uOttawa

L'Université canadienne
Canada's university

FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTOCTORALES

uOttawa

L'Université canadienne
Canada's university

FACULTY OF GRADUATE AND POSDOCTORAL STUDIES

Emily Meadows

AUTEUR DE LA THÈSE / AUTHOR OF THESIS

M.Sc. (Epidemiology)

GRADE / DEGREE

Department of Epidemiology and Community Medicine

FACULTÉ, ÉCOLE, DÉPARTEMENT / FACULTY, SCHOOL, DEPARTMENT

Sexual Networks of Individuals Infected with Sexually Transmitted Infections: Structure and Disease Transmission

TITRE DE LA THÈSE / TITLE OF THESIS

Ann Jolly

DIRECTEUR (DIRECTRICE) DE LA THÈSE / THESIS SUPERVISOR

Rama Nair

CO-DIRECTEUR (CO-DIRECTRICE) DE LA THÈSE / THESIS CO-SUPERVISOR

EXAMINATEURS (EXAMINATRICES) DE LA THÈSE / THESIS EXAMINERS

George Wells

Brenda Wilson

Gary W. Slater

Le Doyen de la Faculté des études supérieures et postdoctorales / Dean of the Faculty of Graduate and Postdoctoral Studies

SEXUAL NETWORKS OF INDIVIDUALS INFECTED WITH SEXUALLY TRANSMITTED INFECTIONS: STRUCTURE AND DISEASE TRANSMISSION

EMILY CHRISTINE MEADOWS

Thesis submitted to the Faculty of Graduate and Postdoctoral Studies in Partial fulfilment of the requirements for the MSc degree in Epidemiology

Epidemiology and Community Medicine
Faculty of Medicine
University of Ottawa

Your file Votre référence
ISBN: 978-0-494-18446-2
Our file Notre référence
ISBN: 978-0-494-18446-2

Acknowledgements:

I would like to thank Dr. Ann Jolly for the many afternoons spent discussing the thesis. Your encouragement and support throughout this challenging period was greatly appreciated. You are truly a wonderful supervisor.
I would like to thank Dr. Rama Nair for his statistical guidance and valuable input.
I would like to thank Dr. Bill Cameron for his guidance and encouragement.
Additionally, I would also like to thank Dr. John Wylie at the University of Manitoba and Cadham Provincial Laboratory for helping me understand the data source and for his overall input in this project.

Abstract:

Increases in the rates of sexually transmitted infections (STIs) suggest that control programs may not be targeting the population responsible for the spread of STIs - core groups. The objective was to examine STI transmission within these groups using both traditional epidemiology and social network analysis. Routine partner notification data, supplemented with more detailed voluntary information, was collected from individuals diagnosed with, or exposed to a STI in Manitoba. Groups of individuals were identified (n=2,508), and their profiles described. Larger groups (size > 15) had more repeat cases, and contacts that were repeatedly named. Three different groups were identified, containing individuals which differed demographically and clinically. This study uniquely identified same-sex partnerships within larger groups. A greater understanding of disease transmission patterns within these groups will clearly aid in the development of targeted education and prevention programs for all STIs.

Abbreviations:

CA	Cluster analysis
CDC	Communicable Disease Control
CPL	Cadham Provincial Laboratory
ICD-9-CM	International Classification of Diseases, Clinical Modification (9th Revision)
MH	Manitoba Health
MSM	Men who have sex with men
PID	Pelvic inflammatory disease
RHA	Regional Health Authority
SNA	Social network analysis
STD	Sexually transmitted disease
STI	Sexually transmitted infection
WSW	Women who have sex with women

Table of Contents:

Appendices:

Figures:

Tables:

1 Introduction

1.1 Disease Manifestation and Burden

Chlamydia trachomatis and *Neisseria gonorrhoeae* are responsible for a sizeable amount of morbidity in humans. Of particular concern are the strains of *C. trachomatis* that are sexually transmitted: serovars L1-L3 cause lymphogranuloma venereum an ulcerative disease with steeply increasing incidence in Canada; and serovars D-K are the primary cause of chlamydia infections of the genital tract and are also responsible for ocular and congenitally acquired infections (1). *N. gonorrhoeae* is also a cause of sexually transmitted, ocular and congenitally acquired infections (2). Both infectious agents preferentially infect columnar epithelial cells causing inflammation, ulceration and epithelial scarring (1;2).

Effective treatments exist for both infections, but greater than 50% of men and 70% of women with chlamydial infection, and greater than 50% of both males and females with gonococcal infection can be asymptomatic, which complicates case finding (3). These asymptomatic carriers have been referred to as "the reservoir for the spread" of both infections, being less likely to be diagnosed and treated (4;5). If symptomatic, the incubation period for *C. trachomatis* is approximately 7 to 21 days or longer, and that for *N. gonorrhoeae* is 2 to 7 days (3;6). Disease manifestations for both infections are similar. In females, infection most often manifests as cervicitis, and less frequently as urethritis, Bartholinitis, and perihepatitis. Genitourinary symptoms include abnormal discharge, burning on urination, lower abdominal pain, abnormal vaginal bleeding, and pain during intercourse. If unrecognized or untreated, infections can lead to complications such as pelvic inflammatory disease (PID), and its sequelae - ectopic pregnancy, tubal factor infertility, and chronic pelvic pain. Estimations indicate that 10% of females who have experienced a single episode of PID will become infertile; this risk more than doubles after three or more episodes (7;8). A previous history of PID increases the risk of ectopic pregnancy, and other adverse pregnancy outcomes (7;8). Specifically, it has been estimated that 5 to 15% of spontaneous abortions, stillbirths, premature labours and congenitally acquired infections are associated with maternal STIs (9).

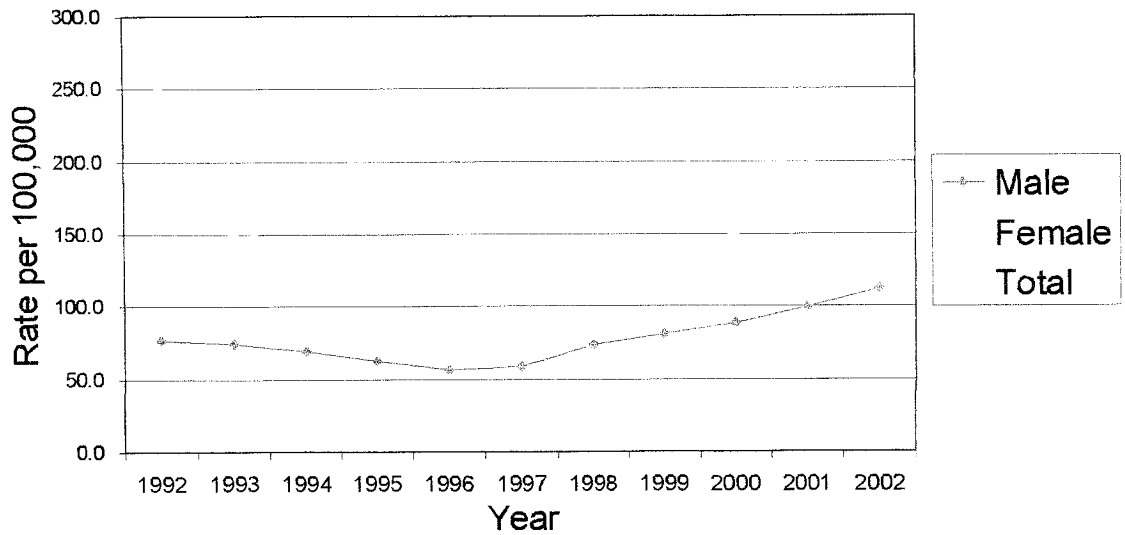
In males, the most common manifestation is urethritis, with symptoms including: urethral discharge, burning on urination and urethral itch. Rare complications in men include epididymitis, and infertility. In both males and females, conjunctival, rectal, pharyngeal and disseminated infections can occur (3;6).

The financial costs accrued from the management of both infections are enormous and are mostly associated with PID, and its sequelae. In Canada, the direct (i.e. hospitalizations, professional billings and antimicrobial therapy) and indirect (lost productivity) costs of chlamydial and gonococcal infections were estimated at between $89 and 123 million, and $54 and 74 million annually (1990 dollars), respectively (10). In addition to these financial costs, bacterial STIs have been reported to increase the risk of both cervical neoplasia and HIV acquisition and transmission (11-14). Specifically, in a recent meta-analyses examining the effects of STIs on HIV susceptibility, combined effect estimates indicated that both chlamydia and gonorrhea increased HIV susceptibility, by 2.2 (95% CI: 1.4, 3.3) and 2.1 (95% CI: 1.7, 2.5) times, respectively (15). Although the authors of the meta-analyses indicated cautious interpretation of the results, one major conclusion was, "STD control programs should be a part of all HIV prevention efforts" (15).

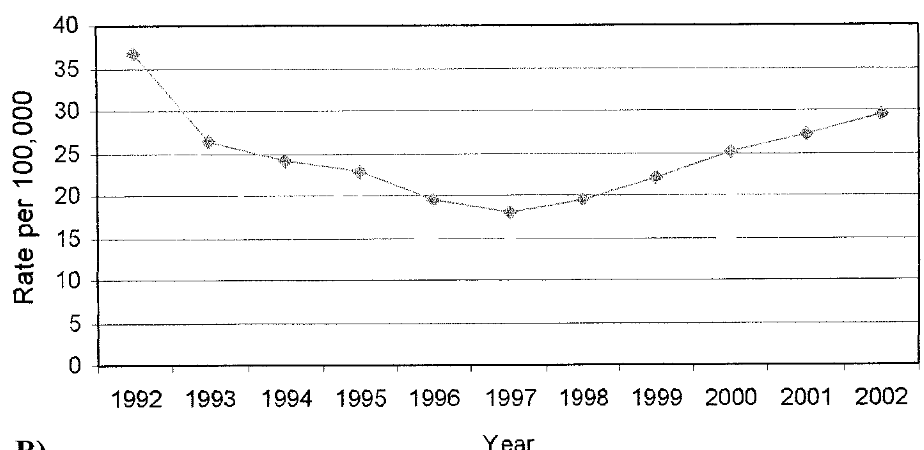
As evidenced from the preceding information, infection with *C. trachomatis* and/or *N. gonorrhoeae* can be devastating to those infected and to society as a whole.

1.2 Distribution and Determinants

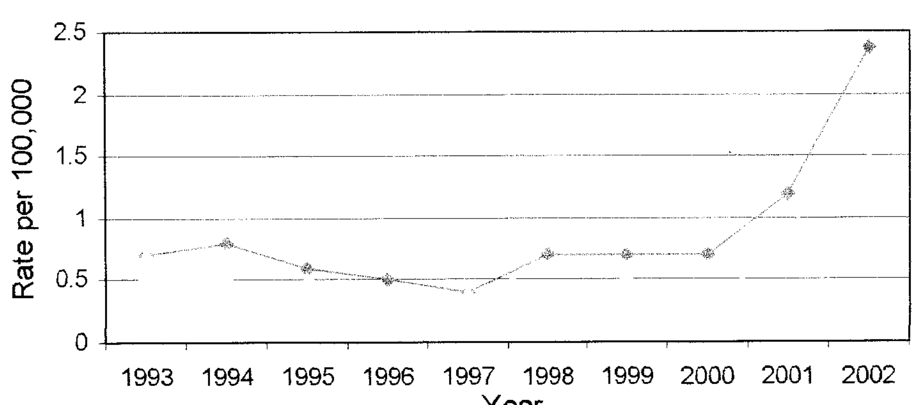
Worldwide, approximately 90 million cases of chlamydia and 62 million cases of gonorrhea are detected annually (16). In attempts to decrease both the incidence and prevalence, and the previously described burdens, population-based control programs were initiated throughout Canada and national goals were established. By the year 2010, Health Canada proposed the rate for chlamydia be less than 50 cases per 100,000, falling below 200 cases per 100,000 in individuals aged 15 to 24. Locally-acquired gonorrhea was to be eliminated at this time (17). Despite major gains, the effectiveness of these control programs has reached a plateau and national goals seem unattainable (18-22). Specifically, since 1997 rates of both infections have increased, refer to Figure 1 (19). Moreover, several provinces have recently experienced outbreaks of locally-acquired syphilis infection. Similar disease patterns have been observed in the province of Manitoba (23).



A)



B)



C)

Figure 1 Incidence rates of chlamydia (A), gonorrhea (B), and syphilis (C) in Canada (21).

These incidence rates are greatly influenced by the disproportionate morbidity in vulnerable populations, including persons of younger age, females, Aboriginal peoples, and other populations.

Young age has been repeatedly found to be a significant risk factor for both chlamydia and gonorrhea infection (24-29). In Manitoba, rates of chlamydia in females (2,468 per 100,000) and males (862 per 100,000) aged 15 to 24 years in 2003 far exceeded provincial rates (316 per 100,000); while the incidence of gonorrhea in persons aged 15 to 24 years was three times greater than the provincial rate (30).

Several factors, both biological and behavioural, have been suggested to increase the risk of STI in persons of younger age (31). Gynecological factors related to physical development have been cited to increase susceptibility to infection, including: the persistence of cervical columnar epithelium; and the composition of vaginal mucus (31-36). Differing sexual behaviours between persons of younger age and adults also contribute to the observed disparity in rates (31).

At even greater risk of infection are street involved youth. (The term street involved refers to individuals with unstable living conditions). In a multi-site study conducted in Canada examining the prevalence of chlamydia infection in street involved youth aged 15 to 24, the authors reported a prevalence rate almost nine-times greater than that reported for the general population (37).

National and provincial surveillance reports for chlamydia indicate the existence of gender disparities in rates. In Manitoba in 2003, 70% (2,546 of 3,640) of all laboratory-confirmed chlamydial cases occurred in females (30). Several risk factors place females at increased risk of acquisition of chlamydia, including cervical ectopy, as a result of oral contraceptive use and young age; and the more efficient transmission of *C. trachomatis* from males to females after one single act of sexual intercourse (38-40). However, this over-representation is most likely due to: a) screening guidelines which predominately target females; b) the greater frequency of contact between females and healthcare professionals; and c) STI treatment and testing patterns in males.

In Canada and the US, higher rates of STIs in Aboriginal populations have been reported (41;42). In Canada in 1999, rates of chlamydia in First Nations people were six times higher than the national rate (43). This national trend has also been observed at a provincial level, in outbreak investigations and through the examination of routine surveillance data (44-49). In Manitoba, 26% of all chlamydia cases and 29% of all gonorrhea cases were among individuals self-identifying as Aboriginal (30).

Although no known biological mechanism exists, it has been suggested that ethnicity may be a risk marker for differences in risk behaviours, and a barrier to healthcare access and impediments to utilization (50). It has also been hypothesized that sexual mixing contributes to the observed disparities in incidence rates. Specifically, Laumann *et al.* (1999) hypothesized that the high rates of STIs observed within African American populations were a result of sexual mixing between high-risk males and low-risk females (51).

Increasing rates of both chlamydia and gonorrhea, combined with their disproportionate occurrence in vulnerable populations inspired the seminal paper by Wasserheit *et al.* (1996). It was in this paper that she expressed the idea that - "STD epidemics evolve through four predictable phases … each phase shaped by interplay among the pathogen, the behaviors of the subpopulations in which it emerges, and the prevention efforts that are developed to limit its impact" (52). She went on to emphasize the need for phase-appropriate STI control programs differentially targeting those subpopulations which maintain and those which spread STIs. Based on the description of these phases, it has been proposed that both chlamydia and gonorrhea are in the low incidence endemic phase (53). In this phase, general population level control programs, effective at reducing incidence and prevalence in the general population, are ineffective at reaching the section of the population responsible for the spread of STIs, known as *core groups* (52).

1.3 Core Groups

The core group concept emerged in the 1970s, when Yorke *et al.* (1978) theorized that a small group of sexually connected individuals, with high contact rates and disease occurrence were responsible for the endemicity of gonorrhea (54). It was further posited that removal of infection from this group would lead to the eradication of disease in the general population.

The core group theory is based on the concept of the basic reproductive number, R_0, a measure of the number of infections produced, on average, by an infected individual in the early stages of an epidemic, when he or she enters a totally susceptible population (55). R_0, is the product of three parameters- the probability of transmission from one infected individual to a susceptible partner (ß); the sexual contact rate between infected and susceptible individuals (c); and the duration of infectivity (D) (55). For an organism to achieve ecological success R_0 needs to be equal to one with a probability of one, as this is a stochastic process (55).

Brunham (1997), calculated the critical threshold of sexual partner change (c) required by various STIs to maintain ecological success (56;57). In an environment with no control activities, in order for *C. trachomatis* to maintain endemicity, c was estimated to be 4 new partners per year (D=1.25, ß=0.2). In the presence of control activities, *N. gonorrhoeae* required 13 new sexual partners a year (D=0.15, ß=0.5). When compared to the sexual behaviours exhibited in the general population, it was evident that these rates are found in only a small section of the population (57-60).

The first empirical evidence to support the existence of core groups, and their disproportionate role in the spread of STIs was documented in 1983 (61). Rothenberg used gonorrhea case reports obtained from 1975 to 1980 to estimate incidence rates for census tracts in upstate New York. The core group was defined as census tracts containing 50% of reported cases. The identified core group consisted of 13% of all census tracts, containing only 9% of the total population. This group was characterized by high incidence rates (1,822 per 100,000) in comparison to non-core areas (83 to 792 per 100,000); and census tracts were typically inner city areas with high population densities and low socioeconomic status. Rothenberg suggested that the characterization of core areas would allow for the most cost-effective control efforts.

Potterat *et al.* (1985) examined various social factors contributing to the spread of gonorrhea in Colorado Springs, Colorado (62). Sexual contacts of laboratory-confirmed cases of gonorrhea, identified within a six-month period, were elicited and "lots" - individuals connected by sexual contact - were constructed. The largest "lot" consisted of 226 individuals and the incidence of gonorrhea was quite high (1,550 cases per 1,000 persons per year). The authors also documented that six "lots" accounted for 20% of gonorrhea cases during the study period.

In 1996, Potterat *et al.* demonstrated, once again, the existence of core groups, and their role in STI transmission while investigating an outbreak of penicillinase-producing

Neisseria gonorrhea in street gangs involved in crack cocaine trade in Colorado Springs (1990-1991) (63). A "lot" consisting of 410 sexually connected men and women, accounted for 22% of all reported cases of gonorrhea, while merely representing 0.1% of the 18 to 44 year old population in this area.

In 1999, Potterat *et al.* used partner notification data for chlamydia cases in Colorado Springs (mid-1996 to mid-1997) to examine factors, both epidemiologic and network, affecting disease transmission (64). Sexual network data was combined with clinical and exposure information enabling the estimation of both individual and group R_0. Although the overall R_0 for all interviewed cases was less than one (0.55) a few subgroups were found to have a R_0 which exceeded unity: males (1.36) and females (1.18) with concurrent partnerships, occurring when sexual relationships overlap in time; co-infected females (1.05); and males with no identified source of infection (1.26). Multivariate analysis revealed that interviewed cases with concurrent partnerships were 3 times more likely to have an R_0 greater than one. In comparison to other key variables associated with STI transmission (number of contacts, repeater status, the absence of symptoms (males) and mixing patterns) concurrency was determined to be the most significant predictor of transmitter status.

However, as core group characteristics differ depending on the population studied, it is invaluable to characterize these populations in order to better understand disease transmission.

1.3.1 Core Groups in Manitoba

In light of the many seminal articles detailing the theoretical importance of core groups, and the recent empirical characterizations - researchers in Manitoba have deemed the latter of great importance.

Blanchard *et al.* (1998) examined the epidemiology of both chlamydia and gonorrhea infections by geographically mapping incidence rates derived from surveillance reports (1991 to 1995) for postal areas in Winnipeg (65). Core postal areas were empirically defined by plotting incidence rates and examining distributions for natural breakpoints (those areas with incidence rate greater than or equal to 887 per 100,000 and 443 per 100,000 for chlamydia and gonorrhea, respectively). In comparison to non-core areas, core areas contained a greater proportion of individuals with repeat infection and co-infection; and had a higher population density, rate of unemployment, and lower average income.

Although the identification of core areas through geographic mapping provides important understanding with respect to the epidemiology of STIs – Wylie and Jolly (2001) suggested that it does not address the connectedness or inherent bridging which occurs between "core" and "non-core" areas (66). This is further emphasized when examining the sexual mixing patterns among chlamydia and gonorrhea cases and their sexual contacts in Winnipeg in 1995 (65). Specifically, Blanchard *et al.* (1998) reported that 15% (95% CI: 11, 19) of all chlamydia cases residing in a "non-core" area of Winnipeg named at least one sexual partner residing in a "core" area. A similar mixing pattern was observed among "non-core" gonorrhea cases and their nominated sexual contacts. Specifically, 20% (95% CI: 12, 31) of all "non-core" gonorrhea cases nominated at least one contact residing in a "core" area.

Jolly *et al.* (2005) attempted to characterize and define the reproductive rate of four clinical groups hypothesized to contain core group members using routinely collected partner notification data (1990 to 1992) (67). The four hypothesized groups, included:

chlamydia, gonorrhea, co-infected repeaters, and repeatedly nominated individuals with no laboratory-confirmed infection. A repeater is an individual with more than one disease episode or nomination during the study period. In comparison to their single episode counterparts, those with repeat infections differed significantly with respect to sociodemographic and clinical characteristics. Although the number of reported partners per disease episode did not differ between repeaters and non-repeater cases, the estimated R_0 values of all repeater groups exceeded unity; while those of non-repeaters were less than unity. Another unique finding was that multiply nominated contacts did not differ significantly from those repeatedly infected, with respect to sociodemographic and clinical variables- emphasizing their role in the spread of infection.

Of particular interest, was the difference within repeater groups (male vs. female), as seen in their estimated R_0 values, 2.61 and 0.71 for co-infected female and male repeaters, respectively. This finding indicates that each group contributes differently to STI endemicity and that there is no absolute core group. More specifically, the underlying network structure highly influences transmission and STI endemicity (67). Recently, the characterization of these interconnected *networks* has garnered much attention.

1.4 Emergence of New Paradigm

Strong empirical evidence in support of core groups in STI epidemiology exists. This has resulted in a shift in STI epidemiology from the traditional focus on individual characteristics to one concentrating on these interconnected networks (52). Although previous studies focused on the characteristics and behaviours of an individual and their risk of acquiring a STI have contributed significantly to the understanding of STI epidemiology, they do not take into consideration that STI risk occurs at multiple levels.

For example, Laumann and Youm (1999) attempted to explain the disproportionate burden of STIs in African Americans (51). Although controlling for individual risk factors associated with STIs, multivariate analysis indicated that this disparity still existed. African Americans were five times as likely to acquire a STI in comparison to other ethnic groups. When network level characteristics were taken into consideration (i.e. partner selection) a clearer explanation emerged. Specifically, the authors found that African American women, perceived to be at low risk of infection as they had a lower number of partners, were more likely to select African American men with a high number of partners. White women were more likely to select partners having a similar number of sex partners. This in combination with the small amount of partnerships occurring across ethnic groups was hypothesized to account for the disparity in prevalence (51).

The shift from the examination of an individual's risk profile to that of their risk environment is captured in the following statement: "the significance of an individual's risk behaviours are increased or decreased depending on the infection status and behaviours of the people with whom they typically interact" (66;68;69).

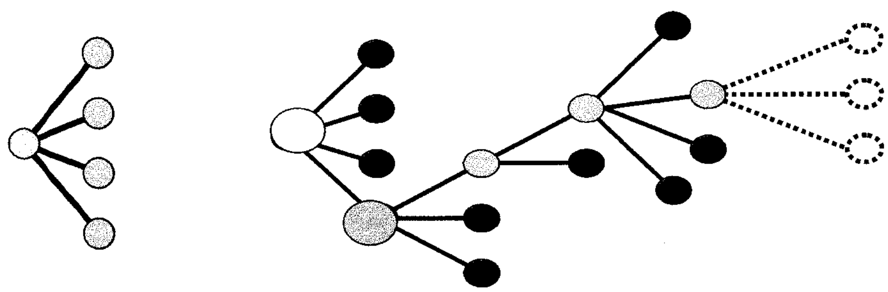
1.5 Introducing Social Network Analysis

1.5.1 Sexually Transmitted Infections and Social Network Analysis

Recently, the use of social network analysis (SNA) has been used as a key tool in understanding the transmission of STIs (51;63;70-75). The basic proposition of SNA being, "the structure of a network has consequences for its individual members and for the network as a whole over and above effects of characteristics and behaviours of the individuals involved" (71). As explained by Laumann *et al.* (2004)—there is a "social component to sex" (76). The selection of sexual partners is not a random process. As such the formation of a sexual relationship is influenced by the social norms and behaviours in which the relationship is embedded.

In 1997, Ghani *et al.* modeled the spread of gonorrhea in heterosexual networks, in order to determine if network structure influenced the establishment and endemic prevalence of gonorrhea (77). The authors concluded that the inclusion of network properties aided in the understanding of STI epidemiology. Specifically, the establishment of infection was directly influenced by the proportion of non-monogamous relations, connectedness of high activity individuals, and the mean component size. (A component is group of individuals directly or indirectly connected through sexual contact).

Key network terminology is detailed in **Appendix A**. A network consists of a set of nodes (vertices, points, actors) connected by a set of lines (edge, ties, links) (78;79). Nodes represent individuals, events or groups, and lines represent relationships, attitudes, affiliations, etc. Two different types of network data can be collected and analyzed: 1) ego-centric (Figure 2A); and 2) sociometric networks (Figure 2B). An ego-centric network consists of one principal node (ego) and the set of nodes he or she is directly connected to (alters). Sociometric networks collect further information on the relations of these alters (**Figure 2**). The binary relations between nodes (dyadic relationships) can be represented in a matrix or a graph, which can be analyzed by matrix algebra (78). Figure 3A shows the most common representation of a social network—the adjacency matrix. An adjacency matrix is a square matrix in which the rows and columns represent the nodes, and the cells indicate if a relation is present or absent, 1 or 0, respectively. The information contained within the adjacency matrix can be visualized as a graph in which all nodes are represented as a dot and the presence or absence of relations is indicated by a line connecting the respective nodes (Figure 3B).



A) **B)**

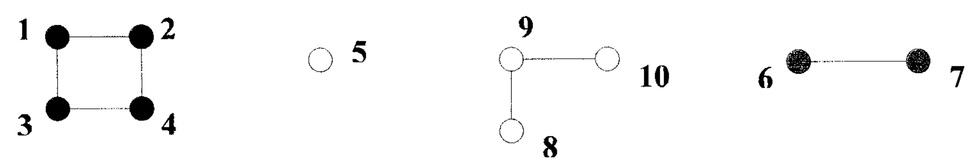
Figure 2: Diagrammatic representation of an ego-centric (A) and sociometric network (B). In section A, the green node represents the ego and the blue nodes represents the ego's alters; while in section B, red indicates nodes which do not have any other relations (degree of 1 or a termination point). The ego-centric network is embedded in the sociometric network depicted in B.

As previously mentioned, a network is a set of **all** nodes and lines representing a group of individuals; a component consists of a group of individuals within the network, which are connected either directly or indirectly (80). For example, in Figure 3, there are four components contained within the overall network of size 10. Components are of size 1 (isolate), 2 (dyad), 3 (triad), and 4, and are indicated by the colours of the nodes, green, blue, yellow and red, respectively. In a sexual network the lines represent sexual relationships (sexual dyads), i.e., individual 1 had sex with individuals 2 and 3.

The degree centrality of a node refers the total number of nodes directly connected to it, for example, in Figure 3B, individual 1 has a degree of 2 (directly connected or adjacent to individuals 2 and 3); whereas individual 6 has a degree of 1. Degree centrality is a measure of a node's local importance or prominence in a network or component, and a prominent node is one that has many direct relations with other nodes, refer to Figure 4.

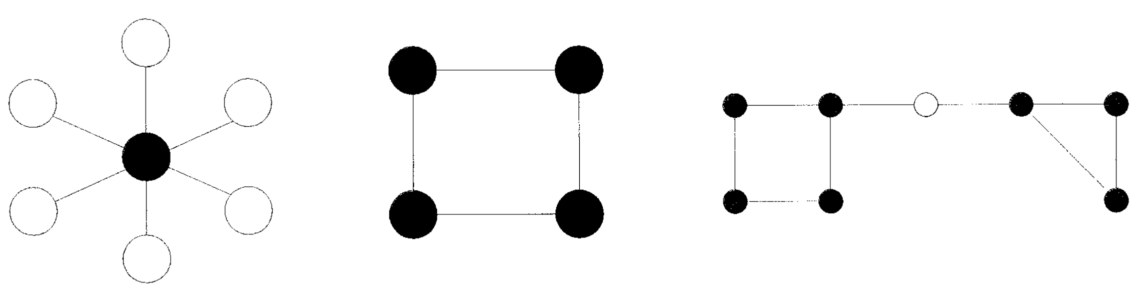
	1	2	3	4	5	6	7	8	9	10
1	-	1	1	0	0	0	0	0	0	0
2	1	-	0	1	0	0	0	0	0	0
3	1	0	-	1	0	0	0	0	0	0
4	0	1	1	-	0	0	0	0	0	0
5	0	0	0	0	-	0	0	0	0	0
6	0	0	0	0	0	-	1	0	0	0
7	0	0	0	0	0	1	-	0	0	0
8	0	0	0	0	0	0	0	-	1	0
9	0	0	0	0	0	0	0	1	-	1
10	0	0	0	0	0	0	0	0	1	-

A)



B)

Figure 3 Adjacency matrix of network consisting of 10 members (A) and the corresponding graph diagram (B). Lines represent the presence of relations and circles represent nodes (unique identification numbers are provided for each node, 1 to 10). Colours indicate identified components.



A) **B)** **C)**

Figure 4 A) The red node is the most central individual, with a degree centrality equal to 6. This is a typical star or radial component which has a degree centralization of 1. Section B, has a degree centralization of 0; the turquoise node is a bridge (Section C).

The density of a network refers to the proportion of relations which exist out of all those which could exist in a network, and ranges from 0 to 1. It is quite rare for a network to have a density equal to 1.

The corresponding formula for density is:

$$\text{Density} = \frac{l}{n(n-1)/2},$$

where l is the number of relations or lines which exist in the network, and n represents the number of nodes in the network. Density is a measure of the connectedness of a network or component; theoretically, denser components facilitate transmission. The density of the network consisting of 10 nodes in **Figure 3** is 0.15.

Degree centralization measures the heterogeneity of degree centralities within a network and measures "the extent to which a single actor has a high centrality" (81).

$$\text{Degree centralization} = \frac{\sum_{i=1}^{g}\left[C_D(n*)-C_D(n_i)\right]}{\left[(g-1)(g-2)\right]}$$

where, $C_D(n_i)$ are the g actor or node degree indices, while $C_D(n*)$ is the largest observed value, with values between 0 and 1 (81). As observed in Figure 4A, a star network has a degree centralization of 1. The extreme (degree centralization of 0) is observed for the network depicted in Figure 4B.

Within a network, one can observe nested subgroups which are defined by specific characteristics. These subgroups are referred to as microstructures - which increase both the density and the potential for disease spread. A k-core is defined as a subgroup in which the members are adjacent or directly connected to a least k other nodes in that same subgroup (78). For example, a 2-core of size 4 is depicted in Figure 5.

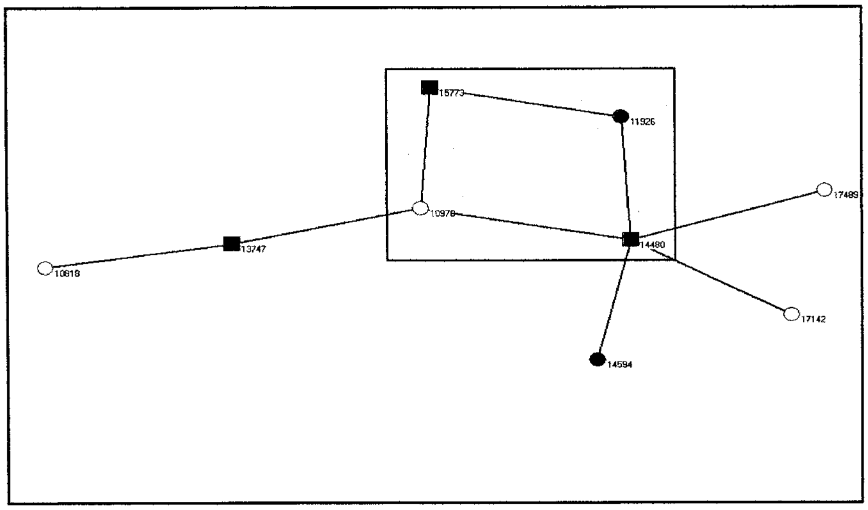


Figure 5 The four nodes contained within the highlighted yellow box are members of a 2-core of size 4.

A bridge or cut-point refers to a link or a node, which connects two subcomponents. One example of a bridge is depicted in Figure 4C. Bridges may connect two subpopulations with different disease rates, risk behaviours, or geographic locations, presenting potential points of intervention.

The application of SNA in infectious disease epidemiology is quite new (71). Klovdahl used it to evaluate whether an infectious agent could be responsible for disease in a network of 40 AIDS patients connected through sexual contact. Based on the estimated number of men who had sex with men (MSM) in the US in 1980 and the highest rate of sexual contact reported by an MSM within this network, the observed pattern of relationships was not likely due to chance. This finding supported the hypothesis that the etiological agent was an infectious disease spread through sexual contact. Since then, studies have indicated that mixing patterns, network structure, and temporal changes all influence the transmission of STIs within a population.

In 1988, Potterat *et al.* commenced the first longitudinal network study examining the influence of network structure on HIV transmission in a high-risk heterosexual population in Colorado Springs (82).

Klovdahl *et al.* (1994) assessed the social network of individuals at high risk for HIV infection: sex trade workers, injection drug users and their personal contacts, in Colorado Springs (72). A component of over 600 individuals was identified. A highly connected region, referred to as the 'core', was embedded within this component and consisted of 106 individuals. The 'core' contained one of three identified HIV seropositive individuals (year one of study). All seronegative individuals, within the 'core', were within six steps of this individual, and on average three steps away. Although the prevalence of HIV suggested minimal risk of HIV infection, the authors indicated that slight behavioural changes could result in the explosive spread of HIV (72).

Furthermore, a network consisting of 5,162 individuals was identified (1988 to 1991); containing 147 components, the largest consisting of 3,658 members (83). Although high-risk behaviours were commonly reported (injection drug use, sharing needles, high number of sexual partners, and low usage of condoms), HIV incidence was quite low, with only 19 HIV-positive individuals identified. SNA aided in the understanding of the observed pattern of disease transmission - revealing that only 11 of the 19 HIV-positive individuals were members of the largest component and the majority of the 11 did not engage in high-risk behaviours (sexual and sharing) (83). The remaining HIV-positive individuals were located in small, isolated components - minimizing the spread of disease.

In 1998, Rothenberg *et al.* used SNA to examine an outbreak of syphilis in a suburban county in the Atlanta metropolitan area (74). The network consisted of 99 individuals (10 cases) linked through sexual contact. Two components were identified, the largest consisting of 95 individuals which contained a k-core (k=8) of 19 individuals (refer to **Figure 5** for explanation). Temporal analysis of network formation during the course of the outbreak highlighted the importance of microstructures as indicators of increased sexual activity and increased syphilis transmission.

Although core groups are essential for the maintenance of STIs within a population, mixing between sub-populations is required for the spread of STIs. As such, the characterization of mixing patterns, "who mixes with whom" with respect to age, ethnicity, sexual activity class, and spatial location; and their influence on the spread of infection has garnered recent attention (84). Mixing patterns vary from assortative (like with like) to disassortative (low level of like with like) (84). Anderson *et al.* (1990) stated that "heterogeneity in behaviour is one of the most important factors, as measured by its influence on projected trends on infection and disease" (84). The authors modeled the influence of mixing patterns, with respect to sexual activity class, on the transmission dynamics of HIV (84). Results indicated that disassortative mixing patterns resulted in a low initial rise in incidence in comparison to the rapid increase observed as a result of assortative mixing, however the former resulted in a larger overall epidemic (84).

Since then, several authors have demonstrated the impact of mixing patterns on disease transmission and on individual risk of infection (73;75;85). Aral *et al.* (1999) examined mixing patterns, with respect to age, education, ethnicity and sexual activity, and their influence on the risk of chlamydia or gonorrhea infection in a population of heterosexual clinic attendees (85). For the most part, sexual partnerships exhibited an assortative mixing pattern. However, individuals in discordant partnerships were at a higher risk of chlamydia or gonorrhea infection- risk being greater for societal level factors (age, education, and ethnicity). Similarly Service *et al.* (1995) examined whether mixing within sexual networks of homosexual men was a risk factor for acquiring HIV (75). Young homosexual men (18 to 29 years of age) who partner with homosexual men over the age of 30 were found to have an increased risk of acquiring HIV, in comparison to young men who partner with individuals of the same age.

Since its first application in 1985, social network analysis has been an invaluable tool used to further our understanding of the transmission and acquisition of HIV and STIs.

1.5.2 SNA and Manitoba

In 1997, Wylie and Jolly used SNA to study transmission patterns of chlamydia and gonorrhea in the province of Manitoba, using routine STI surveillance data collected from the Communicable Disease Control (CDC) Unit of Manitoba Health (MH) (46;53;66;86;87). They identified 1503 components, ranging in size from 2 to 82 individuals differing on the basis of observed mixing patterns between individuals (with respect to number of partners). Linear components exhibited highly assortative mixing while radial components exhibited disassortative mixing (**Appendix B**). More significantly, linear networks had a higher composition of aboriginals, higher positivity rates, and had both infectious agents present in approximately 60% of the networks (whereas chlamydia was the sole STI circulating in radial networks). The authors hypothesized that these linear networks resembled the theoretical definition of a core group defined by Brunham in 1997. Specifically, "core groups are defined by the sexual behaviours of their members and the connectivity among members" (57). Of other significance was the discovery that sexual networks were not localized but spanned several communities throughout the province.

In a comparison of sexual networks from Winnipeg and Colorado Springs, structures of similar size and composition were found to exist – a step towards demonstrating the generalizability of network analysis (46). Additionally, the concordance between chlamydia genotypes and sexual networks constructed during the 1997 study was examined (86). The analysis revealed a high degree of concordance, indicating that networks obtained from partner notification data, in most cases, trace actual transmission routes. Although some of the networks spanned many distinct and geographically separated communities - 59% of these networks for which specimens were available were concordant.

1.6 Project Conception

As previously discussed (**Section 1.5.2**) SNA was used to study transmission patterns of chlamydia and gonorrhea in Manitoba in 1997. Basic demographic information was used to characterize these networks, and significant differences with respect to several demographic, geographic, and clinical variables were discovered.

As this characterization of sexual networks was limited to demographic data maintained within a pre-existing database; it was necessary to further characterize the social forces behind the formation of these interconnected groups by examining the social and sexual relationships of individuals diagnosed with STIs and their sexual contacts.

This thesis project is part of a larger CIHR funded research project, "The Manitoba Sexually Transmitted Disease Network Study", in which all chlamydia and gonorrhea cases identified between June-2002 and October-2003 were asked to participate in a questionnaire-based study. The purpose of this project was to collect more detailed sociodemographic, healthcare and behavioural information about cases and their relationships with sexual contacts, absent in the previous retrospective study. This project concentrated on examining the characteristics of this population, identifying and describing the sexual network, and analyzing the differences between components.

1.7 Potential Contribution and Unique Scope

In a time when public health funding is scarce, a greater understanding of disease transmission patterns within different subpopulations will allow for more precisely targeted, population-specific, and effective education and prevention programs.
With Cadham Provincial Laboratory (CPL), performing approximately 95% of the province's STI testing, using the same techniques for nearly all samples collected, and one centralized reporting facility, this province provides a great opportunity to do this. Although Manitoba reports some of the highest rates of chlamydia and gonorrhea, it is possible that similar networks exist in both Alberta, and Saskatchewan – making the results applicable to STI control not only in Manitoba, but in other prairie provinces with similar demography (44;88).

1.8 Objectives

1. To obtain a detailed epidemiological overview of all participants in the sexual network identified from routine surveillance data collected by the Communicable Disease Control Unit of Manitoba Health from June-2002 to October-2003.
2. To construct components of individuals linked directly or indirectly through sex, and describe basic social network measures.
3. To define groups or clusters of components using cluster analysis, and analyze the differences in sociodemographic characteristics, and responses to social and sexual behavioural questions between the identified groups.

1.9 Hypothesis

It is hypothesized that social and sexual behaviours will differ between components, when subdividing components by size, ethnic composition, or infecting pathogen. Specifically, components with high aboriginal content will display high-risk relationship variables, such as, concurrency, and high rate of partner change.

2 Data Sources

In the province of Manitoba, all positive laboratory reports for chlamydia and gonorrhea are received at the Communicable Disease Control (CDC) Unit of Manitoba Health (MH). A public health nurse then contacts the attending physician, who is required to fill out a sexually transmitted infection (STI) notification form which is forwarded to the CDC and entered into a centralized, computerized STI database. Next, cases are contacted and interviewed in order to elicit locating information of sexual contacts. This information is reported and entered into the contact database. Data from July-2002 to October-2003 was extracted from the CDC STI database for this study.

Additionally, more detailed questions on symptoms, sexual behaviours and proxy information on each client's sex partners were asked of all those cases that consented from five regions representing urban and rural Manitoba, from June-2002 to October-2003 (**Appendix C**). Public health nurses administered the questionnaire to consenting cases 13 years of age or older after contacting them for partner notification procedures.

Prior to acquisition of all data, personnel at the CDC completed the following data processing: de-duplication (the identification of individuals appearing multiple times as cases and/or contacts); assignment of unique identifiers; and removal of identifying information. Upon receipt, all data sources were non-nominal (i.e. databases contained no names). De-duplication was completed using an algorithm similar to that used in the previous study, which involved first and last names, dates of birth, and addresses (66).

Each database, described in the following text, contained essential but different information.

2.1 Manitoba Health (MH) STI Dataset

All individuals legally entitled to be in Canada, who have established permanent residence and live in the province for a minimum of six months per year, are eligible to register for the Manitoba Health Services Insurance Plan, which provides essential physician and hospital (in-patient and out-patient) services to registrants (89). The Manitoba Health Services Insurance Plan requires no payment of premiums and as such, a high proportion of those eligible are registered (90).

STI prevention and control is a collaborative effort between the CDC and the eleven Regional Health Authorities (RHAs) and other healthcare jurisdictions in the province (**Appendix C**). The CDC develops policies, standards and guidelines, which are implemented by each RHA. Regional healthcare jurisdiction is determined on the basis of postal codes for First Nations and municipality for others (91). STI prevention and control is a core healthcare service accessible to all provincial health registrants (92). Screening, diagnosis, management and treatment of STIs is completed by physicians and nurses in clinics and hospitals, and by nurses employed by Health Canada who deliver healthcare services to First Nations Reserves. There are no designated STI clinics in Manitoba (93).

2.1.1 MH STI Case Registry

2.1.1.1 Background, Acquisition, and Description

The Public Health Act's Diseases and Dead Bodies Regulation, legislated by MH, requires that both primary health professionals and laboratories report all laboratory-confirmed cases of chlamydia and gonorrhea to the Director of the CDC at MH within 48 hours (94). MH's CDC Unit defines a confirmed case of chlamydia as "the detection of *C. trachomatis* from any site by culture, antigen detection, DNA probe technique, nucleic acid amplification or fluorescent antibody", and confirmation of gonorrhea case requires "the detection of *N. gonorrhoeae* from any site by culture, antigen detection, DNA probe technique or nucleic amplification" or the presence of intracellular gram-negative diplococci on gram-stains of urethral exudates from males (94). Gonorrhea has been notifiable in the province for many years, and chlamydia was made notifiable in 1987. During the study period (June 1, 2002 to September 30, 2003), laboratory testing for *C. trachomatis* and *N. gonorrhoeae* was recommended for: individuals presenting with chlamydial and gonococcal syndromes, which included urethritis (males), cervicitis (females), and proctitis; individuals who had sexual contact with a laboratory-confirmed case; or individuals who have been sexually assaulted or abused (94). As many chlamydial and gonococcal infections are asymptomatic, in 2001 focused screening was recommended in the following circumstances: 1) testing at least once during pregnancy; 2) annual testing of all sexually active individuals under the age of 25 during presentation to a healthcare professional (chlamydia only); or 3) individuals (independent of age) with one or more of the following risks, i) prior to the insertion of an intrauterine device; ii) prior to an abortion; iii) individuals with more than one sex partner in previous year; iv) individuals with a new sex partner in past 2- to 12-months, for chlamydia and gonorrhea, respectively; v) persons whose partner has had other sex partners; vi) street involved individuals; vi) individuals involved in substance misuse; viii) individuals with a history of STI in the previous year; and/or ix) history of unprotected sex with an individual in one of the above categories (94).

During the study period, it was recommended that individuals presenting to a primary healthcare provider with suspected chlamydia or gonorrhea be treated presumptively with the appropriate antimicrobial therapy; and those identified through focused screening be immediately located and treated (94). Preferred treatment for gonorrhea was a single dose of cefixime and for chlamydia, a single dose of azithromycin, both administered orally. Cadham Provincial Laboratory (CPL) is the central microbiological testing facility performing the majority of chlamydia and gonorrhea testing. The principle laboratory diagnostic test for *C. trachomatis* and *N. gonorrhoeae* at CPL during the study period was the Gen-Probe® PACE®2C DNA probe test (95-98). PACE®2 enables the detection of both *C. trachomatis* and *N. gonorrhoeae* in specimens collected from endocervical, male urethral and conjunctival swabs. For specimens collected from pharyngeal or rectal swabs, *C. trachomatis* detection was completed using MicroTrak® direct fluorescent antibody test (99). For *N. gonorrhoeae* detection, culture was used for pharyngeal and rectal specimens. Alternatively, the Becton Dickinson BDProbeTec ™ ET System was used to detect *C. trachomatis* or *N. gonorrhoeae* in urine specimens (100). This amplification technique was used in locations where swab testing was not ideal or for individuals declining swabs. Additionally, in some instances, culture was used to detect *N. gonorrhoeae* from endocervical, male urethral and conjunctival specimens.

After the reporting of confirmed-cases of chlamydia and gonorrhea, clerks then contact the RHA in the client's area of residence, who assigns a public health nurse to the case. The public health nurse then contacts the attending physician, who is required to fill out a *Notification of Sexually Transmitted Diseases* form (**Appendix E**), which is forwarded to the CDC. Once received, the surveillance staff determines if it is a continuing or new infection. All paper files for the current year in addition to the previous two years are kept. If on file, the client will be given the same unique identifier and assigned a sequence number. If no file is found, a new unique identifier will be generated. The information contained in the notification form is then entered into the *MH STI Case Registry*, which is a centralized, computerized database with the dual purposes of 1) providing information for public health nurses and healthcare professionals for the management of cases, and 2) STI surveillance (93).

Upon acquisition, the *Registry* contained case-specific data which consisted of one unique record for each reported new disease episode of confirmed *C. trachomatis*, *N. gonorrhoeae* and *T. pallidum*, from July 1, 2002 to September 30, 2003 (n=5,078). Each record included: a unique record number, demographic (province of residence, 3-digit Forward Sortation Area, RHA, marital status, number of sexual contacts, treaty code), and laboratory information (specimen date, report date, diagnostic codes (ICD-9-CM), sequence number).

2.1.2 MH STI Contact Database

2.1.2.1 Background, Acquisition, and Description

Partner notification is the method by which sexual partners of confirmed-cases are informed of their possible exposure to an infectious organism. Upon report of a confirmed case of chlamydia and/or gonorrhea, the index case is located and interviewed in order to identify and notify exposed partners. Priority for interviews is given to individuals with repeated infections, women with PID and other individuals determined to be at high-risk. It is recommended that all interviews be completed within five working days (94). The names of sexual partners "exposed 2-months prior to the onset of symptoms, up to and including the interview date" are collected by the public health nurse (94). The names of sexual partners exposed "3-months prior to diagnosis" of an asymptomatic laboratory-confirmed case are collected by the public health nurse (94).

Partner notification is conducted using one of 3 methods: 1) provider referral; 2) patient referral; or 3) "contracting". Provider referral occurs when either the attending healthcare professional or an assigned public health nurse interviews and completes partner notification. Patient referral occurs when the index case is given the sole responsibility of partner notification, so no names or locating information of sex partners are gathered by the healthcare provider. "Contracting" occurs when the names and locating information of sexual partners are obtained from the index case. The index case is given the initial responsibility of contacting partners, and if partners have not sought treatment after a certain time period, the attending healthcare provider assumes responsibility for the notification process (93). Results of the partner notification investigation are forwarded to the CDC and entered into a centralized, computerized *STI Contact Database.*

Upon acquisition, the *STI Contact Database* contained a unique record for each notification event of a sexual partner exposed to a STI entered in the database from October 1, 2002 to September 30, 2003 (n=5,236). To minimize the inclusion of partners with no linked index case, due to the delay in naming and identifying of partners, different extraction periods were used for the *STI Contact Database* in comparison to the *Registry* (66). All records contained a unique identifier, demographic (3-digit Forward Sortation Area, RHA, treaty code); infection (dates of exposure, infection(s) to which the contact was exposed), and contact management information.

2.1.3 Data Management

All databases in the *MH STI Dataset*, were directly imported into SAS, and data cleaning was completed (101). Data cleaning consisted of: examining univariate descriptive statistics; determining the amount and distribution of missing data; and identifying outliers. In the *Registry*, two records had inaccurate entries for the variable sequence number (sequence number = '00'). Both were confirmed to be data entry errors and were later recoded. Five records did not contain a province of residence; one record was missing a report date and two records were missing a specimen collection date. No changes were made to these records at this time.

2.2 Social and Sexual Behaviours within STI Networks in Manitoba

In addition to the routine surveillance data collected by the CDC, a more detailed questionnaire collecting information on symptoms, healthcare behaviours, risk behaviours (social and sexual), interrelationships, and proxy information on sexual partners was administered. The questionnaire was administered to a convenience sample of consenting index cases in five of the eleven RHAs (post July, 2002), representing both urban and rural Manitoba (**Appendix C**). Participating RHAs included Assiniboine (Marquette and South Westman amalgamated July 1, 2002 to form Assiniboine RHA), Brandon, Burntwood and Norman (rural), and Winnipeg (urban), servicing a population of 70,183; 47,677; 44,770; 25,010; and 656,339 respectively as of June 1, 2002 (91). The questionnaire was pilot-tested on ten consenting laboratory-confirmed STI cases. Administration of the pilot-test was completed by public health nurses in the manner intended for the full study; and changes were made based on the responses of the participants. The final questionnaire is found in **Appendix F**.

2.2.1 Study Case Database

2.2.1.1 Background, Acquisition, and Description

After the receipt of a positive chlamydia and/or gonorrhea laboratory report, a blank questionnaire was placed in the file of the index case. Only individuals 13 years of age or older who were not suspected sexual assault or abuse victims were eligible to participate.

All participants were informed of the study objective, and verbal consent was obtained. The questionnaire was administered by public health nurses employed by the RHAs, following routine partner notification procedures. No monetary compensation was given. In the Winnipeg RHA, the blank questionnaires of individuals who did not consent were kept in order to keep count of refusals. The *Rural* questionnaire was implemented from June 1, 2002 to March 31, 2003 and the *Urban* questionnaire was implemented from October 1, 2002 to September 30, 2003. A slightly different questionnaire was administered to the index cases in the rural jurisdictions. Specifically, two questions were added: 1) "If you live outside of an urban area (Winnipeg, Brandon, The Pas, Thompson), where would you prefer to be tested and treated for your STD infection?", and 2) "At the time you were having a sexual relationship with this partner, did he or she live in a different community than the one you lived in?". In a previous assessment of the concordance between genotype data for *C. trachomatis* and sexual network data, results indicated that partner notification efforts within small communities, in some instances, were not actually tracing transmission routes (86). Therefore a non-nominal approach to gathering information on sexual contacts was undertaken in some rural RHAs. Independent of RHA, all participants had the option of remaining anonymous at the end of questionnaire administration.

This *Study Case* database served as a source of data for this analysis. Upon acquisition, the *Study Case* database contained case-specific information which included a record of the supplemental questionnaire by consenting *C. trachomatis* and/or *N. gonorrhoea* cases (n=327).

The original questionnaire was designed to collect nominal data (first and last names) of study participants and their nominated contacts. It was thought that the collection of nominal data would enable the determination of an individual's gender through subsequent linkage to routine surveillance data. If an individual's gender was missing from routine surveillance data, an individual's first name could be used to determine gender. As such, the original questionnaire did not ask the participant to specify their gender or that of their nominated sexual contact(s). This design did not take into account non-nominal data collection for some participants and resulted in the "unknown" gender category.

2.2.2 Study Contact Database

2.2.2.1 Background, Acquisition and Description

In addition to collecting case-specific information (**described in Section 2.2.1**), the supplementary questionnaire obtained details regarding the relationship between confirmed cases and their sexual partners (**Appendix F**). Proxy information was collected on sex partner(s) with whom the case had sexual contact (anal, vaginal, oral) in the previous 12-months. This time frame has been used in previous network studies, as it as been deemed "short enough and recent enough to facilitate accurate recall" (76).

If an individual had more than six partners, information was collected only on the 3 with which he or she had sex with the most frequent, and the 3 which he or she had sex with the fewest times. This would also counteract the bias to include partnerships of longer duration. From the National Health and Social Life Survey, Laumann suggested that the sex life of most individuals is characterized by a pattern of long stable relationships interspersed with more numerous short, casual relationships (76). A separate questionnaire was completed for each partner.

Upon acquisition, the *Study Contact* database contained a unique record for each nominated sexual partner (n=549). All records contained proxy information, which included sociodemographic (place of residence, Forward Sortation Area, ethnicity, education), and social and sexual interrelationship data.

2.2.3 Data Management: Screening, Editing, and Missing Data

All data screening and editing, for the '*Questionnaire*' dataset was completed in SAS after exporting the records from MS Access 97 into Epi Info v6, checking the alignment of columns and cross-checking records against the original file (102). Basic descriptive statistics and range checks were computed; and combinations of variables were verified. If an entry was unclear, the record was verified against the original paper file. All entered place of residences and corresponding Forward Sortation Areas were verified against the Manitoba Postal Code Conversion File (2001) (103). All other nominal entries were verified or recoded appropriately.

In the *Study Case* database, one place of residence included a combination of two distinct regions (i.e. Winnipeg and The Pas); this was recoded to reflect the RHA indicated on the consent form. For many categorical variables, if a participant's response did not fit into the pre-defined categories, the option to select 'other' and further specify, was given. These entries were examined to determine if they could be included within other pre-defined categories. This process was completed for the variables defining main source of income in the previous year, other sources of income in the previous year, highest level of education, ethnicity, and reason for healthcare visit. Many participants specified two 'other' answers for their reason for the healthcare visit, both of which were included in the pre-defined categories. Two new variables were derived, recoding for both reasons.

For example, if a participant had indicated "symptoms & contact of a case", two new variables would be derived: 1) recoding to 'symptoms'; and 2) recoding to 'contact of a case'.

Combinations of the variables pertaining to symptoms, healthcare and sexual behaviours of the individual during the symptomatic period were examined. For example, if an individual indicated no symptoms, but a date of symptom onset was included; the record was checked against the original paper file. If determined to have been entered correctly, the record was recoded as symptoms present. This logic was repeated for other variable combinations. The date of symptom onset was recorded as either a specific date or a length of time passed; as such, entries were verified, and subsequently formatted as a day, month, and year value. No dates were out of range; however, six records contained unclear entries. If both a date and a length of time passed were given, the former took precedence. Two entries specifying the time between healthcare treatment and symptom onset remained unclear after verification against the original paper file; both were recoded as missing. Subsequently, all time periods were converted into days. Four records for age of first consensual sex were out of range (less than 13 years of age); all had been correctly entered and these records were excluded in subsequent analyses. Additionally, eight entries for the recent number of sex partners were verified against the original paper files. These records were either recoded as missing or placed into an appropriate category.

Although there are no firm guidelines as to the amount of missing data acceptable for a given population, there is agreement that the pattern of missing data is more important than the amount (104). The amount of missing data for each variable was determined, and only variables with greater than 10% missing data were examined to determine if patterns existed (i.e. missing data not randomly distributed). In most instances, the amount of missing data was less than 10%. For these variables, the corresponding paper records were checked, and if determined to be truly missing, the decision was to leave as missing. Several other decisions can be made regarding the handling of missing data, which include: 1) deleting cases or variables; 2) imputing a missing value (prior knowledge, means, or regression); and 3) completing a sensitivity analysis (104). The possible effects of these decisions include: diminished statistical power; a biased sample; and biased analyses (105). Specifically, pattern testing was completed by first creating a dummy variable with two groups, missing and non-missing data; and then, examining the differences in the distributions or frequencies of other variables across these two groups. The amount of missing data in the *Study Case* database is presented in **Table 1**. Postal codes and province of residence were missing for more than 10% of the records. As this database would in future be supplemented with information contained in the CDC dataset, no pattern analysis was completed and it was left as missing. The time from symptom onset to healthcare visitation for those indicating the presence of symptoms, was missing for 10.5% of all records. A comparison between individuals missing data and those not missing data was completed – all were left as missing (**Table 2**).

Table 1: Amount of missing data identified for variables contained within the *Study Case* database, collected from the 'Social and sexual behaviours within sexually transmitted disease networks in Manitoba' supplementary questionnaire

Variable	**Amount (%) (n=327)**
Place of residence	22 (6.7)
Province of residence	41 (12.5)
Postal code	103 (31.5)
Ethnicity	2 (0.6)
Level of education	4 (1.2)
Main source of income	3 (0.9)
Injection drug use	3 (0.9)
Drug use before or during sex	1 (0.3)
Alcohol use before or during sex	1 (0.3)
Date of interview	10 (3.1)
Age of first consensual sex	4 (1.2)
Sexual partners (previous 12-months)	2 (0.6)
Lifetime number of sexual partners	2 (0.6)
Reason for healthcare visit	12 (3.7)
Symptomatic	2 (0.6)
Symptomatic sex	6 (3.3)*
Date of symptom onset	9 (5.0)*
Delay in seeking healthcare	19 (10.5)*

* n=181

Table 2: Test for patterns in missing data for the variable examining the time from symptom onset to healthcare visitation. The variable was categorized into two categories, missing (n=19) and non-missing (n=162) and differences in distributions of other healthcare, demographic and behavioural data was examined

Variable		Missing (n=19) Freq (%)	Non-missing (n=162) Freq (%)
Symptomatic sex	Yes	5 (26.3)	100 (61.7)
	No	7 (36.8)	59 (36.4)
	Unsure/refused/missing	7 (36.8)	3 (1.8)
Reason for healthcare visit when diagnosed with STI	Symptoms	11 (57.9)	124 (76.5)
	Other*	4 (21.0)	37 (22.8)
	Unsure/refused/missing	4 (21.0)	1 (0.6)
Alcohol use before or during sex	Yes†	14 (73.7)	124 (76.5)
	Never	4 (21.0)	36 (22.2)
	Unsure/refused/missing	1 (5.3)	2 (1.2)
Drug use before or during sex	Yes†	6 (31.6)	75 (46.3)
	Never	13 (68.4)	84 (51.8)
	Unsure/refused/missing	0	3 (1.8)
Main source of income	Legal‡	19 (100)	151 (93.2)
	Illegal§	0	8 (4.9)
	Unsure/refused/missing	0	3 (1.8)
Education	Primary school	0	6 (3.7)
	Grade 9 to 11	10 (52.6)	91 (56.2)
	Graduated grade 12	7 (36.8)	36 (22.2)
	Post-secondary°	2 (10.5)	27 (16.7)
	Unsure/refused/missing	0	2 (1.2)
Ethnicity	Caucasian	8 (42.1)	72 (44.4)
	Asian; black; Latin American; other	2 (10.5)	20 (12.3)
	Aboriginal	9 (47.4)	67 (41.4)
	Unsure/refused/missing	0	3 (1.8)
Number of lifetime sexual partners	1 to 5	6 (31.6)	41 (25.3)
	6 to 19	7 (36.8)	71 (43.8)
	≥ 20	6 (31.6)	47 (29.0)
	Unsure/refused/missing	0	3 (1.8)

* 'Other' includes the categories: contact of case, prenatal examination, routine examination, and other reason

† 'Yes' includes the categories: always, greater than 1/2 the time, approximately 1/2 of the time, and less than 1/2 of the time

‡ 'Legal sources of income' includes the categories: regular work, government assistance, and money from family and friends

§ 'Illegal sources of income' includes the categories: sex trade, dealing or doing drug runs, panhandling, and other crime

° 'Post-secondary education' includes the categories: university, college, and trade school

In the *Study Contact* database, entries in the 'other' option of the variable were examined to determine if they could be included within other pre-defined categories. This process was completed for the variables defining ethnicity, and the first place of meeting. Additionally, if the first place of meeting was a bar or hotel, participants were asked to further specify the name of the establishment. All entries were verified using phone and/or municipality directories. Three additional categories were created: bathhouse, restaurant and telepersonals.

The period of acquaintance before beginning a sexual relationship, and the duration of the sexual relationship (first and last dates of sexual contact) were recorded as either a specific date or a length of time. All entries were verified; for example, if the date of last sexual contact preceded that of the first sexual contact, neither value was used in future analyses. The period of acquaintance was converted into the appropriate number of days; and where possible, the first and last dates of sexual contact were formatted into a day, month, and year value. No dates were out of range; however, six records contained unclear entries.

The amount of missing data for the *Study Contact* database is presented in **Table 3**. All geographic variables were missing more than 10% of their entries. As this dataset would be supplemented with information contained in the CDC dataset, no pattern testing was completed and it was left as missing.

Table 3: Amount of missing data identified for variables contained within the *Study Contact* database, collected from the 'Social and sexual behaviours within sexually transmitted disease networks in Manitoba' supplementary questionnaire

Variable	**Amount (%) (n=547)**
Record number	2 (0.4)*
Place of residence	265 (48.4)
Province of residence	272 (49.7)
Postal code	488 (89.2)
Length of social acquaintance	4 (0.7)
First time/date of sexual contact	17 (3.1)
Last time/date of sexual contact	14 (2.6)
Ethnicity	2 (0.4)
Highest level of education	14 (2.6)
Community first met partner	6 (1.1)
Location first met partner	6 (1.1)
Type of relationship	3 (0.5)
Received anything for sex	4 (0.7)
Given anything for sex	8 (1.5)
Frequency of sexual contact (previous 3 months)	76 (13.9)
Anyone else ever had sex with partner	4 (0.7)
Injection drug use in previous 12-months	5 (0.9)
Frequency of condom use	8 (1.5)
Discussion of STD	4 (0.7)
Discussion of HIV	4 (0.7)

* N=549

Pattern testing was completed for frequency of sexual contact in the previous 3-months (13.9% of records missing entries) (**Table 4**). Records missing data for the frequency of sexual contact in previous 3-months, had a greater percentage of uncertainty or non-response with respect to their partner's highest level of education compared to those records with complete data, 42.1% and 20.2%, respectively. Additionally, 13.2 % of people with missing data indicated that they met that partner through sex work, on the street, through telepersonals or online, or in a bathhouse, in comparison to 3.4% of those with complete data. A higher proportion of records with missing data indicated that the relationship was of a casual nature in comparison to those records with complete data, 68.4% and 36.7%, respectively. Both HIV status and STI history were indicated to be more frequently discussed when records had complete data (33.8% and 50.7%, respectively) compared to those records missing data (14.5% and 32.9%). Through examination of the differences in distributions presented in **Table 4**, missing data was not randomly distributed. The missing data correlated well with sexual partnership in which partners social data is likely unknown. It was deduced that some records with missing data did not fit into the pre-defined categories, as most appeared to be single instance of casual sex or a relationship that had ended greater than 3 months ago (data not shown). A new category was defined, '0 times in previous 3-months' and where possible, missing data was imputed. Specifically, for each record, a combination of variables was separately assessed by both EM and AJ, and only those records with agreement between the two were recoded. For example, a record with the following information, date of first sex = 14/05/2002, date of last sex = 14/06/2002, date of interview = 13/11/2002, was recoded to reflect '0 times in previous 3-months'.

Table 4: Test for patterns in missing data for variable examining sexual contact in the previous 3-months. The variable was categorized into two categories, missing (n=76) and non-missing (n=471) and differences in distributions of other proxy information and interrelationship variables were examined.

		Missing (n=76)	Non-missing (n=471)
Variable		**Freq (%)**	**Freq (%)**
Ethnicity	Caucasian	35 (46.0)	215 (45.6)
	Asian; black; Latin; other	11 (14.5)	78 (16.7)
	Aboriginal	24 (31.6)	162 (34.4)
	Unsure/refused/missing	6 (7.9)	16 (3.4)
Education	Primary school	0	7 (1.5)
	Grade 9 to 11	21 (27.6)	165 (35.0)
	Graduated grade 12	19 (25)	140 (29.7)
	Post-secondary*	4 (5.3)	64 (13.6)
	Unsure/refused/missing	32 (42.1)	95 (20.2)
Where did you first meet partner	Bar; hotel; restaurant	19 (25)	99 (21.0)
	Private residence; family or friends	30 (39.5)	239 (50.7)
	School	4 (5.3)	51 (10.8)
	Sex trade/street; telepersonals/online; bathhouse	10 (13.2)	16 (3.4)
	Work	2 (2.6)	26 (5.5)
	Other	8 (10.5)	35 (7.4)
	Unsure/refused/missing	3 (3.9)	5 (1.1)
Type of relationship	Spouse; boy/girlfriend	9 (11.8)	174 (36.9)
	Casual sex; "on and off"	52 (68.4)	173 (36.7)
	Ex-lover; ex-spouse	12 (15.8)	115 (24.4)
	Unsure/refused/missing	3 (3.9)	9 (1.9)
Anyone you known ever had sex with this person	Yes	28 (36.8)	170 (36.1)
	No	32 (42.1)	250 (53.1)
	Unsure	14 (18.4)	45 (9.5)
	Refused/missing	2 (2.6)	6 (1.3)
IDU	Yes	1 (1.3)	18 (3.8)
	No	51 (67.1)	367 (77.9)
	Unsure	21 (27.6)	84 (17.8)
	Refused/missing	3 (3.9)	2 (0.4)
Condom use	Never	45 (59.2)	188 (39.9)
	Sometimes (< 1/2 time, ~ 1/2 time, > 1/2 time)	17 (22.4)	200 (42.5)
	Always	9 (11.8)	70 (14.9)
	Unsure/refused/missing	5 (6.6)	13 (2.8)
HIV discussion	Yes	11 (14.5)	159 (33.8)
	No	63 (82.9)	308 (65.4)
	Unsure/refused/missing	2 (2.6)	4 (0.8)
STD discussion	Yes	25 (32.9)	239 (50.7)
	No	49 (64.5)	227 (48.2)
	Unsure/refused/missing	2 (2.6)	5 (1.1)

* 'Post-secondary education' includes the categories: university, college, and trade school

2.2.4 Heaping in Discrete Data

Heaping is a source of measurement error, occurring when responses to open-ended interview questions are rounded-off to preferred terminal digits, or to multiples of a particular unit (106-108). Examples of heaping documented in literature include: reporting of age, and number of sexual partners (109;110). As several of the interview questions were open-ended, distributions of discrete variables were visually examined for heaping. Smoothing and categorizing are methods to deal with heaping, the latter option was used for this study (111;112). Categories were selected based on: the occurrence of natural peaks in the distribution, and epidemiologic meaning (113).

Two of the variables in the *Study Case* database suggested heaping: 1) number of sexual partners in previous 12-months; and 2) number of days passed from the time the individual first noticed symptoms until time visited a healthcare professional. For the former, respondents preferred numbers ending with a zero or five; and the latter, multiples of calendar periods were favoured. One variable in the *Study Contact* database exhibited heaping, period of social acquaintance before commencement of sexual relationship, where multiples of calendar periods were preferred (**Appendix G**).

2.3 Original Edge List Database

2.3.1 Background, Acquisition, Management

The *Original Edge List* database contained information which linked laboratory-confirmed cases to their respective sexual partners (n=10,393). This database was a combination of all sexual relations contained within the *MH STI Dataset* and the supplementary *Questionnaire* dataset. Where possible, study records were condensed into their corresponding CDC record. Preprocessing by MH, in most instances, included the deletion of individuals under the age of 13 (n=8), and contacts with no identified index case (n=112). All records contained the following information: a field indicating if identified in the CDC dataset and corresponding unique record number; a field which designated study participation and corresponding unique record number; a link number defining groups of individuals linked through sexual contact; and demographic data (age, gender, province, Forward Sortation Area). Not all records were used in the analysis, refer to **Section 2.3.3**.

From this database, personnel at the CDC created the *Geocode* database containing a record of all geographic locations within the *Original Edge List* (n=441 records). Records were not unique, as the list included both correctly and incorrectly spelled geographic locations. All place codes were standardized.

Verification of the previous assignment of link numbers and unique identifiers revealed problems. Problems with link numbers included: greater than one index case in a sexual grouping; an index case named as a contact of him- or her-self; and duplicate records. Problems found with the assignment of unique identifiers included: duplicate disease episodes, discordant values of gender; and age discrepancies of greater than +/- 2 years between disease episodes. **For a detailed description refer to Appendix H and I.**

2.3.2 Merging

All previously described databases contained relational fields which permitted subsequent match-merging. The relational fields for each database are described in **Table 5**. After the appropriate data management, the databases were then merged with the *Original Edge List*, upon which additional verification was completed. Diagnostic codes (ICD-9-CM) were checked for consistency with the subject's gender. For example, records with a diagnostic code for PID should be coded as female.

Table 5: Assigned relational fields, allowing for prospective match-merging

Database	**Relational field**	**Value**
The Registry	Unique Record Number	877 to 5,954
MH Contact Database	Unique Record Number	5,955 to 11,190
Study Case Database	Unique Record Number	1 to 327
	Record Number*	10 to 337
Study Contact Database	Unique Record Number	328 to 876
	Record Number*	10 to 337
Geocode	Unique Record Number	1 to 11,190

* Record number is the link which specifies the sexual relations between the study participants

2.3.3 Eligibility

Eligibility of each record was examined. All laboratory confirmed cases of chlamydia and/or gonorrhea and their respective contacts 13 years of age or older were included. Four records did not meet the age criteria. If a confirmed case of syphilis was exposed to either chlamydia and/or gonorrhea, both the case and respective contact(s) were included. Forty-four syphilis cases and 35 of their nominated sexual contacts were not included. Contacts with no corresponding index case were not included (n=126).

As all databases contained records of events and not individuals, a *historical file* had to be created. The historical file was created in SAS, by numbering events in sequential order for each unique identifier; creating separate datasets for event1, event2, event3, … , event*n*; and finally joining all events horizontally by unique identification number. The historical file contained one record for each uniquely identified individual, and consisted of all disease episodes (case records) and all notification events (contact records). There were a total of 8,476 unique individuals.

2.3.3.1 Missing Data

The amount of missing data for each variable was separately determined for cases (n=4,683) and contacts (n=3,793) (**Table 6**). For both cases and contacts, records missing data for place of residence were imputed using an individual's designated RHA (n=11 cases; n=52 contacts). For contacts, the variable age was missing more than 10% of its entries (48.5%). Pattern testing was completed (refer to **Section 2.2.3** for description) (**Table 7**). Records missing age were left as such.

Table 6: Amount of missing data for entire study of laboratory-confirmed cases (n=4,683) and contacts exposed to STI (n=3,793) identified from June 1, 2002 to September 30, 2003

	Cases (n = 4,683) Freq (%)	**Contacts (n = 3,793) Freq (%)**
Gender	35 (0.7)	154 (4.1)
Age	17 (0.4)	1,840 (48.5)
Place of residence	13 (0.3)	447 (11.8)

Table 7: Test for patterns in missing data for variable age for all identified contacts (n=3,793). Age was categorized into two categories, missing and non-missing and differences in distributions of other variables were examined.

		Missing (n=1,840) Freq (%)	**Non-missing (n=1,953) Freq (%)**
Gender	Male	1097 (59.6)	1561 (79.9)
	Female	595 (32.3)	386 (19.8)
	Missing/unknown	148 (8.0)	6 (0.3)
Ethnicity	Treaty	5 (0.3)	385 (19.7)
	Other	1835 (99.7)	1568 (80.3)
Geographic location	Winnipeg	689 (37.4)	903 (46.2)
	Mb (excluding Winnipeg)	566 (30.8)	965 (49.5)
	Canada, outside Mb	129 (7.0)	48 (2.5)
	USA	24 (1.3)	2 (0.1)
	Outside North America	16 (0.9)	4 (0.2)
	Missing/unknown	416 (22.6)	31 (1.6)

2.4 Ethics Approval

Ethical approval was granted by the following institutions: the Health and Research Ethics Board, University of Manitoba; participating RHAs in the province of Manitoba and the Health Information Privacy Committee of Manitoba Health; and the Ottawa Hospital Research Ethics Board (OHREB), refer to **Appendix J** for greater detail.

3 Methods

3.1 Objective 1

The purpose of the first objective was fourfold: 1) to describe the entire study population using basic demographic and clinical variables found in the *MH STI Dataset*; 2) to compare the demographic and clinical variables of those participating and those who were eligible but did not participate in the supplementary questionnaire; 3) to describe the sociodemographic, clinical, healthcare and behavioral characteristics of those identified through the supplementary questionnaire; and 4) to describe the sexual relationship dyads or case-contact pairs (refer to **Section 1.5.1**).

3.1.1 Variables and Definitions

Variables were derived using the historical records for each uniquely identified individual. Refer to **Table 8, and Appendix F and K** for definitions and derivations of variables used in Objective 1.

3.1.2 Statistical analysis

SAS was used to obtain basic statistics, which were subsequently organized into tables and charts (101).

Table 8: Definitions of variables derived from the historical file containing the entire study population of identified cases (4,683) and contacts (3,793) during study period, June 1, 2002 to September 30, 2003.

Variable		**Definition**
Disease episode (case record)		An episode is a separate report of laboratory-confirmed STI; an episode could consist of two infections (i.e. concomitant infection of *C. trachomatis* and *N. gonorrhoeae*)
Notification event (contact record)		An event is a separate notification report as nominated by a specific index case
Case		An individual who ever had a laboratory-confirmed disease episode during the entire length of the study or a case who participated in the supplementary questionnaire but not found in the CDC file
Contact		An individual with no laboratory-confirmed disease episode, but who either had a notification event or was nominated as a study contact
Eligibility (supplemental questionnaire)		Defines if an individual was eligible to participate in the supplemental questionnaire
Aboriginal people(s)		"… descendants of the original inhabitants of North America. The Canadian Constitution recognizes three groups of Aboriginal people: Indians (now known as First Nations people), Métis and Inuit." (114)
First Nations people		"Refers to Indian people in Canada, both Status and Non-Status" (114)
Treaty First Nation		"A Status Indian who belongs to a First Nation that signed a treaty with the Crown" (114)
Reserve		"A reserve is a tract of land, the legal title to which is held by the Crown, set apart for the use and benefit of an Indian band. Some bands have more than one reserve." (114)
<u>Variables collected for entire population</u>		
Age		Age in years at the date of naming or diagnosis
Co-infection		If an individual ever had two laboratory-confirmed STIs reported during one disease episode, over the course of the extraction period (July-2002 to September-2003). Combinations included: 1) chlamydia and gonorrhea, 2) chlamydia and syphilis, or 3) gonorrhea and syphilis
Ethnicity	CDC	Limited to Registered First Nation (Treaty) or other
	Study	Self-identified ethnic group or family background
Infection acquired		Accumulation of laboratory- confirmed STI(s) diagnosed during the extraction period (July-2002 to September-2003)

Variable	Definition
Repeat nominee	An individual with two or more notification reports between the period commencing on October-2002 to September-2003
Repeater	An individual with two or more disease episodes reported to MH during the extraction period of July-2002 to September-2003
Variables collected on voluntary study participants	
Healthcare delay	Number of days passed from the time the individual first noticed symptoms until time visited a healthcare professional
Other source of income (drugs)	If received any additional money, not main source, from dealing or doing drug runs in the 12-months prior to study participation
Other source of income (government support)	If received any other money from welfare, social assistance, employment insurance, or other government support in the 12-months prior to study participation
Other source of income (panhandling and other crime)	If received any additional money, not main source, from panhandling and other crime, except sex work, in the 12-months prior to study participation
Other source of income (personal support)	If received any additional money, not main source, from family or friends in the 12-months prior to study participation
Symptomatic	If an individual experienced any symptoms from current STI episode
Symptomatic sex	If an individual had sex while they were experiencing symptoms, and before prescription of antimicrobial
Variables collected on identified study relationships (partner-specific)	
Given anything for sex	If index case ever had given anything to specific partner in exchange for sex, including money, drugs or other items including shelter
Length of social acquaintance	Length of time (days) index case knew partner before beginning a sexual relationship
Partner ever had sexual relation with acquaintance(s)	If any acquaintance of the index case ever had known sexual contact with nominated partner
Received anything for sex	If ever received anything from partner in exchange for sex, including money, drugs or other items
Component characteristics (Objective 2)	
Network variables (refer to Appendix A)	
Mean % MSM or WSW links	Percentage of total links in component occurring between men who have sex with men (MSM) and women who have sex with women (WSW)
Symptomatic sex	If a component contains an individual who reported having symptomatic sex

3.2 Objective 2

The purpose of Objective 2 consisted of two parts: 1) to identify components of individuals connected directly or indirectly through sex using PAJEK; and 2) describe basic social network measures and component characteristics. PAJEK is a social network computer program specifically designed to handle large datasets (115). Data files used in PAJEK must first be structured into a recognizable input file format (refer to **Appendix L** for example of file format). First the *Original Edge List,* referred to in **Section 2.3**, was restructured into a listing of all sexual relationships or dyads in the sexual network, also referred to as an edge-list. Restructuring included: identifying and selecting the index case in each sexual grouping; linking an index case with their respective nominated sexual partner(s); deleting all records in which the index case was a contact of themselves; deleting duplicate records in the edge-list (based on link number, index case identifier, and contact identifier); and identifying same sex relationships. Variables derived from the *historical file* (gender, repeater, co-infection, ethnicity, and infection acquired) were merged into the edge-list (refer to **Section 2.3**). The file was saved as *.dbf* file, and imported into Epi Info v6.

Finally, a visual basic program (*Syph.bat* © Ann Jolly 2003) was used to structure the edge-list into a PAJEK readable input file. The *Syph.bat* program was modified by the author to fit both the structure and the variables found in this dataset. This program processes the information from an Epi Info v6 *.REC* file, and outputs it as a text file that can be read by PAJEK. Specifically, it worked by: 1) creating a unique actor list for all individuals in the sexual network, 2) assigning a unique PAJEK identification number to all actors in the list; 3) creating a PAJEK format actor- and edge list; and 4) adding in basic descriptors for each actor (gender, infection acquired, treaty status) and edges (same sex links). An actor list is a list of all individual, cases or contacts, in the sexual network.

The resulting text file was imported into PAJEK, and components were identified (refer to **Appendix M** for PAJEK output). The size of each component was determined. Diagrammatic representations of components were created as needed.

3.2.1 Statistical analysis

PAJEK is designed to handle large networks (thousands of nodes). However, calculation of social network measures is very intensive for large networks. For efficiency, PAJEK uses approximate measures for many formulas. As a result, component information was extracted from PAJEK, and imported into UCINET V- a SNA computer program that calculates basic social network measures for small networks (116). Microstructures were identified using UCINET V (refer to **Section 1.5**) and density and degree centralization were calculated in SAS (as per the formulas in **Section 1.5.1**).

A profile of collective responses to sociodemographic, clinical and geographic variables was calculated for each component using SAS. Additionally, study components were identified and collective responses to supplementary questions were calculated for each study component. (Study components are defined as those including at least one individual participating in the supplementary questionnaire). Component profiles were organized into tables and charts.

3.3 Objective 3

3.3.1 Statistical analysis

Cluster analysis (CA) was used to define groups or clusters of components, and differences between identified groupings were compared. All steps of the CA were completed in SAS. Clustering units were components containing 5 or more individuals (n=239), excluding anonymous components. A component is a group of individuals within a network, connected either directly or indirectly through sex. Anonymous components were either: a) components containing a participant who maintained anonymity after completion of the study questionnaire; or b) components containing participants residing in rural Regional Health Authorities (RHAs).

Clustering variables (n=3) were selected on the basis of empirical and theoretical importance, namely: 1) percentage of First Nations (Treaty); 2) degree centralization, and 3) percentage of individuals residing in the same geographic region (refer to **Appendix K**, for derivation of *geographic region*).

In Manitoba, Aboriginal ethnicity is a recognized risk marker for STI. Previous research in this population has demonstrated that components with high Aboriginal composition had significantly different disease patterns in comparison to components with low Aboriginal composition (refer to **Section 1.5.2**).

Mixing patterns, with respect to number of sexual partners, have been shown to strongly contribute to the maintenance of STIs within a population. Theoretically, Moody *et al.* (2004) demonstrated that STI epidemics can occur in a population where the maximum number of sexual partners of an individual, or degree, is three (117). In the previous study, Wylie and Jolly (2001) visually classified components into two structural groups linear and radial (refer to **Appendix B**) (66). Components which exhibited assortative mixing patterns with respect to number of sexual partners (linear components) had significantly higher STI positivity rates than components with predominantly disassortative mixing patterns (radial components) (66). In the current study, degree centralization, a measure of the variance in degree distribution within a component, was used to objectively distinguish between linear and radial components.

Large intracommunity networks have been shown to increase the likelihood of STI persistence within a component. Consequently, the maximum percentage of individuals within the same geographic region was considered to be an appropriate clustering variable (66).

Univariate statistics were calculated, variable distributions were examined, and bivariate scatter-plots and Pearson product-moment correlations were computed. It has been shown that extreme observations may compromise the recovery of true clusters (118). Outliers were carefully examined to determine a reason for their existence and if clustering should be completed in the presence or absence of these observations. Only a few components had extreme observations, and all were included in the CA.

The approximate covariance estimation for clustering (ACECLUS) procedure, based on Mahalanobis' generalized distance, was used to measure the distance between pairs of components. The *ACECLUS* procedure is an iterative process which uses pair-wise differences to estimate the pooled within-cluster covariance matrix (**A**) (refer to **Appendix N**) (119;120). Iterations continue until successive estimates of **A** are sufficiently close. The basic assumption of the ACECLUS procedure is that all clusters have multivariate normal distributions with equal covariance matrices.

Prior to initiation of this procedure, several criteria needed to be specified. The initial estimate for the pooled within-cluster covariance matrix (**A**) was the total-sample covariance matrix (**S**) obtained from the component by variable input matrix (239 components by 3 clustering variables). The distance cut-off (μ) was obtained by specifying the proportion of pair-wise differences between observations that would be less than μ, and transforming this proportion (p) into the appropriate distance value. Literature suggests that a series of values of p should be tried (120). As such, three values were chosen for p: 0.2, 0.10, and 0.05. Values of 0.2 or lower were chosen in attempts to minimize the inclusion of between-cluster pairs (119). Finally, closeness was determined using the following formula:

$$e_i = \frac{1}{v}\left\| Z'(A_i - A_{i-1})Z \right\|, \text{ where } \mathbf{Z} = \mathbf{S}^{-(1/2)}$$

As suggested in literature, a value of 0.001 or smaller was used to indicate sufficient closeness between successive estimates of the pooled within-cluster covariance matrix (120).

The ACECLUS procedure was used, as previously published experiments examining the utility of this metric demonstrated its use improved cluster recovery, when compared against the Euclidean metric (119;120). Even in the presence of moderately heterogeneous within-cluster covariance matrices cluster recovery was improved when this metric was used.

The resulting standardized canonical coefficients were used to linearly transform the raw data. The three resultant datasets, which differed with respect to the initial specification of p, were used as input matrices for clustering.

Cluster analysis was carried out using the *Cluster* procedure in SAS (121). Ward's Minimum Variance method, which forms clusters on the basis of loss of information calculated as the sum of squares, was selected as the clustering algorithm (122). Ward's method was selected as it has continually performed well in multiple experiments and has often shown superior recovery in comparison to other algorithms (123-126). In brief, the algorithm begins with each component (n=239) representing a single cluster; a situation in which the largest amount of information is available. Next, the algorithm considers all possible clusters that could be formed joining the two that result in a minimal increase in the error sum of squares, resulting in *n*-1 clusters. This algorithm is repeated until the data is systematically reduced into one cluster containing all 239 components. The basic assumptions of this algorithm are: a multivariate normal distribution of variables in the clusters; equal spherical covariance matrices; and equal sampling probabilities.

The optimal clustering level, restricted to the final 10 hierarchical clustering levels, was determined by examining both dendrograms and two numeric stopping rules. The two stopping rules examined were: 1) the Pseudo F test, a variance ratio criterion; and 2) the Pseudo T^2, related to a rule developed by Duda *et al.* (1973) which compares within-cluster variance when two clusters are joined to that when the two clusters are separated (refer to **Appendix N**) (127;128). These two stopping rules were chosen on the basis of previously published experiments, where both performed superior to multiple other rules, when assumptions were violated (129;130). The suggested stopping point for the Pseudo F and T^2 tests were observed absolute or local maxima, and observed local minima, respectively. Finally, the resultant clustering solutions were examined to determine if logical clusters were formed, given the context of the research question, by comparing the clusters on the variables used to generate the solution and those external to the solution; and through the comparison of supplementary information, elicited from the study questionnaire.

4 Results and Discussion

Due to the large amount of data presented, the results will be discussed as they appear in the text. Furthermore, pertinent findings will be summarized in a concluding statement found in **Section 5**. The author feels that a comprehensive results section is appropriate, as to: 1) minimize repetitiveness; and 2) to provide a basic explanation of the application of social network analysis as it appears in the text.

4.1 Objective 1

4.1.1 Description of the Entire Study Population

The entire study population contained 8,476 uniquely identified individuals consisting of 4,683 cases and 3,793 contacts (**Table 9 and 10**) (refer to **Table 8** for definitions). Of the 4,683 cases, 32.6% were males and 66.7% were females.

Even though a greater emphasis has been placed on screening asymptomatic males; the observed gender gap is most likely an artifact of screening asymptomatic females during routine physical and prenatal examinations and other circumstances built into the current guidelines. Additionally, males are often under diagnosed, due to: an unwillingness to be tested; presumptive treatment without corresponding laboratory testing; poor test sensitivities; and spontaneous clearance (66;93;131).

Table 9: Demographic and clinical characteristics of laboratory-confirmed cases with sexually transmitted infections identified during the "Manitoba Sexually Transmitted Disease Study" in Manitoba, Canada, June 1, 2002 to September 30, 2003 (n=4,683)

		Cases (n=4,683)		
		Male (n=1,525)	**Female (n=3,123)**	**Unknown (n=35)***
		Freq (%)	**Freq (%)**	**Freq (%)**
Age (years) Median (Iqr)†		23 (19, 28) (n=1,523)	20 (18, 24) (n=3,121)	20 (18, 23) (n=22)
Geographic region	Winnipeg	827 (54.2)	1,565 (50.1)	3 (8.6)
	Southern rural Manitoba	286 (18.7)	592 (19.0)	20 (57.1)
	Northern remote Manitoba	398 (26.1)	938 (30.0)	12 (34.3)
	Outside Manitoba	13 (0.8)	27 (0.9)	0
	Missing/Unknown	1 (0.1)	1 (0.03)	0
CDC Ethnicity‡	First Nations (Treaty)	439 (28.8)	916 (29.3)	10 (28.6)
	Other	1,086 (71.2)	2,207 (70.7)	25 (71.4)
Infection acquired	Chlamydia	1,064 (69.8)	2,636 (84.4)	N/A
	Gonorrhea	285 (18.7)	220 (7.0)	
	Syphilis	5 (0.3)	6 (0.2)	
	Chlamydia, gonorrhea	168 (11.0)	253 (8.1)	
	Chlamydia, syphilis	0	2 (0.1)	
	Gonorrhea, syphilis	2 (0.1)	2 (0.1)	
	Chlamydia, gonorrhea, syphilis	1 (0.1)	0	
	Unknown	0	4 (0.1)	
Repeaters		100 (6.6)	246 (7.9)	N/A
Co-infection present		152 (10.0)	221 (7.1)	N/A
Repeat nominee		96 (6.3)	55 (1.8)	N/A

* Unknown refers to those individuals who remained anonymous after study completion or those whose gender could not be determined from given names

† Iqr = interquartile range

‡ Ethnicity for those with unknown gender is derived from variable Study Ethnicity (refer to Table 8 for definitions)

N/A Data not available

The observed age distribution is most likely a result of screening practices (refer to **Section 2.1.1.1**). Chlamydia testing is recommended for all sexually active individuals under the age of 25 during presentation to a healthcare professional (94). Males were older than females (median; 23, 20, respectively). The difference in age distributions is most likely due to differing partner selection patterns between genders (84;132).

Specifically, Anderson *et al.* (1990) reported that in both European and North American heterosexual populations, males are typically two to three years older than their female partners (84).

The percentage of First Nations Treaty enrolled in the provincial health insurance plan was approximately 6% in 2002 (91). This number is likely a conservative estimate of the actual number, due to the self-report nature of this field. In this study, individuals identifying as First Nations Treaty were over represented as laboratory-confirmed cases, consisting of 29% of all cases independent of gender. This discrepancy has been reported in previous studies within this population; even while controlling for socioeconomic status (45;46;66). Compounding this disparity is the high proportion of laboratory-confirmed cases residing in northern remote Manitoba; an area containing only 9% of the population. Part of this disparity may be attributed to the STI control strategies implemented by nurses employed by Health Canada delivering direct services to individuals on reserves. It could also be due to differences in sexual or healthcare seeking behaviours. However, differences in sexual mixing within and between ethnic groups may better explain the high rates of STIs observed within the Aboriginal population (51).

The majority of males and females acquired chlamydia during the study period, 69.8% and 84.4%, respectively. However, males were more likely to have acquired gonorrhea (18.7%) than females (7.0%). This finding agrees with national surveillance statistics. Specifically, in 2002, 7,185 cases of gonorrhea were reported nationally, of these 62% were in males (133). The higher proportion of infections identified in males could be a result of multiple factors, including: males are more likely to experience symptoms in comparison to females; efficient transmission between men who have sex with men

(MSM); and the increased risk of gonorrhea infection in males frequenting sex trade workers. The latter item would result in a greater number of men being infected by a smaller number of females.

An outbreak of locally acquired infectious syphilis occurred in Manitoba from January-2003 to March-2004 (49). Grassly *et al.* (2005) suggested that repeated syphilis outbreaks in the United States may result from the natural cycling of host immunity (134). Furthermore, evidence indicated the existence of a "connected sexual network", as repeated outbreaks were often "synchronized" across distinct geographic regions (134).

The current findings emphasize the importance of characterizing sexual networks. Repeat infections, co-infection, and repeat nomination have been suggested to be markers of core group membership - all features observed in the current study (67). Over the course of the study, a small proportion of both males and females were repeatedly diagnosed with STIs, 6.6% and 7.9%, respectively. This number was comparable to that observed in sexual networks formed in Winnipeg in 1997, in which 8% of identified laboratory-confirmed cases had repeated chlamydial infections (46).

A small proportion of repeat infections may actually be continuing infections resulting from treatment failure. However, previous studies have suggested that repeat infections, specifically in females, have been associated with continued sexual contact with an untreated partner. This is of concern for many reasons: 1) females are at greater risk of sequelae following repeated genital chlamydia infection; 2) repeater status may indicate a dependence on sexual activities for livelihood or drugs; and 3) it can be an indicator of power differentials in a relationship (135-138). Re-infection also indicates a need for improvement in STI prevention and control activities; particularly with respect to the partner notification process. It has been suggested that nurses may neglect to gather names from nominated contacts with laboratory-confirmed infection (in addition to that of the index case); ignoring that they may be exposed to another infected partner (93). Other strategies to combat recurrent infections include re-screening and patient-delivered partner therapy.

A higher proportion of males (6.3%) had more than one notification event during the study period in comparison to females (1.8%). This figure emphasizes the importance of men in disease transmission. Specifically, Potterat *et al.* (1987) examined the transmission of gonorrhea in three clinical sub-groups (symptomatic, sub-symptomatic and asymptomatic) of heterosexual males (139). Although asymptomatic males consisted of only 13% of all gonorrhea cases identified during this time period, they contributed to 35% of all spread cases in laboratory-confirmed female cases (139).

Of the 3,793 contacts, 70.1% were males and 25.9% were females (**Table 10**). Contacts were geographically dispersed, spanning many regions within Canada and outside North America. This emphasizes that partner notification and treatment must incorporate information gathered from a wide geographic area (as suggested in previous study) (66).

Table 10: Demographic characteristics of contacts identified during the "Manitoba Sexually Transmitted Disease Study" in Manitoba, Canada, June 1, 2002 to September 30, 2003 (n=3,793)

		Contacts (n=3,793)	
	Male (n=2,658)	**Female (n=981)**	**Unknown (n=154)***
	Freq (%)	**Freq (%)**	**Freq (%)**
Age (years) Median (Iqr)†	23 (20, 28) (n=1,561)	21 (18, 27) (n=386)	20 (18, 24) (n=6)
Geographic region Winnipeg	1,189 (44.7)	400 (40.8)	18 (11.7)
Southern rural Manitoba	468 (17.6)	149 (15.2)	4 (2.6)
Northern remote Manitoba	672 (25.3)	242 (24.7)	5 (3.2)
Outside Manitoba	172 (6.5)	78 (7.9)	1 (0.6)
Missing/Unknown	157 (5.9)	112 (11.4)	126 (81.8)
CDC ethnicity‡ Treaty Code	318 (12.0)	72 (7.3)	18 (11.7)
Repeat nominee	107 (4.0)	15 (1.5)	N/A

* Unknown refers to those individuals who remained anonymous after study completion or those whose gender could not be determined from given names

† Iqr = interquartile range

‡ Ethnicity for those with unknown gender is derived from variable Study Ethnicity (refer to Table 8 for definitions)

N/A Data not available

This is further emphasize on detailed examination of component 430 (**Figure 6**). Component 430 spanned 10 geographic locations, including northern remote, southern rural and urban communities. Of interest was the existence of a bridge, connecting residents in two distinct geographic subpopulations. Specifically, the same sex relationship occurring between unique identifier 15519 and 14016. Further questions emanate from this finding, including: 'what is the relationship between these two individuals?', 'where did they meet?', and 'what is the frequency of sexual contact?'. Due to the large proportion of records with missing data for age and geographic location, numerical summary measures should be interpreted cautiously (**Table 6**). This problem is commonly observed, and results from individuals not knowing, or not wanting to disclose information on their sexual partners. However, where data are present, they are of high quality and completeness. As such, results obtained from this study will only further our understanding of the transmission of STIs within this population.

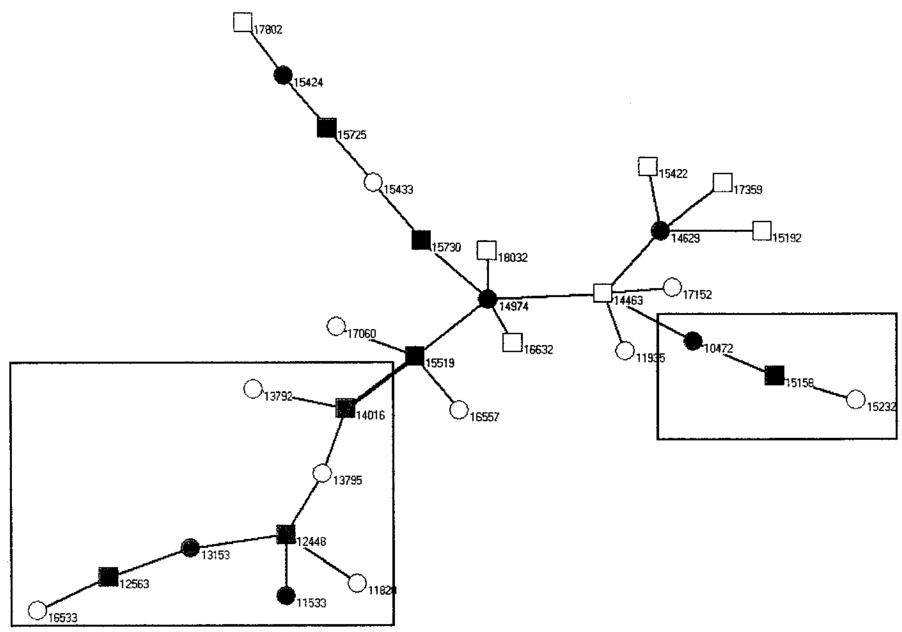


Figure 6 Component 430. Males and females are designated by circles and squares, respectively. Colours indicate infection, chlamydia (red), gonorrhea (green), both chlamydia and gonorrhea (plum), and unknown infection (white). Lines represent sexual links; red lines indicate same sex links and black indicate heterosexual links. Individuals residing in northern remote Manitoba are highlighted by a yellow box; all others are those residing in Winnipeg or southern Manitoba

Additionally, previous studies have indicated that other information (i.e. frequent places of socialization) could be obtained which would facilitate specific targeting of these high-risk individuals; a point which will be discussed in the following text.

4.1.2 Comparison of Participants to Eligible Non-Participants

Out of 836 individuals asked to participate in the Winnipeg RHA, 272 consented (32.5%). This low response rate could be due to: 1) the invasiveness of the questionnaire; 2) lack of compensation for participation; 3) recruitment after initial healthcare visit (i.e.

diminished concern); 4) public health nurse acting as the interviewer; or 5) perceived unimportance of academic research.

As no information was kept on all 836 individuals, it was decided to use the total *eligible* population in this comparison. Specifically, eligibility was determined by including all chlamydia and/or gonorrhea cases presenting to a participating RHA during the period of questionnaire administration (refer to **Appendix K** for complete definition). The calculated eligible population consisted of 2,831 unique individuals (**Table 11**). This number indicates that individuals were either: 1) not asked to participate, or 2) their eligibility status was misclassified.

The gender breakdown of participants and non-participants was similar. Greater than 65% of both participants and non- participants, excluding those with unknown gender, resided in Winnipeg. However, a greater number of both male and female non-participants resided in northern remote Manitoba (23.2% and 27.1%, respectively), in comparison to less than 1.0% of participating males and females. A lower proportion of male and female participants were registered as First Nations Treaty, 16.5% and 13.2%, respectively, in comparison to non-participant males (27.6%) and females (28.4%).

The lower participation rate among individuals residing in northern remote Manitoba and First Nations is most likely due to differences in the appropriation of healthcare services to First Nations (refer to **Section 2.1**). Health Canada was not an active partner in this study. As such, laboratory-confirmed cases residing on reserve would not have been asked to participate. Additionally, many First Nations have had negative encounters with the healthcare system. As a result, they are often distrustful of academic research.

Although data derived from questionnaire-based surveys on the sociodemographic, and healthcare and sexual behaviours of individuals infected with STIs, may not be comparable to other populations, the reported estimates provide a better understanding of the sociodemographic profile and risk behaviours of this population.

Table 11: Comparison of demographic and clinical characteristics between interviewed cases who participated and those who were eligible but did not participate in the in the "Social and sexual behaviours within sexually transmitted disease networks in Manitoba" supplementary questionnaire (n=2,831)

		Cases (n=2831)				
		Male (n=937)		**Female (n=1859)**		**Unk (n=35)***
		Study (n=85)	**Non (n=852)**	**Study (n=205)**	**Non (n=1654)**	**Study (n=35)**
Age (Years) Median (Iqr)†		21 (17,30) (n=85)	23 (20,29) (n=851)	20 (18,24) (n=205)	20 (18,24) (n=1654)	20 (18,23) (n=22)
Geographic region	Winnipeg	83 (97.6)	606 (71.1)	193 (94.1)	1103 (66.7)	3 (8.6)
	S. rural Mb	2 (2.3)	48 (5.6)	10 (4.9)	103 (6.2)	20 (57.1)
	N. remote Mb	0	198 (23.2)	2 (1.0)	448 (27.1)	12 (34.3)
	Outside Mb	0	0	0	0	0
CDC ethnicity‡	Treaty Code	14 (16.5)	235 (27.6)	27 (13.2)	470 (28.4)	10 (28.6)
	Other	71 (83.5)	617 (72.4)	178 (86.8)	1184 (71.6)	25 (72.4)
Infection acquired	Chlamydia	55 (64.7)	579 (68)	176 (85.8)	1385 (83.7)	N/A
	Gonorrhea	19 (22.3)	176 (20.7)	12 (5.8)	117 (7.1)	
	Syphilis	0	0	0	0	
	Chlamydia, gonorrhea	11 (12.9)	95 (11.2)	12 (5.8)	149 (9)	
	Chlamydia, syphilis	0	0	1 (0.5)	1 (0.1)	
	Gonorrhea, syphilis	0	1 (0.1)	0	2 (0.1)	
	Chlamydia, gonorrhea, syphilis	0	1 (0.1)	0	0	
	Unknown	0	0 (0)	4 (1.9)	0	

* Unknown (Unk) refers to those individuals who remained anonymous after study completion or those whose gender could not be determined from given names
† Iqr = interquartile range
‡ Ethnicity for those with unknown gender is derived from variable Study Ethnicity (refer to Table 8)
N/A Data not available

4.1.3 Description of the Supplementary Questionnaire Population

The supplementary questionnaire population (refer to **Section 2.2**) consisted of 857 unique individuals; of which 46.8% were cases and 53.2% were contacts. Of the 401 cases, 31.9% were males and 59.3% were females (**Table 12**). The majority of males and females reported not graduating from secondary school. A higher proportion of females (44.9%) reported receiving their primary source of income through government or personal support, than males (25.9%).

Further breakdown of ethnic background revealed that the majority of both males and females were of an ethnic minority, 60.9% and 53.8%, respectively. The majority of cases with unknown gender identified as Caucasian (65.7%). When comparing the distribution of males (16.4%) and females (13%) who self-reported as being First Nations Treaty when registering for the Manitoba Health Services Insurance Plan (*CDC ethnicity*, **Table 12**), to males (21.1%) and females (16%) self-identifying as such during participation in the supplementary questionnaire (*study ethnicity*), it is apparent that the CDC designation undercounts registered treaty First Nations.

The infections acquired were similar to those reported for the entire study population (**Table 9**).

Table 12: Demographic and clinical characteristics of laboratory-confirmed cases with sexually transmitted infections identified from the supplementary questionnaire population during the "Manitoba Sexually Transmitted Disease Study" in Manitoba, Canada, June 1, 2002 to September 30, 2003 (n=401)

		Cases (n=401)		
		Male (n=128)	**Female (n=238)**	**Unknown (n=35)***
Age (Years) Median (Iqr)†		21 (18, 28) (n=128)	20 (17, 24) (n=238)	20 (18, 23) (n=22)
Geographic region	Winnipeg	118 (92.2)	224 (94.1)	3 (8.6)
	Southern rural Manitoba	8 (6.2)	12 (5.0)	20 (57.1)
	Northern remote Manitoba	2 (1.6)	2 (0.8)	12 (34.3)
Education	Less than secondary	69 (53.9)	137 (57.6)	13 (37.1)
	Completed secondary	30 (23.4)	57 (23.9)	14 (40.0)
	College, university, trade school	22 (17.2)	35 (14.7)	7 (20.0)
	Unsure, refused, missing	7 (5.5)	9 (3.8)	1 (2.9)
Income (n=325)		n=85	n=205	n=35
	Regular work, other	55 (64.7)	105 (51.2)	22 (62.9)
	Support	22 (25.9)	92 (44.9)	13 (37.1)
	Illegal	7 (8.2)	4 (1.9)	0
	Unsure, refused, missing	1 (1.2)	4 (1.9)	0
Study ethnicity	Caucasian	44 (34.4)	105 (44.1)	23 (65.7)
	Asian	9 (7.0)	12 (5.0)	0
	Other	6 (4.7)	8 (3.4)	0
	Black	11 (8.6)	9 (3.8)	1 (2.9)
	First Nations (Treaty)	27 (21.1)	38 (16.0)	10 (28.6)
	Aboriginal (Non-Treaty)	7 (5.5)	23 (9.7)	0
	Metis	18 (14.1)	38 (16.0)	1 (2.9)
	Unsure, refused, missing	6 (4.7)	5 (2.1)	0
CDC ethnicity‡	Treaty Code	21 (16.4)	31 (13.0)	10 (28.6)
Infection acquired	Chlamydia	82 (64.1)	200 (84.0)	N/A
	Gonorrhea	25 (19.5)	15 (6.3)	
	Syphilis	0	0	
	Chlamydia, gonorrhea	21 (16.4)	18 (7.6)	
	Chlamydia, syphilis	0	1 (0.4)	
	Gonorrhea, syphilis	0	0	
	Chlamydia, gonorrhea, syphilis	0	0	
	Unknown	0	4 (1.7)	
Repeater		17 (13.3)	34 (14.3)	N/A
Co-infection present		19 (14.8)	12 (5.0)	N/A
Repeat nominee		13 (10.2)	9 (3.8)	N/A

* Unknown refers to those individuals who remained anonymous after study completion; those whose gender could not be determined from given names

† Iqr = interquartile range

‡ Ethnicity for those with unknown gender is derived from variable Study Ethnicity (refer to Table 8)

N/A Data not available

Of the 456 contacts, 53.3% were males, 18.2% were females, and 28.5% had unknown gender (**Table 13**). Again, the majority resided within the Winnipeg RHA; however, those with unknown gender were more likely to have no known address (96.1%). A higher proportion of males were reported to have completed secondary school and obtained some level of postsecondary education in comparison to females; however, as this was a proxy variable, there was a large proportion of index cases who reported that they were unsure of the highest level of education obtained by their sexual partner.

Table 13: Characteristics of contacts identified from the supplementary questionnaire population during the "Manitoba Sexually Transmitted Disease Study" in Manitoba, Canada, June 1, 2002 to September 30, 2003 (n=456)

		Contacts (n=456)		
		Male (n=243)	**Female (n=83)**	**Unknown (n=130)***
Age (Years) Median (Iqr)†		23 (19.5,28) (n=160)	21 (17,26) (n=29)	N/A
Geographic region	Winnipeg	181 (74.5)	49 (59.0)	5 (3.8)
	Southern rural Manitoba	20 (8.2)	8 (9.6)	0
	Northern remote Manitoba	10 (4.1)	5 (6.0)	0
	Outside Manitoba	15 (6.2)	7 (8.4)	0
	Missing/Unknown	17 (7.0)	14 (16.9)	125 (96.1)
Education	Less than secondary	84 (34.6)	31 (37.3)	28 (21.5)
	Completed secondary	82 (33.7)	15 (18.1)	42 (32.3)
	College, university, trade school	25 (10.3)	7 (8.4)	30 (23.1)
	Unsure, refused, missing	52 (21.4)	30 (36.1)	30 (23.1)
Study ethnicity	Caucasian	108 (44.4)	31 (37.3)	83 (63.8)
	Asian	16 (6.6)	6 (7.2)	0
	Other	13 (5.3)	2 (2.4)	4 (3.1)
	Black	19 (7.8)	6 (7.2)	5 (3.8)
	First Nations (Treaty)	34 (14.0)	15 (18.1)	18 (13.8)
	Aboriginal (Non-Treaty)	18 (7.4)	7 (8.4)	8 (6.1)
	Metis	32 (13.2)	10 (12.0)	5 (3.8)
	Unsure, refused, missing	3 (1.2)	6 (7.2)	7 (5.4)
CDC ethnicity‡	Treaty Code‡	15 (6.2)	1 (1.2)	18 (13.8)
Repeat nominees		22 (9.0)	6 (7.2)	N/A

* Unknown refers to those individuals who remained anonymous after study completion or those whose gender could not be determined from given names
† Iqr = interquartile range
‡ Ethnicity for those with unknown gender is derived from variable Study Ethnicity (refer to Table 8 for definitions)
N/A Data not available

4.1.4 Sociodemographic and Behavioural Characteristics of Participants in the Supplementary Questionnaire

Over 55% of all study cases reported experiencing symptoms (**Table 14**). The observed proportions exceeded estimates reported in *Canadian STD Guidelines* stating that greater than 50% of males and 70% of females with chlamydial infection, and greater than 50% of both males and females with gonococcal infection may be asymptomatic (3). However, in comparison to other literature sources, the proportion of individuals experiencing symptoms is not on the high-end. For example, in men, incidence of asymptomatic gonorrhea infections has been reported to range from 1 to 3% (140).

More importantly, the presence of symptoms seems to be influenced by the population studied; number of days after exposure; presence of co-infection; and, arguably, the infecting strain. Previous estimates in this population indicated high proportions of individuals experiencing symptoms. Specifically, in a retrospective study conducted by Jolly (1998) the proportion of symptomatic laboratory-confirmed cases identified from 1990 to 1992 was compared between repeaters and non-repeaters of the following infection categories: chlamydia, gonorrhea and co-infected (93). Greater than 53% of all individuals in all groups experienced symptoms. Reporting of symptoms was lowest for chlamydia non-repeaters and highest for co-infected repeaters, 53% and 80%, respectively. However, one can not discount that participation bias may have resulted in a higher proportion of individuals reporting symptoms. Specifically, individuals with symptoms may have experienced a greater sense of disease burden, as such, were more willing to participate in this research. Additionally, the questionnaire incorporated reading off gender-specific symptoms to participants, which might have influenced the proportion of individuals reporting symptomatic infections.

Table 14: Behaviours of interviewed persons with STIs identified during the "Manitoba Sexually Transmitted Disease Study" (n=325)

		Male (n=85)	Female (n=205)	Unknown (n=35)*
Symptoms	Yes	61 (71.8)	113 (55.1)	21 (60.0)
	No	24 (28.2)	91 (44.4)	13 (37.1)
	Unsure	0	1 (0.5)	1 (2.9)
Symptomatic Sex†	Yes	24 (39.3)	76 (66.7)	14 (63.6)
	No	33 (54.1)	33 (28.9)	6 (27.3)
	Unsure, missing	4 (6.6)	5 (4.4)	2 (9.1)
Healthcare delay†	0-7	26 (42.6)	36 (31.6)	6 (27.3)
	8-30	21 (34.4)	39 (34.2)	4 (18.2)
	≥31	7 (11.5)	17 (14.9)	6 (27.3)
	Unclear, missing	7 (11.5)	22 (19.3)	6 (27.3)
Reason for healthcare visit‡	Symptoms	54 (63.5)	77 (37.6)	10 (28.6)
	Contact, prenatal, routine, other	29 (34.1)	117 (57.1)	25 (71.4)
	Unsure, missing	2 (2.3)	11 (5.4)	0
Alcohol during sex (previous 12-months)	Always	5 (5.9)	10 (4.9)	1 (2.9)
	Sometimes	57 (67.1)	144 (70.2)	30 (85.7)
	Never	21 (24.7)	48 (23.4)	3 (8.6)
	Unsure, refused, missing	2 (2.3)	3 (1.5)	1 (2.9)
Drugs during sex (previous 12 months)	Always	12 (14.1)	10 (4.9)	1 (2.9)
	Sometimes	32 (37.6)	63 (30.7)	10 (28.6)
	Never	38 (44.7)	128 (62.4)	23 (65.7)
	Unsure, refused, missing	3 (3.5)	4 (1.9)	1 (2.9)
Injection drug use (previous 12-months)	Yes	0	1 (0.5)	0
	No	83 (97.6)	200 (97.6)	35 (100)
	Unsure, refused, missing	2 (2.3)	4 (1.9)	0
Age first sex (years)	**Median (Iqr)§**	15 (14, 17) (n=70)	15 (14, 16) (n=184)	16 (15, 17.5) (n=32)
Lifetime sexual partners	1 to 5	15 (17.6)	75 (36.6)	15 (42.9)
	6 to 19	32 (37.6)	87 (42.4)	15 (42.9)
	20 to 99	33 (38.8)	36 (17.6)	3 (8.6)
	≥ 100	4 (4.7)	3 (1.5)	0
	Unsure, refused, missing	1 (1.2)	4 (1.9)	2 (5.7)
Sexual partners in previous 12-months	1	16 (18.8)	65 (31.7)	9 (25.7)
	2	18 (21.2)	49 (23.9)	10 (28.6)
	3 to 4	20 (23.5)	44 (21.5)	9 (25.7)
	5 to 9	12 (14.1)	24 (11.7)	5 (14.3)
	≥10	16 (18.8)	18 (8.8)	1 (2.9)
	Unclear, missing	3 (3.5)	5 (2.4)	1 (2.9)

* Unknown refers to those individuals who remained anonymous after study completion or those whose gender could not be determined from given names

† Percentages calculated from n=197 individuals who participated in the questionnaire who answered 'yes' or 'unsure' to presence of symptoms (61 males, 114 females, 22 unknowns)

‡ Refer to variable derivations, **Appendix K**

§ Iqr = Interquartile range

Of those who were symptomatic, a high proportion continued to have sex before antimicrobial treatment. This proportion was higher in females and those with unknown gender, 66.7% and 63.6%, respectively, in comparison to males (39.3%). This behaviour has been commonly reported in other populations. Specifically, 38% of men and 46% of women who attended an STD clinic in Baltimore, reported continuing sexual activities (141). Furthermore, Yorke *et al.* (1978), theorized that the *core group* consisted of members who continue to have sex while experiencing symptoms (54).

Over 10% of all participants reported a delay in seeking healthcare of 31 days or more. This, combined with the continuance of sexual activity, could result in the continued transmission of STI. Additionally, it increases the likelihood of infected individuals experiencing adverse sequelae. This is a common behaviour reported in many studies. In both men and women, self-treatment is one such reason for delaying diagnosis and treatment. However, the often mild and non-specific symptoms experienced by females, often mistakenly attributed to the menstrual cycle or pregnancy, may explain both the high frequency of symptomatic sex and the delay in seeking healthcare in this population. This is corroborated by findings which indicate that Canadian adolescents have limited knowledge of sexual health. Specifically, Langille *et al.* (1998) examined the sexual health knowledge of high-school youth aged 14 to 19 years and found that fewer than one-half of participants realized that chlamydia infection is often sub-clinical or asymptomatic; and only one-quarter believed chlamydia infection to be a common problem experienced by their peers (142).

Greater than 60% of males visited the doctor because of symptoms; whereas females were more likely to visit for other reasons, such as prenatal examinations (greater than

50%). This reinforces the need for adherence to provincial STI screening guidelines, in which all females are to be tested at least once during pregnancy.

Over 70% of all participants reported some alcohol use before or during sex, while greater than 30% reported drug use before and during sexual activities. This is of concern, as it is thought that alcohol and drug consumption is an impediment to safe sex practices, specifically condom use and may be associated with risky sexual behaviours (143;144). In an extensive review of the relationship between alcohol use and risky sexual behaviour, Halpern-Felsher *et al.* (1996) concluded that alcohol use at first-time sexual intercourse increases risky sexual behaviour; however, evidence was limited for other types of sexual relationships (145). The establishment of an estimate of alcohol and drug use before and during sexual activities in this high-risk population provides a greater understanding of the frequency of engagement in risk behaviours.

The age of first consensual sex was quite low. It should be noted that age at first sexual intercourse implies that an individual initiated sexual relations at this point in time. However, other high-risk sexual behaviours could have preceded this act (146). Nonetheless, age of first consensual sex is an important risk factor, as it affects an individual's susceptibility to infection. Additionally, it is a risk marker for other high-risk sexual behaviours, such as the number of partners in previous 12-months.

Males reported both higher numbers of recent and lifetime sex partners, in comparison to females. However, for both genders, the distribution of recent sex partners differed from that reported in the Sexual Behaviour Module of the 2000/2001 Canadian Community Health Survey, in which 91% of individuals reported only having one sex partner in the previous 12-months (146).

Further characterization of the population by dichotomizing ethnicity into Caucasian and Aboriginal groupings was completed (**results not shown**). Aboriginal males reported higher numbers of recent partners in comparison to Aboriginal females, while Caucasian males and females had lower numbers of partners than Aboriginals. This difference may reflect the age difference between the two populations; Aboriginals tend to be younger than Caucasians. This disassortative mixing pattern observed in the Aboriginal population, with respect to partner number, has also been reported in other ethnic groups, specifically African American (51). Additionally, a higher proportion of sexual relationships occurred within these two ethnic groups; possibly concentrating STIs within the Aboriginal population.

Once an individual consented to participate, item non-response was quite low. In previous studies, this has been attributed to the use of experienced interviewers - a factor considered in the current study (147). However, it could also in part be due to differences in sexual attitudes between those participating and those refusing to participate; the former having more liberal attitudes, hence more likely to disclose sensitive information. Biases can not be discounted, including reporting and social desirability bias. Techniques were used to minimize bias, including: limiting the period of recall; using trained interviewers; asking questions with definite answers; provision of information regarding the purpose of the study and information to be collected; and ensuring and emphasizing confidentiality.

Concerns have been raised regarding the reliability of: 1) participants' proxy reports on contacts' sexual risk behaviours and demographic characteristics; and 2) participants' self-reports about their sexual relationships. In a study by Neaigus *et al.* (1996)

examining the personal networks of recent injection drug users, participants' reports on network member's sociodemographic and risk behaviours were reported to be quite reliable (73). Additionally, studies have indicated that self-reported sexual histories can be reliable (148-153).

4.1.5 Characterization of Case-Contact Pairs

The 325 participating laboratory-confirmed chlamydia and/or gonorrhea cases reported on a total of 547 case-contact pairs or dyadic relationships (refer to **Section 1.5.1**). This information permitted the characterization of an individual's risk network, and is presented in **Table 15** (82). Overall these estimates are influenced by individuals reporting a large number of partners in the preceding 12-months. However, of the 325 individuals only a small proportion reported greater than six recent sex partners.

Further examination of the case-contact pairs revealed several high-risk behaviours, which increase the risk of STI and HIV transmission. Specifically, 38.6% of cases met their sexual contacts through family or friends, an indicator of the importance of the social context of relationship formation. The second most frequently reported place of meeting was at a bar, hotel or restaurant (21.6%). However, sexual partners were met outside of an individual's social group, for example through telepersonals or online (1.1%), or through sex trade work (3.7%). Although small in number, these partnerships would enable bridging between different social groups. The social context in which a sexual relationship occurs influences behaviours associated with the risk of STI. For example, individuals meeting their partners through sex trade work are less likely to negotiate condom use and more likely to engage in high-risk sexual behaviours, whereas casual partnerships imply decreased familiarity with a specific partner. Of interest was the high proportion of individuals reporting that their sexual activities occurred within the context of a steady relationship (spouse or boy- or girlfriend) (33.5%). This is of concern as it indicates that some individuals may have been indirectly exposed to a STI - i.e. through a partner involved in concurrent (overlapping) sexual relationships. Frequency of sex in the previous 3-months paralleled the distribution of relationship types. The two highest ranked frequencies were less than once per month (34.5%) and 2 to 3 times per week (24.3%).

Table 15: Characteristics of case-contact pairs as reported by participating laboratory-confirmed cases (n=325) of chlamydia and/or gonorrhea, Manitoba, Canada from June 1, 2002 to September 30, 2003 (n=547)

Variable		**Freq (%)**
Place of first meeting	Bar/hotel/restaurant	118 (21.6)
	Private residence	58 (10.6)
	Family/friends	211 (38.6)
	School	55 (10.1)
	Sex trade, on the street, bathhouse	20 (3.7)
	Work	28 (5.1)
	Telepersonals/online	6 (1.1)
	Other	43 (7.9)
	Unsure/missing	8 (1.5)
Relationship type	Spouse	18 (3.3)
	Boy/girlfriend	165 (30.2)
	Casual sex	190 (34.7)
	"On and off"	35 (6.4)
	Ex-lover/spouse	127 (23.2)
	Client/trick	1 (0.2)
	Unsure/missing	11 (2.0)
Received anything for sex	Drugs	3 (0.5)
	Money	8 (1.5)
	Food, shelter, clothing, other	6 (1.1)
	No	526 (96.2)
	Unsure/missing	4 (0.7)
Given anything for sex	Drugs	0
	Money	0
	Food, shelter, clothing, other	1 (0.2)
	No	537 (98.2)
	Unsure/missing	9 (1.7)
Length of acquaintanceship before sex	0 days	67 (12.2)
	1 to 7 days	74 (13.5)
	8 to 30 days	102 (18.6)
	31 to 91 days	74 (13.5)

Variable		Freq (%)
	92 to 365 days	107 (19.6)
	> 1 year	107 (19.6)
	Unclear/missing	16 (2.9)
Frequency of sex in previous 3-months	Everyday	68 (12.4)
	2-3 times/week	133 (24.3)
	About once/week	56 (10.2)
	1-3 times/month	56 (10.2)
	Less than once/month	189 (34.5)
	0 times	26 (4.7)
	Unsure/refused/missing	19 (3.5)
Sex with acquaintance	Yes	198 (36.2)
	No	282 (51.5)
	Unsure	59 (10.8)
	Refused/missing	8 (1.4)
Condom use	Never	233 (42.6)
	Less than 1/2 time	115 (21)
	About 1/2 time	50 (9.1)
	More then 1/2 time	52 (9.5)
	Always	79 (14.4)
	Unsure/missing	18 (3.3)
Discussed HIV	Yes	170 (31.1)
	No	371 (67.8)
	Unsure/refused/missing	6 (1.1)
Discussed STI	Yes	264 (48.3)
	No	276 (50.5)
	Unsure/refused/missing	7 (1.3)

A small, but concerning number of relationships involved the exchange of money, drugs, or other items for sex. This percentage is most likely a conservative estimate.

Over one-half of participants knew their sexual partner for more than one month before beginning a sexual relationship; while 19.6% were socially acquainted for a period of one-year or greater. A similar distribution was reported in a previously published study conducted by Potterat *et al.* (1985) examining the social context of gonorrhea transmission (62). Specifically, 45% of case-contacts pairs had an acquaintance period of two months or greater and 18% had an acquaintance period greater than one-year. This data indicates that high-risk behaviours occur between individuals who are socially acquainted.

Participants were asked if their sexual partner ever had sexual relations with anyone else they personally knew. This information is an indicator of the density or connectedness of a participant's sexual network, and the extent to which it is embedded within the larger social network. A large proportion of individuals reported that sexual relations had occurred between their partner and known acquaintances (36.2%).

Only 14.4% of participants reported that they always used a condom with their sexual partner. Communication between case-contact pairs was quite low. Specifically, less than 50% of participants ever discussed their partner's history of STIs; and only 31.1% discussed their partner's HIV status. However, these numbers may be even lower, as it is more socially desirable to report frequent use of condoms, and discussion of HIV status and STI history.

4.2 Objective 2

To construct components of individuals linked directly or indirectly through sex, and describe basic social network analysis measures and component summaries.

The characterization of case-contact pairs is important in understanding the transmission of STIs. However, these relationships are influenced by the behaviours and social norms of the sexual network in which they are embedded. Hence, further characterization of the sexual network was necessary. Routinely collected partner notification data was used to construct the sexual network. Many studies, both theoretical and empirical, support the use of this type of data source. Specifically, Ghani *et al.* (1998) evaluated three methods of sampling a simulated network of sexual contacts (154). The authors reported that networks sampled through partner notification resulted in the least biased estimates of network measures. Furthermore, this method was preferred over others as: 1) it captures the characteristics of individuals most at risk of acquiring STIs; and 2) it provides both a reason and technique for the collection of sensitive information (154).

4.2.1 Size and Frequency of the Components Identified from the Entire Sexual Network

The study data combined with the complete notifiable disease data revealed a large sexual network consisting of 2,508 connected components of size two or greater (**Table 16**). Sixty percent of the components were dyads, and 23% were triads (for definitions refer to **Section 1.5.1**). Thirty-five components of size 10 or greater were identified – the largest containing 33 individuals.

The frequency distribution of network components identified during the current study was compared with results obtained during the previous study (November-1997 to May-1998) (66). Both studies yielded comparable variation in component size. Specifically, in the previous study 61% and 24% of the 1,503 components were dyads and triads, respectively (**Figure 7**). This observed distribution is a common feature to this data. Potterat *et al.* (1985) speculated that the observed segmentation results from the tendency of individuals to choose sexual partners similar to themselves which leads to distinctive component boundaries (62). It has also been suggested that a population with a low incidence of infection is characterized by this component distribution (155)

Table 16: Frequency distribution of network components identified during the "Manitoba Sexually Transmitted Disease Study", Manitoba, Canada, June 1, 2002 to September 30, 2003 (n=2,508)

Component size	Freq (%)	Cumulative percent of individuals
2	1512 (60.3)	3,024 (41.5)
3	577 (23)	4,755 (65.3)
4	174 (6.9)	5,451 (74.9)
5	89 (3.6)	5,896 (81)
6	53 (2.1)	6,214 (85.4)
7	34 (1.4)	6,452 (88.7)
8	20 (0.8)	6,612 (90.9)
9	14 (0.6)	6,738 (92.6)
10	6 (0.2)	6,798 (93.4)
11	7 (0.3)	6,875 (94.5)
12	5 (0.2)	6,935 (95.3)
13	1 (0)	6,948 (95.5)
14	2 (0.1)	6,976 (95.9)
15	1 (0)	6,991 (96.1)
16	2 (0.1)	7,023 (96.5)
17	2 (0.1)	7,057 (97)
18	1 (0)	7,075 (97.2)
19	1 (0)	7,094 (97.5)
20	1 (0)	7,114 (97.8)
22	1 (0)	7,136 (98.1)
25	1 (0)	7,161 (98.4)
28	1 (0)	7,189 (98.8)
30	1 (0)	7,219 (99.2)
32	1 (0)	7,251 (99.6)
33	1 (0)	7,284 (100.1)
TOTAL	**2,508**	**7,284 (100)**

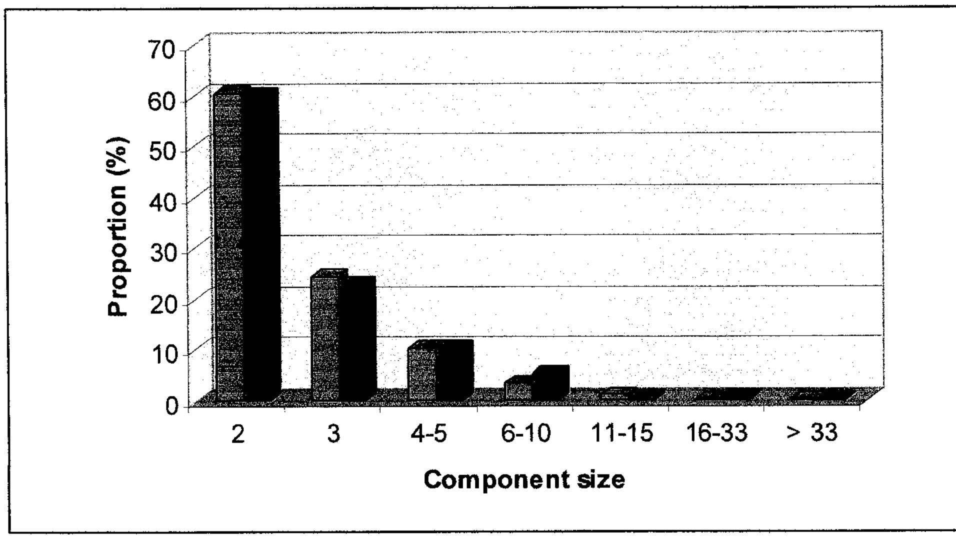


Figure 7: Comparison of the size of network components identified during the previous network study, November-1997 to May-1998 (n=1,503) and present study "Manitoba Sexually Transmitted Disease Study", Manitoba, Canada, June 1, 2002 to September 30, 2003 (n=2,508) using PAJEK, lighter and darker colours, respectively.

Additionally, the present study identified 1,192 laboratory-confirmed cases nominating no partners (**Table 17**). Unnamed partners are likely a result of: illicit behaviours, or fear of violence or abandonment; no previous partners within the tracing period (reflecting prevalent cases picked up by screening); concerns regarding anonymity; refusal to nominate partners; names not asked by healthcare provider; re-infection; and/or a lack of knowledge about untreated partners.

Isolates were further examined to determine if gender differences existed (**Table 17**). Of these 1,192 isolates, 33.4% were males and 66.6% were females. A higher proportion of isolate females (31.4%) were First Nations Treaty, in comparison to males (23.9%). Additionally, a larger proportion of isolate females (30.9%) resided in the health jurisdictions contained within northern remote Manitoba, than isolate males (21.4%).

Table 17: Summary demographic, geographic and clinical characteristics of isolate cases (n=1,192) identified from complete sexual network (study data combined with complete notifiable data) in Manitoba, Canada, June 1, 2002 to September 30, 2003.

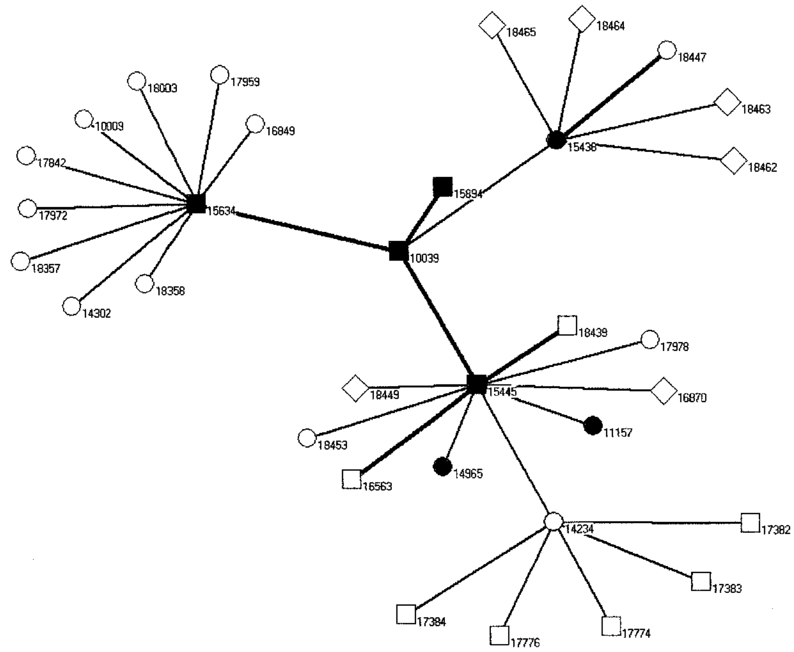
		Males (n=398)	**Females (n=794)**
		Freq (%)	**Freq (%)**
Age (Years) Median (Iqr)†		24 (20,31) (n=398)	20 (18,24) (n=793)
Treaty code		95 (23.9)	249 (31.4)
Geographic region	Winnipeg	230 (57.8)	362 (45.6)
	Southern rural Manitoba	75 (18.8)	169 (21.3)
	Northern remote Manitoba	85 (21.4)	245 (30.9)
	Outside Manitoba	8 (2)	17 (2.1)
	Missing	(0	1 (0.1)
Infection acquired	Chlamydia	269 (67.6)	688 (86.7)
	Gonorrhea	99 (24.9)	57 (7.2)
	Chlamydia and gonorrhea	30 (7.5)	49 (6.2)
Co-infection		29 (7.3)	48 (6.1)
Repeater		10 (2.5)	19 (2.4)

Problems with partner notification exist. One such problem is the likely underestimation of network size, occurring when: cases are undetected; clients fail to report all contacts; contacts are not located and subsequently tested; or when the interview period does not capture epidemiologically relevant contacts. However, the use of provider referral and experienced public health nurses during the partner notification procedure may serve to increase network completeness. Additionally, some contacts will become cases, so multiple interviews result in a more complete network structure.

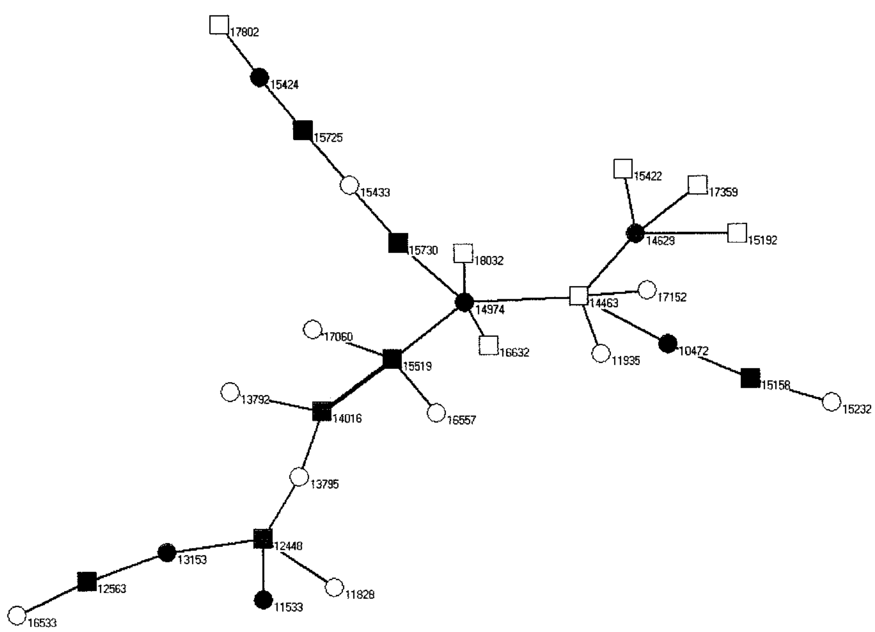
Additionally, Brewer *et al.* (1999) found that there is no significant difference between recalled and forgotten partners with respect to sociodemographic, clinical and behavioural variables (156).

4.2.2 Basic Visualization and Characterization of Identified Components

Further examination of the five largest components revealed the existence of same sex partnerships, a network feature never observed in previous studies of this population. Of particular interest was component 9 consisting of 33 individuals. Eighteen percent of its members identified as First Nations Treaty; the age range within this component was 34 years; and both chlamydia and gonorrhea were being transmitted (**Figure 8A**). Three of the 5 study cases reported that their main source of income was from sex trade work, or that they had exchanged sex for either money or drugs. Component 9 spanned three geographical locations in Manitoba, but was primarily focused in Winnipeg. Component 430 consisted of 30 members. A high proportion of its members identified as First Nations Treaty (40%). It spanned 10 locations in Manitoba (including Winnipeg), and had the unique feature of same sex relationships. Both chlamydia and gonorrhea were circulating within this component.



A)



B)

Figure 8: Component 9 and 430, A and B, respectively. Each symbol represents an individual; males and females are designated by circles and squares, respectively. Diamonds represent individuals with unknown gender. Colours indicate infection, chlamydia = red, gonorrhea = green, both chlamydia and gonorrhea = plum, and white indicates unknown infection. Lines represent sexual links; red lines = same sex links and black = heterosexual links.

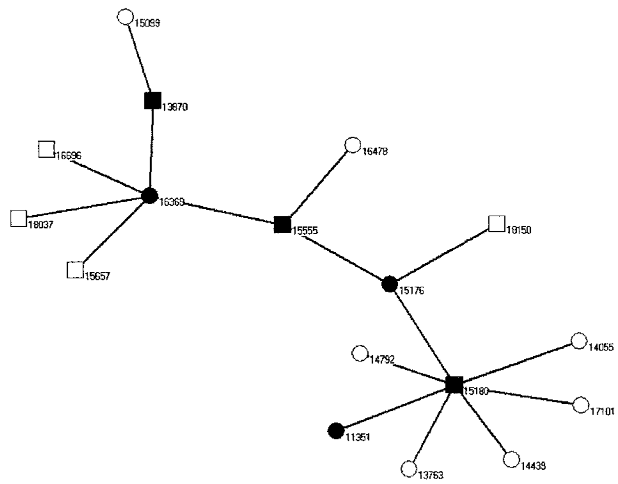
4.2.3 Characterization of Identified Components

The size and structure of a network influences both the transmission and maintenance of chlamydia and gonorrhea infections. As such, the characterization of components relative to size was completed using network, demographic, geographic and clinical summary measures, **Tables 18, 19, 20 and 21**, respectively.

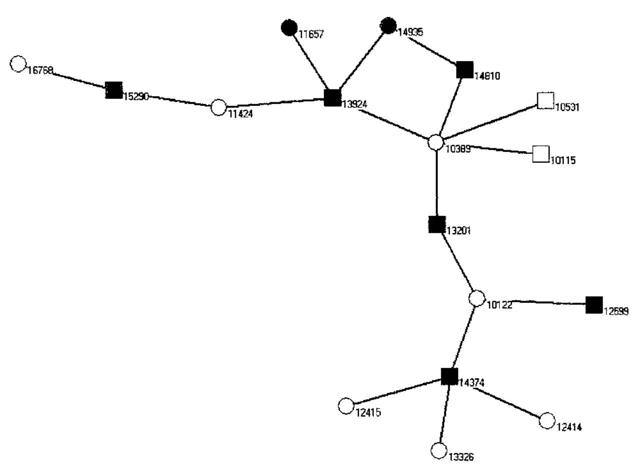
Table 18: Summary of network measures relative to component size for components of size 2 through 33 (n=2,508) identified from complete sexual network (study data combined with complete notifiable data) in Manitoba, Canada, June 1, 2002 to September 30, 2003.

Size	Freq	Mean density (range)	Mean degree centralization (range)	# k-cores (k = 2) (size=4)	Mean % MSM or WSM links (range)	# of components w MSM or WSM links (%)
2	1512	1 (1,1)	0 (0,0)	-	3.17 (0,100)	48 (3.2)
3	577	0.67 (0.67,0.67)	1 (1,1)	-	2.08 (0,100)	19 (3.3)
4-5	263	0.47 (0.4,0.5)	0.75 (0.17,1)	0	2.19 (0,75)	15 (5.7)
6-10	127	0.29 (0.2,0.33)	0.62 (0.1,1)	1	1.75 (0,100)	8 (6.3)
11-15	16	0.17 (0.13,0.18)	0.41 (0.14,0.88)	0	0.62 (0,10)	1 (6.2)
16-33	13	0.1 (0.06,0.12)	0.32 (0.11,0.59)	3	2.55 (0,18.75)	4 (30.8)

Scott (2000) indicated that there is an upper limit to the number of relations that can be sustained by an individual as a result of both time and relationship constraints (79). As a result, the maximum number of relationships that could occur in a network would always be greater than the number of relationships observed. As such, the maximum density (equal to 1) of a graph would most likely never be observed. In the current study, the highest observed density of a component of size 3 or greater was 0.67. Additionally, as component size increased the density of a component decreased. Specifically, larger components had the lowest proportion of people with links to each other (**Table 18**). In a comparison of components 1091 and 109, both consisting of 17 members - component 109 had a greater density (0.125) than component 1091 (0.118). This difference is attributed to the presence of a 2-core of size 4 in component 109 (**Figure 9**). Although similar in size, component 109 had nine members infected with chlamydia and/or gonorrhea, while component 1091 had six members infected with chlamydia alone. Mean degree centralization, a measure of how centered a component is around one individual was used to objectively determine the structure of a component - i.e. if it has a linear or radial structure. Linear components, as previously described, were characterized by a low variation in the number of sexual partners within a component, or a component with no central individual. Whereas, a radial component is characterized by one central individual with a high degree centrality, connected to many individuals with one or two partners. As such, radial components have a large variance in degree centrality. Degree centralization was inversely related to component size (for those of size three or greater). This is an indication that larger components (mean degree centralization; 0.32) have a linear or branching structure in comparison to those of smaller sizes. However, variability was observed within each size grouping. For example, three distinct structures were observed in components of size 16 or greater - linear, composite and radial (**Figure 10**). It should be noted that time is a major limitation of this study. For example, components containing a case identified near the beginning of the study period are more likely to be larger those containing a case identified near the end of this period. As such, time may contribute to the identification of linear and radial component types.

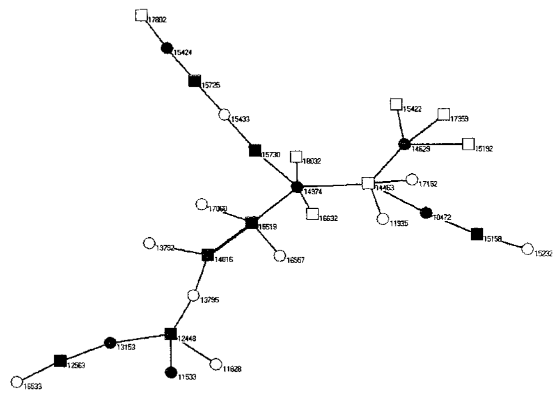


A)

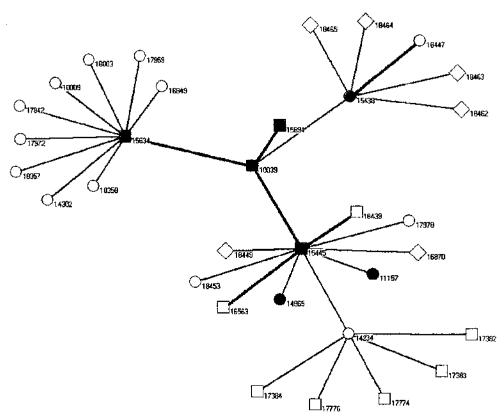


B)

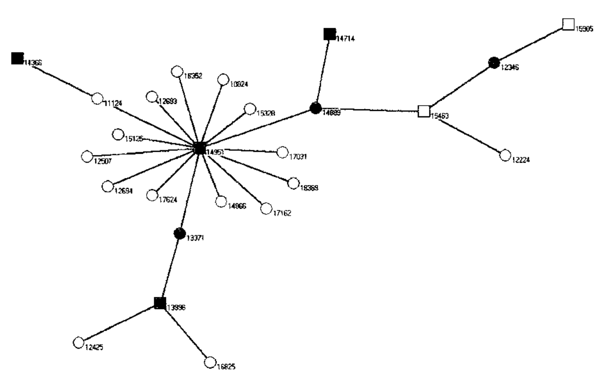
Figure 9 Component 1091 and Component 109, A and B respectively. Males and females are designated by circles and squares, respectively. Colours indicate infection, chlamydia = red, gonorrhea = green, both chlamydia and gonorrhea = plum, and white indicates unknown infection. Lines represent sexual links; red lines = same sex links and black lines = heterosexual links. Component 1091 has a density of 0.118 and 109 has a density of 0.125.



A) Component 430: Degree centralization = 0.11 (n=30)



B) Component 9: Degree centralization = 0.27 (n=33)



C) Component 714: Degree centralization = 0.59 (n=25)

Figure 10 Component 430, 9, and 714; A, B and C, respectively. Each symbol represents an individual; males and females are designated by circles and squares, respectively. A diamond represents an individual with unknown gender. Colours indicate infection, chlamydia = red, gonorrhea = green, both chlamydia and gonorrhea = plum, and white indicates unknown infection. Lines represent sexual links; red lines = same sex links and black lines = heterosexual links.

UCINET identified four 2-cores of size 4, in components of size 9, 17, 20 and 32 (**Figure 11**). The identification indicates an area of the component, with a greater density. No higher order microstructures were observed - most likely due to the content of data (primarily heterosexual relations). Additionally, as argued by Potterat *et al.* (2002), the absence of higher order microstructures may be representative of the epidemic phase— "low endemic rather than epidemic spread" (157). These structures have been associated with increased cohesion within a network or component, thus facilitating the transmission of STIs. In a radial network, if the central individual is targeted during control efforts, transmission to all others stops; while if a k-core is present targeting one individual will not stop the transmission of infection as the infection can travel through other paths. Unique to this dataset was the identification of same sex partnerships; 30.8% of the components of size greater than 15 exhibited this feature; contributing to a mean percentage of 2.5% of the links within these components. In total 123 same sex relationships were reported, 75 between men who have sex with men (MSM) and 48 between women who have sex with women (WSW). Further examination of the characteristics of the 102 unique individuals reporting same sex relations (59 males and 43 females) revealed that gender differences existed (**Appendix O**). Males had a lower mean degree, were members of smaller components, and were less often identified as repeaters (1.9, 3.8 and 12%, respectively), in comparison to females (2.6, 7.4, and 28%, respectively). Additionally, components including MSM (81%) had a higher proportion of same sex links than components containing WSW (56%); indicating that women reporting same sex relations often have a history of recent sex with men. This is of concern, as recent studies have indicated that females reporting both same and opposite sex relations are at increased risk of HIV in comparison to females reporting only opposite sex relations. The former groups are more likely to report a greater number of recent partners, injection drug and crack cocaine use, sex with high risk partners, and sex for money or drugs (158-160).

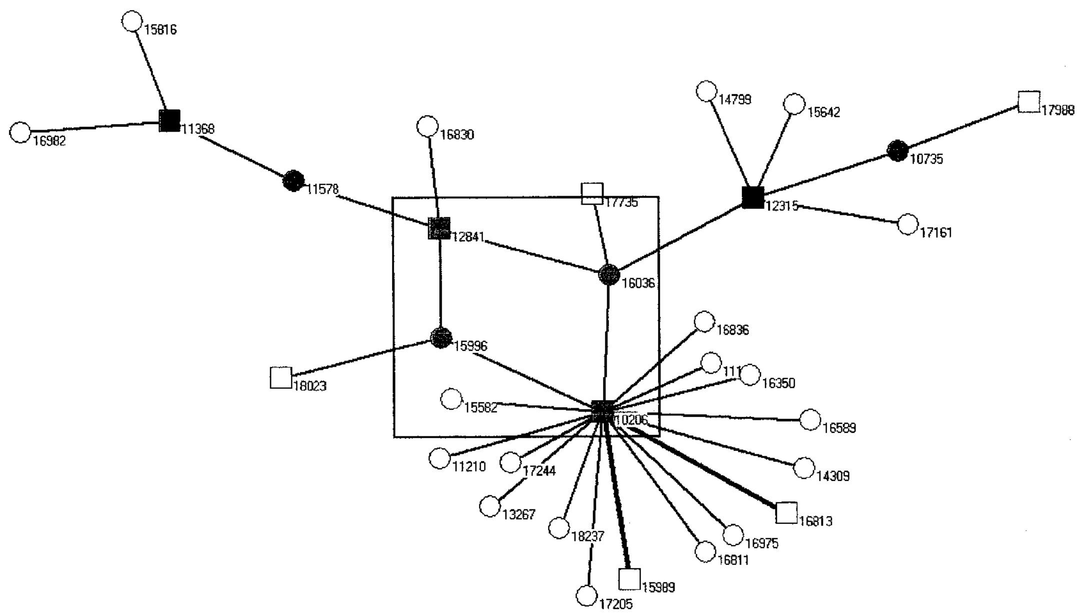


Figure 11: Component 190. Each symbol represents an individual; males and females are designated by circles and squares, respectively. Colours indicate infection, chlamydia = red, gonorrhea = green, both chlamydia and gonorrhea = plum, syphilis = orange, both gonorrhea and syphilis = brown, and white = unknown infection. Lines represent sexual links; red lines = same sex links and black lines = heterosexual links. The yellow box outlines a 2-core of size 4.

Table 19: Summary of demographic measures relative to component size for components of size 2 through 33 (n=2,508) identified from complete sexual network (study data combined with complete notifiable data) in Manitoba, Canada, June 1, 2002 to September 30, 2003.

Size	Mean % female (range)	Mean % Aboriginal (range)	Mean % Treaty (range)	Mean age (range)*	Mean age range (range)†
2	48.9 (0,100)	22.5 (0,100)	20.4 (0,100)	23.7 (13,64) n=1,506	2.6 (0,47)
3	42.7 (0,100)	19.2 (0,100)	15.9 (0,100)	23.2 (14,56) n=574	2.7 (0,30)
4-5	42.2 (0,80)	21.2 (0,100)	17.0 (0,100)	22.7 (14,45) n=259	5.2 (0,33)
6-10	42.7 (0,87.5)	25.9 (0,100)	20.3 (0,100)	21.4 (14,37.7) n=127	7.3 (0,40)
11-15	49.4 (18.2,72.7)	20.9 (0,64.3)	17.7 (0,64.3)	21.8 (15.8,39.2) n=16	12.8 (2,28)
16-33	45.4 (24,64.3)	43.3 (3.6,82.3)	35.7 (0,82.3)	21.5 (15.7,43.4) n=13	18.7 (7,52)

* Mean age = the mean of the mean age of participants in a component. For example, if there were 4 components of size 3 and the mean age of members in each component = 15, 16, 17 and 18—the mean of the mean age for components of size 3 = 16.5

† Mean age range = the mean of the mean age range of participants in a component. For example, if there were 4 components of size 3 and the mean age range of the members in a component = 1 (minimum = 14, maximum = 15), 3, 15, 5—the mean of the mean age range for components of size 3 = 6.

Demographic characteristics of components relative to network size are presented in **Table 19**. Larger components (16 to 33) contained the highest mean proportion individuals identifying as First Nations Treaty (35.7%); and the age range within these components was wide. Upon further examination of the individuals within the largest components, the wide age range was a result of a few individuals whose age was 10 years or greater than the mean age. For example, the mean age of component 9 was 23 (range: 13, 47), however, two individuals were 47 years of age (18357 and 14302) (**Figure 12**). This disassortative mixing pattern has been thought to be a marker of sex work, survival sex or relationship power differentials, and has been associated with the transmission of STIs and HIV in a population (73;75;85). As component 9 was an identified study component, supplementary information was examined which indicated that members of this component were involved in injection drug use (10039) and three individuals (15438, 15445 and 15634) reported their main source of income was through sex work. Condoms were seldom used in all reported relationships; and either alcohol and/or drugs were used before or during every sexual act.

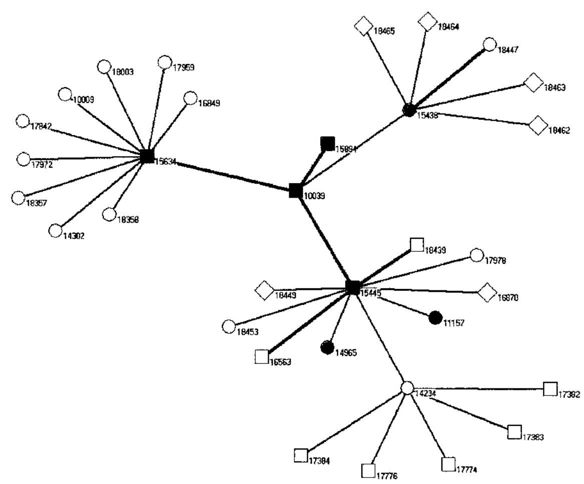


Figure 12 Component 9. Each symbol represents an individual; males and females are designated by circles and squares, respectively. Diamonds represent individuals with unknown gender. Colours indicate infection, chlamydia = red, gonorrhea = green, both chlamydia and gonorrhea = plum, both gonorrhea and syphilis = brown, and white indicates unknown infection. Lines represent sexual links; red lines indicate same sex links and black indicate heterosexual links.

Additionally, due to the large proportion of contacts with missing age, the summary measures for this variable should be interpreted with caution. Nonetheless, most data of this nature suffers from this problem and the data found in this database is most likely more complete than other sources. Hence, these measures do provide insight into the transmission dynamics.

However, other methods have been suggested to deal with casual or anonymous partnerships. For example, network approaches may be used to identify where partners met or where sexual encounters occur. Specifically, the role of social venues in the spread of STIs has been empirically exhibited by both Potterat *et al.* (1985) and De *et al.* (2004) (62;88). In addition to the collection of routine epidemiological data in the examination of an outbreak of gonorrhea, De *et al.* (2004) collected information on preferred social establishments (88). Using social network analysis, the authors used the bar as a node in the network, and connected all cases identifying this specific establishment as a preferred site of socialization. The resultant network connected 8 separate components resulting in a total of 89 individuals (49% of the study population).

In an attempt to examine the importance of social venues in the spread of disease, components with limited information were connected via a commonly named social venue. Specifically, in seven distinct components, one member in each named a common social venue as their first point of contact with a sexual partner. The size of the seven components ranged from 3 to 7 members. When the social venue, in this case a bar, was incorporated into the network as a node - the network consisted of 34 individuals (**Figure 13**). This indicates that specific social venues may be an ideal location for prevention efforts. Furthermore, even when there is limited data, for example, when a partner's name or locating information is unknown, the place where the met may be invaluable.

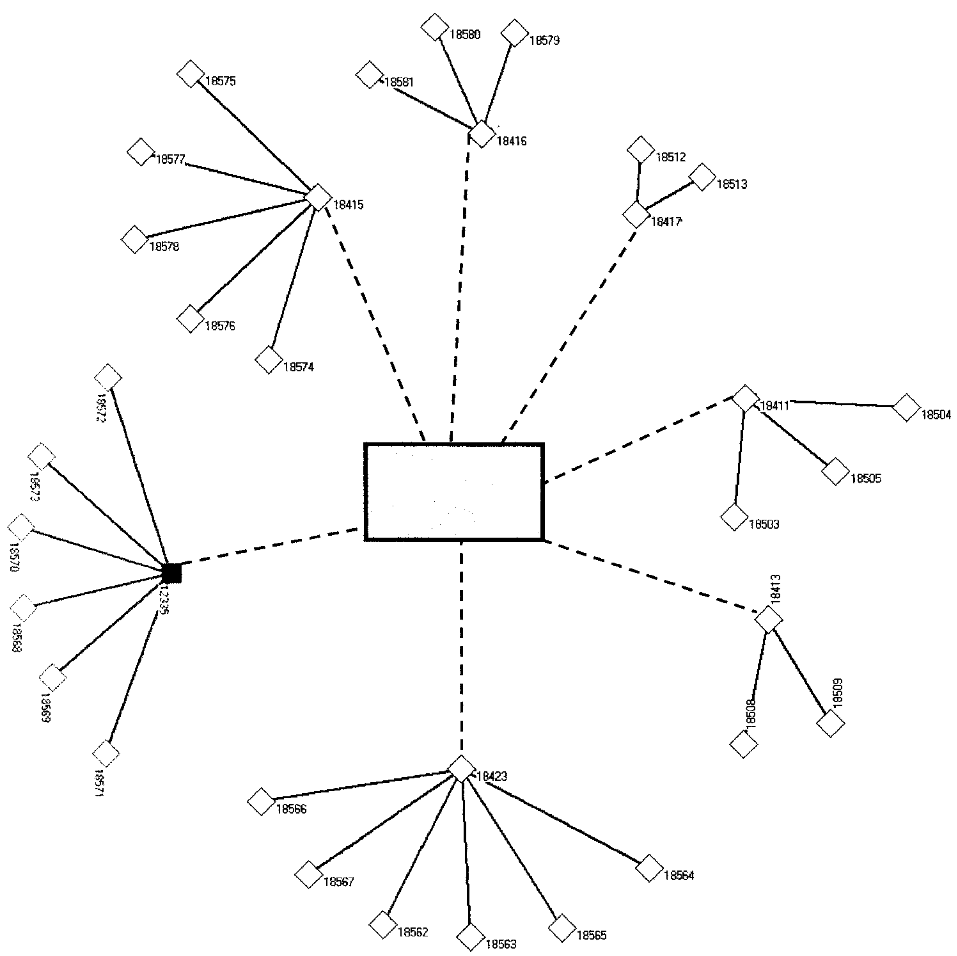


Figure 13 A network (n=34) consisting of seven components connected by a common place of meeting. Each symbol represents an individual; males and females are designated by circles and squares, respectively. A diamond indicates unknown gender. Colours indicate infection, chlamydia = red, and white indicates unknown infection. Lines represent sexual links, black = heterosexual links and dashed = connection to the bar. The bar is represented by the turquoise box.

Elliott *et al.* (2002) examined the geographical variation of chlamydia and gonorrhea infection in Manitoba, the authors identified two high-risk areas: northern remote Manitoba; and the inner city "core" in Winnipeg (161). The geographical summary of the components is presented in **Table 20**.

Table 20: Summary of geographic measures relative to component size, for components of size 2 through 33 (n=2,508) identified from complete sexual network (study data combined with complete notifiable data) in Manitoba, Canada, June 1, 2002 to September 30, 2003.

Size	Mean # distinct geographic locations (range)	Mean % in same region (range)	Mean % in Winnipeg (range)	Mean % in S. Rural (range)	Mean % in N. Remote (range)	Mean % outside MB (range)
2	1.3 (1,2)	86.9 (0,100)	51.0 (0,100)	17.1 (0,100)	24.9 (0,100)	3.4 (0,100)
3	1.6 (1,3)	82.2 (33.3,100)	46.6 (0,100)	19.7 (0,100)	23.6 (0,100)	4.7 (0,100)
4-5	1.8 (1,5)	79.6 (20,100)	42.8 (0,100)	17.1 (0,100)	28.6 (0,100)	3.4 (0,50)
6-10	2.3 (1,6)	79.7 (14.3,100)	40.2 (0,100)	17.3 (0,100)	30.8 (0,100)	4.1 (0,85.7)
11-15	3.5 (1,8)	73.2 (36.4,100)	42.1 (0,100)	18.6 (0,91.7)	26.4 (0,93.33)	2.1 (0,18.2)
16-33	3.8 (1,9)	84.6 (50,100)	43.0 (0,100)	12.0 (0,88.2)	41.8 (0,100)	0.7 (0,5.6)

A direct relationship is observed between component size and the mean number of distinct geographic locations within a component; components of size two have a mean number of 1.3 geographic locations, and those larger than 15 individuals had a mean of 3.8 geographic locations. Paralleling ethnic distribution, components with a higher proportion of individuals identifying as First Nations Treaty had a higher proportion of their members residing in jurisdictions contained within northern remote Manitoba (mean percent; 41.8%). The smaller components had a higher mean proportion of members outside of Manitoba; suggesting termination of the partner notification process. Further examination revealed that the majority of components spanned northern remote Manitoba and Winnipeg. As observed, a high percentage of individuals resided within the same geographic region. Potterat *et al.* (1985) referred to this as the "neighbourhood nature of sexual choices" (62). However, mixing between two regions has been indicated in the spread of infections. Specifically, Calvazara *et al.* (1998) reported that the most important discovery in her assessment of sexual partnering in First Nations communities in Northern Ontario was the discovery of inter-community mixing (162).

The clinical characteristics, relative to component size are presented in **Table 21.** Components sizes of 16 or more had the lowest proportion of cases with chlamydia (only) (47.8%). The largest component had the highest mean proportion of cases who acquired gonorrhea alone (13.9%); whereas only components of size six or greater contained individuals diagnosed with syphilis (data not shown). Both the mean proportion of cases with more than one infection during one disease episode (co-infected), and the proportion of those who were repeatedly infected with an STI during the course of the study were directly related to component size. Additionally, the proportion of individuals repeatedly nominated from 2002 to 2003 increased with component size.

Table 21: Summary of clinical characteristics relative to component size, for components of size 2 through 33 (n=2,508) identified from complete sexual network (study data combined with complete notifiable data) in Manitoba, Canada, June 1, 2002 to September 30, 2003.

Size	Mean % case (range)	Mean % CT cases (range)*	Mean % GC cases (range)*	Mean % CT&GC cases (range)*	Mean % co-infected cases (range)*	Mean % repeat cases (range)*	Mean % repeat nominee (range)*
2	55.9 (50,100)	83.0 (0,100)	9.5 (0,100)	6.6 (0,100)	5.6 (0,100)	5.8 (0,100)	1.0 (0,100)
3	43.7 (33.3,100)	82.8 (0,100)	9.1 (0,100)	6.3 (0,100)	5.0 (0,100)	7.5 (0,100)	3.1 (0,100)
4-5	42.8 (20,100)	75.7 (0,100)	8.7 (0,100)	10.8 (0,100)	10.3 (0,100)	13.6 (0,100)	8.3 (0,66.7)
6-10	38.9 (12.5,77.8)	69.5 (0,100)	10.2 (0,100)	16.2 (0,100)	14.3 (0,100)	15.2 (0,100)	12.0 (0,66.7)
11-15	45.6 (16.7,72.7)	73.6 (0,100)	6.1 (0,44.4)	14.0 (0,100)	8.6 (0,50)	23.4 (0,66.7)	19.8 (0,40)
16-33	42.6 (24.2,62.5)	47.8 (0,100)	13.9 (0,45.4)	30.6 (0,62.5)	23.8 (0,62.5)	22.9 (0,57.1)	20.9 (6.7,41.2)

* CT = chlamydia; GC = gonorrhea; CT&GC = chlamydia and gonorrhea

These characteristics are all markers of core group membership - however, they are also associated with component size. Theoretically, gonorrhea and syphilis require higher rates of sex partner change for transmission (56). Additionally, gonorrhea components are more likely to trace transmission routes, due to the higher proportion of infected individuals experiencing symptoms. Those with repeated infection have greater contact with the healthcare system. Through repeated contact, public health nurses are more likely to collect more detailed information regarding an individual's sexual network. For example, if a sexual partner was missed during the first partner notification process, if contact with the healthcare system occurs again, they may be nominated at the second appearance. However, those with repeated infection have been shown to be more likely to delay seeking healthcare. This delay, in combination with the high proportion of individuals continuing sexual activity while symptomatic, would potentially lead to a greater potential for the transmission of STIs to sexual partners and a larger network. Additionally, larger networks contained a higher proportion of First Nations Treaty, as previously stated, nurses employed by Health Canada who serve reserves are more likely to adhere to STI control procedures, therefore collecting more detailed information on sexual partners. However, a large proportion of isolates were registered First Nations Treaty or resided in northern remote Manitoba and tracing was not completed.

4.3 Objective 3

To define groups or clusters of components using cluster analysis, and analyze the differences in responses to social and sexual behavioural questions between the identified groupings.

4.3.1 Characteristics of Study Components

In total, 308 components containing at least one person from the study were contained within the observed 2,508 components (**Table 22**). Again, 60.7% of the components were dyads and 23.7% were triads. Identified study components of size 2, were embedded within larger components ranging in size from 2 to 28 members. The largest study component contained 19 individuals (embedded within a component of size 33). Components of size three or less contained 69.2% of the study network members. Further characterization of these 308 identified components is found in **Table 23**. Only 28.6% of the 308 components contained one or more individual(s) who obtained some level of post-secondary education. A total of 78 components (25.3%) contained individuals of an ethnic minority or non-Aboriginal ethnic background.

Table 22: Frequency distribution of network components identified during the supplementary questionnaire, June 1, 2002 to September 30, 2003 (n=308)

Size	Freq (%)	Cumulative freq of individuals in study network (%)	Size of complete identified component (Freq)
2	187 (60.7)	374 (43.6)	2 (141)
			3 (21)
			4 (11)
			5 (4)
			6 (3)
			7 (1)
			8 (1)
			9 (1)
			10 (1)
			11 (1)
			12 (1)
			28 (1)
3	73 (23.7)	593 (69.2)	3 (47)
			4 (8)
			5 (9)
			6 (5)
			7 (2)
			8 (2)
4	20 (6.5)	673 (78.5)	4 (15)
			5 (1)
			6 (1)
			7 (1)
			8 (1)
			18 (1)
5	8 (2.6)	713 (83.2)	5 (5)
			6 (1)
			7 (1)
			11 (1)
6	11 (3.6)	779 (90.9)	6 (4)
			7 (2)
			8 (1)
			9 (2)
			12 (1)
			30 (1)
7	6 (2)	821 (95.8)	7 (4)
			8 (1)
			22 (1)
8	1 (0.3)	829 (96.7)	11 (1)
9	1 (0.3)	838 (97.8)	16 (1)
19	1 (0.3)	857 (100)	33 (1)

Only 2.9% of components had one or more individuals reporting that their main source of income was from illegal activities; while 39.6% of components contained one or more individuals reporting that their main source of income was from support. Further examination of relationship characteristics found that 17.9% of the components contained individuals with concurrent partnerships. A small proportion of the components were economic, exchanging money, drugs or other items. A large proportion of the index cases within the components indicated that one or more of their sex partners had known sexual relations with an acquaintance. A large percentage of the 308 components contained individuals who did not use condoms during sexual activities.

Table 23: Study component summary using supplementary questionnaire information on both the attributes of the participating index case, and those of the sexual interrelationships (n=308)

Variable	Freq (%)
Individual attributes	
Postsecondary education	88 (28.6)
Ethnicity (other)	78 (25.3)
Injection drug use	16 (5.2)
Main source of income illegal	9 (2.9)
Main source of income support	122 (39.6)
Symptomatic sex	110 (35.7)
Healthcare delay greater than 7 days	64 (20.8)
Relationship attributes	
Concurrency	55 (17.9)
Received anything for sex	6 (1.9)
Given anything for sex	1 (0.3)
Partner ever had sex with acquaintance(s)	127 (41.2)
Condom use (never)	153 (49.7)
Discussed STI history (yes)	194 (63)
Discussed HIV status (yes)	129 (41.9)
First place of meeting (sex trade/street)	20 (6.5)
First place of meeting (public)	104 (33.7)
First place of meeting (private)	229 (74.3)
Type of relationship (steady)	164 (53.2)
Relationship type (casual)	202 (65.6)
Relationship type (sex trade)	1 (0.3)
Length of social acquaintance (0 days)	49 (15.9)
Sex trade component	19 (6.2)

4.3.2 Cluster analysis

In 1997, Wylie and Jolly successfully used SNA to further the understanding of chlamydia and gonorrhea transmission within this population (46;53;66;86;87). One finding was the existence of two structurally distinct component types - linear and radial components (**Appendix B**). More significant were the observed differences, including: higher aboriginal content; higher positivity rates; and both chlamydia and gonorrhea circulating within the former type. As this characterization was limited to the examination of basic demographic data, it was necessary to further explore the social determinants behind network formation. To further understand network formation within this population, a more detailed supplementary questionnaire was administered. Cluster analysis was appropriate as its purpose is to group together objects, in this case components, with a similar profile, thus enabling the differentiation of component types on the basis of variables external to the clustering procedure. Ethnicity was deemed an appropriate clustering variable. More specifically, as significant disparities with regards to STI infection exist between those of Aboriginal ethnicity and the general population, and it is known that being Aboriginal is not in itself a risk factor for STI infection, it was important to gain a better understanding of disease transmission within these networks. Additionally, as structural properties were observed in the previous study, degree centralization was used as a more formal approach to the distinction of these component types.

Univariate descriptions of the three variables used in the clustering algorithm are presented in **Table 24**. Plots and numerical information indicate that all 3 variables have a skewed distribution.

Table 24: Univariate description of three variables used in clustering algorithm (n=239 components)

Variable	Mean (Median)	SD	Range	Kurtosis	Skewness	Shapiro-Wilk (p-value)
Percent Treaty	0.20 (0.09)	0.25	0,1.0	0.35	1.19	0.78 (<0.0001)
Percent in same region	0.81 (0.83)	0.19	0.14,1.0	0.11	-0.88	0.88 (<0.0001)
Degree centralization	0.62 (0.58)	0.28	0.1,1.0	-1.13	0.02	0.91 (<0.0001)

Bivariate correlations were weak: a negative linear relationship was observed between percent First Nations Treaty and degree centralization, and between percent in same region and degree centralization (Pearson correlation coefficient (r): -0.27 and -0.17, respectively); and a positive linear relationship was observed between percent First Nations Treaty and percent in same region (r = 0.25).

The optimal clustering level, restricted to the final 10 hierarchical clustering levels, was determined by examining dendrograms (**Figure 14**) and two numeric stopping rules: the Pseudo F test, and the Pseudo T^2 (**Table 25**). The suggested stopping point for the Pseudo F and T^2 tests were observed absolute or local maxima, and observed local minima, respectively. When *p* was set at 0.20, no local maxima were observed for the Pseudo F test. When *p* was set at 0.10, two local maxima for the Pseudo F existed at clustering levels 5 and 3; whereas the Pseudo T^2 test indicated optimal clustering at levels 9, 7, 6, 5, and 3. When *p* was set at 0.05, local maxima were observed. However, the pooled within cluster covariance matrix was unstable. Upon examination of the resultant dendrogram (*p* = 0.10), large differences between consecutive r-squared (R^2) values occurred during the last 3 clustering levels; where R^2 represents the proportion of variance attributed to clustering (**Figure 14**). The dendrogram starts with all components in one cluster, and joins components or clusters one by one until all components are in one cluster. Based on the consensus between the stopping rules when p was set at 0.10, the dendrogram and past literature, the optimal number of clusters was determined to be 3. A scatter-plot of degree centralization versus percent First Nations (Treaty) displays the three identified clusters (**Figure 15**).

Table 25 Pseudo F and T^2 stopping rules obtained for clustering levels 10 through 1, when using Ward's Minimum Variance clustering method with three different input matrices corresponding to the three different starting values of p (0.2, 0.1, 0.05) used during the ACECLUS procedure. Clustering level 1 refers to point at which all n components (n=239) are grouped into one cluster.

	P = 0.2		P = 0.1		P = 0.05	
Clustering Level	**Pseudo F**	**Pseudo T^2**	**Pseudo F**	**Pseudo T^2**	**Pseudo F**	**Pseudo T^2**
10	187	33	485	28.2	677	17.1
9	191	14.4	506	18.9	698	22
8	191	31.3	518	25.1	710	35.2
7	191	28.9	520	30.7	716	70.9
6	196	75.5	546	63.2	706	53.6
5	210	27.5	564	66	716	30.4
4	238	48.7	535	89.3	732	99.7
3	240	94.7	595	67	489	132
2	319	66.3	540	259	543	423
1	.	319	.	540	.	543
Optimal level		9, 7, 4	5, 3	9, 7, 6, 5, 3, 2		

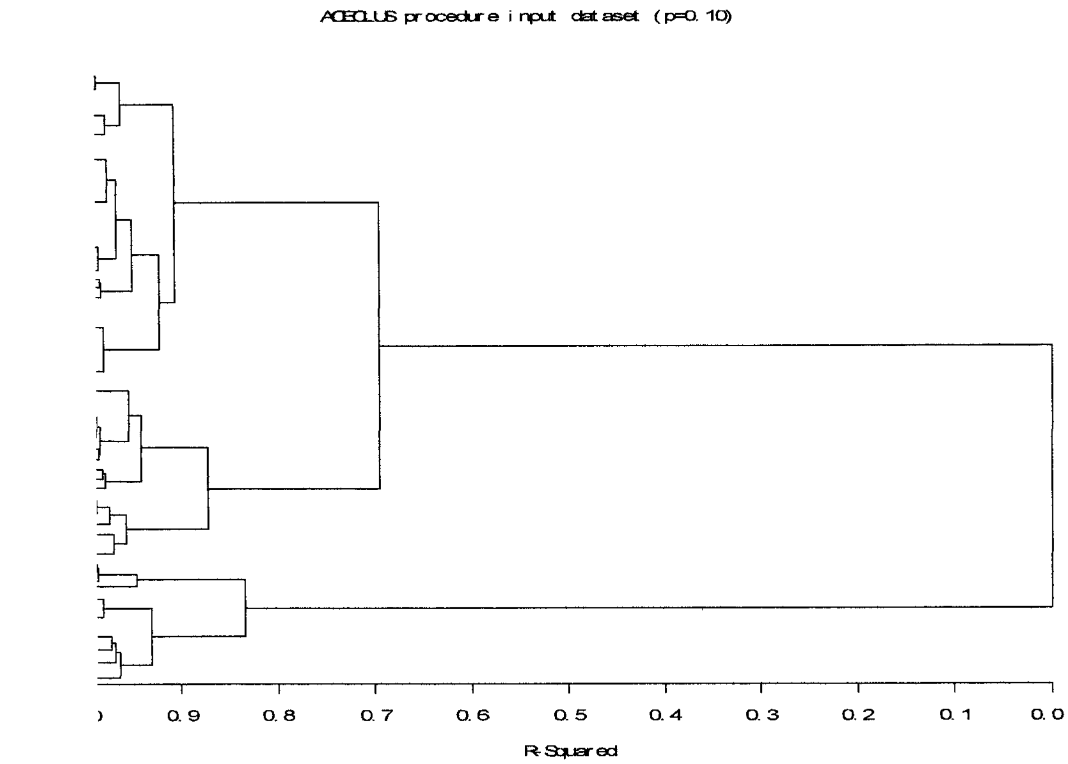


Figure 14 Dendrogram using input dataset variables from ACECLUS procedure, where *p* was set equal to 0.1. The hierarchical algorithm proceeds from left to right, n=239 clusters to n=1 cluster, respectively; each vertical line corresponds to the joining of two clusters and the respective R^2 is indicated on the horizontal axis.

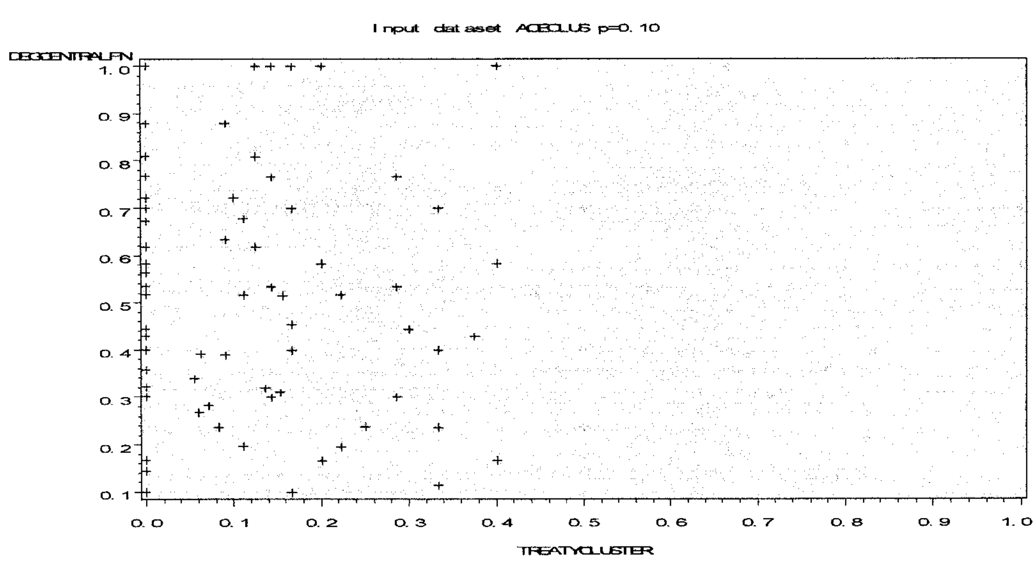


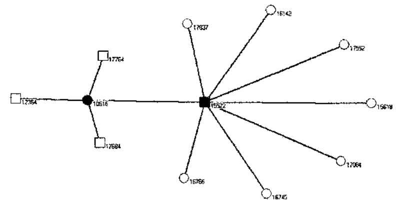
Figure 15 Scatter plot of degree centralization against percent First Nations (Treaty) where colours of points indicate cluster assignment.

Clusters one through three consisted of 123, 69, and 47 components, respectively; the mean number of individuals in the components contained in clusters 1 through 3 was 7.1, 7.6, and 8.6, respectively (**Table 26**). Visualizations of components within these three clusters are included in **Figure 16**.

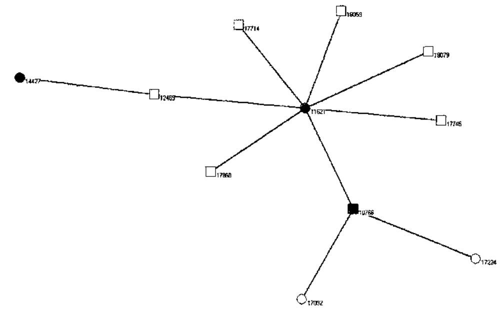
Table 26: A comparison of the demographic, infection and network characteristics of the three identified clusters using both variables used to generate the clusters and those external to the clustering (n=239 components).

Variables Mean (range)		**Cluster 1** (n=123)	**Cluster 2** (n=69)	**Cluster 3** (n=47)
Clustering variables		**Mean (range)**	**Mean (range)**	**Mean (range)**
% First Nations (Treaty)		0.7 (0,11.1)	23.5 (11.1,40)	63.9 (44.4,100)
Degree centralization		0.7 (0.1,1)	0.6 (0.1,1)	0.5 (0.1,1)
% in same region		77.3 (14.3,100)	80.2 (25,100)	90.3 (44.4,100)
Demographic variables		**Mean (range)**	**Mean (range)**	**Mean (range)**
Mean age (years)		21.9 (14,40)	21.6 (14.3,43.4)	21.4 (15.3,33.6)
Mean age range (years)		7.5 (0,40)	7.4 (0,52)	10 (2,39)
% females		41.3 (0,85.7)	46.4 (12.5,87.5)	45.7 (16.7,80)
% unknown gender		3.4 (0,85.7)	1.8 (0,75)	0.4 (0,20)
No. of distinct geolocations		2.1 (1,8)	2.7 (1,9)	2.7 (1,6)
% in Winnipeg		56.2 (0,100)	36.2 (0,100)	9.8 (0,83.3)
% in Southern Rural Mb		15.2 (0,100)	21.6 (0,100)	19.9 (0,100)
% in Northern Rural Mb		14 (0,100)	35 (0,100)	69.1 (0,100)
% outside Mb		6.4 (0,85.7)	2.6 (0,40)	1.0 (0,20)
Clinical variables		**Mean (range)**	**Mean (range)**	**Mean (range)**
% lab-confirmed cases		35.6 (12.5,80)	42 (12.5,80)	50.7 (16.7,80)
% cases with CT*		82.9 (0,100)	60.3 (0,100)	68.6 (0,100)
% cases with GC*		6.9 (0,100)	14.9 (0,100)	9.4 (0,50)
% cases with CT and GC*		9.9 (0,100)	20.2 (0,100)	22 (0,80)
% cases with repeat infections		15.3 (0,100)	18.3 (0,100)	16.4 (0,100)
% repeat nominees		9.8 (0,66.7)	11.4 (0,44.4)	20.6 (0,66.7)
Network variables		**Mean (range)**	**Mean (range)**	**Mean (range)**
Component size		7.1 (5,33)	7.6 (5,32)	8.6 (5,25)
% same-sex links		3.2 (0,100)	1.4 (0,25)	1 (0,20)
		Freq (%)	**Freq (%)**	**Freq (%)**
Degree centralization	0 to 0.25	12 (9.8)	9 (13.0)	11 (23.4)
	0.26 to 0.50	20 (16.3)	15 (21.7)	14 (29.8)
	0.51 to 0.75	40 (32.5)	25 (36.2)	16 (34.0)
	>= 0.76	51 (41.5)	20 (29.0)	6 (12.8)

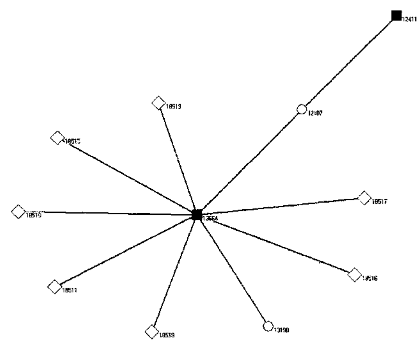
* CT = chlamydia; GC = gonorrhea; CT&GC = chlamydia and gonorrhea



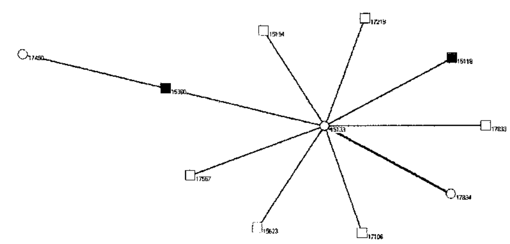
Component 556



Component 671

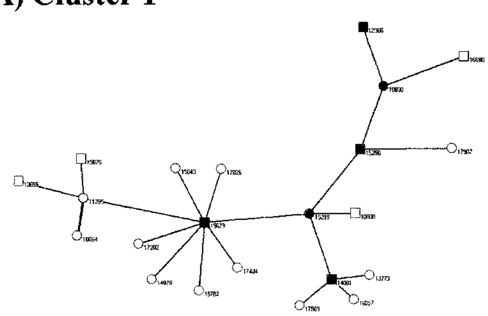


Component 1610

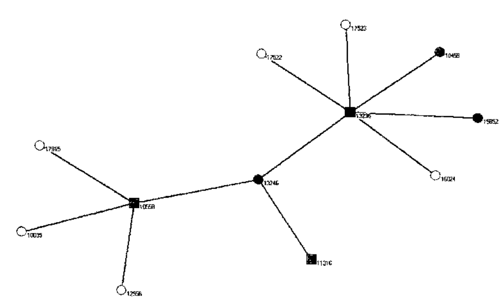


Component 3129

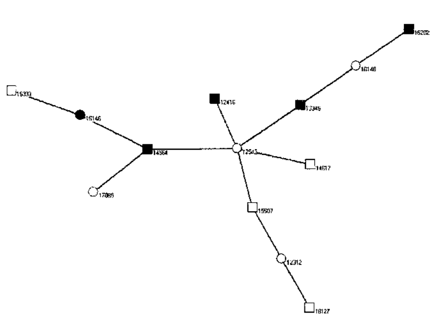
A) Cluster 1



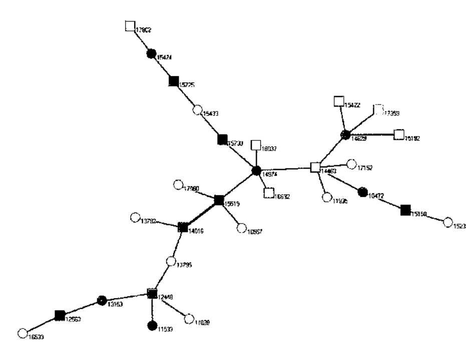
Component 8



Component 35

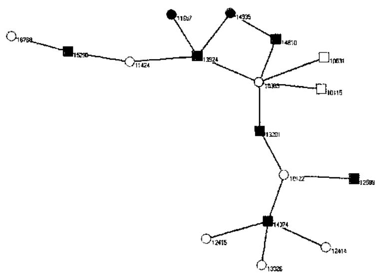


Component 1728

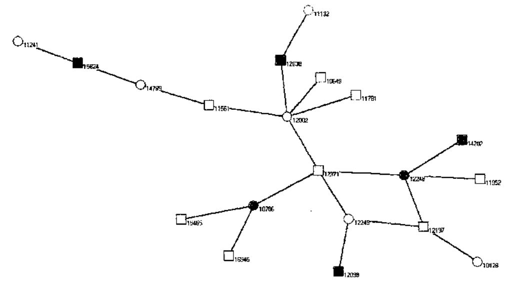


Component 430

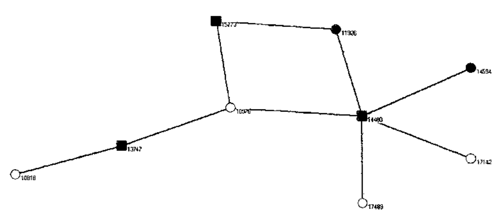
B) Cluster 2



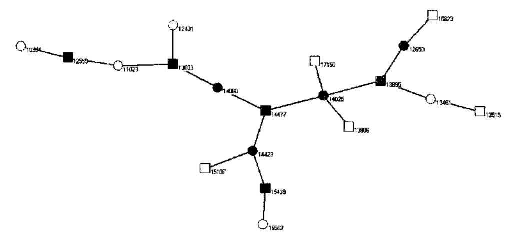
Component 109



Component 119



Component 709



Component 844

C) Cluster 3

Figure 16: Visualization of components contained within the resultant three clusters, A-C, respectively. Each symbol represents an individual; males and females are designated by circles and squares, respectively. Colours indicate infection, chlamydia = red, gonorrhea = green, both chlamydia and gonorrhea = plum, syphilis = orange, both gonorrhea and syphilis = brown; and white = unknown infection. Lines represent sexual links; red lines = same sex links and black lines = heterosexual links.

The three clusters differed primarily on the basis of their First Nations Treaty content, increasing from cluster 1 to 3, from 0.7% to 63.9%, respectively. Cluster 3 contained components with wider ranges of age (mean; 10 years) than clusters 1 and 2. Cluster 1 contained more components with individuals with unknown gender (mean percent; 3.4%). The mean percentage of individuals residing within the jurisdiction of the Winnipeg RHA within the components decreased from Cluster 1 to Cluster 3, 56.2% and 9.8%, respectively. Mean percentage of individuals within the components from the northern rural region increased from Cluster 1 to Cluster 3. Cluster 2 contained the highest mean percentage of components with gonorrhea alone (14.9%); was the only cluster to include syphilis cases (data not shown). Cluster 1 contained the largest mean proportion of components with same sex links. Cluster 3 contained the components with the highest mean proportion of repeat nominees. In terms of degree centralization, 41% of components in cluster 1 had degree centralizations greater than or equal to 0.76, indicating the prominence of one individual within these components. The greatest proportion of components contained within cluster 2 had degree centralizations between 0.51 and 0.75 indicating more branched structures in comparison to components of cluster 1. While the fewest number of components contained in cluster 3 had degree centralizations greater than or equal to 0.76, and 23% had values of 0 to 0.25, indicating that no one individual is prominent within these components.

Two-hundred thirty nine components were used in the clustering procedure, of these 59 included at least one supplementary questionnaire participant (refer to **section 4.3.1**). Cluster 1 through 3 included 40, 15, and 4 study components, respectively. For example, 40 study components were contained within the 123 components in cluster 1. An additional comparison of supplementary information, elicited from the study questionnaire is presented in **Table 27**. Thirteen of the 40 components within cluster 1 contained at least one individual with postsecondary education; 17 components contained at least one individual of an ethnic minority group; 14 components contained at least one individual reporting that their main source of income was from government or personal support; and 2 components within cluster 1 contained at least one individual reporting that their main source of income was from illegal activities.

Table 27: A comparison of the sociodemographic, and interrelationship information obtained from the 'Social and sexual behaviours within sexually transmitted disease networks in Manitoba' questionnaire. Cells contain counts of specific variables in components.

	Cluster 1 (n=40)	Cluster 2 (n=15)	Cluster 3 (n=4)
Individual attributes	**Freq (%)**	**Freq (%)**	**Freq (%)**
Postsecondary education	13 (32.5)	1 (6.7)	0
Ethnicity (other)	17 (42.5)	3 (20)	1 (25)
Injection drug use	4 (10)	1 (6.7)	0
Main source of income illegal	2 (5)	1 (6.7)	0
Main source of income support	14 (35)	11 (73.3)	0
Symptomatic sex	20 (50)	6 (40)	1 (25)
Delay in seeking healthcare > 7 days	9 (22.5)	6 (40)	1 (25)
Relationship attributes			
Concurrency	17 (42.5)	7 (46.7)	2 (50)
Received anything for sex	4 (10)	1 (6.7)	0
Partner ever had sex with acquaintance(s)	23 (57.5)	9 (60)	3 (75)
Condom use (never)	21 (52.5)	9 (60)	3 (75)
Discussed STI history (yes)	29 (72.5)	10 (66.7)	3 (75)
Discussed HIV status (yes)	19 (47.5)	8 (53.3)	1 (25)
First place of meeting (sex trade/street)	4 (10)	2 (13.3)	0
First place of meeting (public)	16 (40)	6 (40)	1 (25)
First place of meeting (private)	34 (85)	13 (86.7)	3 (75)
Type of relationship (steady)	17 (42.5)	7 (46.7)	1 (25)
Relationship type (casual)	38 (95)	14 (93.3)	4 (100)
Relationship type (sex trade)	0	1 (6.7)	0
Length of social acquaintance (0 days)	8 (20)	6 (40)	0
Sex trade component	5 (12.5)	3 (20)	0

In terms of healthcare behaviour, 9 out of 40 components in cluster 1 contained at least one individual who delayed seeking healthcare for more than 7 days after first noticing symptoms. Forty-three percent of the study components in cluster 1 contained individuals with concurrent (overlapping) sexual relationships. Over 50% of these components contained at least one individual who reported never using condoms; and 47.5% of components contained at least one individual who reported never discussing HIV status with his or her sexual partner. Twelve percent of components contained at least one individual participating in sex trade, i.e. exchanging or giving money, drugs or other items for sex; meeting their partner through sex trade; or reporting that a partner was a client.

Components in cluster one had the lowest mean proportion of members with laboratory-confirmed infection, despite the moderately high proportion of individuals with repeat infections and those repeatedly nominated – suggesting that partner notification and follow-up may not be adequately completed. Quite possibly as a result, the highest proportion of components within this cluster had degree centralization values greater than 0.75 (where a degree centralization of one represents a star).

Components within cluster two were geographically dispersed throughout the entire province. This cluster was characterized by a high proportion of gonorrhea cases and individuals with repeat infection—it also contained all components containing laboratory-confirmed syphilis cases (data not shown). Based on this profile, this cluster appears to contain marginalized populations currently missed by current STI prevention practices- as suggested by the high proportion of individuals with gonorrhea and syphilis. This is suggested by examination of the supplementary information, in which 40% of the study components within this cluster had individuals who delayed seeking healthcare for more than seven days after noticing symptoms. Additionally, disassortative mixing patterns with respect to ethnicity have been shown to increase disease transmission—a possible risk marker for future targeting of these networks.

The final cluster included components whose members primarily resided in northern remote Manitoba. Its components exhibited a higher degree of intra-regional and intra-racial mixing. In terms of infection, a high proportion of the members of the components were diagnosed with both chlamydia and gonorrhea; while a large proportion of the members of the components were repeatedly nominated. Although the proportion of laboratory-confirmed cases was quite high, repeat nomination indicates a high frequency of contact between members. Furthermore, the majority of the components had degree centralizations less than 0.75, which is more indicative of a linear type component as previously seen.

As minimal number of components had complete data on social and sexual behaviours—statistically, clusters could not be differentiated on the bases of these variables. However, as observed in the previous objective, they do provide insight into the formation of these components and should be collected in future education and counseling and partner notification procedures.

Although differences did exist upon examination of the three cluster solution, clustering was quite weak. Several problems were observed with the data including the presence of ties when clustering. However, as all ties occurred within the initial stages of the clustering, it was not as much as a problem than if they had occurred at the final stages of clustering. Additionally, as all variables were non-normally distributed—the assumptions of both the ACECLUS procedure and Ward's clustering method may have been violated. However, studies have indicated that these methods are quite robust to violations in assumptions (119;120).

5 Concluding Statement

Since 1997, Canada has witnessed an upsurge in the incidence of bacterial STIs. Rates of chlamydial and gonococcal infection have increased greatly. Additionally, many provinces have incurred recent outbreaks of locally acquired infectious syphilis, although there is evidence suggesting that these outbreaks may result from the natural cycling of host immunity (134). Although a proportion of the increase may be attributed to the growing use of less invasive testing procedures for chlamydia and gonorrhea, with greater diagnostic sensitivity, the increase in incidence of these bacterial STIs as a group indicates that there are other contributing factors.

It has been hypothesized that over time, within a population in which control programs have been initiated, the reservoir of infection that supports an epidemic becomes concentrated within marginalized, hard-to-reach subpopulations who have limited contact with the healthcare system, 'core groups' (52). There is a need for prevention strategies which target these subpopulations which may have an impact beyond the core group itself due to reduced secondary transmission (163). However, prevention strategies can not be implemented without proper characterization of these subpopulations.

This idea has been further iterated by Manitoba Health's Provincial STD Control Strategy, which states: 'if prevention efforts are to be successful then, research most focus on: the identification of these groups, and the social and sexual dynamics of identified core groups' (23). The current study attempted to both identify and to characterize these groups using social network analysis. This study was unique in its scope, as it combined both traditional epidemiologic and social network analytic tools to examine the transmission of STIs in Manitoba.

The project consisted of three objectives. The first objective was to characterize the study population. The study population consisted of individuals diagnosed with, or exposed to a STI. The study population was characterized using basic demographic information collected through mandatory communicable disease control procedures. Additional characterization was completed using more detailed information gathered by public health nurses after routine partner notification procedures.

It was shown that men are being missed by current STI control procedures. Specifically, only 33% of all cases identified during the study period were men. Although less invasive diagnostic procedures have been implemented in this province, statistics indicate that they are not commonly used. As such, there is a need for more widespread usage in high-risk male populations.

Of even greater concern was the disproportionate number of Aboriginals diagnosed with and/or exposed to a STI. This finding is further compounded by the disproportionate burden of STI in Aboriginals residing in northern Manitoba. These results indicate that STI prevention and control efforts are not adequately reaching Aboriginal populations. Furthermore, the results suggest a need for culturally sensitive, population-specific education and prevention programs.

Further characterization of the supplementary information indicated that this population engaged in several high-risk behaviours. Specifically, a large proportion of participants reported that they continued to have sexual relations while experiencing symptoms. Additionally, a large proportion of individuals waited eight or more days after symptom onset to seek healthcare treatment. These findings, in combination with the high number of recent sex partners reported, contribute significantly to the continuing transmission of

STIs within this population. These behaviours could result from a lack of knowledge regarding the clinical presentation of STIs, stigma or barriers to healthcare utilization. These issues should be specifically addressed in education programs.

Characterization of sexual relationships revealed some interesting findings. Communication between sexual partners was quite low. Less than half of all participants discussed their partner's HIV status or STI history. Healthy communication skills need to be addressed in prevention efforts- specifically, educating and counseling on negotiating and communication skills. Furthermore, healthy communication needs to be stressed for all types of relationships, as even those reporting long periods of socialization, acquired an STI.

The second objective was to identify and characterize groups of individuals diagnosed with and/or exposed to a STI. These groups consisted of individuals connected directly or indirectly through sexual relationships. A large number of groups were identified, ranging in size from 2 to 33 individuals. Several interesting findings were observed. Unique to this study, high proportions of same sex relationships were observed, between both men and women. The latter group is also of concern, as previous studies identified that these women are at risk of acquiring HIV. As such, prevention and education efforts should target this specific group of women.

Thirteen groups contained 16 or more individuals. These groups resembled the previously discussed marginalized subpopulations in which infection is thought to be concentrating. Specifically, these groups had high levels of chlamydia, gonorrhea and syphilis, and the members were often repeatedly exposed and repeatedly infected. Furthermore, members were likely to be of Aboriginal background, with a large proportion residing in northern

Manitoba. These networks should be targeted for prevention, for example through the dissemination of educational materials, prophylaxis or repeat screening. These groups were also characterized by a wide range in the ages of its members. In one instance a difference of 52 years between the oldest and youngest members was observed. A wide age range may be a marker of sex trade involvement or coercion, and control programs should specifically target such groups.

Characterization of these groups using more detailed information gathered by public health nurses after routine partner notification procedures enhanced our understanding of these high-risk groups. Specifically, the largest group which contained 33 individuals had members involved in sex trade work– exchanging both sex for both drugs and money. Members reported rarely using condoms; and alcohol and/or drugs were used before or during every sexual act. All of these are behaviours can be addressed in control initiatives.

Additional findings indicate that collecting data on specific social venues should be common practice in partner notification. Several segmented, smaller groups had at least one member naming a common venue in which they met a specific sexual partner.

The purpose of the third objective was to cluster together groups with similar profiles. Three distinct clusters were identified. Clusters differed on the basis of the clinical characteristics and the geographic distribution of the group members. For instance, one cluster was characterized by groups with a high concentration of gonococcal infection. Additionally, it was the only cluster containing groups infected with syphilis. Another cluster contained groups primarily concentrated in northern Manitoba, and exhibited a large amount of intra-regional and intra-racial mixing. This finding, that distinct clusters of groups do in fact exist, further indicates a need for diverse control initiatives. For instance, control programs targeting groups residing in northern remote Manitoba may need to incorporate traditional medicine practices or teachings; while programs targeting the former groups may need to take into consideration the accessibility of healthcare services, and other barriers to service utilization.

These results have several implications with respect to public health programming. In terms of testing, the data indicates a need to increase the overall use of less invasive testing procedures, for example, urine-based screening. A targeted approach may also be needed, for example, scheduled testing at specific locations of socialization. The disproportionate burden of STIs in individuals of Aboriginal ethnicity indicates that culturally appropriate programming is necessary. This may include having control messages and education in appropriate languages, with messages appropriate to specific cultures. Peer-outreach by community members may be another approach in education and prevention.

The findings also indicate that sexual education should place a greater emphasis on the natural history and prevalence of bacterial STIs. Education should emphasize that infected individuals may have very mild to no symptoms. Additionally, adolescents should be educated about the risks of using alcohol and/or drugs during sexual activities. High proportions of same sex relationships were observed in this study. Appropriate prevention tools need to be developed for this population. For example, a targeted approach may be necessary for women who have sex with women and who are involved in sex trade work or injection drug use.

In terms of network composition, larger components were identified to be more high risk. These components could be targeted using partner delivered therapy, re-testing and peer-outreach. Of interest, three distinct types of components were identified using cluster analysis. Clusters differed on the basis of clinical, geographic and sociodemographic variables. Cluster analysis could be used as a tool in other health regions in order to better target diverse sexual groups, which require differing control programs. In this study, the findings indicate that the same control program may not be effective for all groups. For example, in Winnipeg, components were often star-like or radial indicating that partner notification may not be as effective. One suggestion is to collect the names of places of socialization for this region.

The quality of previous research conducted in this population has been acknowledged to be quite high (164). Previous studies have proven the utility of using social network analysis on this population, and notifiable disease data is likely the most complete source available for this type of analysis. The reasons for its completeness include a centralized reporting procedure, a computerized database, the use of one central laboratory for the majority of diagnostic and screening tests, and the use of trained staff committed to disease prevention. Although participation in the supplementary questionnaire was low, a limited number of studies have collected data of this nature. This study will further our understanding of STI transmission within this population.

In conclusion, in a time when public health funding is scarce, a greater understanding of sexual mixing and disease transmission patterns within these subpopulations will aid in development of targeted, population-specific education and effective prevention programs. Use of enhanced partner notification procedures enables the initial characterization needed for the development of these programs.

6 References

1. Schachter J. Biology of *Chlamydia Trachomatis*. In: Holmes KK, Mardh P-A, Sparling PF et al., eds. Sexually Transmitted Diseases. New York: McGraw-Hill, 1999:391-405.
2. Sparling PF. Biology of *Neisseria gonorrhoeae*. In: Holmes KK, Mardh P-A, Sparling PF et al., eds. Sexually Transmitted Diseases. New York: McGraw-Hill, 1999:433-50.
3. LCDC Expert Working Group on Canadian Guidelines for Sexually Transmitted Diseases. Canadian STD Guidelines. 1998. Ottawa, Division of STD Prevention and Control, Bureau of HIV/AIDS, STD and TB, Health Canada.
4. Government of Canada, Public Health Agency of Canada. Notifiable Diseases On-Line, Gonorrhea. http://dsol-smed.phac-aspc.gc.ca/dsol-smed/ndis/diseases/gono_e.html . 2003.
5. Government of Canada, Public Health Agency of Canada. Notifiable Diseases On-Line, Genital Chlamydia. http://dsol-smed.phac-aspc.gc.ca/dsol-smed/ndis/diseases/chlm_e.html . 2003.
6. Weir E. Upsurge of genital *Chlamydia trachomatis* infection. CMAJ 2004; 171:855.
7. Westrom L. Incidence, prevalence, and trends of acute pelvic inflammatory disease and its consequences in industrialized countries. American Journal of Obstetrics and Gynecology 1980; 138:880-892.
8. Westrom LV. Sexually transmitted diseases and infertility. Sexually Transmitted Diseases 1994; 21:S32-S37.
9. Ronald AR, Peeling RW. Sexually transmitted infections: Their manifestations and links to infertility and reproductive illness. In: Minister of Supply and Services Canada, ed. Understanding infertility: Risk factors affecting fertility. Ottawa: 1993:1-131.
10. Goeree R, Gully PR. The burden of chlamydial and gonococcal infection in Canada, research studies of the Royal Commission on New Reproductive Technologies. Prevention of Infertility. Ottawa: Minister of Supply and Services, 1993:29-76.
11. Anttila T, Saikku P, Koskela P et al. Serotypes of *Chlamydia trachomatis* and risk for development of cervical squamous cell carcinoma. Journal of the American Medical Association 2001; 285:47-51.
12. Koskela P, Anttila T, Bjorge T et al. *Chlamydia trachomatis* infection as a risk factor for invasive cervical cancer. International Journal of Cancer 2000; 85:35-39.
13. Paavonen J, Karunakaran KP, Noguchi Y et al. Serum antibody response to the heat shock protein 60 of *Chlamydia trachomatis* in women with developing cervical cancer. American Journal of Obstetrics and Gynecology 2003; 189:1287-92.

14. Wasserheit JN. Epidemiological synergy: interrelationships between HIV infection and other STDs. In: Chen LC, ed. AIDS and Women's Reproductive Health. New York: Plenum Press, 1991:47-73.
15. Rottingen J-A, Cameron DW, Garnett GP. A systematic review of the epidemiologic interactions between classic sexually transmitted diseases and HIV - How much really is known? Sexually Transmitted Diseases 2001; 28:579-97.
16. Gerbase AC, Rowley JT, Mertens TE. Global epidemiology of sexually transmitted diseases. Lancet 1998; 351:2-4.
17. Health Canada. National Goals for the Prevention and Control of Sexually Transmitted Diseases in Canada. Canada Communicable Disease Report 1997; 23S6:2-1.
18. Public Health Agency of Canada. 2002 Canadian Sexually Transmitted Infections (STI) Surveillance Report. Canadian Communicable Disease Report 31S2, 1-39. 2005.
19. Division of STD Prevention and Control. 1998/1999 Canadian Sexually Transmitted Disease (STD) Surveillance Report. Canada Communicable Disease Report 2000; 26S6:1-35.
20. Sarwal S, Wong T, Sevigny C, Ng L-K. Increasing incidence of ciprofloxacin-resistant *Neisseria gonorrhoeae* infection in Canada. CMAJ 2003; 168:872-73.
21. Public Health Agency of Canada. STI Data Tables. http://www.phac-aspc.gc.ca/std-mts/stddata_june05/index.html . 2003.
22. Public Health Agency of Canada. Reported cases and rates of notifiable STI from January 1 to June 30, 2004 and January 1 to June 30, 2003. http://www.phac-aspc.gc.ca/std-mts/stdcases-casmts/index.html . 2005.
23. Manitoba Health. Discussion paper for the development of a provincial sexually transmitted diseases control strategy. 1-36. 2000.
24. Allard R, Robert J, Turgeon P, Lepage Y. Predictors of asymptomatic gonorrhea among patients seen by private practitioners. CMAJ 1985; 133:1135-39.
25. Aral SO, Soskoline V, Joesoef RM, O'Reilly KR. Sex partner recruitment as risk factor for STD: Clustering of risky modes. Sexually Transmitted Diseases 1991; 18:10-17.
26. Hook III EW, Reichart CA, Upchurch DM, Ray P, Celentano D, Quinn TC. Comparative behavioral epidemiology of gonococcal and chlamydial infections among patients attending a Baltimore, Maryland, sexually transmitted disease clinic. American Journal of Epidemiology 1992; 136:662-72.
27. Lacey CJ, Merrick DW, Bensley DC, Fairley I. Analysis of the sociodemography of gonorrhoea in Leeds, 1989-93. British Medical Journal 1997; 314:1715-18.
28. Low N, Daker-White G, Barlow D, Pozniak AL. Gonorrhoea in inner London: results of a cross sectional study. British Medical Journal 1997; 314:1719-23.
29. Navarro C, Jolly A, Nair R, Chen Y. Risk factors for genital chlamydia infection. The Canadian Journal of Infectious Diseases 2002; 13:195-207.
30. Communicable Disease Control Unit, Manitoba Health. The descriptive epidemiology of sexually transmitted infections and blood-borne pathogens in Manitoba: 2002-2003. 2005. Winnipeg, Manitoba Health.

31. Berman SM, Hein K. Adolescents and STDs. In: Holmes KK, Sparling PF, Mardh P-A et al., eds. Sexually Transmitted Diseases. New York: McGraw-Hill, 1999:129-42.
32. Blythe MJ, Katz BP, Orr DP, Caine VA, Jones RB. Historical and clinical factors associated with *Chlamydia trachomatis* genitourinary infection in female adolescents. Journal of Pediatrics 1988; 112:1000-1004.
33. Chacko MR, Lovchik JC. *Chlamydia trachomatis* infection in sexually active adolescents: Prevalence and risk factors. Pediatrics 1984; 73:836-40.
34. Johnson BA, Poses RM, Fortner CA, Meier FA, Dalton HP. Derivation and Validation of a Clinical Diagnostic Model for Chlamydial Cervical Infection in University Women. Journal of the American Medical Association 1990; 264:3161-65.
35. Quinn TC, Gaydos C, Shepherd M et al. Epidemiologic and microbiologic correlates of Chlamydia trachomatis infection in sexual partnerships. Journal of the American Medical Association 1996; 276:1737-42.
36. Stergachis A, Scholes D, Heidrich FE, Sherer DM, Holmes KK, Stamm WE. Selective screening for *Chlamydia trachomatis* infection in a primary care population of women. American Journal of Epidemiology 1993; 138:143-53.
37. Shields SA, Wong T, Mann J et al. Prevalence and correlates of chlamydia infection in Canadian street youth. Journal of Adolescent Health 2004; 34:384-90.
38. Critchlow CW, Wolner-Hanssen P, Eschenbach DA et al. Determinants of cervical ectopia and of cervicitis: age, oral contraception, specific cervical infection, smoking, and douching. American Journal of Obstetrics and Gynecology 1995; 173:534-43.
39. Lycke E, Lowhagen G-B, Hallhagen G, Johannisson G, Ramstedt K. The risk of transmission of genital Chlamydia trachomatis infection is less than that of genital Neisseria gonorrhoeae infection. Sexually Transmitted Diseases 1980; 7:6-10.
40. Washington AE, Gove S, Schachter J, Sweet RL. Oral contraceptives, Chlamydia trachomatis infection, and pelvic inflammatory disease. A word of caution about protection. Journal of the American Medical Association 1985; 253:2246-50.
41. Orr P, Sherman E, Blanchard JF, Fast M, Hammond G, Brunham RC. Epidemiology of infection due to *Chlamydia trachomatis* in Manitoba, Canada. Clinical Infectious Diseases 1994; 19:876-83.
42. Toomey KE, Rafferty MP, Stamm WE. Unrecognized high prevalence of *Chlamydia trachomatis* cervical infection in an isolated Alaskan Eskimo population. Journal of the American Medical Association 1987; 258:53-56.
43. Health Canada. A statistical profile on the health of First Nations in Canada. http://www.hc-sc.gc.ca/fnih-spni/pubs/gen/stats_profil_e.html . 2003. 10-10-2004.
44. De P, Singh AE, Wong T, Yacoub W. Outbreak of Neisseria gonorrhoeae in Northern Alberta, Canada. Sexually Transmitted Diseases 2003; 30:497-501.
45. Jolly AM, Orr PH, Hammond G, Young TK. Risk Factors for Infection in Women Undergoing Testing for Chlamydia trachomatis and Neisseria gonorrhoeae in Manitoba, Canada. Sexually Transmitted Diseases 1995; 22:289-95.

46. Jolly AM, Muth SQ, Wylie JL, Potterat JJ. Sexual networks and sexually transmitted infections: A tale of two cities. Journal of Urban Health 2001; 78:433-45.
47. Lee CB, Brunham RC, Sherman E, Harding GK. Epidemiology of an outbreak of infectious syphilis in Manitoba. American Journal of Epidemiology 1987; 125:277-83.
48. Romanowski B, Sutherland R, Love EJ, Mooney D. Epidemiology of an outbreak of infectious syphilis in Alberta. International Journal of STD & AIDS 1991; 2:424-27.
49. Winnipeg Regional Health Authority. Outbreak of infectious syphilis in the Winnipeg Health Region. Final Report Summary- March 2004. www.wrha.mb.ca/caring/preventill/files/Syphilis_080604.pdf. 2004. 1-3-2005.
50. Jolly AM, Orr PH, Hammond G, Young TK. Risk factors for infection in women undergoing testing for Chlamydia trachomatis and Neisseria gonorrhoeae in Manitoba, Canada. Sexually Transmitted Diseases 1995; 22:289-95.
51. Laumann EO, Youm Y. Racial/ethnic group differences in the prevalence of sexually transmitted diseases in the United States: A network explanation. Sexually Transmitted Diseases 1999; 26:250-261.
52. Wasserheit JN, Aral SO. The dynamic topology of sexually transmitted disease epidemics: Implications for prevention strategies. Journal of Infectious Diseases 1996; 174:S201-S213.
53. Jolly AM, Wylie JL. Gonorrhoea and chlamydia core groups and sexual networks in Manitoba. Sex Transm Infect 2002; 78:i145-i151.
54. Yorke JA, Hethcote HW, Nold A. Dynamics and control of the transmission of gonorrhea. Sexually Transmitted Diseases 1978; 5:51-56.
55. May RM, Anderson RM. Transmission dynamics of HIV infection. Nature 1987; 326:137-42.
56. Brunham RC, Plummer FA. A general model of sexually transmitted disease epidemiology and its implications for control. Medical Clinics of North America 1990; 74:1339-52.
57. Brunham RC. Core group theory: A central concept in STD epidemiology. Venereology 1997; 10:34-39.
58. ACSF Investigators. AIDS and sexual behaviour in France. Nature 1992; 360:407-9.
59. Anderson JE, Dahlberg LL. High-risk sexual behaviour in the general population: Results from a national survey, 1988-1990. Sexually Transmitted Diseases 1992; 19:320-325.
60. Turner R. Landmark French and British studies examine sexual behavior, including multiple partners, homosexuality. Family Planning Perspectives 1993; 25:91-92.
61. Rothenberg RB. The geography of gonorrhea: Empirical demonstration of core group transmission. American Journal of Epidemiology 1983; 117:688-94.
62. Potterat JJ, Rothenberg RB, Woodhouse DE, Muth JB, Pratts CI, Fogle JS. Gonorrhea as a social disease. Sexually Transmitted Diseases 1985; 12:25-32.

63. Potterat JJ, Muth SQ, Bethea RP. Chronicle of a gang STD outbreak foretold. Free Inquiry in Creative Sociology 1996; 24:11-16.
64. Potterat JJ, Zenilman JM, Muth SQ et al. Chlamydia transmission: Concurrency, reproduction number, and the epidemic trajectory. American Journal of Epidemiology 1999; 150:1331-39.
65. Blanchard JF, Moses S, Greenaway C, Orr P, Hammond GW, Brunham RC. The evolving epidemiology of chlamydial and gonococcal infections in response to control programs in Winnipeg, Canada. American Journal of Public Health 1998; 88:1496-502.
66. Wylie JL, Jolly AM. Patterns of chlamydia and gonorrhea infection in sexual networks in Manitoba, Canada. Sexually Transmitted Diseases 2001; 28:14-24.
67. Jolly AM, Moffatt ME, Fast MV, Brunham RC. Sexually transmitted disease thresholds in Manitoba, Canada. Annals of Epidemiology 2005; 15:781-88.
68. Koopman JS. Comment: Emerging objectives and methods in epidemiology. American Journal of Public Health 1996; 86:630-632.
69. Shiboski S, Padian NS. Population- and individual-based approaches to the design and analysis of epidemiologic studies of sexually transmitted disease transmission. Journal of Infectious Diseases 1996; 174:188-200.
70. Friedman SR, Neaigus A, Jose B et al. Sociometric risk networks and risk for HIV infection. American Journal of Public Health 1997; 87:1289-96.
71. Klovdahl AS. Social networks and the spread of infectious diseases: The AIDS example. Social Science and Medicine 1985; 21:1203-16.
72. Klovdahl AS, Potterat JJ, Woodhouse DE, Muth JB, Muth SQ, Darrow WW. Social networks and infectious disease: The Colorado Springs Study. Social Science and Medicine 1994; 38:79-88.
73. Neaigus A, Friedman SR, Jose B et al. High-risk personal networks and syringe sharing risk factors for HIV infection among new drug injectors. Journal of Acquired Immune Deficiency Syndromes and Human Retrovirology 1996; 11:499-509.
74. Rothenberg RB, Sterk C, Toomey KE et al. Using social network and ethnographic tools to evaluate syphilis transmission. Sexually Transmitted Diseases 1998; 25:154-60.
75. Service SK, Blower SM. HIV transmission in sexual networks: An empirical analysis. Proceedings of the Royal Society of London B Series 1995; 260:237-44.
76. Laumann EO, Mahah J, Paik A, Youm Y. Network Data Collection and its Relevance for the Analysis of STDs: The NHSLS and CHSLS. In: Morris M, ed. Network Epidemiology: A Handbook for Survey Design and Data Collection. New York: Oxford University Press, 2004:25-41.
77. Ghani AC, Swinton J, Garnett GP. The role of sexual partnership networks in the epidemiology of gonorrhea. Sexually Transmitted Diseases 1997; 24:45-56.
78. Rothenberg RB, Potterat JJ, Woodhouse DE. Personal risk taking and the spread of disease: Beyond core groups. Journal of Infectious Diseases 1996; 174:S144-S149.
79. Scott J. Social network analysis: a handbook. London: SAGE, 2000.

80. Potterat JJ, Rothenberg RB, Muth SQ. Network structural dynamics and infectious disease propagation. International Journal of STD and AIDS 1999; 10:182-85.
81. Wasserman S, Faust K. Social network analysis. Methods and applications. Cambridge: Cambridge University Press, 1994.
82. Potterat JJ, Woodhouse DE, Muth SQ et al. Network dynamism: History and lessons of the Colorado Springs Study. In: Morris M, ed. Network Epidemiology: A Handbook for Survey Design and Data Collection. New York: Oxford University Press, 2004:87-114.
83. Woodhouse DE, Rothenberg RB, Potterat JJ et al. Mapping a social network of heterosexuals at high risk for HIV infection. AIDS 1994; 8:1331-36.
84. Anderson RM, Gupta S, Ng W. The significance of sexual partner contact networks for the transmission dynamics of HIV. Journal of Acquired Immune Deficiency Syndromes 1990; 3:417-29.
85. Aral SO, Hughes JP, Stoner B et al. Sexual mixing patterns in the spread of gonococcal and chlamydial infections. American Journal of Public Health 1999; 89:825-33.
86. Cabral T, Jolly AM, Wylie JL. Chlamydia Trachomatis *omp1* genotypic diversity and concordance with sexual network data. The Journal of Infectious Diseases 2003; 187:279-86.
87. Jolly AM, Wylie JL. Sampling individuals with large sexual networks: An evaluation of four approaches. Sexually Transmitted Diseases 2001; 28:200-207.
88. De P, Singh AE, Wong T, Yacoub W, Jolly AM. Sexual network analysis and gonorrhea outbreak. Sex Transm Infect 2004; 80:280-285.
89. Government of Manitoba, Manitoba Health. Questions and Answers about Health Coverage. http://www.gov.mb.ca/health/mhsip/index.html#eligible . 2003. 3-2-2005.
90. Souque-Boehmer, Marie. Personal Communication. 2005.
91. Manitoba Health. Manitoba Health Annual Statistics 2002-2003. 2004. Winnipeg, Manitoba Health.
92. Manitoba Health. Core Health Services in Manitoba. 1997. Winnipeg, Manitoba Health.
93. Jolly, A. M. Sexually transmitted disease core group membership in Manitoba; strategies for definition and description of risk markers. 1-7-1998. University of Manitoba.
94. Communicable Disease Control and Manitoba Health. Communicable Disease Control Management Protocol Manual. 2001. Winnipeg, Manitoba Health.
95. Wylie JL. Personal Communication. 2005.
96. Gen-Probe Incorporated. Gen-Probe ® Pace ® 2C Chlamydia trachomatis and Neisseria gonorrhoeae. 2001. San Diego, Gen-Probe Incorporated.
97. Gen-Probe Incorporated. Gen-Probe ® Pace ® 2 Chlamydia trachomatis Probe Competition Assay. 2001. San Diego, Gen-Probe Incorporated.
98. Gen-Probe Incorporated. Gen-Probe ® Pace ® 2 Neisseria gonorrhoeae Probe Competition Assay. 2001. San Diego, Gen-Probe Incorporated.

99. Trinity Biotech plc. MicroTrak ® *Chlamydia trachomatis* Direct Specimen Test. 2001. Ireland, Trinity Biotech PLC.
100. Becton, Dickinson and Company. BD ProbeTec ™ ET *Chlamydia trachomatis* and *Neisseria gonorrhoeae* Amplified DNA Assays. 2003. Maryland, Becton, Dickinson and Company.
101. SAS Institute Inc., Cary NC. SAS V8. 1999. Cary, NC.
102. Dean AG, Dean JA, Coulombier D, Brendel KA, Smith DC, Burton AH, Dicker RC, Sullivan K, Fagan RF, and Arner TG. Epi Info, Version 6: A word processing, database, and statistics program for public health on IBM-compatible microcomputers. 1996. Atlanta, Georgia, Centers for Disease Control and Prevention.
103. Wilkins, R. PCCF+ Version 4C User's Guide. Automatic Geographic Coding Based on the Statistics Canada Postal Code Conversion File, Including Postal Codes to June 2003. [82F0086-XDB]. 2004. Ottawa, Health Analysis and Measurement Group, Statistics Canada.
104. Tabachnik BG, Fidell LS. Using multivariate statistics 1. Northridge, California: Harper Collins Publishers, 1989.
105. McCleary L. Using multiple imputation for analysis of incomplete data in clinical research. Nursing Research 2002; 51:339-43.
106. Heitjan DF, Rubin DB. Inference from coarse data via multiple imputation with application to age heaping. Journal of the American Statistical Association 1990; 85:304-14.
107. Heitjan DF. Ignorability in general incomplete-data models. Biometrika 1994; 81:701-8.
108. Roberts JM, Brewer DD. Measures and tests of heaping in discrete quantitative distributions. Journal of Applied Statistics 2001; 28:887-96.
109. Stockwell EG, Wicks JW. Age heaping in recent national censuses. Social Biology 1974; 21:163-67.
110. Wiederman MW. The truth must be in here somewhere: Examining the gender discrepancy in self-reported lifetime number of sex partners. The Journal of Sex Research 1997; 34:375-86.
111. Adrade C. Categorizing continuous variables. Canadian Journal of Psychiatry 2002; 47:886.
112. Hobson R. Properties preserved by some smoothign functions. Journal of the American Statistical Association 1976; 71:763-66.
113. Rothman KJ, Greenland S. Modern Epidemiology. Philadelphia: Lippincott Williams & Wilkins, 1998.
114. Communications Branch and Indian and Northern Affairs Canada. Words first: An evolving terminology relating to Aboriginal Peoples in Canada. http://www.ainc-inac.gc.ca/pr/pub/wf/pdf_e.html . 2002. 13-3-2005.
115. Batagelj V. PAJEK: Program for large network analysis. 1999.
116. Borgatti, S. P. and Everett, M. G. Freeman L. C. UCINET V. 1999. Analytic Technologies, Columbia, SC.
117. Moody J, Morris M, Adams J, and Handcock M. Epidemic potential in human sexual networks. 2003. 29-8-2003.

118. Milligan, G. W. and Cooper, M. C. A study of variable standardization. WPS 87-63. 1987. Columbus, OH, The Ohio State University. *College of Administrative Science Working Paper Series*.
119. Art D, Gnanadesikan R, Kettenring JR. Data-based metrics for cluster analysis. Utilitas Mathematica 1982; 21A:75-99.
120. SAS Institute Inc. SAS OnlineDoc Version 8. http://v8doc.sas.com/sashtml . 2003.
121. SAS Institute Inc., Cary NC. SAS V9.1. 2003. Cary, NC.
122. Ward JH. Hierarchical grouping to optimize an objective function. Journal of the American Statistical Association 1963; 58:236-44.
123. Blashfield RK. Mixture model test of cluster analysis: Accuracy of four agglomerative hierarchical methods. Psychological Bulletin 1976; 83:377-88.
124. Kuiper FK, Fisher L. A Monte Carlo comparison of six clustering procedures. Biometrics 1975; 31:777-83.
125. Milligan GW. An examination of the effect of six types of error perturbation on fifteen clustering algorithms. Psychometrika 1980; 45:325-42.
126. Mojena A. Hierarchical grouping methods and stopping rules: An evaluation. The Computer Journal 1977; 20:359-63.
127. Calinski T, Harabasz J. A dendrite method for cluster analysis. Communications in Statistics 1974; 3:1-27.
128. Duda RO, Hart PE. Pattern Classification and Scene Analysis. New York: John Wiley & Sons, Inc., 1973.
129. Cooper MC, Milligan GW. The effect of measurement error on determining the number of clusters in cluster analysis. Decision Support and Expert Knowledge Representation in Marketing and Related Areas of Research. Berlin, Germany: P Springer-Verlag, 1988:319-28.
130. Milligan GW, Cooper MC. An examination of procedures for determining the number of clusters in a data set. Psychometrika 1985; 50:159-79.
131. Moses S, Elliott L. Sexually transmitted diseases in Manitoba: Evaluation of physician treatment practices, STD drug utilization, and compliance with screening and treatment guidelines. Sexually Transmitted Diseases 2002; 29:840-846.
132. Laumann EO, Gagnon JH, Michael RT, Michaels S. The social organization of sexuality: Sexual practices in the United States. Chicago: University of Chicago Press, 1994.
133. Public Health Agency of Canada. 2002 Canadian Sexually Transmitted Infections (STI) Surveillance Report: Pre-Release. http://www.phac-aspc.gc.ca/std-mts/stddata_pre06_04/index.html . 2004. 16-3-2005.
134. Grassly NC, Fraser C, Garnett GP. Host immunity and synchronized epidemics of syphilis across the United States. Nature 2005; 433:417-21.
135. Blythe MJ, Katz BP, Batteiger BE, Ganser JA, Jones RB. Recurrent genitourinary chlamydial infections in sexually active female adolescents. Journal of Pediatrics 1992; 121:487-93.
136. Fortenberry JD, Evans DL. Routine Screening for Genital Chlamydia trachomatis in Adolescent Females. Sexually Transmitted Diseases 1989; 16:168-72.

137. Hillis SD, Owens LM, Marchbanks PA, Amsterdam LF, Mac Kenzie WR. Recurrent chlamydial infections increase the risks of hospitalization for ectopic pregnancy and pelvic inflammatory disease. American Journal of Obstetrics and Gynecology 1997; 176:103-7.
138. Whittington WL, Kent C, Kissinger P et al. Determinants of persistent and recurrent Chlamydia trachomatis infection in young women: Results of a multicenter cohort study. Sexually Transmitted Diseases 2001; 28:117-23.
139. Potterat JJ, Dukes RL, Rothenberg RB. Disease transmission by heterosexual men with gonorrhea: An empiric estimate. Sexually Transmitted Diseases 1987; 14:107-10.
140. Control of Communicable Diseases Manual Washington, DC: American Public Health Association, 2000.
141. Upchurch DM, Brady WE, Reichart CA, Hook III EW. Behavioral contributions to acquisition of gonorrhea in patients attending an inner city sexually transmitted disease clinic. Journal of Infectious Diseases 1990; 161:938-41.
142. Langille DB, Andreou P, Beazley RP, Delaney ME. Sexual health knowledge of students at a high school in Nova Scotia. Canadian Journal of Public Health 1998; 89:85-89.
143. Flora JA, Thoresen CE. Reducing the risk of AIDS in adolescents. American Psychology 1988; 43:965-70.
144. Leigh BC. Alcohol and unsafe sex: An overview of research and theory. In: Seminara A, Pawlowski A, Watson R, eds. Alcohol, immunomodulation, and AIDS. New York: Alan R. Liss, 1990:35-46.
145. Halpern-Felsher BL, Millstein SG, Ellen JM. Relationship of alcohol use and risky sexual behavior: A review and analysis of findings. Journal of Adolescent Health 1996; 19:331-36.
146. Hansen L, Mann J, Wong T, McMahon S. Sexual health. BMC Womens Health 2004; 4:S1-S10.
147. Potterat JJ, Meheus A, Gallwey J. Partner notification: Operational considerations. International Journal of STD & AIDS 1991; 2:411-15.
148. Ellish NJ, Weisman CS, Celentano D, Zenilman JM. Reliability of partner reports of sexual history in a heterosexual population at a sexually transmitted diseases clinic. Sexually Transmitted Diseases 1996; 23:446-52.
149. Harvey SM, Bird ST, Henderson JT, Beckman LJ, Huszti HC. He said, she said: Concordance between sexual partners. Sexually Transmitted Diseases 2004; 31:185-91.
150. Padian NS, Aral S, Vranizan KM, Bolan G. Reliability of sexual histories in heterosexual couples. Sexually Transmitted Diseases 1995; 22:169-72.
151. Sison JD, Gillespie B, Foxman B. Consistency of self-reported sexual behavior and condom use among current sex partners. Sexually Transmitted Diseases 2004; 31:278-82.
152. Upchurch DM, Weisman CS, Shepherd M et al. Interpartner reliability of reporting of recent sexual behaviors. American Journal of Epidemiology 1991; 134:1159-66.

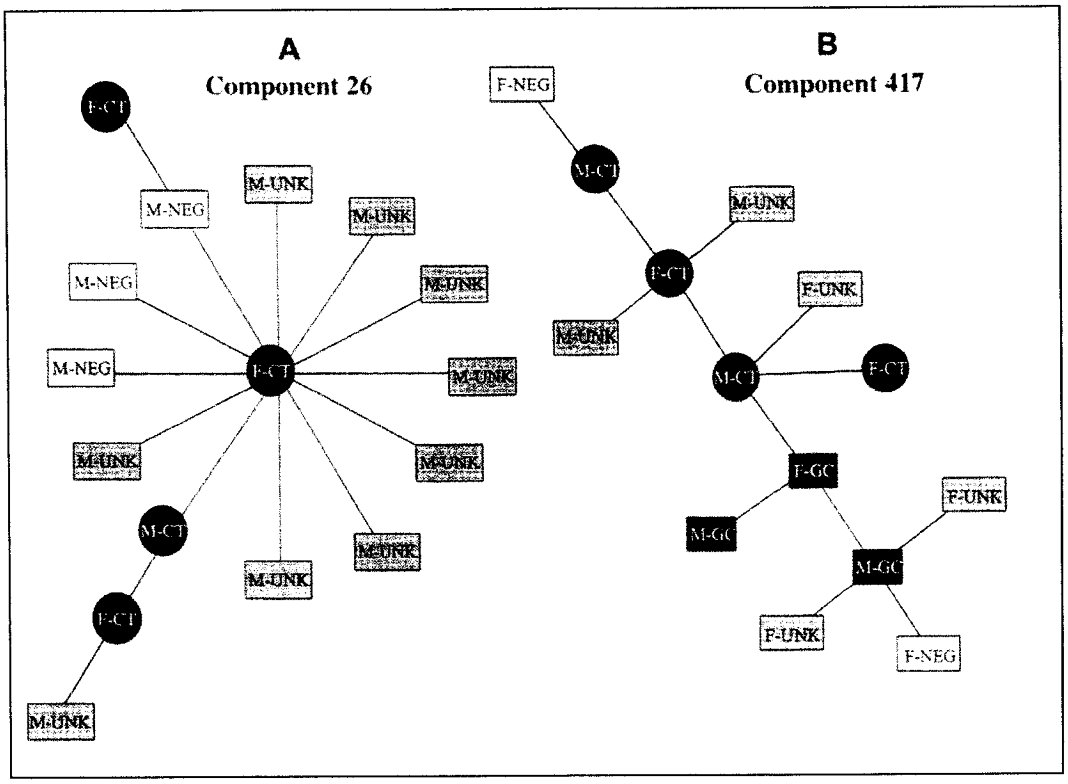
153. van Duynhoven YTHP, Nagelkerke NJD, van de Laar MJW. Reliability of self-reported sexual histories: Test-retest and interpartner comparison in a sexually transmitted diseases clinic. Sexually Transmitted Diseases 1999; 26:33-42.
154. Ghani AC, Donnelly CA, Garnett GP. Sampling biases and missing data in explorations of sexual partner networks for the spread of sexually transmitted diseases. Statistics in Medicine 1998; 17:2079-97.
155. Doherty IA, Padian NS, Marlow C, Aral SO. Determinants and consequences of sexual networks as they affect the spread of sexually transmitted infections. Journal of Infectious Diseases 2005; 191:S42-S54.
156. Brewer DD, Garrett SB, Kulasingam S. Forgetting as a cause of incomplete reporting of sexual and drug injection partners. Sexually Transmitted Diseases 1999; 26:166-76.
157. Potterat JJ, Muth SQ, Rothenberg RB et al. Sexual network structure as an indicator of epidemic phase. Sexually Transmitted Infections 2002; 14:107-10.
158. Fethers K, Marks C, Mindel C, Estcourt CS. Sexually transmitted infections and risk behaviours in women who have sex with women. Sex Transm Infect 2000; 76:345-49.
159. Foxman B, Aral SO, Holmes KK. Heterosexual repertoire is associated with same-sex experience. Sexually Transmitted Diseases 1998; 25:232-36.
160. Marrazzo JM, Koutsky LA, Handsfield HH. Characteristics of female sexually transmitted disease clinic clients who report same-sex behaviour. International Journal of STD & AIDS 2001; 12:41-46.
161. Elliott L, Blanchard J, Beaudoin C et al. Geographical variations in the epidemiology of bacterial sexually transmitted infections in Manitoba, Canada. Sex Transm Infect 2002; 78:i139-i144.
162. Calzavara LM, Bullock SL, Myers T, Marshall VW, Cockerill R. Sexual partnering and risk of HIV/STD among Aboriginals. Canadian Journal of Public Health 1999; 90:186-91.
163. Garnett GP, Bartley LM, Cameron DW, Anderson RM. Both a 'magic bullet' and good aim are required to link public health interests and health care needs in HIV infection. Nature Medicine 2000; 6:261-62.
164. Klovdahl AS. Networks and pathogens. Sexually Transmitted Diseases 2001; 28:25-28.

Appendix A: Social Network Analysis Terminology

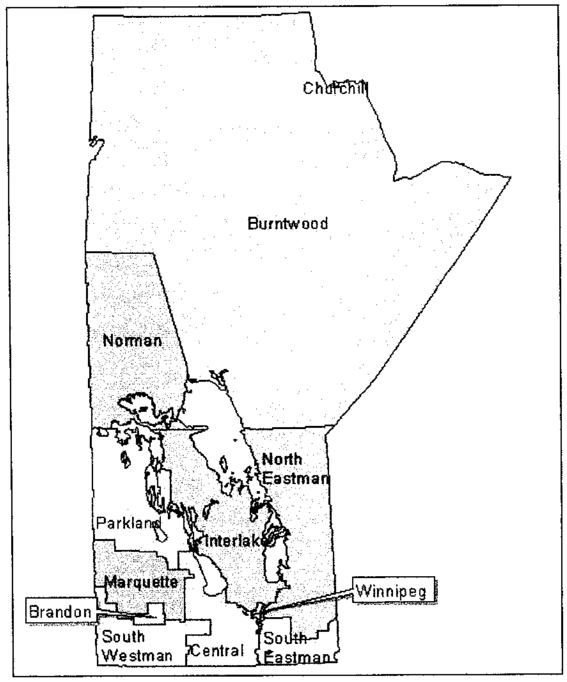
Term	**Definition***
Network	A set of nodes and the resultant set of lines that connect these nodes
Sexual Network	A network, where the lines that connect nodes represent sexual contact
Component	A group of nodes directly or indirectly connected within a network; multiple components can exist within a given network
Degree or degree centrality	Total number of nodes to which a particular node is directly connected to (also referred to as a neighbourhood)
Density	Comparison of the actual number of lines in a network to the total number of lines that would be present if the network were complete (one in which all points are directly connected to each other)
Concurrency	If sexual relationships, index case specific, overlapped in time
Centrality	The importance of a node within a network, can be examined by looking at local or global centrality
Local centrality	The relative importance of a node within its neighbourhood, ex. degree, relative measure of local centrality
Global centrality	When an node has a position of great significance in the overall structure of the network
Betweeness	Extent to which a particular node lies on the path that connects two nodes
Centralization	Measures the heterogeneity of actor centralities within a network
Degree centralization	Measures the heterogeneity of degree centralities within a network
Microstructures	
k-core	The number of people with k contacts who are connected to each other
Clique	A group in which everyone is connected to everyone else
n-clique	N is the maximum pathway at which members of the clique will be regarded as connected
k-plex	A set of points in which each point is adjacent (directly connected) to all except k of the other points

* Definitions are taken out of Rothenberg et al. (1998), Scott (2000), and Wasserman and Faust (1994) (78;79;81)

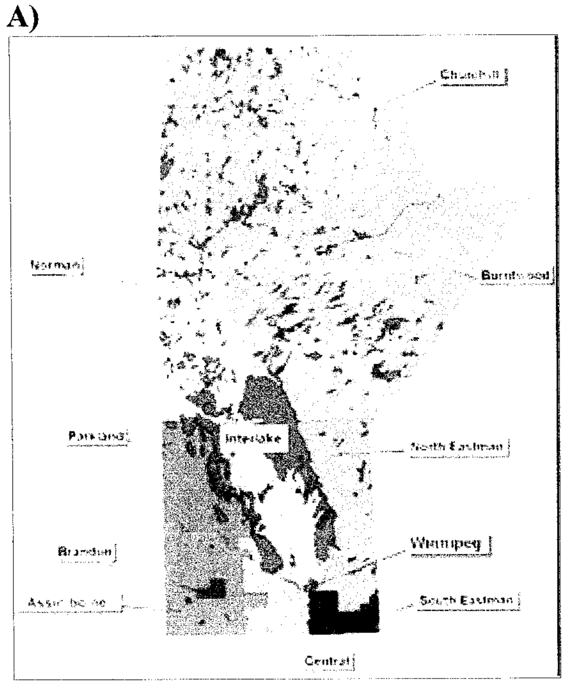
Appendix B Examples of radial (A) and linear (B) components identified in Manitoba by network analysis. Each symbol represents an individual. The symbol and code for an individual indicates either positive results for chlamydia (CT), gonorrhoea (GC), or chlamydia and gonorrhoea (CT/GC); negative test result (NEG); or unknown test result (UNK). The sex of an individual is also indicated (66).



Appendix C: Map of Regional Health Authorities (RHAs) in the Province of Manitoba, (A) before July 1, 2002; (B) as of July 1, 2002 the Marquette and South Westman RHAs amalgamated, forming the Assiniboine RHA

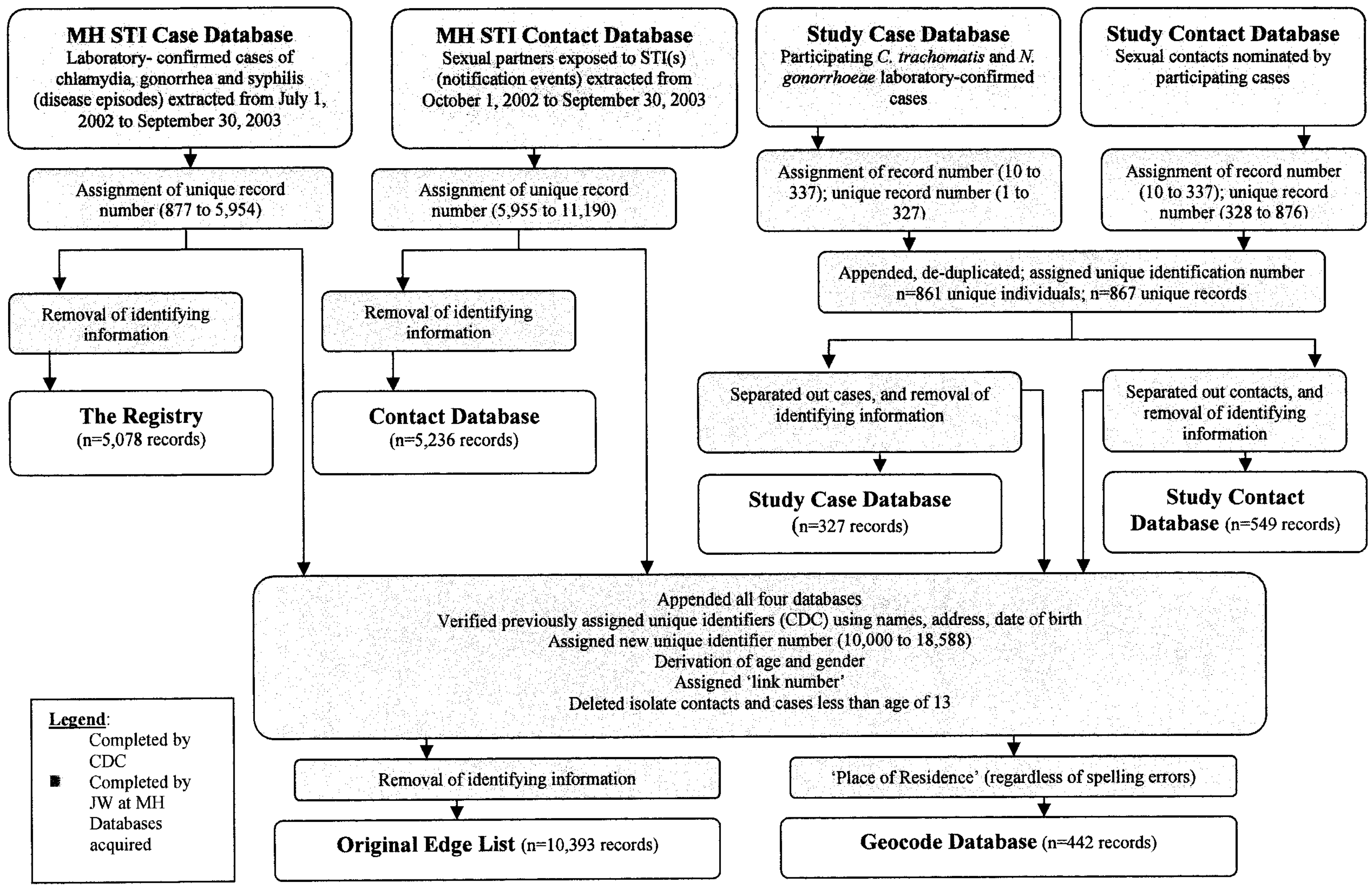


A)



B)

Appendix D: Flowchart of data editing by Manitoba Health's Communicable Disease Control Unit



Appendix E: Manitoba Health Sexually Transmitted Disease Control Notification of Sexually Transmitted Disease Form

Notification of Sexually Transmitted Disease – Confidential

Manitoba Health
Sexually Transmitted Disease Control
4th Floor - 300 Carlton Street
Winnipeg, Manitoba
R3B 3M9
(204) 788-6736

Regn. No. (Leave Blank) ______

NAME (SURNAME) | ADDRESS | TELEPHONE HOME

GIVEN NAME | POSTAL CODE | WORK

DATE OF BIRTH __/__/__ (Y M D) | SEX ☐ FEMALE ☐ MALE | ☐ MARRIED/COMMON LAW SPOUSE'S FIRST NAME ▶ | ☐ MAIDEN NAME (IF APPLICABLE) ▶ | DATE TESTED __/__/__ (Y M D) ONSET __/__/__ (Y M D)

☐ **GONORRHEA** ☐ **CHLAMYDIA**

TYPE: ☐ GENITO-URINARY ☐ OTHER (SPECIFY) ______

CLINICAL FINDINGS (DESCRIBE) ______

______ DURATION OF SYMPTOMS ______

LABORATORY TESTS:

SMEAR ☐ POS ☐ NEG CULTURE ☐ POS ☐ NEG

GONOZYME ☐ POS ☐ NEG CHLAMDIAZYME ☐ POS ☐ NEG

HAS A BLOOD TEST BEEN TAKEN FOR SYPHILIS ☐ YES ☐ NO

TREATMENT GIVEN (SPECIFY) ______

DATE __/__/__ (Y M D)

PREVIOUS TREATMENT FOR: GONORRHEA ☐ YES ☐ NO

CHLAMYDIA ☐ YES ☐ NO

DATE __/__/__ (Y M D)

☐ **SYPHILIS**

TYPE: ☐ PRIMARY ☐ SECONDARY ☐ EARLY LATENT ☐ LATE LATENT

☐ CARDIOVASCULAR ☐ NEUROSYPHILIS ☐ CONGENITAL

TREATMENT GIVEN (SPECIFY) ______

DATE __/__/__ (Y M D)

DIAGNOSIS BASED ON:

DARKFIELD DFA-TP ☐ POS ☐ NEG SCREEN ☐ POS ☐ NEG

CONFIRMATORY ☐ POS ☐ NEG REFERENCE ☐ POS ☐ NEG

CEREBROSPINAL FLUID ☐ POS ☐ NEG ☐ NOT DONE

CLINICAL FINDINGS (DESCRIBE) ______

PREVIOUS TREATMENT FOR SYPHILIS ☐ YES ☐ NO

BY WHOM ______ DATE __/__/__ (Y M D)

OTHER SEXUALLY TRANSMITTED DISEASES

☐ AIDS ☐ CHANCROID ☐ LGV

DO YOU WISH: ☐ CONSULTATIVE SERVICE ☐ NOTIFICATION FORMS ☐ PATIENT LITERATURE

PHYSICIAN'S SIGNATURE: ADDRESS

PLEASE RETURN ALL FOUR COPIES

CONFIDENTIAL

Contact Information

Manitoba Health
Sexually Transmitted Disease Control
4th Floor - 300 Carlton Street
Winnipeg, Manitoba
R3B 3M9
(204) 788-6736

NAME | MAIDEN NAME | TELEPHONE HOME WORK

ADDRESS | POSTAL CODE | DATE OF BIRTH / AGE | FEMALE ☐ MALE ☐

MARITAL STATUS: ☐ MARRIED/CL ☐ SINGLE ☐ OTHER ______ WHERE LIVING: ☐ PARENTS ☐ INFORMANT ☐ OTHER ______

OCCUPATION ______ PLACE OF EMPLOYMENT ______

RELATIONSHIP: ☐ MARITAL ☐ PICK-UP ☐ FRIEND ☐ PROSTITUTE FEE ______

CHARACTERISTICS: HEIGHT ______ WEIGHT ______ EYE COLOUR ______ HAIR ______ COMPLEXION ______ OTHER ______

PLACE OF MEETING ______ PLACE OF EXPOSURE ______ DATE OF EXPOSURE (FIRST) ______ TO ______

FREQUENCY OF SEX CONTACT ______ ☐ TYPE O-G ☐ G-G ☐ R-G

REMARKS: ______

FOR PUBLIC HEALTH USE ONLY DATE ______ MONTH ______

NAME OF INFORMANT ______ PAGE ______

ADDRESS ______ DATE/TYPE SPECIMEN ______

☐ GONORRHEA ☐ CHLAMYDIA ☐ SYPHILIS ☐ OTHER ______

RESULT OF EXAMINATION

DATE	STS BLOOD/CSF	GC SMEAR / CULTURE	CHLAMYDIA TEST	CLINICAL FINDINGS	TREATMENT

EXAMINED BY: ______ SUBMITTED BY ______

COMMENTS:

CONFIDENTIAL

Contact Information

Manitoba Health
Sexually Transmitted Disease Control
4th Floor - 300 Carlton Street
Winnipeg, Manitoba
R3B 3M9
(204) 788-6736

NAME | MAIDEN NAME | TELEPHONE HOME WORK

ADDRESS | POSTAL CODE | DATE OF BIRTH / AGE | FEMALE ☐ MALE ☐

MARITAL STATUS: ☐ MARRIED/CL ☐ SINGLE ☐ OTHER ______ WHERE LIVING: ☐ PARENTS ☐ INFORMANT ☐ OTHER ______

OCCUPATION ______ PLACE OF EMPLOYMENT ______

RELATIONSHIP: ☐ MARITAL ☐ PICK-UP ☐ FRIEND ☐ PROSTITUTE FEE ______

CHARACTERISTICS: HEIGHT ______ WEIGHT ______ EYE COLOUR ______ HAIR ______ COMPLEXION ______ OTHER ______

PLACE OF MEETING ______ PLACE OF EXPOSURE ______ DATE OF EXPOSURE (FIRST) ______ TO ______

FREQUENCY OF SEX CONTACT ______ ☐ TYPE O-G ☐ G-G ☐ R-G

REMARKS: ______

FOR PUBLIC HEALTH USE ONLY DATE ______ MONTH ______

NAME OF INFORMANT ______ PAGE ______

ADDRESS ______ DATE/TYPE SPECIMEN ______

☐ GONORRHEA ☐ CHLAMYDIA ☐ SYPHILIS ☐ OTHER ______

RESULT OF EXAMINATION

DATE	STS BLOOD/CSF	GC SMEAR / CULTURE	CHLAMYDIA TEST	CLINICAL FINDINGS	TREATMENT

EXAMINED BY: ______ SUBMITTED BY: ______

COMMENTS:

MG 6036 (Rev. 90)

CONFIDENTIAL

Appendix F: 'Social and Sexual Behaviours within Sexually Transmitted Disease Networks in Manitoba' Supplemental Questionnaire

At this point we do not have to put anything on the answer sheet that could be used to identify you. At the end of the questionnaire, once you've had a chance to see what the questions are like, you will be given a choice of putting your name with your answers or remaining anonymous.

The first questions are about yourself:

R1) Did you experience any symptoms from your current STD?

A. Yes
B. No
C. Unsure/don't know
D. Refused to answer

Read from list:

Males
Redness at tip of penis
Discharge from penis
Pain on passing urine
Swelling or pain of testicles/balls
Itching of genitals

Females
Vaginal discharge different from normal
Pain on passing urine
Pain in lower back or abdomen
Pain during sexual intercourse
Abnormal vaginal bleeding

If *No* go to question R5.

R2) Did you have sex while you were symptomatic?

A. Yes
B. No
C. Unsure/don't know
D. Refused to answer

R3) When did you first notice these symptoms?
(record as specific date or length of time such as *x* days, weeks, or months ago)

R4) From the time you first noticed your symptoms, how many days passed before you saw a doctor or nurse? **(record as length of time)**

R5) When you were diagnosed with an STD, why had you gone to see the doctor or nurse?

A. Symptoms
B. Contact of case
C. Prenatal
D. Routine
E. Unsure/don't know
F. Refused to answer
G. Other

R6) In the past 12 months how often have you used alcohol before or during sexual activities?

A. Always
B. More than 1/2 the time
C. About 1/2 the time
D. Less than 1/2 the time
E. Never
F. Unsure/don't know
G. Refused to answer

R7) In the past 12 months how often have you used drugs, other than alcohol, before or during sexual activities?

A. Always
B. More than 1/2 the time
C. About 1/2 the time
D. Less than 1/2 the time
E. Never
F. Unsure/don't know
G. Refused to answer

R8) In the past 12 months have you ever injected street drugs, or been injected with drugs?

A. Yes
B. No
C. Unsure/don't know
D. Refused to answer

R9) Over the last year what was your main source of income (How do you get money to live on)? (Choose only one).

A. Regular work (full or part time)
B. Welfare/social assistance/EI or other gov't support
C. Money from family/friends
D. Sex trade/prostitution
E. dealing or doing drug runs
F. Panhandling
G. Other (specify)
H. Unsure/don't know
I. Refused to answer

R10) Over the last year what other types of income have you had? (List as many as necessary).

A. Regular work (full or part time)
B. Welfare/social assistance/EI or other gov't support
C. Money from family/friends
D. Sex trade/prostitution
E. dealing or doing drug runs
F. Panhandling
G. Other (specify)
H. Unsure/don't know
I. Refused to answer

R11) What is the highest level of education you have completed?

A. Primary school (grade 6 or less)
B. Secondary school (between grade 9 and grade 11)
C. Graduated secondary school (graduated grade 12)
D. Trade school
E. University
F. College
G. Other
H. Unsure/don't know
I. Refused to answer

R12) What ethnic group or family background do you most identify yourself with:
(Do not read choices)

A. Caucasian/White
B. Chinese
C. Filipino
D. South-Asian (e.g. Indian, Pakistani)
E. Other Asian (e.g. Vietnamese, Japanese)
F. Latin American, specify country:
G. Middle Eastern
H. Black-African
I. Black-Caribbean
J. Other black, (specify)
K. First Nations (Treaty)
L. Aboriginal (Non-treaty)
M. Metis
N. Inuit
O. Other, specify
P. Unsure/don't know
Q. Refused to answer

R13) How old were you when you first willingly had sex with a partner?

R14) In your lifetime, how many sexual partners have you had?

A. 1-5
B. 6-19
C. 20-99
D. 100-499
E. >500
F. None
G. Unsure
H. Refused to answer

R15) If you live outside of an urban area (Winnipeg, Brandon, The Pas, Thompson), where would you prefer to be tested and treated for your STD infection:

A. Within my own community by a nurse
B. Within my own community by a doctor
C. Outside of my community in Thompson
D. Outside of my community in The Pas
E. Outside of my community in Brandon
F. Outside of my community in Winnipeg
G. Other (specify)
H. Unsure/don't know
I. Refused to answer

Go to section II.

Date: ____________

RHA in which questionnaire is being completed: ____________

Interviewer's name: ____________

Answer sheet:
Section I: Case data

R1		R6		R11	
R2		R7		R12	
R3		R8		R13	
R4		R9		R14	
R5		R10		R15	

R16) Number of sex partners: ____________

Section II. Partner data

	Partner 1	Partner 2	Partner 3	Partner 4	Partner 5	Partner 6
R17 Ethnicity						
R18 education						
R19 Community						
R20 Circumstances						
R21 Other comm.						
R22 Relationship						
R23 Exchange-Rec						
R24 Exchange-give						
R25 Social contact						
R26 First						
R27 Last						
R28 Frequency						
R29 Others						
R30 Drugs						
R31 Protection						
R32 H status						
R33 S status						

(enter the interviewed person's nominal information here or enter "anonymous" if they do not want their name used; as outlined on the questionnaire they are given this choice at the end of the questionnaire)
Name of person interviewed (include aliases): ____________

Date of birth: ____________

Address: ____________

(house/apt. no. and street) (city/town/reserve, province) (postal code)
(use last known address as recorded on notification form)

Section II. Sexual partner information for named partners:

Now I would like to ask you some questions about the relationship between you and your sexual partner or partners. This information will help us to understand what behaviours put individuals at risk for sexually transmitted infections. In the following questions, I will not ask you to provide any information which could be used to identify these individuals.

R16) How many sexual partners have you had in the past year?

I would now like to ask you some questions about that partner (or each of those partners). The questions will be about where you met your partner(s), and how long you have known them.

(For individuals with more than partner).
We will go through the questionnaire for each partner **(Use first names or initials to keep track of multiple partners)**. If you have had more than 6 partners in the past year, I would like to ask you questions only about the 3 that you had sex with the most often and the 3 which you had sex with the fewest times in the previous year.

R17) How would you describe your partner's ethnic background?
(If the case answers "Canadian" read the choices)

A. Caucasian/White
B. Chinese
C. Filipino
D. South-Asian (e.g. Indian, Pakistani)
E. Other Asian (e.g. Vietnamese, Japanese)
F. Latin American, specify country:
G. Middle Eastern
H. Black-African
I. Black-Caribbean
J. Other black, (specify)
K. First Nations (Treaty)
L. Aboriginal (Non-treaty)
M. Metis
N. Inuit
O. Other, specify
P. Unsure/don't know
Q. Refused to answer

R18) To the best of your knowledge, what is the highest level of education partner (#) has completed?

A. Primary school (grade 6 or less)
B. Secondary school (between grade 9 and grade 11)
C. Graduated secondary school (graduated grade 12)
D. Trade school
E. University
F. College
G. Other
H. Unsure/don't know
I. Refused to answer

R19) In what community (city, town, or reserve) did you first meet partner (#)?

R20) Where did you first meet partner (#)?

A. Bar/Hotel - if yes, name of bar/hotel
B. Private residence (e.g. house party)
C. Through family or mutual friends
D. School
E. College/University
F. Sex trade
G. On the street
H. Work
I. Other (specify)
J. Unsure
K. Refused to answer

R21) At the time you were having a sexual relationship with this partner, did he or she live in a different community than the one you lived in?

A. Yes
B. No
C. Unsure/don't know
D. Refused to answer

R22) What is your relationship to this partner?

A. Spouse
B. Boy/girlfriend (lover)
C. Casual sex (friend/acquaintance/pick-up)
D. "On and Off"
E. Ex-lover
F. Ex-spouse
G. Sex trade worker/prostitute
H. Client/trick
I. Unsure
J. Refused to answer

R23) Have you ever received anything from this partner in exchange for sex?

A. drugs.
B. money.
C. food, shelter, clothing or something other than drugs or money.
D. No, never received anything
E. Unsure/Don't know
F. Refused to answer

R24) Have you ever given anything to this partner in exchange for sex?

A. drugs.
B. money.
C. food, shelter, clothing or something other than drugs or money.
D. No, never received anything
E. Unsure/Don't know
F. Refused to answer

R25) How long did you know this partner before beginning a sexual relationship?
enter as a specific date or length of time (e.g. two months)?

R26) When did you begin a sexual relationship with this partner?
enter as a specific date or length of time (e.g. two months)?

R27) When was the last time you had sex with this partner?
enter as a specific date or length of time (e.g. two months)?

R28) Over the past 3 months, how often did you have sex with this person?

A. Every day
B. 2-3 times per week
C. About once per week
D. 1-3 times per month
E. Less than once per month
F. Unsure
G. Refused to answer

R29) Has anyone else you known ever had sex with this person?

A. Yes
B. No
C. Unsure
D. Refused to answer

R30) To the best of your knowledge, has this person ever injected drugs in the last 12 months?

A. Yes
B. No
C. Unsure/don't know
D. Refused to answer

R31) How often was a condom (either male or female condom) used with this partner?

A. Never
B. Less than 1/2 the time
C. About 1/2 the time
D. More than 1/2 the time
E. Always
F. Unsure
G. Refused to answer

R32) Have you ever discussed with this partner whether or not they have HIV?

A. Yes
B. No
C. Unsure/don't know
D. Refused to answer

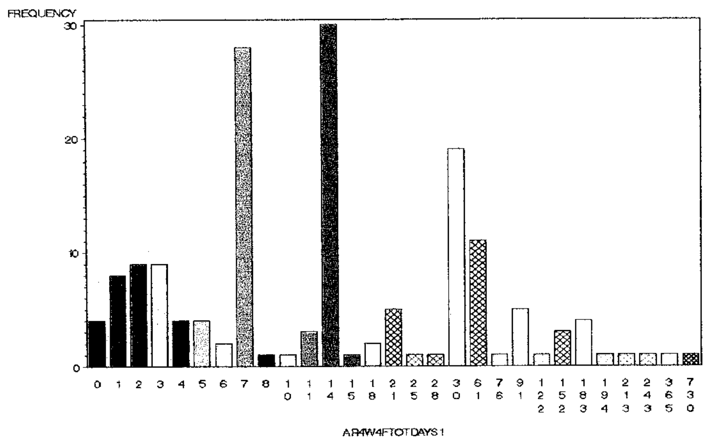
R33) Have you ever discussed with this partner whether or not they have ever had an STD?

A. Yes
B. No
C. Unsure/don't know
D. Refused to answer

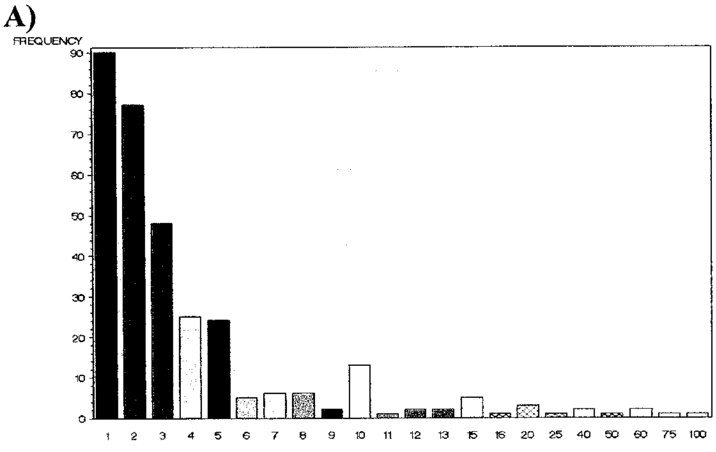
This completes the interview. Now that you've had a chance to see what the questions were like, I would like to ask you if you would be willing to put your name with the answers you provided or if you would like to remain anonymous. If you put your name on the sheet it will help us to better understand your risk of infection. We are able to draw pictures of how people are indirectly connected to groups of other people through sexual contact. You and your partner's activities might not normally have put you at high risk of infection, but possibly because of the activities of previous sexual partners' behaviours or their partners' behaviour, you might have been indirectly connected to some high risk individuals. This information helps us to understand where we can better focus future STD prevention efforts.

Would you like to provide your name at this point, or would you like to remain anonymous in the study?

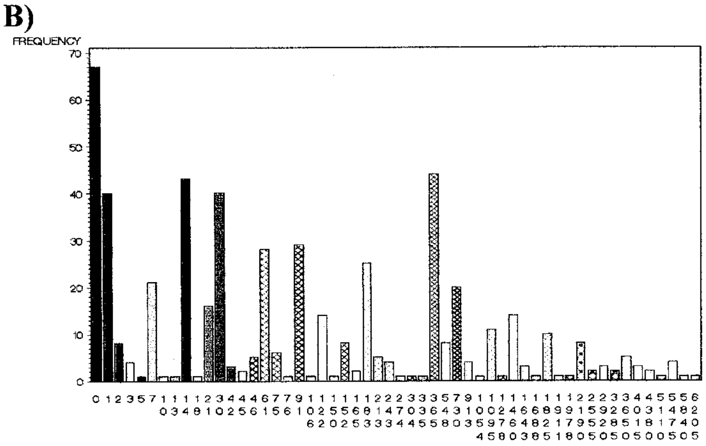
Appendix G Distribution of variables with observed heaping: A) number of days passed from the time the individual first noticed symptoms until the time healthcare professional was visited; B) number of sexual partners in previous 12-months, and C) length of time (days) index case knew partner before beginning a sexual relationship



A)



B)



C)

Appendix H: Management of the *Original Edge List* database

Original Edge List
n=10,393 records

↓

Manual verification of previously assigned links groupings and unique identifiers by scanning the *Original Edge List* for discrepancies such as: greater than 1 index case per grouping; index case who is also a named contact; duplicate records; etc.

↓

Imported Final links non-nominal.mdb file into SAS

↓

Creation of a history file (by unique identifier) to continue verification of previous assignment of links groupings and unique identifiers by checking for differences in age (greater than 2 years) and discordant gender; and searching for duplicate records (study records that were not condensed into corresponding CDC record); duplicate entry of CDC case or contact record in links groupings
(n=8,574 unique individuals)

↓

Revert to original links file and make changes corresponding to results of investigation in Step A: 1) discordant gender, 2) age difference > 2 years, 3) problems with previously assigned links groupings, 4) duplicate records

↓

Merge in cleaned databases including: the *Registry*; the *CDC Contact*; *Geocode*; and *Study Case* and *Contact* database

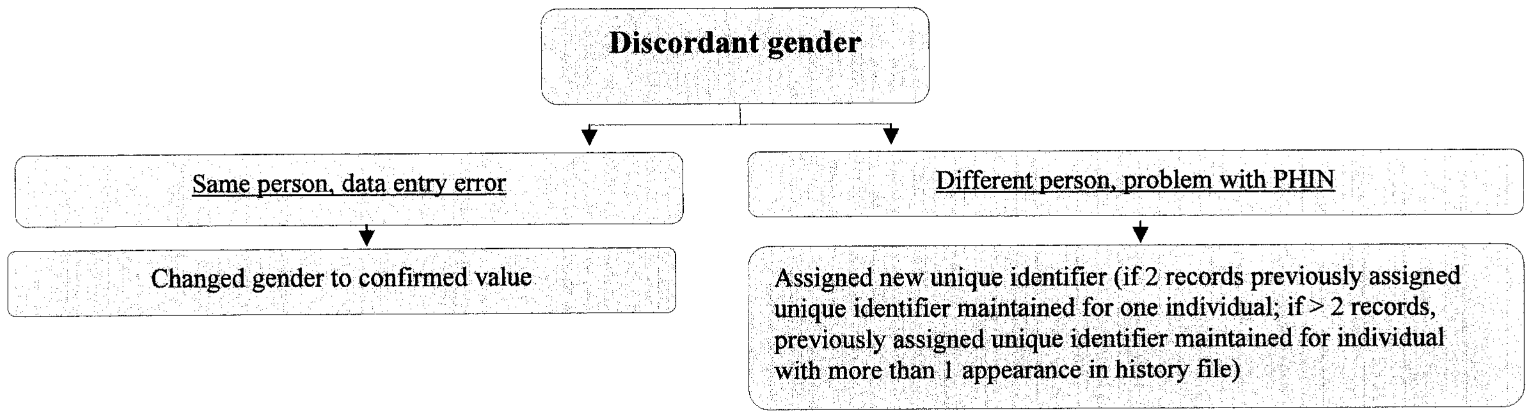
↓

Eligibility criteria (entire population):
- if less than 13 years of age delete (n=4)
- if syphilis case (and their respective contacts), and not linked directly or indirectly to gonorrhea and/or chlamydia case then delete (n=44; n=35)
- if isolate contact, then delete (n=126)

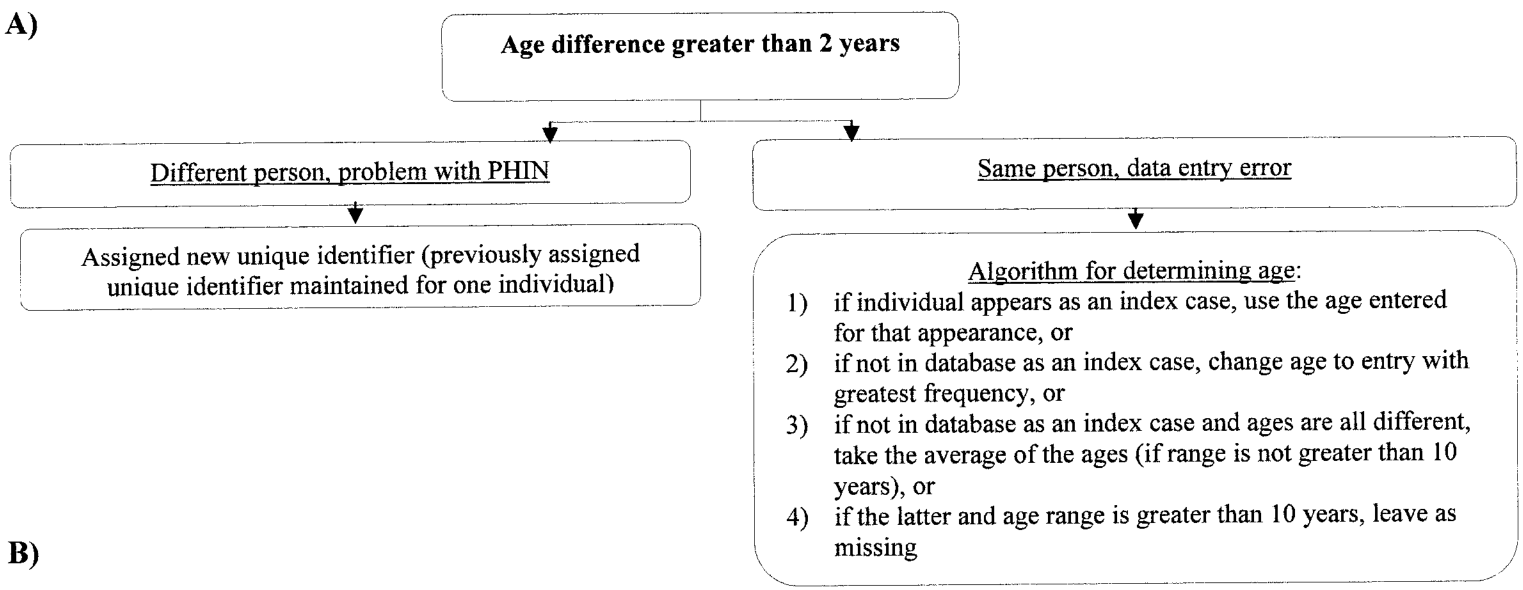
↓

N=8,476 unique individuals
(4,683 cases and 3,793 contacts)

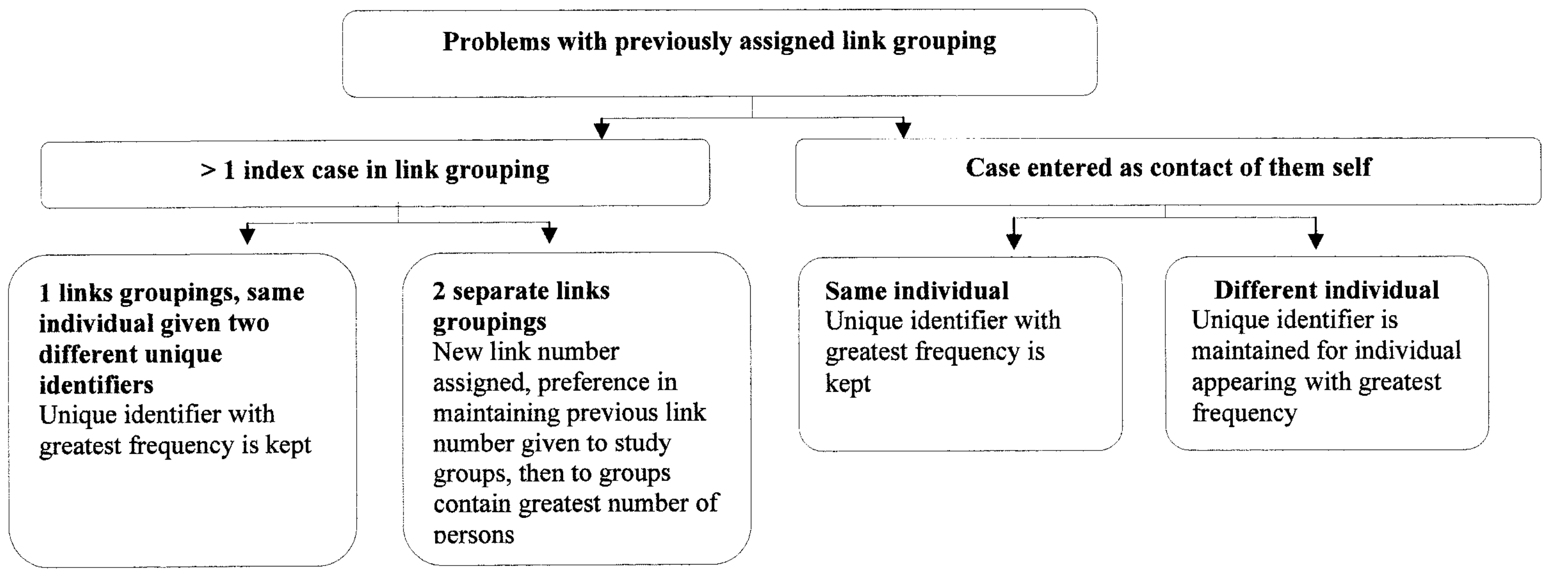
Appendix I: Management of *Original Edge List* database, algorithms for handling discordant gender, age differences, relational fields and duplicate records, Section A, B, C and D, respectively



A)

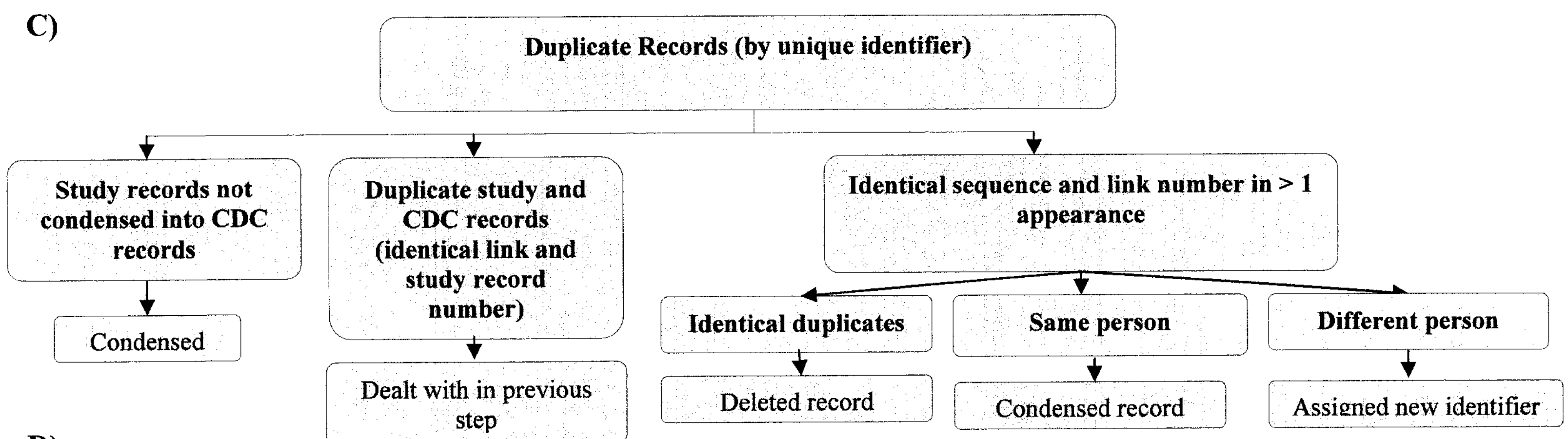


B)



Problems with previously assigned link grouping
> 1 index case in link grouping
Case entered as contact of them self
1 links groupings, same individual given two different unique identifiers
Unique identifier with greatest frequency is kept
2 separate links groupings
New link number assigned, preference in maintaining previous link number given to study groups, then to groups contain greatest number of persons
Same individual
Unique identifier with greatest frequency is kept
Different individual
Unique identifier is maintained for individual appearing with greatest frequency

C)



Duplicate Records (by unique identifier)
Study records not condensed into CDC records
Condensed
Duplicate study and CDC records (identical link and study record number)
Dealt with in previous step
Identical sequence and link number in > 1 appearance
Identical duplicates
Deleted record
Same person
Condensed record
Different person
Assigned new identifier

D)

Appendix J: Ethics

The research protocol was sent for ethics review to the Health and Research Ethics Board, Bannatyne Campus; University of Manitoba, participating RHAs in the province of Manitoba and the Health Information Privacy Committee of Manitoba Health; and the Ottawa Hospital Research Ethics Board (OHREB). The protocol was granted approval by the Health and Research Ethics Board, Bannatyne Campus, University of Manitoba "for the ethics of human use only" on June 14, 2000, and annually on August 1, 2001, September 20, 2002. Changes made to the consent forms were also approved on two occasions: September 18, 2000 and December 11, 2001. Ethics approval was granted by 6 of the RHAs in Manitoba: Brandon, Burntwood, Marquette, Norman, South Westman and Winnipeg and overall approval was granted by the Health Information Privacy Committee at Manitoba Health. (As of July 1, 2002 the Marquette and South Westman RHAs have amalgamated forming the Assiniboine RHA). The protocol was given expedited review by the OHREB and approved on 15 January 15, 2004. The protocol was up for renewal with the OHREB on January 14, 2005 and an 'Annual Renewal Report' was forwarded to the OHREB which was approved on December 22, 2004.

The Ottawa Hospital | L'Hôpital d'Ottawa

Research Ethics Board
Conseil d'éthique en recherches
798-5555 ext 14146, 14902 or 15072
Fax No. ~ 761-4311
http://www.ohri.ca/ohreb/

Thursday, January 15, 2004

Ms. Emily Meadows
Epidemiology and Community Medicine
University of Ottawa
451 Smyth Road
Ottawa, ON
K1H 8M5

Dear Ms. Meadows:

Re: Protocol # 2004029-01H **Sexual Networks of Individuals Infected with Sexually Transmitted Infections: Structure and Disease Transmission**

Protocol approval valid until - **Friday, January 14, 2005**

I am pleased to inform you that your study (listed above) was given expedited review by the Ottawa Hospital Research Ethics Board (OHREB) and is approved. No changes, amendments or addenda may be made in the protocol without the OHREB review and approval.

Approximately two months prior to the expiration date listed above, a single renewal form should be sent to the OHREB office.

The Tri-Council Policy Statement requires a greater involvement of the OHREB in studies over the course of their execution. You must inform the Board of adverse events encountered during the study, here or elsewhere, or of significant new information which becomes available after the Board review, either of which may impinge on the ethics of continuing the study. The OHREB will review the new information to determine if the protocol should be modified, discontinued, or should continue as originally approved.

Yours sincerely,

Raphael Saginur, M.D.
Chairman
Ottawa Hospital Research Ethics Board

Appendix K: Variable derivations

Variable	**Type (Values)**	**Derivation**
Eligibility	Binary	Used combination of RHA, ICD-9-CM codes, and report date to determine eligibility for supplementary questionnaire Eligibility criteria: 1. Laboratory-confirmed case of gonorrhea and/or chlamydia; AND 2. Reside in jurisdiction of participating RHA- Winnipeg, Assiniboine, Norman, Burntwood, Brandon; AND 3. Laboratory-confirmed case reported to participating RHA during respective study dates: June 1, 2002 to March 31, 2003 (Rural) October 1, 2002 to September 30, 2003 (Urban)
Variables collected for entire population		
Age	Discrete	Derived from variable 'Age' found in *Original Edge List*; MH calculated the age of cases by subtracting date of birth from the specimen date, entered data for contacts.
Co-infection	Binary	Each disease episode was examined for the concomitant occurrence of two different infections in one disease episode, using the ICD-9-CM codes
Ethnicity **CDC Treaty**	Binary	Recoded 'Treaty code', found in both the *Registry* and the *MH STI Contact* database, into a binary variable in Treaty or other ethnicity **Note:** Manitoba Health Services Insurance Plan registrants self-report having a Band or Treaty number, and it is recorded in their file. As this is a self-report system, MH estimates that this designation undercounts Aboriginal Manitobans by anywhere from 35 to 52 percent.
Geographic location	Categorical	Derived from the variable 'Coded', a numerically coded geographic residence found in the *Geocode* database.

Variable	Type (Values)	Derivation
		Derivation was based on the following algorithm: If case: 1. If study participant, use residence entered in corresponding record of *Study Case* database 2. If non-participant or no residence entered in *Study Case* database, use residence from first appearance as case in the *Registry* 3. If no residence entered in first appearance as case and a repeater, then chronologically go through disease episodes proceeding through to contact notification events, where possible If contact: 1. Use first notification event in the *Study Contact* database 2. If not identified through the *Study Contact* database or no residence entered, then use first notification event in *MH STI Contact* database 3. If no residence entered in first notification event in database and a repeat nominee, then chronologically go through notification events **Note**: The algorithm for geographic location was created on the basis that, 1) it would use the residence known to be most accurate (that given directly by a case), and 2) use the address at first appearance in database
Geographic region	Categorical 1=Winnipeg 2=Southern rural Mb (Brandon, South Eastman, Interlake, Central, Assiniboine, Parkland RHAs); 3=Northern remote Mb (North Eastman, Norman, Burntwood, Churchill RHAs) 4=Outside Mb	Further categorized geographic location into three regions of Manitoba using RHA jurisdiction
Infection acquired	Categorical	A cumulative summary of the ICD-9-

Variable	Type (Values)	Derivation
	1=chlamydia (CT) 2=gonorrhea (GC) 3=syphilis (TP) 4=CT & GC 5=CT & TP 6=GC & TP 7=CT & GC & TP 8=Unknown	CM codes entered for each disease episode in the *Registry*
Repeat nominee	Binary	Historical records were examined for the occurrence of two or more notification events from October 1, 2002 to September 30, 2003
Repeater	Binary	Historical records were examined for the occurrence of two or more reported disease episodes from July 1, 2002 to September 30, 2003
Variables collected on voluntary study participants		
Concurrency	Binary	Derived from examining the two variables the first date of sexual contact and last date of sexual contact for all sexual relations of a specific study participant in order to determine if relations overlapped in time
Frequency of sex in previous 3 months	Categorical (Refer to question for listing of categories)	Refer to question R28 (**Appendix F**)
Length of social acquaintance	Categorical (Refer to question for listing of categories)	Refer to question R25 (**Appendix F**)
Reason for healthcare visit 1, 2, 3	Categorical (Refer to question for listing of categories)	Refer to question R5 (**Appendix F**) 2 & 3 - If specified 'Other', and symptoms was entered, recoded to fit record into pre-defined 'Symptoms' category Refer to **Section 2.2.3**

Appendix L: A sample PAJEK readable file created from the *Syph.bat* computer program, © Ann Jolly 2003.

```
*Vertices 8511
  1      '10000'   ic Red     ellipse
  2      '10001'   ic White   ellipse
  3      '10003'   ic White   ellipse
  4      '10004'   ic White   ellipse
  5      '10005'   ic White   ellipse
  .
  .
  .
8508     '50031'   ic White   box
8509     '50032'   ic White   box
8510     '50033'   ic White   box
8511     '50034'   ic White   box
*Edges
  1 5464     1     c Black
  2 8127     1     c Black
  10 8066    1     c Black
  11 6447    1     c Black
  11 6926    1     c Black
  .
  .
  .
8313 8479    1     c Black
8347 8360    1     c Black
8347 8361    1     c Black
8347 8362    1     c Black
8347 8481    1     c Black
```

Appendix M: An output from PAJEK, of the number of components by number of individuals, calculating cumulative percent of persons.

```
--------------------------------------------------------------------------
Reading Network   ---   C:\Documents and Settings\Emily
Meadows\Desktop\Pajek\Pajek 28 Oct 04\pajek_infection_study.txt
--------------------------------------------------------------------------
 Working...
   13683 lines read.
 Time spent:  00:00
--------------------------------------------------------------------------
Weak Components
--------------------------------------------------------------------------
 Time spent:  00:00
--------------------------------------------------------------------------
1. Weak Components of N1 [>=1] (8511, comp.=3700)
--------------------------------------------------------------------------
Dimension: 8511
The lowest value: 1
The highest value: 3700

Frequency distribution of cluster numbers:
```

Cluster	Freq	Freq%	CumFreq	CumFreq%	Representative
1	4	0.0470	4	0.0470	10000
2	5	0.0587	9	0.1057	10001
3	2	0.0235	11	0.1292	10003
4	2	0.0235	13	0.1527	10004
5	7	0.0822	20	0.2350	10005
6	9	0.1057	29	0.3407	10006
7	3	0.0352	32	0.3760	10007
8	22	0.2585	54	0.6345	10008
9	33	0.3877	87	1.0222	10009
10	3	0.0352	90	1.0575	10010
11	5	0.0587	95	1.1162	10011
12	2	0.0235	97	1.1397	10012
13	2	0.0235	99	1.1632	10013
14	2	0.0235	101	1.1867	10014
15	3	0.0352	104	1.2219	10015
16	2	0.0235	106	1.2454	10016
17	4	0.0470	110	1.2924	10017
18	3	0.0352	113	1.3277	10018
19	3	0.0352	116	1.3629	10019
20	3	0.0352	119	1.3982	10020
21	5	0.0587	124	1.4569	10021

Appendix N: Equations used in cluster analysis

Approximate covariance estimation for clustering (ACECLUS) procedure

The sample covariance matrix ($\mathbf{S} = s_{jk}$) was defined as:

$$s_{jk} = \frac{1}{\eta(n-1)}\sum_{i=2}^{n}\sum_{h=1}^{i-1}(x_{ij} - x_{hj})(x_{ik} - x_{hk}),$$

where $\mathbf{X} = (x_{ij})$ is the data matrix with n observations (rows) and v variables (columns)

The pooled within-cluster covariance matrix ($\mathbf{W}$) was approximated by $\mathbf{A} = (a_{jk})$:

$$d_{ih} \begin{cases} 1 & \text{if} \quad \sum_{j=1}^{v}\sum_{k=1}^{v} m_{jk}(x_{ij} - x_{hj})(x_{ik} - x_{hk}) \le \mu^2, \\ 0 & \text{otherwise} \end{cases}$$

where μ is an appropriately chosen value and $\mathbf{M} = (m_{jk})$ is an appropriate metric.

$$a_{jk} = \frac{\sum_{i=2}^{n}\sum_{h=1}^{i-1} d_{ih}(x_{ij} - x_{hj})(x_{ik} - x_{hk})}{2\sum_{i=2}^{n}\sum_{h=1}^{i-1} d_{ih}},$$

P was transformed into a value of t:

$$t^2 = 2v\left\{\left[F^{-1}{}_{v,n-v}(p)\right]^{[(n-v)/(n-1)]}\right\},$$

where $F^{-1}{}_{v,n-v}$ is the quantile (inverse cumulative distribution) function of an F random variable with v and n-v degrees of freedom

t is than multiplied by $1/\sqrt{2i}$ times the root mean square (RMS) distance between observations

Convergence metric:

$$e_i = \frac{1}{v}\left\| Z'(A_i - A_{i-1})Z \right\|,$$

where ||…|| indicates the square root of the sum of the squares of the elements in the matrix, and $\boldsymbol{Z}$ is a user-specified matrix.

Ward's minimum variance method

$\mathbf{x_{ijk}}$ = value of the j^{th} variable for the i^{th} element of cluster k

$(\bar{x}_{.ik}, \bar{x}_{.2k}, \bar{x}_{.3k}, \ldots, \bar{x}_{.nk})$ = the vector of the centroid of cluster k,

where $\bar{x}_{.jk} = \frac{1}{m_k} * \sum_{i=1}^{m_k} x_{ijk}$, m_k is the number of elements in cluster k

Error sum of squares of cluster k:

$$E_k = \sum_{j=1}^{n}\sum_{i=1}^{m_k}(x_{ijk} - \bar{x}_{.jk})^2$$

$$E_{Total} = \sum_{k=1}^{K} E_k$$

Stopping rules, Pseudo F and T^2 tests

$$\text{Pseudo F test} = \frac{\left[traceB \big/ (k-1) \right]}{\left[traceW \big/ (n-k) \right]},$$

where B and W are between and the pooled within-cluster sums of squares and cross products matrices, and *n* and *k* are the total number of items and the number of clusters in the solution, respectively.

$$\text{Pseudo } T^2 \text{ test for joining } C_K \text{ and } C_L = \left[\frac{(B_{KL})}{\left(\left[(W_K + W_L) \big/ (N_K + N_L - 2) \right] \right)} \right]$$

$$\text{R-squared } (R^2) = 1 - \left(pooled\ within-cluster \big/ T \right)$$

Appendix O Characteristics of 102 individuals reporting same-sex relationships

	Male (n=59)	**Female (n=43)**
	Freq (%)	**Freq (%)**
Age (median, IQR)	24 (19,34)	20 (17,25)
Treaty	11 (18.64)	11 (25.58)
Winnipeg	40 (67.8)	26 (60.47)
Repeater	7 (11.86)	12 (27.91)
Infection Acquired		
Chlamydia	36 (61.02)	38 (88.37)
Gonorrhea	20 (33.9)	0 (0)
Syphilis	0 (0)	1 (2.33)
Chlamydia and Gonorrhea	3 (5.08)	4 (9.3)
Degree (mean, range)	1.9 (1,9)	2.6 (1,17)
Component size (mean, range)	3.8 (2,33)	7.4 (2,33)
% same-sex relations in component* (mean, range)	81.5 (4.8,100)	56.0 (3.4,100)

* Lines = sum of degrees/2