheet. During the study period, 3,499 chlamydia cases and 475 gonorrhea cases were diagnosed. The total number of individuals with chlamydia in that period was 3,060, including 157 co-infections with gonorrhea. Differences between those with reinfection and those with single infection were evaluated using univariate and multivariate (Cox proportional hazards model) methods.

Results: Four hundred and ninety-nine (16.30%) individuals experienced CR 28 days from initial infection, of which 328 (65.73%) occurred within 2 years and 211 (42.28%) within 1 year. The median time to CR was 276 days, consistent with existing Canadian CR literature. Independent risk factors for CR included being female, 25 years old or younger, and not receiving

recommended treatment for initial and/or subsequent infection.

Discussion: These findings suggest that inadequate treatment plays a significant role in CR, while accounting for young age and female gender, both of which have been previously noted.

Introduction

Chlamydia is a sexually transmitted infection (STI) caused by the bacterium *Chlamydia trachomatis*.[6] Chlamydia−as other STIs−negatively impacts the Canadian health system due to increased health care costs and loss of productivity. As initial infection does not guarantee immunity, reinfection is a common occurrence, especially in high-risk populations. Chlamydial reinfection (CR) increases the risk for the sequelae–pelvic inflammatory disease (PID), ectopic pregnancy, and infertility in women and urethritis, epididymitis and orchitis in men.[7,15] CR may occur as a result of failure to perform partner notification (contact tracing), inappropriate or absence of treatment, and continuation of risky sexual behaviors.[10,15,24-26] Additionally, Jolly et. al. (2005) showed that people who become reinfected, and especially those with gonococcal-chlamydial coinfection are likely to be core group members, with higher partner change rates. These groups are also more likely to be less wealthy, younger, and come from ethnic minorities.

There is little information available in the Canadian context regarding bacterial STIs, especially CR, despite significant research from Brunham (2005) and from other countries.[3,5,10,43] Therefore, not much is understood about how CR contributes to overall rates of chlamydia in a Canadian population, which inhibits advances in the management and control of both infection and reinfection. As research highlights the link between CR, partner notification, and positivity/prevalence rates, it is of interest to determine to what extent CR contributes to the resurgence of chlamydia in Canada.[9,21]

Brant County has one of the highest chlamydia rates across Ontario. Previous research showed that people with chlamydia infections were concentrated in "core" areas in the centre of the city, marked by poverty, teen pregnancy and single parent families.[1] We hypothesize that individuals who experience a CR within two years after initial infection differ from individuals who only have had one infection in terms of demography, risk factors, reasons for testing, treatment, and contact tracing indicators. Our analysis will reveal characteristics that may contribute to the high incidence of chlamydia in Brant County, including those factors associated with core group membership.

Methods

Setting

The population of Brant County was approximately 125,099, in 2006 and increased by 3.8% in 2011 to 129,288. Data were extracted from the integrated public health information system (iPHIS), for the period between January 1st, 2006 and December 31st, 2015. iPHIS is used by public health units to report cases of notifiable diseases−such as chlamydia−to Public Health Ontario, in accordance to the Health Protection and Promotion Act.[44] Information available for each case includes age at time of infection, residential address, as well as more specific clinical information such as: diagnosis date, record of past infections, treatment information, and contact tracing indicators.

Study Population

We developed a retrospective cohort including individuals who had a laboratory confirmed *C. trachomatis* infection within the study period. These individuals were followed until they experienced a CR or until the end of the study period. Data on individuals with a

laboratory confirmed *Neisseria gonorrhoeae* infection were also included in order to study the extent of co-infection. Individuals with any other STI diagnoses were excluded. Clients using the sexual health services at the Brant County Health Unit (BCHU) volunteered all the information used in the present analyses. During the study period, 3,499 chlamydia cases and 475 gonorrhea cases were diagnosed. After merging client information of separate records representing only cases of gonorrhea to remove any duplicate observations, the total sample size was 3,060 (including 167 cases of co-infection). Details on this process can be found in *Appendix A*. Although there was contact data linked to several cases, these were removed as we only required an indicator of whether or not partner notification was carried-out. Statistical analyses involving all available information on contacts of the cases described here can be found elsewhere.[45]

Measures

A CR was defined as a subsequent infection 28-730 (2yrs) days after the initial infection and only the first CR was included in analysis. Individuals without a second infection, who did not return with a positive diagnosis, or moved out of Brant, were censored at the end of the study period. Treatment information was available from mandatory data fields in iPHIS. Approximately 93% of chlamydia cases in Brant County were treated according to national recommendations (azithromycin: 1gm or doxycycline: 100mg), with only 6% of data on drug, dose and administration date missing. As there were various treatments listed for different clients, the treatment variable was created by grouping recommended treatment into "recommended", any other treatment into "other" and any blank spaces into "missing". In addition to treatment, other baseline characteristics investigated in the final model as predictors of CR included age, sex, evidence of partner notification, behavioural risk factors, reasons for testing (including routine screening and symptoms) and postal code (used to determine whether

cases had a postal code within the previously identified core group of chlamydia in downtown Brantford). The age of cases was analyzed as both a continuous variable and a categorical variable using meaningful cut-off points. The risk factors were grouped into three categories: high risk-behavioural (including anonymous sex, more than one sex contact in the last 6 months, new contact in past 2 months, no condom used, alcohol and injection/inhalation drug use, sex trade worker and homeless), high risk-medical (including repeat STI, pregnant or HIV), low risk and other factors (including various factors such as bath house use, travel, having been in a correctional facility and condom breakage).

Analysis

Records containing missing data were compared with those that were complete to investigate if data was missing at random, as most information for risk factors and reasons for testing were missing (n=3,060). The Mantel-Haenszel X^2 test was used for bivariate analysis to determine the relationship between various variables under consideration. The Cox proportional-hazards model was used to determine the influence of multiple covariates on time to CR, defined as the number of days between the first and second infections. The proportional-hazards assumptions, that survival curves have hazard functions that are constant over time and the censoring of an individual is not related to the probability of an event occurring, were not violated. Details on this process can be found in *Appendix B*. All analyses were conducted using SAS version 9.3 (SAS Institute Inc., Cary, NC).

Results

Descriptive Statistics

In Brant County, 7,654 index cases of chlamydia and gonorrhea were identified between January 1st 2006 and December 31st 2015. This number decreased to 3,499 (3000 unique individuals) after removing cases with gonorrhea and superfluous data. When gonorrhea and chlamydia cases were joined to determine the extent of co-infection, the final sample size was 3,060 cases (2,829 unique individuals). Sixteen percent (499) of the 3,000 individuals experienced CR after 28 days from initial infection, of which 65.73% (328) experienced it within 2 years of their initial chlamydial infection and 42.28% (211) within 1 year. However, characteristics of CR cases that occurred at least one year or any time after the initial infection were not significantly different from those who experienced CR after two years (not shown). Although most CR cases occurred in females (75%), the rate of CR was nearly the same between sexes as 12% of females and 8% of males experienced a CR within two years.

The median age of the study population was 23 (the median age for co-infected and reinfected individuals were 21.50 and 21.15, respectively). As expected, individuals within the age range 18-25 (categorical variable as described in the methods above) accounted for the most initial chlamydia cases and CR cases (see *Table 1*). The majority of cases were females (75%) and received the recommended treatment of azithromycin and/or doxycycline (92%). The proportion of individuals in the entire sample who had no treatment information was not extensive (n= 177, 6.3%); therefore no further analysis was done to compare these individuals with individuals who were reported as having (being prescribed) appropriate treatment.

Table 1: Baseline characteristics of individuals with chlamydia reinfections and co-infection in comparison to those with single infections, 2006-2015, Brant County, Ontario (n=2,829)

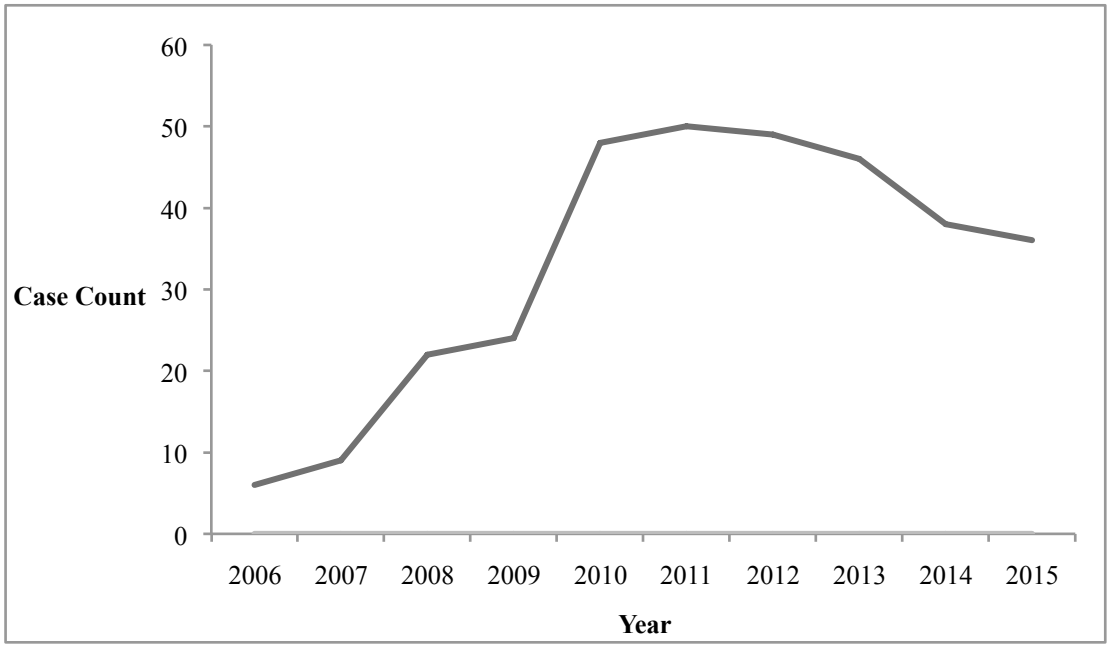
Baseline Characteristics	Chlamydia One Episode **n= 2,501 (%)**	Chlamydia Reinfection **n= 328 (%)**	Co-Infected **n= 167 (%)**
Age			
13-17	126 (5.04)	21 (6.40)	15 (8.98)
18-21	760 (30.39)	140 (42.69)	66 (39.52)
22-25	754 (30.15)	80 (24.39)	48 (28.74)
26-29	357 (14.27)	41 (12.50)	19 (11.38)
30-40	384 (15.35)	37 (11.29)	15 (8.98)
>40	120 (4.80)	9 (2.74)	4 (2.40)
Sex			
Male	882 (35.27)	83 (25.30)	77 (46.11)
Female	1619 (64.73)	245 (74.70)	90 (53.89)
Treatment			
Recommended	2317 (92.72)	301 (91.77)	157 (94.01)
Other	31 (1.24)	1 (0.30)	4 (2.40)
Missing	151 (6.04)	26 (7.93)	6 (3.59)
Documented Evidence of Partner Notification			
Yes	302 (12.08)	36 (10.94)	21 (12.57)
No	2199 (87.92)	292 (89.06)	146 (87.43)
Place of Residence			
Known	2424 (96.92)	322 (97.87)	146 (97.01)
Missing	77 (3.08)	7 (2.13)	5 (2.99)
Risk Factors			
High Risk Behaviors	353 (14.11)	37 (11.28)	20 (11.98)
High Risk Medical	30 (1.20)	16 (4.88)	5 (2.99)
Low Risk and/or Other	36 (1.44)	5 (1.52)	4 (2.40)
Missing	2082 (83.25)	270 (82.32)	138 (82.63)
Reason for Testing			
Symptoms	501 (20.03)	79 (24.09)	45 (26.95)
Contact Tracing	499 (19.95)	34 (10.37)	37 (22.16)
Routine Screening	564 (22.55)	74 (22.56)	28 (16.77)
Prenatal Screening	61 (2.44)	12 (3.66)	6 (3.59)
Missing	876 (35.03)	129 (39.33)	51 (30.54)
Median Time to Reinfection		276 Days	
Residing Within Brantford Core Area			
Yes		181 (55.18)	24*
No		140 (42.69)	20*
Missing		7 (2.13)	1*

*This represents the 45 individuals who were both co-infected and reinfected

Partner notification was documented as being completed for only 364 (338 if not including all 499 CR cases) individuals in the entire sample (12.13%) and for 36 (10.94%) of CR cases. Of the CR cases, 60 (12.02%) were also co-infected with gonorrhea and of those, 45 (13.72%) occurred in individuals who experienced a CR within two years. Chi-square tests indicated that co-infection was significantly associated with CR (Mantel-Haenszel p-value= <.0001) and partner notification was not (Mantel-Haenszel p-value= 0.4420). Of CR cases, 7 (2.13%) had missing residential information and 2 (0.613%) had a residential address outside of Brant County, in surrounding areas. The latter individuals were removed for final analysis, but there was no significant change in the final model.

The CR count rose sharply between 2009 and 2010, reaching a peak in 2011. Even though it began to fall following 2012, it remained high in comparison to other areas in Canada until most recently, the end of 2015 (See *Figure 1*).[21,46] The median time to CR was 276 days.

Figure 1. Counts of CR within two years of a first chlamydia infection in Brant County, Ontario, 2006-2015 (n= 328)



Overall, routine screening (22.27%) was the most common reason for testing, followed by having symptoms (20.87%). A chi-square test indicated that providing information for "reason for testing" was significantly associated with CR (Mantel-Haenszel p-value= 0.0027). Since a large proportion of the study population did not provide risk factor information (missing accounted for 83.13% of the study population), this variable was excluded from the final model. Of those who did provide risk factor information, the most common was "no condom use". When comparing cases with missing data to those without, there were no evident demographic differences between the two groups of individuals (see *Table 2*). However, CR occurred sooner in those with missing reasons for being tested. The chi-square test showed that risk factors were not significantly associated with CR in this sample population (Mantel-Haenszel p-value= 0.3512).

Table 2: Characteristics of chlamydia infections by risk factor information (available vs. unavailable), 2006-2015, Brant County, Ontario (n=3,060)

Baseline Characteristics	Risk Factor Information Available, n= 525 n (%)	Risk Factor Information Unavailable, n= 2535 n (%)
Age		
13-17	29 (5.52)	123 (4.85)
18-21	158 (30.10)	801 (31.60)
22-25	160 (30.48)	760 (29.98)
26-29	71 (13.52)	372 (14.67)
30-40	83 (15.81)	368 (14.52)
>40	24 (4.57)	111 (4.38)
Sex		
Male	165 (31.43)	861 (33.96)
Female	360 (68.57)	1674 (66.04)
*Treatment**		
Recommended	508 (96.76)	2328 (91.83)
Other	9 (1.71)	25 (0.99)
Missing	8 (1.52)	182 (7.18)
*Partner Notification**		
Yes	317 (60.38)	54 (2.13)
No	208 (39.62)	2481 (97.87)
Place of Residence		
Known	520 (99.05)	2452 (96.73)
Missing	5 (0.95)	83 (3.27)
*Reason for Testing**		
Symptoms	162 (30.86)	472 (18.62)
Contact Tracing	105 (20.00)	468 (18.46)
Routine Screening	137 (26.10)	555 (21.89)
Prenatal Screening	36 (6.86)	48 (1.89)
Missing	85 (16.19)	992 (39.13)
Reinfection		
Yes	106 (20.19)	453 (17.87)
No	419 (79.81)	2082 (82.13)
*Mean Time to CR**	165 Days	108 Days
Co-infection		
Yes	9 (1.71)	51 (2.01)
No	516 (98.29)	2484 (97.99)

* Indicates statistically significant differences

Survival Analysis

Table 3 shows the multivariate Cox regression model. Independent risk factors for CR included being female, 25 years old or younger, and not receiving recommended antimicrobial treatment for initial and/or subsequent infection. When age was analyzed as a continuous

variable, it indicated that with every year increase in age, an individual had a 6.1% decrease in risk of CR (p= <.0001).

Table 3. Multivariate Cox regression analysis of time to CR, by baseline characteristics, two-year follow-up, 2006-2015, Brant County, Ontario (n= 3,060)

Baseline Characteristics	Chlamydial Reinfection Adjusted HR (95% CI)	Co-Infection Adjusted HR (95% CI)
Age	0.94 (0.92, 0.96)¶	0.93 (0.90, 0.96)¶
Age		
13-17	3.19 (2.46, 3.92)¶	9.19 (7.69, 10.69)¶
18-21	2.77 (2.16, 3.38) ¶	5.22 (3.79, 6.64)¶
22-25	1.61 (1.00, 2.23)¶	3.39 (1.97, 4.82)¶
26-29	1.15 (0.54, 1.77)	2.30 (0.83, 3.76)
30-40	1.17 (0.55, 1.79)	1.96 (0.49, 3.44)
>40	1.00 (Referent)	1.00 (Referent)
Sex		
Male	1.00 (Referent)	1.00 (Referent)
Female	0.79 (0.57, 1.00)¶	0.46 (0.10, 0.81)¶
Treatment		
Recommended	1.00 (Referent)	1.00 (Referent)
Other	0.58 ([-0.59], 1.74)	3.12 (2.10, 4.13)¶
Missing	1.44 (1.09, 1.79)¶	0.78 ([-0.04], 1.61)
Partner Notification		
Yes	1.00 (Referent)	1.00 (Referent)
No	1.126 (0.844-1.408)	0.983 (0.488-1.478)
Reason for Testing		
Routine Screening	1.00 (0.75, 1.26)	0.68 (0.17, 1.19)
Symptoms	1.00 (Referent)	1.00 (Referent)
Contact Tracing	0.86 (0.57, 1.16)	0.94 (0.48, 1.40)
Prenatal Screening	0.89 (0.43, 1.35)	0.88 ([-0.07], 1.84)
Missing	1.17 (0.93, 1.42)	0.88 (0.44, 1.32)
Co-infection		
Yes	0.93 (0.63, 1.21)	
No	1.00 (Referent)	

¶Indicates statistical significance (p<0.05)

Although co-infection was not a statistically significant risk factor for CR (Mantel-Haenszel p= 0.5206), because co-infection and CR were found to be associated with one another (Mantel-Haenszel p-value= <0.01), the final survival analysis model was run a second time with co-infection as the outcome of interest, rather than a risk factor. The statistically significant risk factors were similar to those for CR, being female, 25 or younger, and not receiving the recommended antimicrobial treatment (*Table 3*).

Discussion

In Canada, this is the first recent study to present information on the reason individuals were coming in to a health care professional to be tested for chlamydia. Routine screening was the most common reason for testing in this population, which is important, as previous literature reports that incorporating testing for young women during routine check-ups, such as pap smears, may greatly impact positivity rates.[28] However, it should be noted that screening is not generally distinguished from diagnostic testing, so it is possible that people who had symptoms were reported as "screened" when in fact if they had symptoms, the test would have been diagnostic and not a screening.[47]

The 10.9% CR rate–within two years of initial infection–and the 9 month median time to CR is consistent with results in two other Canadian studies.[5,21] Findings are also consistent with American literature, where reviews of CR studies indicated that approximately 13.9% of females[4] and 11.3% of males[31] experienced a CR within 12 months. However, unlike previous studies, neither partner notification nor co-infection was significantly associated with time to CR. This may be a result of the fact that partner notification procedures were only completed for approximately 10% of cases and most individuals in this particular sample were identified as coinfected after their first episode of chlamydia.

Although Brant County collects information on risk factors, this variable was excluded from the final model due to the high proportion of missing data. Individuals who were missing risk factor information were more likely to be inadequately treated and less likely to have partners notified or to have reasons for testing recorded. Last and most important, individuals with missing risk factor data had a shorter mean time to CR compared to those for whom risk

factors were recorded (108 and 165 days, respectively). This may mean that these patients are more reluctant to provide information to healthcare professionals; health professionals may not be interviewing patients thoroughly; or are unaware of the importance of this information for program evaluation.

The key finding of the present study support the hypothesis that individuals who experience a CR within two years after initial infection differ from individuals who have only had one infection in terms of demographics. In addition, the greatest proportion of CR cases resided in the previously identified core area, which also accounts for the highest percentage of births to teenage mothers, single parent families, and lower median household income than the rest of Brant County.[46]

These findings are vital, as individuals at high risk of CR, live in the core area, and disproportionately impact the infection rate within and outside the core area.[33,48,49] In addition, members of core areas are often socioeconomically disadvantaged, highly mobile (moving in and out of the downtown core), and have high numbers of partners, evident in this population. Therefore, the success of any preventative effort depends on having a firm understanding of the community.[29] The present study findings further the understanding of high-risk groups in Brant County by creating a baseline description of STI epidemiology before a specialized STI clinic is established in the core area. Continuing these measures after implementation will be an important part of evaluating this multisectoral intervention (clinic, education, partner notification promotion), which has been found to be the most effective of preventative methods.[8]

Although a program targeting core group members would have the most impact on decreasing chlamydia prevalence, it is not possible to identify and screen every core group

member. This is due to the fact that some highly active people may reside outside of the core, which creates long distance connections that filter infection in and out of the core area, maintaining infection.[29,37,50,51] However, determining the most ideal method to target these subgroups will optimize the overall effectiveness of the program[52] and understanding what types of people are in that subgroup (high school students, college students, or older individuals) will impact the type of methods required to control infection. The present findings support the fact that Brant County exhibits similar chlamydia trends and patterns as other parts of Canada (Montréal and Vancouver), which indicates that findings from this study may be generalizable to other Canadian areas. Also, it is likely that Brant County would benefit greatly from a more targeted approach such as what was implemented in those areas due to the similar findings between studies in those areas and the present study.

There were some limitations of the study that should be acknowledged. First, the Canadian Guidelines on STIs were updated in 2008 and again in 2013, both within the study period. This may have changed the patterns of transmission in this population to some extent, as evidence suggest higher positivity rates in men and higher infection rates in females due to these changes.[11] We also did not have specific information on screening coverage, which could be an avenue worth exploring using the methods of the present study in order to get a better understanding of the proportion of individuals who are receiving screening, positivity rates, and the risk of CR. Strengths of the present study include the use routinely collected data to make inferences about the high-risk population in Brant County who make up a core group. This assists in our ability to make effective and sustainable changes to the behaviour of people in the underlying sexual network. Further, as mentioned there were the parallels with other studies, which may make these findings applicable to the region, in addition to other areas of Canada.

These findings are novel as they suggest that no treatment record could be a marker for inadequate or no treatment, playing a significant role in CRs. This accounts for young age and female gender, which has been previously noted.[50,53] However, we cannot be certain whether "missing" treatment was due to failure to receive treatment or if it was received elsewhere, such as a physician's office. Whether or not this is the case can be determined by conducting a treatment audit of a small proportion of clients in the future. We directly analyze sexual risk behaviors and the reasons for testing using routinely collected data so as to provide the basis for collecting better information on these as a matter of better practice. This is key, as previous work has stated that differences in risk for CR due to age can be a result of differences in re-testing (included in "reasons for testing" variable) and sexual risk behaviors.[10] These findings begin to illustrate the relationship between single chlamydia infections, CR, and geographic location in Brant County and provide the basis for improved strategies to decrease rates of chlamydia in the region.

Transmission Patterns of *Chlamydia Trachomatis*: Sexual Networks & Spatial Clustering in Brant County, Ontario

Jenny P. Santos, BPH, MSc; Ann Jolly, PhD

Abstract

Introduction: The present study aims to elucidate the social and geographic relationships between individuals with chlamydia and illustrate how this information can be utilized in program planning.

Methods: Routinely collected notifiable disease information from the Brant County Health unit was used to conduct sexual network and spatial analysis to identify subgroups and areas with an excess of infection using the freeware, *SaTScan*. A total of 2,739 cases and contacts were included.

Results: Eighty eight percent of cases had no sexual partners elicited or reported. We identified twelve hotspots with high chlamydia infection rates. A total of 717 sexually active women within 29 dissemination areas in Brant were included within those hotspots. Fifty eight percent of hotspots occurred within the previously identified core group, in the urban core of Brant.

Discussion: Findings highlight the significant geographic clustering of infection in Brant due to over active adolescents and young adults both within and outside the core group. This provides the required information to target these individuals to control infection all over Brant County.

Introduction

Social network analysis of gonorrhoea first emerged in public health research in 1985, when John J. Potterat and colleagues constructed a disease transmission network for gonorrhea.

This study led to crucial public health advances, as it revealed the relationships between cases, venues, and their clustering in areas with certain social and demographic characteristics.[55] However, population-based studies since have focused primarily on individual risk behaviors (such as condom use) rather than the relationships between individuals and the spatial distribution of cases.[56,57]

Although individual level factors play a role, social factors that impact individuals at the population level can better describe transmission patterns, which can be pinpointed using spatial analyses. Direct transmission routes can be approximated using network analysis of contact tracing data. When geographic areas with higher incidence have been described, social network analysis has added important refinements which greatly improved the understanding of geographical core group areas.[50,51,58] In the absence of sexual network data, the spatial distribution of cases can not only describe core areas of infection but also indicate key players in the underlying social network, providing an evidence base for the development of targeted interventions.[49,54]

Due to the complexity of sexual network analysis and scarcity of data, it has been difficult to demonstrate the importance of sexual networks.[8,50] Previous literature analyzed the concept of core groups and spatial patterns of STIs in either urban or rural environment.[49] Brant County is unique as it is a region characterized by being split between urban and rural areas . The high-risk population was characterized elsewhere.[59]

It is hypothesized that Brant has a complex network of individuals operating within the previously identified core area, which facilitates *Chlamydia trachomatis* transmission. We used a social network and geospatial approach to investigate the core groups. This allows for the

identification of large components of cases in space, as well as potential hot spots of infection in dissemination areas that reflect the presence of a sexual network. We address whether these geographic clusters coincide with previous analysis of single infected, reinfected, cases and contacts of chlamydia. As the incidence rate of infection (287 per 100,000 and 311.8 per 100,000 in the urban core) and the density of reinfection (10.9%) are high in Brant in comparison to Ontario (274.8 per 100,000), this analysis may provide valuable insight into chlamydia epidemiology in this region.[1]

Methods

Study Area and Geographical Data

Brant County is centrally located in the province of Ontario, Canada, with surrounding areas including Hamilton to the east, London west, and Kitchener-Waterloo north. Census data revealed that the population in Brant County increased by 3.8% between 2006 and 2011, increasing from 125,099 (90,192 Brantford) to 129,288 (93,650 Brantford), excluding the Six Nations reserve.[60] The proportion of the population aged 65 and over increased by 8.2%, making up 5.9% of the overall Brant population. Young people aged 15-24, who are at high risk for STI account for only 13% of the population, whereas older individuals (over 50 years) account for approximately 38% of the population.

Laboratory-confirmed notifications of *C. trachomatis* from 2006 to the end of 2015 were extracted from the integrated public health information system (iPHIS). Records of individuals with other diagnoses were excluded from the study. Available case information included basic demographic characteristics as well as residential address and contact tracing indicators (including whether the contact was contacted, tested, and treated). Names of contacts along with

as much additional information as index cases could provide was utilized to match contacts with their case file if they tested positive for an infection at a later date. Names were then removed to ensure anonymity. A previous study defined differences in age, sex, and date of diagnosis between reinfected individuals and those who had only been infected once.[59] As chlamydial reinfection (CR) is an indicator of core group membership, cases were mapped separately in this spatial analysis.[29] A CR was defined as a subsequent or recurrent infection 28-730 (2yrs) days after the clearing the initial infection (identified using client ID) and only the first CR was included.

Brant County Core Group

Previous analyses in Brant County illustrated that the downtown core of Brantford had comparatively high incidence of chlamydia.[1] The neighbourhoods that make up the core area included the Core (642.7 per 100,000), Eagle Place (411.0 per 100,000), Terrace Hill (388.9 per 100,000), Homedale-William (374.3 per 100,000) and East Ward (347.4 per 100,000).[46] With this in mind, addresses and postal codes for single chlamydia cases, CR cases, and contacts were utilized to determine whether these different groups of individuals were likely to reside near one another or within the identical core area. Details on this process can be found in *Appendix C*. Using only client ID numbers to link individuals to their location, ensured clients' confidentiality.

Social network and Spatial Analysis

The freeware Pajek (Ljubljana, Slovenia) was utilized to conduct social network analysis and determine component size in which each person is connected by at least one path (sexual contact) with another. Residential information was transformed from address to dissemination

area (DA) codes using a postal code conversion file.[60] A DA is an area that comprises of one or more city blocks containing between 400 and 1000 individuals; it is the smallest geographic unit for which population estimates are available from the 2011 census data. ArcGIS was used to determine the coordinates associated with each DA making up Brant County, which DAs included a client's location and which included none.

SaTScan (Boston, MA, USA) is freeware commonly used in public health research to define clusters of infection. It uses a spatial scan statistic to scan the total study area using circular windows, comparing the probability of cases occurring inside the circle or within the surrounding areas outside of the circle. A null Poisson model was chosen with the assumption that cases of infection would be randomly distributed across space, giving an estimation of the number of cases that would be the expected in a DA based on the population density.

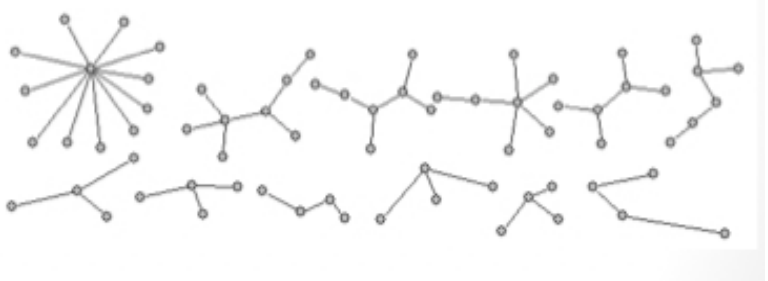
The program was set to scan for areas of "high" rate clusters only, where the observed number of cases exceeds the number of expected cases for the DA in question (*Appendix D*). If any statistically significant circles (clusters) are identified, they are then ranked according to the magnitude of their likelihood ratio test statistic and labeled as either a primary or secondary cluster. The null hypothesis is that the number of infections does not different across the region; therefore the significant clusters indicate that the number of infections within them is significantly different from other areas in the region. No temporal analysis was conducted due to data restrictions. SaTScan produces a shape file of geographic boundaries, which read into ArcGIS (Redlands, CA, USA) to see spatial clusters on a map of Brant County by DA (*Figure 1b*).

Results

Descriptive Epidemiology

There were 2,294 individuals in the single infection file (reduced from 2,501 to include only those who were diagnosed and treated in Brant County); 328 in the CR file, and 116 in the contact file. Of CR cases, 56.7% resided within the Brantford core area, (only 2.8% of individuals had missing residential information), 75% were female, and most occurred among those aged 18 to 25 years. Of the 18 (5.5%) contacts named by CR cases, 14 were located within the Core/East Ward Neighbourhoods, one was missing, and the other three were located in Paris, Shellard Lane, and Mayfair neighbourhoods. Of single cases**,** 98 (4.3%) were named contacts. These individuals were located all over Brant County in every neighbourhood. Partner notification was overall very poor in this region, with only 1-2 partners nominated for every case who provided information (n= 338), resulting in 116 contacts for the total 3,499 cases. This made the components extremely small (*Figure 1*) with only 12 components of 4 or more individuals and about 40 dyads and triads. Therefore, we focused primarily on the spatial analysis component.

Figure 1: Social Network Components of Size 4 or More, Brant County, 2006-2015

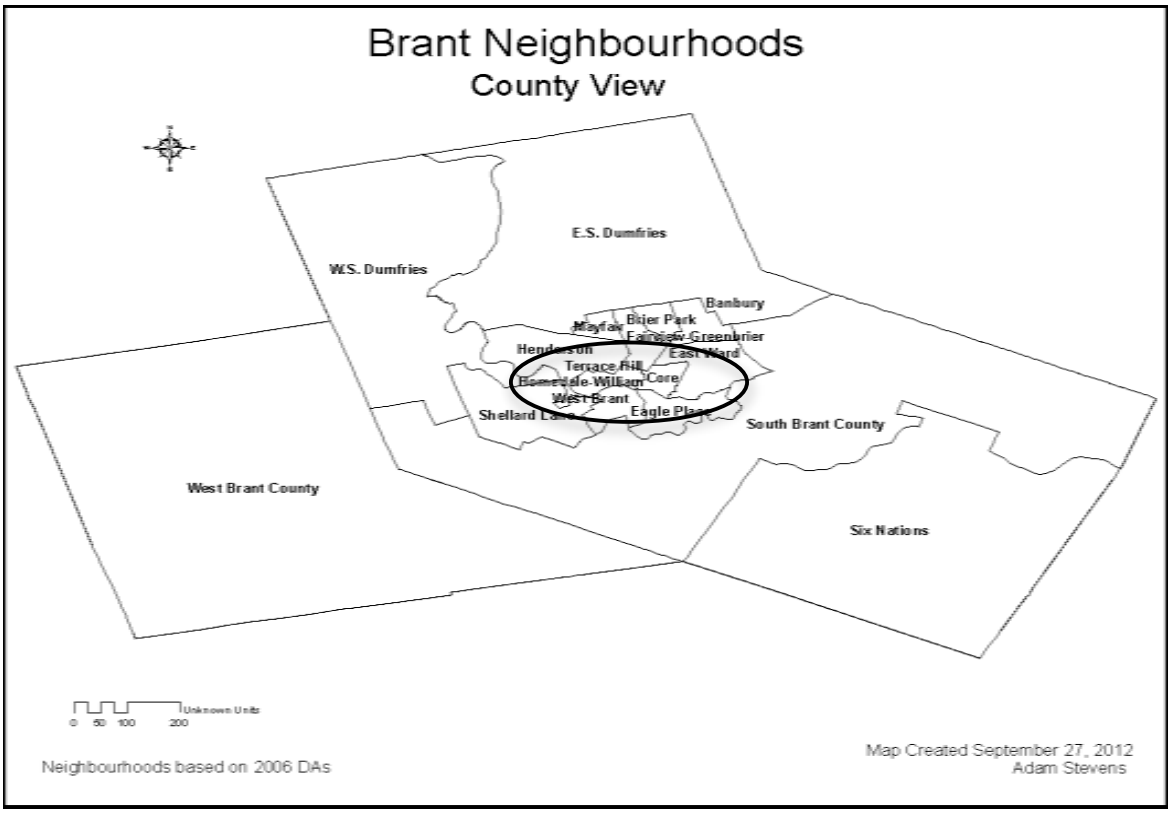


The core (downtown) neighbourhood was at the centre of the identified core group and was the only neighbourhood in which the infection rate consistently increased between 2005 and 2014 and continued to increase in the present analysis that included 2015. The core area is

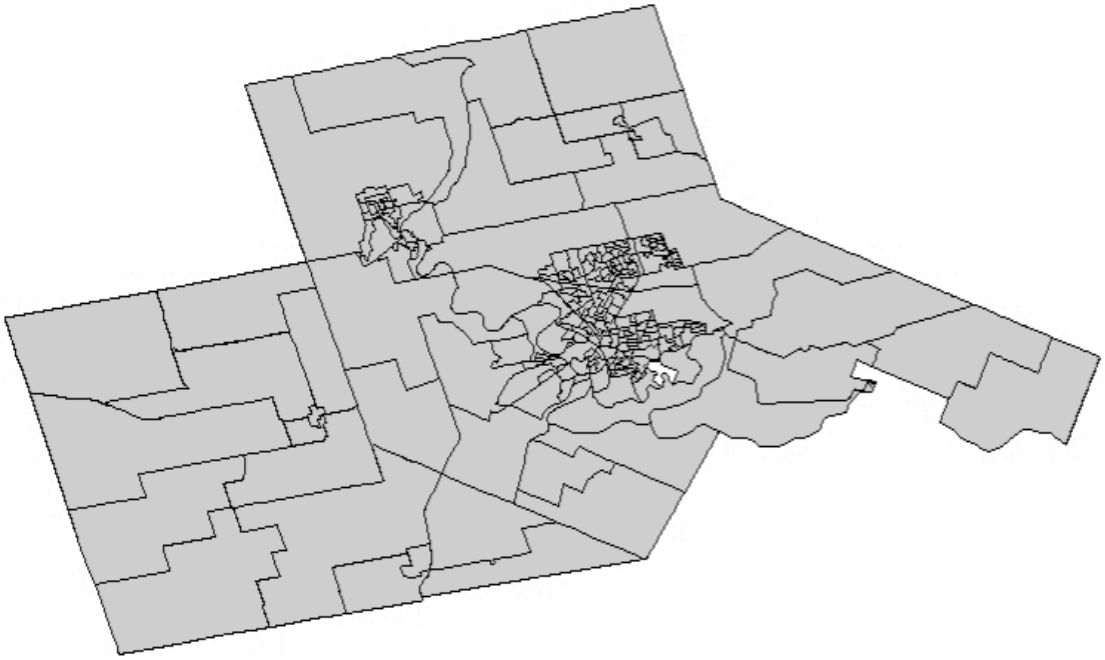
characterized by greater unemployment rates and lower mean household income than surrounding areas.[11] The neighbourhoods Eagle Place (13.4%), Homedale-William (13%), Core (11.9%) and Terrace Hill (9.2%) have the highest percentages of babies born to teen mothers (under 20 years old) over 2012-2014. In addition, Brantford has higher rates of lone-parent families (20.3% versus 16.3% in Canada).[12] *Figure 2a* illustrates the location of these neighbourhoods and indicates where the bulk of the Brant County population resides, in the urban core. *Figure 2b* illustrates Brant County by dissemination area.

Figure 2: (a) Brant County by Neighbourhood (b) Brant County by Dissemination Area

A



B



Hotspots of Increased Chlamydia Infections

Tables 1 and *2* show the results from the SaTScan tests for significant spatial clustering of single chlamydia infections and CR cases, respectively, after adjusting for the population at risk in each dissemination area. *Figure 3* shows the location of significant clusters for single infections and *Figure 4* shows significant clusters among CR cases. Of the 15 clusters identified in the spatial analysis for CR, only 1 was statistically significant (including 7% of CR cases; 23/328). Of the 23 clusters identified for single infections, 11 were statistically significant (included 26% of single cases 585/2294). These clusters were of relatively small radius, ranging from 0 - 4.52km.

Of the 230 DAs that make up Brant County, 29 (12.61%) did not include a single case of infection in any of the three groupings (single, CR, contacts). Of the remaining 201, 29 (14.43%) made up one of the 12 significant clusters identified in this study, of which 7 (58.33%) were

located within the previously identified core area of chlamydia infection (*Appendix E*). *Figure 5* illustrates all significant clusters in the urban core of Brant, upon a map showing the population at risk in each DA. All cases consisted of 3,000 different individuals (reduced to 2,739 when including only individuals who met the study definition of a CR and only those with diagnosis and treatment records within Brant County). Overall, 717 (26.18%) individuals were within one of the 12 significant clusters.

Table 1: Distribution of single chlamydia infection by dissemination area (age and population adjusted)

Clusters	Radius (km)	Observed Cases	Relative risk of excess cases	Test Statistic[a]	p-value	Total DAs
Cluster 1	0.22	80	4.64	59.10	<0.000	2
Cluster 2	0	36	5.41	31.24	<0.000	1
Cluster 3	0.50	80	2.26	20.19	<.0.000	4
Cluster 4	0.68	66	2.39	18.87	<0.000	2
Cluster 5	0	30	3.97	18.82	<0.000	1
Cluster 6	0	32	3.20	15.08	<0.000	1
Cluster 7	0	28	3.46	14.77	<0.000	1
Cluster 8	0	29	3.20	13.72	<0.000	1
Cluster 9	4.52	95	1.72	11.51	<0.002	5
Cluster 10	0.47	81	1.81	11.43	<0.002	5
Cluster 11	0	27	2.84	10.62	<0.004	1

[a]From likelihood-ratio test

Table 2: Distribution of CR cases by dissemination area (age and population adjusted)

Clusters	Radius (km)	Observed Cases	Relative Risk of excess cases	Test Statistic[a]	p-value	Total DAs
Cluster 1	0.48	23	4.00	14.18	<0.000	5

[a]From likelihood-ratio test

Figure 3: Brant County, Map of Single Infection Clusters and Zoom on Urban Core

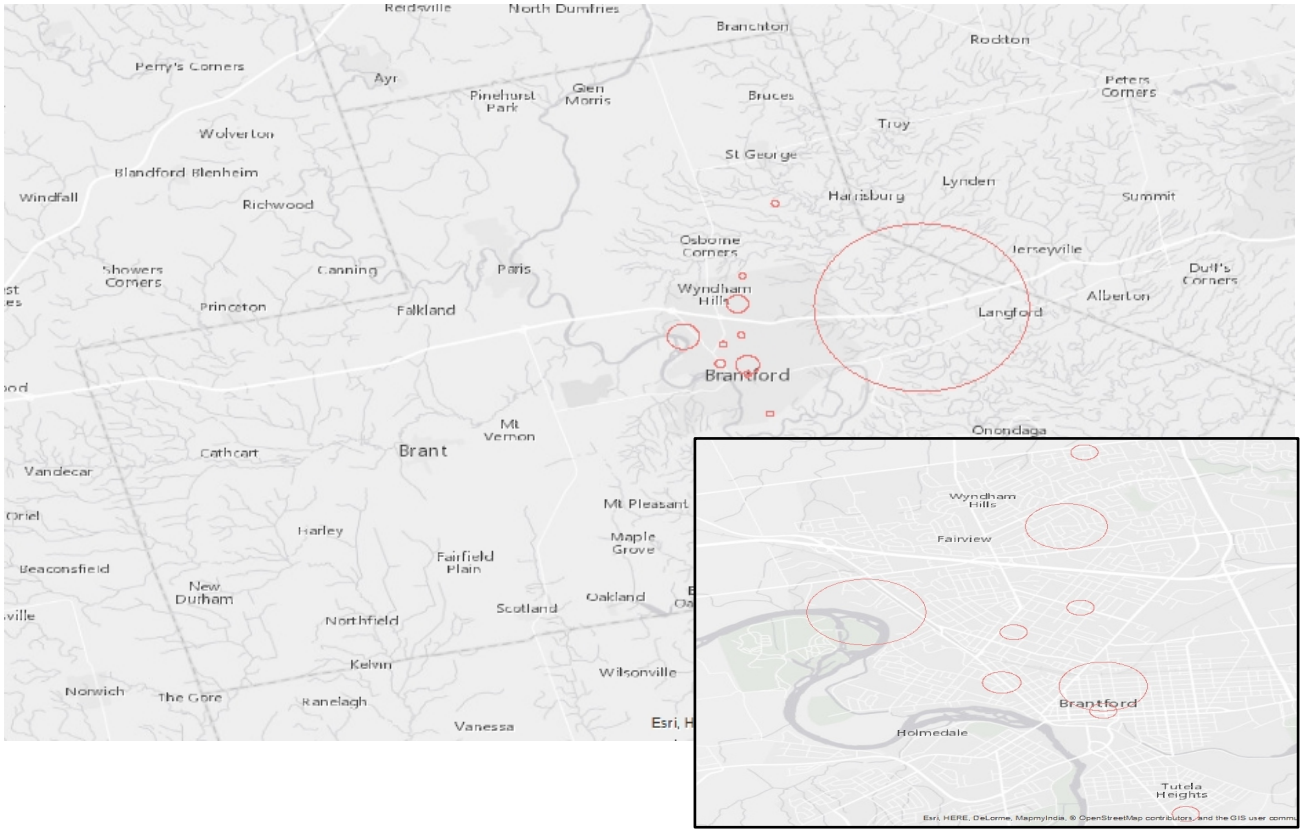
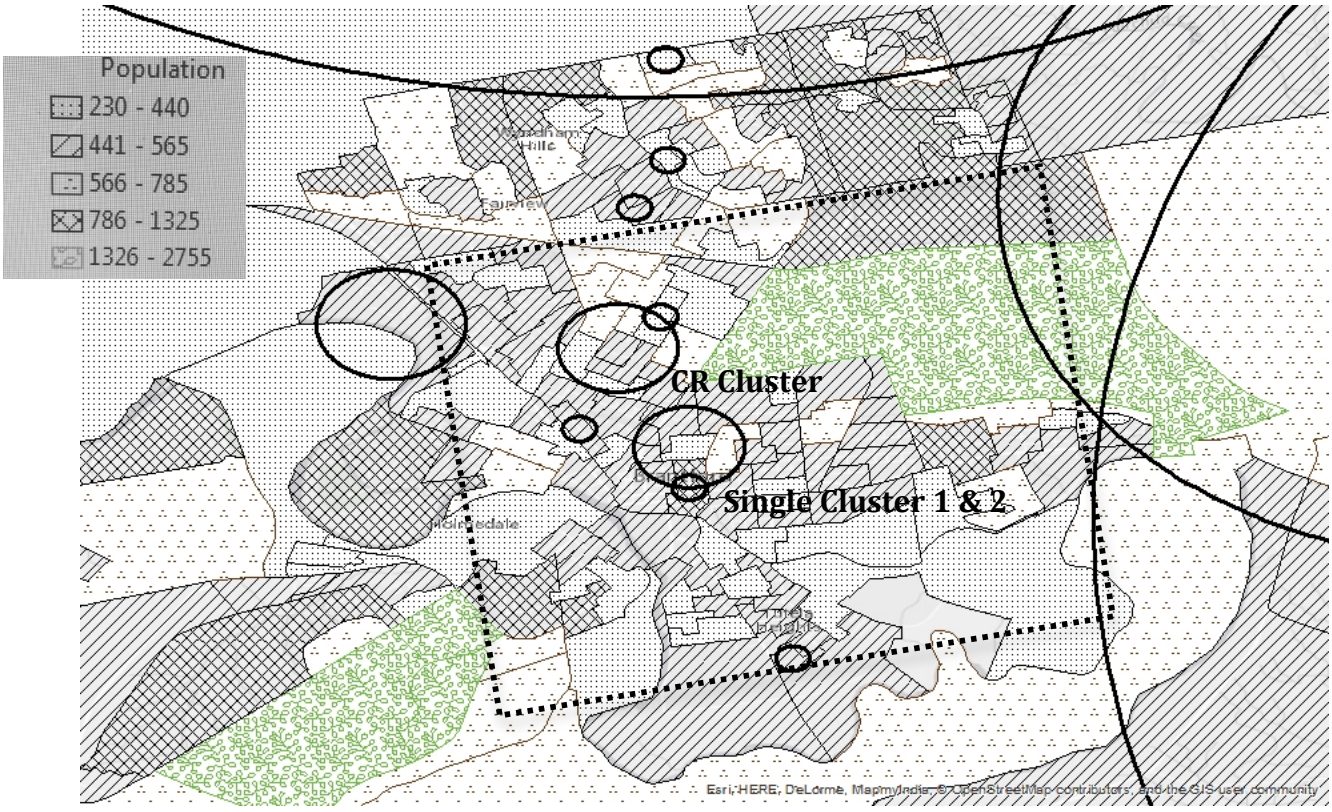


Figure 4: Brant County, Map of CR Cluster and Zoom on Urban Core

The significant CR cluster is located within the Homedale-William neighbourhood, which has a population of approximately 6,812 (3,325 males and 3,515 females) as of 2011. Individuals living within this cluster were four times more likely to experience a CR than those in surrounding areas (RR= 4.00, p= <0.000). Cluster 1 and Cluster 2 are two areas where individuals were more likely to have a single chlamydia infection than the other 9 (RR= 4.64 p= <0.000 and RR= 5.41 p= <0.000, respectively). These two clusters overlap one another and are located within the core neighbourhood. These three clusters are indicated in *Figure 5*, which also shows where the previous identified core area was (dashed box). The segment of larger arcs depicted on the outer limits of the map represent larger clusters of infection outside the core area, not shown in their entirety.

Figure 5: All cases, by dissemination area and population size



Discussion

The main finding of the present study is that clusters of infection when looking at single infection and CR occurred separately in neighbouring areas and both occurred within the previously identified core area. We made some interesting observations including the fact that significant clusters of infection were very small in size ranging from 0 to 4.52km. There is a lot of mixing and partner selection within close proximity, with some cases even sharing the exact same postal code. However, there are also partnerships in all neighbourhoods that make up Brant County. This supports the finding that core members facilitate transmission both inside and outside the core.[33,37,48,61] Access to the STI clinic can be difficult as these core individuals may have transportation difficulties. This is an issue that is currently being considered, to determine whether providing bus passes, or childcare, would increase program uptake.

The implications of these findings for prevention are that targeting groups for screening, or additional interventions in only one geographic area(s) may be helpful but only up to a certain point. Sexual network members who live outside of the core areas can also transmit infection back into the core area as well as continue transmission out of it. Therefore, having enhanced prevention strategies and partner notification with these individuals is also crucial. The fact that we found geographical clusters of up to 4.5 km is consistent with research by Hippe and Jolly (2012), which showed that 60% of cases and contacts of chlamydia reside within 4.5 or 5 km of each other.

Second, the aging population is of particular interest in this region (median age is 43 years old, which is higher than the national median of 40 years old), as those 40 plus account for approximately 54% of the overall population. The age group who are at the highest risk of

chlamydia infection and CR are those in the 15 to 24 year old age group. This group is increasingly diminishing in this population, yet they still account for the highest rates of infection in the region, indicating increased risk activity of the younger population and the increased the need for a targeted intervention.

The spatial analysis is particularly useful for planning interventions for chlamydia, as it highlights which groups in the population would benefit the most from an intervention.[57] Targeting interventions at identified individuals will have the most impact on the control of chlamydia as these individuals are more likely to engage in high-risk behaviors. Brant has developed a plan for a specialized sexual health clinic in downtown Brantford, which was identified as the core area of infection. It is aimed at these high-risk individuals who cluster in this area of the region. After implementation, it will be essential to combine findings from this study and results from post implementation to determine the impact of the clinic on the control of chlamydia in this region.

Contact tracing overall was neglected in this population, which was a great limitation of the study as we were unable conduct a full social network analysis as planned. Contact tracing results in many incomplete networks as cases often forget sex partners they had in the last 2 years.[62] This does not necessarily mean that those contacts are different from the contacts that do get reported. Although we cannot definitively describe the complexity of the sexual networks in Brant County, we can say that the methods used here can be used in the future with a temporal component included. Doing so would not only provide further context for targeted interventions but could potentially predict outbreaks.[49] As contacts and other cases were located outside the core areas, it seems prudent to formulate a systematic partner notification strategy. The strategy

also has the advantage of discharge of obligations of investigation as to source, and treatment under the Ontario Health Promotion and Protection Act and regulations.

Strengths of the present study include the fact that contact information was gathered via contact tracing information from the study population. Contact tracing is one of three ways to construct sexual networks (others includes snowball sampling of patients and asking patients to estimate the number of partners they have had). As there is very little spatial research on STIs in Canada, the present study sheds light on transmission patterns of chlamydia in Canada using methodologies similar to those used in other studies.[63-65]

The notifiable disease database includes only confirmed cases of STIs, which is also a limitation as the sexual network constructed is likely an underestimate of what is actually in place. Asymptomatic, unscreened, and unreported cases are not accounted for. Second, the self-report nature of the contact data will serve to underestimate the size of the sexual networks. This may lead to bias due to the high potential for information errors, as cases that were reinfected were more likely to provide information because they had a second opportunity to provide contact data. Finally, data limitations restricted our ability to test whether there were significant differences within clusters, therefore these findings should be interpreted with caution.

In conclusion, although the findings of the present study provide valuable insight into the current situation in Brant County, the overall effectiveness of the surveillance programs greatly depend on the completeness of surveillance data. This was a significant issue within Brant and leaves the question of just how complex the networks indicated by this spatial analysis are as well as what can be done to dissect them. Future directions for Brant include to work towards more complete surveillance data and partner notification, and incorporating demographic

information in with the spatial analysis to get a better description of the sexual network and where intervention efforts could be most effective.

Conclusion

The first manuscript in this series outlined the need for more research around chlamydia infections, specifically in Brant County, who has one of the highest rates of infection in Ontario. It also highlights the role of CRs in maintaining chlamydia resurgence, suggesting this is where we should focus our efforts. This was our first hypothesis: that CR significantly contributes to Brant County's high chlamydia incidence ad prevalence rates.

The second manuscript looks at whether our hypothesis is true in Brant County by first determining the extent of CR in the region. As the rate of CR was in fact very high and likely significantly contributing to chlamydia resurgence in the region, further analysis was conducted to determine characteristics of individuals significantly related to CR. We were able to determine that on average, a CR within the first two years of initial infection in Brant County constitute 10.9% of new chlamydia cases per year, varying from 5-12% within the study period.

The third manuscript follows up the second to look at our second hypothesis, that there were complex sexual networks in play that cluster in space, contributing to CR and chlamydia incidence. First, we determined where chlamydia cases are located in relation to CR cases within Brant County. It was determined that they clustered in close proximity to one another, supporting that CR contributes to chlamydia resurgence in Brant County and make up a core group in the urban core of the region. It was also determined that cases were located all over the region, while significant clusters of infection were quite small in size, suggesting high activity of the at risk population.

Overall, chlamydia has many sequelae that have a greater impact when CR rates are high. Brant County could tackle their high rate of chlamydia by focusing on the individuals that make

up the CR group. Brant County made a move to execute an intervention program such as this, implementing a specialized sexual health clinic in downtown Brantford. However, to ensure that interventions are the most effective, knowing both the characteristics of those individuals (and their contacts) as well as where they cluster in space is required to tailor programs. Incomplete information on the population caused some safety issues within the clinic, requiring the public health nurses to pull out of the operation.

In the future, information such as what is provided in the present study will be necessary to better understand the needs of the at-risk population. In addition, this information would also inform on what precautions should be planned ahead of time for the success of the program. If successful, intervention programs such as these will impact rates of infection both inside and outside the core area and overtime, decrease prevalence in Brant County.

References

1. Brant County Health Unit. (2013). Multifaceted solutions for complex public health problems: Population-focused interventions to decrease the rate of chlamydia. Brantford, ON: L. Hyseni, M. Huang & S. Edwards.
2. Public Health Ontario. (2012). Chlamydia in Focus. *Monthly Infectious Diseases Surveillance Report; 1(8):* 1-13.
3. Fortenberry, J.D., Brizendine, E.J., Katz, B.P., Wools, K.K., Blythe, M.J., & Orr, D.P. (1999). Subsequent sexually transmitted infections among adolescent women with genital infection due to chlamydia trachomatis, neisseria gonorrhoeae, or trichomonas vaginalis. *Sexually Transmitted Diseases; 26(1):* 26-32.
4. Hosenfeld, C.B., Workowski, K.A., Berman, S., Zaidi, A., Dyson, J., Mosure, D., Bolan, G., & Bauer, H.M. (2009). Repeat infection with chlamydia and gonorrhea among females: A systematic review of the literature. *Sexually Transmitted Diseases; 36(8):* 478-489.
5. Genereux, M., Leclerc, P., Bedard, L., & Allard, R. (2010). Upsurge of chlamydial reinfection in a large Canadian city: An indication of suboptimal chlamydia screening practices? *Can J Public Health; 101(5)*: 420-4.
6. Public Health Agency of Canada. (2013). *Sexually transmitted infections—A continued public health concern.* Retrieved from: http://www.phac-aspc.gc.ca/cphorsphc-respcacsp/2013/sti-its-eng.php
7. Malhotra, M., Sood, S., Mukherjee, A., Muralidhar, S., & Bala, M. (2013). Genital Chlamydia trachomatis: An update. *Indian J Med Res; 138:* 303-16.

8. Blanchard, J.F., Moses, S., Greenaway, C., Orr, P., Hammond, G.W., & Brunham, R.C. (1998). The evolving epidemiology of chlamydial and gonococcal infections in response to control programs in Winnipeg, Canada. *American Journal of Public Health; 88(10):* 1496-1502.
9. Kretzschmar, M., Satterwhite, C., Leichliter, J., & Berman, S. (2012). Effects of screening and partner notification on chlamydia positivity in the United States: A modeling study. *American Sexually Transmitted Diseases Association; 39(5)*: 325-331.
10. Liu, B., Guy, R., Donovan, B., & Kaldor, J.M. (2011). Chlamydia trachomatis re-infections in a population-based cohort of women. Sex Transm Infect; 89: 45-50.
11. Pedersen, K.S. & Andersen, J.S. (2014). Social-, age- and gender differences in testing and positive rates for Chlamydia trachomatis urogenital infection−a register-based study. Family Practice; 31(6): 699-705.
12. Stamm, W.E. (1998). Expanding efforts to prevent chlamydial infection. N Engl J Med; 339: 768-770.
13. Ontario. Ministry of Health and Long-Term Care. Ontario public health standards. Toronto, ON: Queen's Printer for Ontario; 2008 [revised 2014 May 1; cited 2014 May 1]. Available from: http://www.health.gov.on.ca/en/pro/programs/publichealth/oph_standards/docs/ophs_2008.pdf (Assessed June 15th, 2015)
14. Public Health Agency of Canada. (2013). Canadian guidelines on sexually transmitted infections. Retrieved from: http://www.phac-aspc.gc.ca/std-mts/sti-its/cgsti-ldcits/section-5-2-eng.php (Assessed June 15th, 2015)

15. Harpa, K. & Isfeld-Kiely, M.A. (2014). A review of evidence on partner notification practices for Chlamydia. National Collaborating Centre for Infectious Disease; 1-21.
16. Shaw, K., Coleman, D., O'Sullivan, M., & Stephens, N. (2011). Public health policies and management strategies for genital chlamydia trachomatis infection. Risk Management and Healthcare Policy; 4: 57-65.
17. Wang, S.A., Papp, J.R., Stamm, W.E., Peeling, R.W., Martin, D.H., & Holmes, K.K. (2005). Evaluation of antimicrobial resistance and treatment failures for chlamydia trachomatis: A meeting report. The Journal of Infectious Diseases; 191: 917-923.
18. Ginocchio, R.H.S., Veenstra, D.L., Connell, F.A., & Marrazzo, J.M. (2002). The clinical and economic consequences of screening young men for genital chlamydial infection. *Sexually Transmitted Diseases; 30(2):* 99-106.
19. Yu, Y., Frasure-Williams, J.A., Dunne, E.F., Bolan, G., Markowitz, L., & Bauer, H.M. (2011). Chlamydia partner services for females in California family planning clinics. *Sex Transm Dis; 38(10):* 913-18.
20. Cook, R.L., Hutchison, S.L., Ostergaard, L., Baithwaite, S., & Ness, R.B. (2005). Systematic review: Noninvasive testing for chlamydia trachomatis and neisseria gonorrhoeae. Annals of Internal Medicine; 142(11): 914-925.
21. Brunham, R.C., Pourbohloul, B., Mak, S., White, R., & Rekart, M.L. (2005). The unexpected impact of a Chlamydia trachomatis infection control program on susceptibiltity to reinfection. *J Infect Dis; 192(10):* 1836-92.
22. Brunham, R.C., & Rekart, M.L. (2008). The arrested immunity hypothesis and the epidemiology of chlamydia control. *Sexually Transmitted Diseases; 35(1)*: 53-54.

23. Wasserheit, J.N. (1992). Epidemiological Synergy: Interrelationships between human immunodeficiency virus and other sexually transmitted diseases. *Sexually Transmitted Diseases; 19(2):* 61-77.

24. Trelle, S., Shang, A., Nartey, L., Cassell, J.A., & Low, N. (2007). Improved effectiveness of partner notification for patients with sexually transmitted infections: Systematic review. BMJ; 334: 354-360.

25. Saperstein, A. & Firnhaber, G. (2010). Should you test or treat partners of patients with gonorrhea, chlamydia or trichomoniasis? *J Fam Pract; 59:* 46-48.

26. Althaus, C.L., Heijne, J.C.M., Herzog, S.A., Roellin, A., & Low, N. (2012). Indivudal and population level effects of partner notification for chlamydia trachomatis. PLoS ONE; 7(12): e51438.

27. Low, N., McCarthy, A., Roberts, T., &… Egger, M. (2005). Partner notification of chlamydia infection in primary care: Randomised controlled trial and analysis of resource use. *BMJ; 332:* 14-19.

28. Tao, G., Tian, L.H., & Peterman, T.A. (2007). Estimating Chlamydia screening rates by using reported sexually transmitted disease tests for sexually active women aged 16 to 25 years in the United States. *Sex Transm Dis; 34(3):* 180-2.

29. Jolly, A.M., Moffatt, A.E.K., Fast, M.V., & Brunham, R.C. (2005). Sexually transmitted disease threshold in Manitoba, Canada. *Ann Epidemiol, 15(10):* 781-8.

30. Wylie, J.L., Shaw, S., DeRubeis, E., & Jolly, A. (2010). A network view of the transmission of sexually transmitted infections in Manitoba, Canada. *Sex Transm Infect; 86(Suppl 3):* iii10-iii16.

31. Fung, M., Scott, K.C., Kent, C.K., & Klausner, J.D. (2007). Chlamydial and gonococcal reinfection among men: A systematic review of data to evaluate the need for retesting. *Sex Transm Infect; 83:* 304-309.

32. Geisler, W.M.,Wang, C., Morrison, S.G., Black, C.M., Bandea, C.I., & Hook III, E.W. (2008). The natural history of untreated chlamydia trachomatis infection in the interval between screening and returning for treatment. *Sexually Transmitted Diseases; 35(2):* 119-123.

33. Yorke, J.A., Hethcote, H.W., & Nold, A. (1978). Dynamics and control of the transmission of gonorrhea. *Sexually Transmitted Diseases; 5(2):* 51-6.

34. Stoner, B.P., Whittington, W.L., Hughes, J.P. Aral, S.O., & Holmes, K.K. (2000). Comparative epidemiology of heterosexual gonococcal and chlamydial networks: Implications for transmission patterns. *Sexually Transmitted Diseases; 27(4)*: 215-223.

35. Crichton, J., Hickman, M., Campbell, R., Batista-Ferrer, H., & Macleod, J. (2015). Socioeconomic factors and other sources of variation in the prevalence of genital chlamydia infections: A systematic review and meta-analysis. *BMC Public Health; 15(729): 1-10.*

36. Klovdahl, A.S., Potterat, J., Woodhouse, D., Muth, J., Muth, S., & Darrow, W.W. (1992). HIV infection in an urban social network: A progress report. Bulletin de Methodologie Sociologiquqe, 36: 24-33.

37. Jolly, A.M., & Wylie, J.L. (2013). Sexual networks and sexually transmitted infections; "The strength of weak (long distance) ties". In Aral, S.O., Fenton, K.A., & Lipshutz, J.A. (Eds.), The New Public Health and STD/HIV Prevention: Personal, Public and Health Systems Approaches (pp.78-111). New York, NY: Springer Science+Business Media.

38. Hocking, J.S., Vodstrcil1, L.A., Huston, W.M., Timms, P., Chen, M.Y., Worthington, K., McIver, R., & Tabrizi, S.N. (2013). A cohort study of chlamydia trachomatis treatment failure in women: A study protocol. *BMC Infectious Disease; 13(379):* 1-9.

39. Bilardi, J.E., Hopkins, C.A., Fairley, C.K., Hocking, J.S., Tomnay, J.E., Pavlin, N.L., Parker, R.M., &…Chen, M.Y. (2009). Innovative resources could help improve partner notification for chlamydia in primary care. *American Sexually Transmitted Diseases Association; 36(12)*: 779-783.

40. Golden, M.R., Whittington, W.L.H., Handsfield, H.H., Hughes, J.P., Stamm, W.E., Hogben, M., Clark, A., & … Holmes, K.K. (2005). Effect of expedited treatment of sex partners on recurrent or persistent gonorrhea or chlamydial infection. *N Engl J Med; 352(7):* 676-685.

41. Kissinger, P. & Hogben, M. (2011). Expedited partner treatment for sexually transmitted infections: An update. *Curr Infect Dis Rep; 13:* 188-195.

42. Mercer, C.H., Sutcliffe, L., Johnson, A.M., White, P.J., Brook, G., Ross, J.D.C., Dhar, J., &… Cassell, J.A. (2007). How much do delayed healthcare seeking, delayed care provision, and diversion from primary care contribute to transmission of STIs? *Sex Transm Infect; 83*: 400-405.

43. Gaydos, C.A., Wright, C., Wood, B.J., Waterfield, G., Hobson, S., & Quinn, T.C. (2008). Chlamydia trachomatis reinfection rates among female adolescents seeking rescreening in school-based health centers. *Sexually Transmitted Diseases; 35(3):* 233-237.

44. Ontario Public Health Index of Databases. (2015). Integrated Public Health Information System (iPHIS) [Canada]. Retrieved from: http://ophid.scholarsportal.info/details/view.html?q=can&uri=/phirn/iphis_PHIRN_e.xml

45. Santos, J.P. & Jolly, A. (2016). The Contribution of Reinfections to Chlamydia Resurgence in Brant County, Ontario. (Unpublished master thesis). University of Ottawa, Ottawa, Ontario.

46. Brant Health Atlas. (2014). *Reproductive Profile.* Retrieved from: http://www.bchu.org/StatsAndReports/BCHU-Reports/Pages/repro_profile.aspx

47. Boardman, L.A. & Peipert, J.F. (1998). Screening and diagnostic testing. *Clinical Obstetrics & Gynecology; 41(2): 267-274.*

48. Peterman, T.A., Tian, L.H., Metcalf, C.A., Satterwhite, C.L., Malotte, C.K., DeAugustine, N., Paul, S.M.,…& Douglas Jr., J.M. (2006). High incidence of new sexually transmitted infections in the year following a sexually transmitted infection: A case for rescreening. *Ann Intern Med; 145*: 564-572.

49. Gesink, D. C., Sullivan, A. B., Miller, W. C., & Bernstein, K. T. (2011). Sexually transmitted disease core theory: Roles of person, place, and time. American Journal of Epidemiology, 174(1), 81-89.

50. Wylie, J.L. & Jolly, A. (2001). Patterns of chlamydia and gonorrhea infection in sexual networks in Manitoba, Canada. *Sexually Transmitted Diseases; 28(1):* 14-24.

51. Hippe, J. & Jolly, A.M. (2012). STI phase and the geography of sexual partnerships: Prevalence of long-distance sexual contacts among chlamydia, gonorrhea, and coinfected STI cases in Manitoba, Canada. *Spatial and Spatio-temporal Epidemiology; 3:* 255-263.

52. Kretzschmar, M., van Duynhoven, Y.T., & Severijnen, A.J. (1996). Modelling prevention strategies for gonorrhea and chlamydia using stochastic network simulations. *American Journal of Epidemiology; 144(3)*: 306-317.

53. Jolly, A. (1998). Sexually transmitted disease core group membership in Manitoba; Strategies for definition and description of risk markers. (Unpublished doctoral thesis). University of Manitoba, Winnipeg, Manitoba.

54. Doherty, I.A., Padian, N.S., Marlow, C., & Aral, S.O. (2005). Determinants and consequences of sexual networks as they affect the spread of sexually transmitted infections. *The Journal of Infectious Diseases; 191(Suppl 1):* S42-54.

55. Potterat, J.J., Muth, S.Q., Rothenberg, R.B., Zimmerman-Rogers, H., Green, D.L., Taylor, J.E., Bonney, M.S., & White, H.A. (2001). Sexual network structure as an indicator of epidemic phase. *Sex Transm Infect; 78(Suppl 1):* i152-i158.

56. Liljeros, F., Edling, C.R. & Nunes Amaral, L.A. (2003). Sexual networks: Implications for the transmission of sexually transmitted infections. *Microbes and Infection; 5:* 189–196.

57. Luke, D.A. & Harris, J.K. (2007). Network Analysis in Public Health: History, Methods, and Applications. *Annu. Rev. Public Health; 28:* 69-93.

58. Wylie, J.L., Cabral, T. & Joly, A.M. (2005). Identification of networks of sexually transmitted infection: a molecular, geographic, and social network analysis. *J Infec Dis; 191(6):* 899-906.

59. Santos, J.P. & Jolly, A. (2016). The Contribution of Reinfections to Chlamydia Resurgence in Brant County, Ontario. (Unpublished master thesis). University of Ottawa, Ottawa, Ontario.

60. Statistics Canada. (2016). *Focus on Geography Series, 2011 Census: Census subdivision of Brant, CY-Ontrario*. Retrieved from:

https://www12.statcan.gc.ca/censusrecensement/2011/as-sa/fogs-spg/Facts-csdeng.cfm?LANG=Eng&GK=CSD&GC=3529005 (Assessed June 15, 2016)

61. Gesink, D. C., Sullivan, A. B., Miller, W. C., & Bernstein, K. T. (2011). Sexually transmitted disease core theory: Roles of person, place, and time. American Journal of Epidemiology, 174(1), 81-89.
62. Brewer, D.D., Garrett, S.B., & Kulasingam, S. (1999). Forgetting as a cause of incomplete reporting of sexual and drug injection partners. *Sex Transm Dis*; *26(3):* 166–76.
63. Wand, H. & Ramjee, G. (2010). Targeting the hotspots: Investigating spatial and demographic variations in HIV infection in small communities in South Africa. *Journal of The International AIDS Society; 13(41):* 1-9.
64. Westercamp, N., Moses, S., Agot, K., Ndinya-Achola, J.O., Parker, C., Amolloh, K.O. & Bailey, R.C. (2010). Spatial distribution and cluster analysis of sexual risk behaviours reported by young men in Kisumu, Kenya. *International Journal of Health Geographics; 9(24).*
65. Ramjee, G. & Wand, H. (2014). Geographical clustering of high risk sexual behaviors in "hot-spots" for HIV and sexually transmitted infections in Kwazulu-Natal, South Africa. *AIDS Behav; 18:* 317-322.

Appendices

Appendix A: Field Notes on Data Cleaning

Summary:

Originally, the plan was to match case and contact data (as well as case-case data, as the focus was on re-infection) extracted from the iPHIS database. It proved to be difficult to do so as contact data in the database was poor.

Chlamydia Cases: 6478
Gonorrhea Cases: 1187
Co-Infections: 169 (157 single case of coinfection)
Case Contacts: ≈ 389 (reinfection subsample) 374 (co-infection subsample) 371 (merged dataset)

Total: 7665

Data Cleaning Notes:

Step 1: Consider whether time-dependent covariates are a concern in the present data set. This determines both whether the data should be set up in a way that presented two rows of information for each subject or just one, as well as exactly how data is analyzed in the final model.

- In this case, there are none that are a concern

Step 2: Perform data cleaning of all data

Need to:

- Check for *duplicate records*.
    - Both duplicate ID/Case Number and any duplicate observations for each ID/Case Number
- *Verify postal code*/address information (for proper Network Analysis)
    - If given address, ensure address is in Brant County
    - If in Brant County, pinpoint associated postal code (given in this case)
    - If any address or postal code is not located in Brant County, delete those cases
- Check values for *categorical variables* to ensure no other values other than the ones specified in the data dictionary (the form filled out in sexual health clinic) are present
    - Done using PROC FREQ to list responses, followed by PROC FORMAT with a WHERE statement to identify ID/Case Number associated with any anomalies

  - If associated with one particular aspect of the data, explain it. If not possible to do so, likely a data entry error, therefore anomalies are listed as INVALID
  - Lists outliers
  - Use discretion to either remove or analyze any abnormalities (if many, not wise to remove because it is unknown whether they are date entry errors or what caused them)
- Check values for *numerical variables* to ensure no values fall outside the pre-specified range
  - Similar to what is done for categorical variables, except using PROC MEANS and PROC UNIVARIATE, including ID statement to identify anomaly location
  - Lists outliers
- Determine the extent of *"MISSING"*
  - If extensive, determine its' relationship with each of the variables under investigation in order to decide whether they will be included in analysis on its own or whether it will be grouped with "INVALID" values
    - Contact Tracing, Reason for Testing, & Risk Factor Information have a great proportion of missing data
- Check range for *dates* and invalid dates
  - There were none in this case, other than dates that were missing (these cases were removed as it would not be possible to calculate time to event without date of case being provided)

Original Dataset:

n=7665
Variables= 62

n=7665
Variables= 24 (26 after creating new age and risk factor variables)

Exclude Age and Gender Data Entry Errors
n=7665−>7654

Exclude Missing Data (cases with no information associated with it)
n=6675

Exclude Gonorrhea infections (for now)
n=5656

Date Function (in order to extract only the year from the encounter date) & delete some invalid case entries (all information on that case is missing)
n= 5656
Variables=27

Proc Sort NODUPKEY (remove people with duplicate STD_Encounter_IDs, should all have unique IDs for subsequent infections associated with their Client_ID, unless there is a co-infection in question)
n= 3631

Variables=28 (created reinfection variable using client_ID)

Delete Multiple Reinfections (n=3499)
Final Reinfection n= 499 (down from the original n=631)

Time-to-Event Calculation…33 Variables (n=3000)
Remove reinfection time greater than 2 years (reinfection n= 329)

-Remove the one CR case located outside of Brant, in Hagersville (**n=328**)

Co-infection:

Proc Sort *NODUPKEY* by disease…18 Variables (7654–> 4130)

-Delete invalid case entries (all information on case is missing)
n=4106

-Delete Multiple Co-infections (n=3367) Co-Infection (n=167)

Final Dataset:

Merge Co-infection data and Reinfection data

n=3167 Obervations (3060 when coinfected without CR are removed) with 34 variables (including partner notification, treatment, and residence variables)

Appendix B: Field Notes on Reinfection Analysis

As the main data analysis method of choice was to be the Cox regression proportional hazards model, which requires certain pieces of information in order to run, two new variables (other than those extracted from iPHIS) were made:

1. The "Independent Variable"= *Censoring Variable*
    a. Created by determining whether subjects experienced the outcome under investigation (reinfection=1) or not by the end of the follow-up period (no reinfection=0)
2. The "Dependent Variable"= *Length of Follow-up Variable*
    a. Created by determining the number of days from the subjects' index case to repeat infection or until the end of the study period if the case is censored
    b. Done by using date functions in SAS

Data analysis **begins** with *data exploration*:

- Necessary for a basic description of the study population
    - Data for Table 1 generated using various SAS procedures such as Proc Freq, Proc Means and Proc Univariate
- Allows to see whether covariates are correlated with the outcome as well as each other, which in turn allows for the identification of potential interactions
    - Chi-Square and regression tests (where appropriate) were used to determine whether variables were related with one another
- Data on additional contacts of cases (other than the first partner a case named) was removed for the re-infection analysis as they created duplicate records for some cases
    - These were removed using the nodupkey/noduprecs options in SAS
- The final same size for this analysis was n=3499 with 32 variables.
- Newly created variables: reinfection, co-infection, age, partner notification, dependent variable, year (yr)

Data analysis will **continue** with *creating all the necessary variables* needed for the final model:

1. **Dependent Variable:** Using Date Function

Retains the value of the date associated with the first infection (using RETAIN function in SAS) and subtracts the date associated with the second infection (REMINDER: infections after the 1st reinfection were not considered in this research) from the first.

> -Output illustrates the number of days between the two infections
> -After this, as I noticed that some reinfections occurred much later than the initial infection, another data cleaning step was conduced to restrict the output to include only the reinfections that occurred within the first 28 days to 2 years after the first infection (28-730 days) in order to allow for the findings to be generalizable, as this was commonly done in other reinfection literature

Using PROC UNIVARIATE, the median time to reinfection was calculated as 276 (days)

2. **Independent Variable:** Using client ID after duplicates were removed.
    a. If a client ID appeared a second time (with a different disease and/or contact ID listed with the ID), it was categorized as a reinfection case (given a value of '1' if it occurred, or a value of '0' if it did not occur)

Final Model: Cox Regression Proportional Hazards Model

- Ties were not accounted for by the "exact" method. The Breslow method was used, which keeps the log partial likelihood methods used in analysis the same
- Run with the main outcome under consideration as the dependent variable (reinfection) as well as the other outcome that was also of interest (co-infection)
    - This was done only to see whether there were any differences in independent risk factors as co-infection was not a significant risk factor for reinfection, which raised some questions

Missing Data:

As ignoring missing data in analysis also ignores possible systematic differences between the complete cases and the incomplete cases, and the resulting inference might not be applicable to the population of all cases, especially with a small number of complete cases.

- Some initial evaluation was performed when data cleaning was being conducted:
    - In regards to most of the variables under consideration, none had very much missing data and were not of concern
    - Risk factor data was mostly missing and therefore, excluded from the final model but further analysis evaluating whether individuals with this piece of information missing were different from individuals who didn't was conducted and there were some major differences (aside from demographics which were very similar)

Appendix C: Field Notes on Network and Spatial Analysis

For proper network analysis, cases must be accurately linked to any contacts they may have named while receiving testing and treatment for their index (or second) infection

- Done via case ID records
- Every contact ID was entered in a joining file with the case ID that identified them as a partner (therefore, every case ID that engaged in contact tracing procedures were able to be directly linked to the individuals they named)

Data Preparation:

- As some contacts would have returned as cases, resulting in being assigned a case ID and a contact ID, these individuals had to be identified
    a. These individuals were matched to their respective case ID if they had identical names (as much of the other information for contacts was missing, making it difficult to match on the basis of other information)
- Once an individual was identified as being both a contact and a case, their case ID was changed to match their contact ID to ensure that all possible links to the individual were all accounted for
    a. Names were stripped, leaving only Client ID as the identifying information
- Once this was done, an Excel spreadsheet that outlined the address and postal code of single infection cases, CR cases and contacts was created in order to be read into LinkageWiz. This software was used in order to formally check the work that was done to match cases to their contact ID to reduce the amount of error
- The geographic software GIS was utilized to get more accurate coordinates, these were imported into Excel and matched to the appropriate case using the Case_ID
    a. These coordinates were necessary in order to use SaTScan (Spatial Analysis Freeware) to identify any geographic locations within Brant County that had higher than expected rates of infection and CR than surrounding areas

SaTScan:

1. In order to run the software, in addition to geographic coordinates, population estimates by location of cases was necessary since a Poisson model was chosen
2. To get population estimates, a postal code conversion file was utilized as well as the 2011 census information for Brant and Brantford.
    a. Using the postal code of cases, it was indicated which dissemination area each case was located within
    b. Dissemination areas in Brant and Brantford that did not include any cases were still listed but included “0” to indicate no cases in that given area
    c. There are a total of 230 dissemination areas that make up the whole of Brant County

3. Singly infected, CR cases, and contacts were separated into separate spread sheets in order to determine clustering within groups as well as in an overall sample of infected or not infected

Appendix D: SaTScan Parameterization for Cluster Detection

Time precision: None
Type of analysis: Purely spatial
Probability model: Discrete poisson
Scan for areas with: High rates
MC replications: 999
Advanced analysis features:

-Maximum spatial cluster size: 50%
-Spatial window shape: Circular

Appendix E: Previous Core Area Findings in Brant