Social Network Analysis in Tuberculosis Control Among the Aboriginal Population of Manitoba

BY

Assaad Al-Mouaiad Al-Azem

A Thesis submitted to
the Faculty of Graduate Studies
In Partial Fulfillment of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY

Department of Community Health Sciences
University of Manitoba
Winnipeg, Manitoba

THE UNIVERSITY OF MANITOBA

FACULTY OF GRADUATE STUDIES

*****

COPYRIGHT PERMISSION

Social Network Analysis in Tuberculosis Control among the Aboriginal Population of Manitoba

BY

Assaad Al-Mouaiad Al-Azem

A Thesis/Practicum submitted to the Faculty of Graduate Studies of The University of Manitoba in partial fulfillment of the requirement of the degree

OF

DOCTOR OF PHILOSOPHY

To My Parents, Family, Friends, and Coworker

Acknowledgments

Advisor: Dr. Lawrence Elliott

Committee Members: Dr. Amin Kabani
Dr. Earl Hershfield
Dr. Kue Young

Others: Dr. Tom Hassard
Dr. Ann Jolly
Dr. Bob Tate
Dr. Shiva Halli
Ms. Mary Cheang

Funding: Dr. Amin Kabani (Medical Microbiology)
First Nation and Inuit Health Branch
Faculty of Medicine (Associate Dean, Research)

Support: Manitoba Central Tuberculosis Registry Office
Manitoba Institute of Child Health
J.A. Hildes Northern Medical Unit
Community 1:
*Band Office
*Advisory Board
*Nursing Station
*Participants

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Tuberculosis (TB) in the province of Manitoba (population 1.15 million) is a multi-faceted problem, with trends that parallel the current national trends. Canadian-born people with treaty status and foreign-born individuals make up the majority of new TB cases and have sustained the steady level of TB incidence in Manitoba in the last decade. Therefore, enhanced TB control measures are required for these population subgroups, if the incidence of this important disease is to be further reduced.

The present study is composed of three parts. The first part involves a description of TB in Manitoba between 1992 and 1999. The study found that Canadian-born people with treaty status form the largest single group of all TB cases in Manitoba (44%), although they represent only 8.9% of the provincial population. A single TB strain, FP1, dominates all TB cases with isolates in Manitoba (25.8%). Canadian-born people with treaty status represent 75% of all FP1 cases. Certain Northern First Nation communities (reserves) have TB incidence as high as 496 per 100,000 person-years, and 64% of Canadian born with treaty status TB cases originated from only eight communities.

The second part involved a pilot study of the feasibility of using Social Network Analysis (SNA) in TB epidemiology and control in Manitoba. Contact tracing investigation data kept at the Manitoba TB Registry office was employed. We confirmed that it is feasible to use SNA using the Manitoba TB data, and found that TB outbreak networks in Manitoba have varying degrees of compactness, with boundaries that extend beyond the geopolitical boundaries of the communities under investigation. The pilot study findings

raised important questions for further study, in particular the question as to whether TB outbreak networks reflect the social structure of the affected community.

The third part was a case-control study in which SNA was used in conjunction with primary interview data and contact investigation data from one high TB incidence Northern Manitoba Aboriginal community (Community 1). The main objectives of this part were to determine the feasibility and utility of using primary interview data for constructing and analyzing the TB transmission network in a high incidence community, and to determine whether or not the TB network is a reflection of the affected community's social structure. Eighty-three individuals were successfully recruited for the study in this remote community (62% of the target sample). The case-control study confirmed that the TB network reflected the underlying social structure of the community. The TB network of Community 1 was found to be a scale free network and very compact, with an obvious centrally located and densely connected group of network members. The study demonstrated that as k increases in the Seidman k-core collapse sequence, the probability of active or latent tuberculosis infection increases. Those in the network who were exposed to 7 or more TB cases had a 100% probability of being TB cases themselves. Multivariate analysis confirmed that the TB group members (in comparison to the control and contact groups) lived in houses which were significantly more crowded (mean of 1.6 people per room), and were socially connected to a greater number of other TB patients (to 4.0 other TB cases vs. 0.6 for controls). The study enabled the creation of profiles of the TB, contact and control (general population) groups, respectively.

SNA was successfully employed in this study to visualize TB networks, to identify characteristics of key individuals in TB outbreaks, and to delineate outbreak boundaries. It

is concluded that SNA is a useful epidemiological tool for TB control. SNA methods can improve contact tracing performance and potentially improve TB control and prevention programs.

SECTION I: INTRODUCTION and OBJECTIVES

Introduction

Tuberculosis (TB) is considered the most common source of death due to an infectious organism in the world. Approximately one-third of the world's population is infected with the bacillus *Mycobacterium tuberculosis,* causing approximately three million deaths annually(6, 37, 92). The future incidence and mortality rates of TB are expected to increase in developing countries due to several factors, including famine, war, natural disasters, the HIV/AIDS epidemic, and the emergence of multi-drug resistant strains (MDR) of *M. tuberculosis*(6, 96). Industrialized countries face the same risk as developing countries of increased TB rates due to changing immigration patterns, homelessness, decrease in public health funding, and greater volumes of overseas travel(6, 96).

Migration from countries with high TB incidence rates continues to sustain the incidence of TB in developed countries, presenting a major health problem for health authorities. Large numbers of studies in Canada illustrate that though there has been a significant decrease in the number of reported TB cases among the population excluding Aboriginals (Aboriginal defined as status and non-status Indian, Metis, Inuit and Innu), the number of reported TB cases among foreign-born (FB) individuals increased markedly during the same period(26, 39, 46, 52, 56, 59, 60, 77, 86, 90, 99). Similar trends noted in USA, New Zealand, and Western Europe illustrate that FB TB patients are major contributors in slowing down the decrease of morbidity rates in these industrialized countries(12, 35, 60, 81, 82, 92, 97, 109, 115, 123, 129).

In Canada over the last century, TB incidence first declined and then leveled off. The decline has slowed to the point of stopping, particularly in the past ten years(13, 26, 36, 39, 46, 52, 56, 77, 90, 91, 112). Currently, any prospective immigrant to Canada is required to undergo medical examination to determine the presence or absence of active TB. Active cases are not admissible to Canada, and inactive cases are usually put under public health surveillance after arriving to Canada to monitor their TB status closely(29, 129). The proportion of TB cases reported from FB individuals in Canada has increased over the years, from 17% in 1970, to 35% in 1980, and to 64% in 1998(56, 76). The majority of TB cases (75%) in Canada reside in Ontario, Quebec, and British Columbia(6, 76). These three provinces contain 90% of FB TB cases, as they represent the most attractive Canadian provinces for new immigrants. In contrast, FB TB represents 29% of TB patients in Manitoba compared to 81%, 48%, and 60% in Ontario, British Colombia, and Quebec respectively(6, 12).

Study Rationale

As described above, in the Canadian-born non-aboriginal population of Manitoba, TB is close to elimination. However, TB incidence in some northern Manitoba First Nations communities range from 6 to 150 times higher than that in the non-aboriginal population, and exceeds the rates in many developing nations. Conventional TB control methods have been unable to eliminate disease spread within First Nations communities, and incidence has remained at hyper-endemic levels with frequent outbreaks in some of these communities. Given this persistent hyper-endemic transmission of TB in Manitoba, it is important that new approaches to the investigation and control of TB be explored. One new approach, social network analysis (SNA), has been demonstrated to be useful in the

understanding and control of other infectious diseases, including sexually transmitted infections and HIV(69, 72, 130).

SNA is a set of techniques and measures used to understand the interactions between people(3, 5, 69-72, 107, 117, 122). The current study is an exploration of the potential applicability of this new approach to the understanding and control of TB transmission in Manitoba, particularly among the First Nation population.

Objectives

The overall aim of the study is to determine the feasibility and utility of applying the methods of social network analysis (SNA) in understanding the epidemiology of tuberculosis in Manitoba.

Specific objectives of the study include:

1. To determine the feasibility and utility of applying SNA in a retrospective manner to existing TB case and contact data from the Manitoba TB Registry on eight high incidence communities;
2. To determine the feasibility and utility of collecting primary (interview) data for constructing and analyzing TB transmission networks for one high-incidence Manitoba First Nations community;
3. To compare the characteristics of the TB transmission networks in this community with the characteristics of:
    a. The social networks of the contacts of the TB cases; and
    b. The social networks of age- and sex-matched non-TB controls from the same community.

SECTION II: LITERATURE REVIEW

Chapter 1: Epidemiology of Tuberculosis

1.1 Historical background

Tuberculosis is a disease that has coexisted with humanity since time immemorial, and evidence of this disease has been found in Egyptian mummies dating as far back as 4000 BC(14). There are conflicting theories about TB history in Canada before the arrival of Europeans, and whether they brought it with them or it was in Canada before the arrival of Europeans. Proponents of the latter possibility suggest that TB existed in Canada prior to the arrival of Europeans, but did not reach high endemic levels until the Aboriginal people came into contact with European people and culture, compromising their well-being and immunity and making them vulnerable to TB(36, 52, 60, 126). As supporting evidence, these authors point to the discovery of TB evidence in the remains of First Nations people dating as far back as 700 – 1490 AD, indicating the existence of Mycobacterium tuberculosis infections in North America before the arrival of European settlers(36, 52, 60, 126).

In the seventeenth century when European traders arrived in Canada, the tuberculosis epidemic in Europe was nearing the end of its cycle (52, 126). It has been hypothesized that there was a shift in the genetic makeup of the European population as a consequence of the survival of those with moderate resistance to TB infection and the death of the susceptible population(52, 60). By comparison, the Aboriginal people of Canada

were immunologically naïve to tuberculosis and highly susceptible to infection and disease. This proposal is known as "virgin soil theory"(132). Exposure did not occur at the same time for all Aboriginal populations in Canada; in fact, the inhabitants of the East Coast were exposed almost 100 years earlier than the prairie Aboriginals due to their earlier exposure to European sailors(21, 50, 126).

The long history of tremendous suffering from the ravages of tuberculosis disease continues up to recent Canadian history. In the 19th century, for example, one fifth of all Canadians were infected with TB and large numbers died as a result of this affliction. By the beginning of the 20th century, the mortality rate among the Aboriginal people of Saskatchewan and Alberta reached as high as 9,000 per 100,000(52, 125, 126). Throughout the years, tuberculosis was recognized as a great killer that primarily affected less fortunate individuals(40), and disproportionately affected Aboriginal people. The high-risk status of the latter group is attributed to a combination of factors including the colonization history described above, socioeconomic factors, and a possibly increased biological susceptibility(132).

Improvements in socioeconomic conditions, public health services, drug therapy and medical surveillance have markedly reduced the mortality rate (0.4 per 100,000 per year by 1987) and overall incidence (5.9 per 100,000 per year by 1998) of TB in Canada (21, 22, 55, 104). Nevertheless, certain groups remain at higher risk for acquiring the disease, including the Aboriginal population, foreign-born individuals from countries with high TB incidence rates, the homeless, and the HIV-infected population(2, 4, 7-9, 21, 22, 55).

1.2 Tuberculosis in Canada

Tuberculosis is primarily a socioeconomic disease and is related to poverty, poor nutrition, overcrowded conditions, inadequately heated and ventilated housing, stress, poor underlying health, and homelessness(22). The majority of TB patients in Canada (84%) reside in the provinces of Ontario, Quebec, British Columbia and Alberta. Presently, the majority of TB cases within these provinces occur among foreign-born individuals(12, 55, 76). Conversely, in Saskatchewan, Manitoba, and the Northern Territories, the majority of cases occur in Aboriginals. In Manitoba, Canadian-born individuals with treaty status and the communities they live in have the highest TB incidence: 48.4 per 100,000 per year overall, with rates as high as 496 per 100,000 per year in some communities(11, 13).

A large number of TB cases (60%) reside in major Canadian Cities (Quebec City, Montreal, Ottawa-Hull, Toronto, Hamilton, Winnipeg, Calgary, Edmonton, and Vancouver)(76). Gaudett and Ellis(46) conclude in their paper that Census divisions with the highest TB incidence (over 20/100,000) are located in northern regions, with Aboriginal people representing 80% of these cases. Census divisions with moderately high TB incidence of 10-19/100,000 are located both in the northern regions and major cities, with Aboriginal people and foreign-born individuals as the main component of these TB cases(46). Further, Grzybowski (1980)(50)and others(58, 99) have observed a high incidence of TB among inner city residents. To describe inner city TB cases briefly, most of the cases are seniors, single males, substance abusers, low-income earners, unemployed individuals, and individuals with lower levels of education(22, 128). A high proportion of inner city TB cases are alcoholics, drugs abusers, or homeless(46).

In 1998, the incidence among non-Aboriginal Canadians was 1.5 per 100,000 per year. In contrast, status Indian, Aboriginal (defined as status and non-status Indian, Metis, Inuit and Innu), and foreign-born population groups had incidence of 35.4, 22.5 and 21.4 per 100,000 per year, respectively(76). Furthermore, the proportion of all TB cases reported among foreign-born individuals residing in Canada has increased from 35 percent in 1980 to 64 percent in 1998(76).

In an attempt to eliminate tuberculosis in Canada, the Tuberculosis Prevention and Control Unit of Health Canada had implemented guidelines recommended by the World Health Organization in 1997, which stated that tuberculosis in Canada should be eliminated within the next few years. However, despite the adoption of these guidelines, the incidence of TB remains high in certain geographic and demographic zones, making elimination of the disease within the projected timeframe an increasingly difficult goal (55). In addition, some authors note that although TB control policies have been similar for all populations, the decrease of incidence in tuberculosis varies among different population subgroups, particularly the Aboriginal subgroup (36, 58, 60). These authors suggest that Aboriginal populations may require alternative TB control policies if incidence is expected to come down to that observed in the Canadian-born non-Aboriginal population.

1.2.1 Tuberculosis Among Foreign Born (FB) People

Every year, Canada receives over two million visitors and accepts as many as 250,000 landed immigrants(36). Due to the high mobility and diverse health care access and cultural practices of these people, it is difficult to establish accurate tuberculosis rates among the foreign-born population(84). However, national statistics show that the proportion of foreign-born tuberculosis cases among Canadian total cases is increasing(76,

89). In 1998, foreign-born people formed the majority of tuberculosis cases in Canada (64%)(76, 89). Vietnam, China, Philippines, India and Hong Kong, were the top five countries from which foreign-born tuberculosis cases originated(76).

1.2.2 Aboriginal and Treaty people

Aboriginal is a term usually used to describe the indigenous inhabitants of Canada and their descendants. It embraces those people registered as status Indians living on and off reserves, as well as non-status Indians, Metis and Inuit (76).

Under the Federal Indian Act, the federal government have implemented an Indian Registry System to keep records and information about each person with legal treaty status(136). Treaty Indians (with Treaty status under the Federal Indian Act) are people eligible to receive education, health and social services under treaties concluded between the Canadian government and Indian bands.

1.2.3 Tuberculosis Among Canadian Born with Treaty Status

Following the arrival of European settlers in Canada and exposure of the Aboriginal people to TB, the Aboriginal population began to decline(126). Though the precise Aboriginal population was unknown, it is estimated that in 1871, approximately 200,000 Aboriginal people lived in Canada(126). By the end of the nineteenth century, this number had dwindled to 100,000(126). Tuberculosis and other important risk factors, such as malnutrition and alcoholism, are thought to be responsible for this decrease(126). In the twentieth century, the Aboriginal population began to rebound to the point where the birth rate exceeded the death rate(132). As a result of government control programs, deaths due to tuberculosis dropped from 550 to 100 per 100,000 per year between 1945 and 1955.

However, despite these efforts, tuberculosis incidence in the Aboriginal population remain 16 to 24 times higher than those of non-aboriginal Canadians(46, 104). Incidence and death rate due to tuberculosis among Aboriginal people in Canada declined until 1980; since that year incidence did not decrease but have plateaued(26, 39, 60).

1.2.4 Tuberculosis in Manitoba

The population of Manitoba can be divided into Canadian-born and foreign-born people; and the Canadian-born population can be further divided into Canadian-born with treaty (First Nations) status and Canadian-born without treaty status. Manitoba has a population of 1.15 million with 859,593 Canadian-born (CB) non-treaty people (78.1% of the province's population), 98,197 Canadian-born treaty people (8.9%), and 142,505 Foreign-born (FB) people (13%)(11-14, 113). The province has one major city, Winnipeg, where 60% of the population resides. The rest of the province is composed of scattered small urban areas, rural areas, and reserves. All diagnosed cases of TB in Manitoba are reported and registered in the Manitoba Central Tuberculosis Registry (MCTR).

TB in Manitoba is a multifaceted problem and it differs from TB in other Canadian provinces, mostly in its composition of patients. Foreign-born TB cases represent only 29% of TB patients in Manitoba, compared with 81%, 60% and 48% in Ontario, Quebec and BC, respectively(11, 12). TB incidence trends in Manitoba are parallel to current national trends (Figure1)(55, 77) and although TB rates have been declining in every population subgroup, comparative rates among Canadian-born with treaty status individuals in Manitoba (48.4 per 100,000 per year) remain more than fourteen times that of the Canadian born with non-treaty status subgroup (3.3 per 100,000 per year) and over

two times that of the foreign-born subgroup (22.0 per 100,000 per year)(12, 13, 77). It is apparent that Canadian-born with treaty status and FB population subgroups sustain the steady level of TB incidence in Manitoba and are consequently in need of new measures for more effective TB control(11-13).

Treaty status individuals form a disproportionate number of current TB cases in Manitoba. Despite the fact they represent only 8.9 percent of the province's population, this population subgroup represent 44.4 percent of all tuberculosis cases in Manitoba. Remarkably, this figure still under-represents the actual number of cases within this group, due to the unknown number of Aboriginal people in Manitoba who have not yet claimed treaty status. Using treaty status population numbers gives more accurate results(11).

The incidence of tuberculosis among the treaty status population (48.4 per 100,000 per year) is higher than both the provincial rate (9.2 per 100,000 per year) and the national rate (5.9 per 100,000 per year) by five and eight times, respectively. Moreover, certain northern First Nation communities have TB incidence as high as 496.3 per 100,000 person-years(11).

In summary, TB cases in Manitoba consist mainly of Canadian-born individuals (with and without treaty status). Of these individuals, Canadian-born TB cases with treaty status (living both in Winnipeg and in northern reserves) are dominant. TB in Canadian-born non-treaty residents is no longer a serious problem since all indicators point to this disease being near elimination in this population group. In contrast to the Canadian-born subgroup, foreign-born individuals represent only 29 percent of all TB cases in Manitoba. Most FB TB cases in Manitoba reside in Winnipeg and comprise the dominant portion of unique isolated fingerprint patterns.

Chapter 2: Tuberculosis Prevention and Control

2.1 Sanatorium era

By the end of the eighteenth century and the beginning of the nineteenth century, one in every 13 deaths (7.7%) in Canada was the result of tuberculosis, with a mortality rate of nine percent among the Aboriginal Canadian population (33, 60, 126). High mortality rates necessitated the establishment of a new treatment program to combat this disease. Since there was no effective treatment at the time to combat TB, treatment was instead based on reducing the suffering of the patients through symptom relief (126), starting from initial family support efforts and home nursing and graduating later to the creation of the sanatorium and the new treatment techniques(36, 37, 126) of the so-called "sanatorium era". However, as treatment developed, financial burden increased, becoming especially prohibitive as patients were expected to pay for their own sanatorium costs (126). After the discovery of anti-tuberculous drugs and the improvement of living conditions, the sanatoria were slowly phased out, and by 1980 almost all of these facilities in North America were closed(21, 22, 50, 126).

2.2 Principles of Tuberculosis Control and Prevention

Tuberculosis is a social disease and any successful control strategy must incorporate both medical and social action (22, 77). The basic prevention of TB relies on two elements:

1) Stopping the spread of tuberculosis from active cases to new persons

2) Averting the relapse of latent tuberculosis infections (LTBI)(86).

Tuberculosis control in Canada ultimately depends on case finding, contact tracing, case holding, and treatment of latent infection(59, 133).

Anti-tuberculous drug treatment became the cornerstone of tuberculosis control and had tremendous positive influence on tuberculosis mortality rates(21, 50). Between 1926 and 1946, death rates from TB dropped by 3%; over the next five years, mortality rates dropped a further 12% and decreased another 20% following the discovery and introduction of streptomycin in TB management. Streptomycin was discovered by professor Selman A. Waksman in 1944(21, 22, 50, 60, 76).

TB control efforts have virtually eliminated tuberculosis from the Canadian-born non-Aboriginal population. However, this success has not been uniformly distributed among the remaining segments of Canadian society(36). The rapid changes in the TB spread populations required a change in control priorities(36), with focus mainly on high-risk groups(22, 36). In 1988, Young et al. (133) suggested shifting TB control policies in Canada from population-based screening to high-risk screening, with a concomitant shift from mass screening to screening techniques targeting high risk groups(133).

2.3 BCG Vaccine

BCG is a live attenuated strain of *M. bovis* that serves as a vaccine for *M. tuberculosis*(52, 60, 86). Although there has been some controversy around the administration of BCG, the majority of researchers agree that BCG provides some degree of protection among Aboriginal children against the more serious forms of TB, such as TB meningitis and miliary TB(38, 112). In the mid-1960's, the Canadian government adopted a policy of routine BCG vaccination for newborns of First Nations mothers. Studies conducted in Manitoba have shown that BCG offers 60% protection among Aboriginal

children against TB development(134), and it continues to used for TB prevention on First Nation reserves in Manitoba(133). However, use of BCG has been suspended in some other First Nations regions in Canada due to increased rates of complications among infants with inherited or acquired immune deficiencies (140, 141)

2.4 Treatment of Latent Tuberculosis Infection

The main focus of tuberculosis control among high-risk groups lies in detection of latent tuberculosis infections (LTBI) before they develop into active cases. Screening of high-risk groups with the tuberculin skin test is an essential step in identifying and treating individuals with latent infection. High-risk groups include:

1) Individuals with recent exposure to TB
2) HIV-infected individuals
3) Injection drug users
4) People with special medical risk factors (e.g. diabetes mellitus, renal failure)
5) Residents and employees of high-risk sites (e.g. prisons, nursing homes, homeless shelters, overcrowded, poorly-heated, and poorly-ventilated houses)
6) Health care workers
7) First Nations people
8) Foreign Born people from high prevalence countries
9) The economically disadvantaged(4, 21, 25, 39).

For more than 30 years, a 6-12 month regimen of Isoniazid (INH) has remained the choice for the treatment of latent tuberculosis infections (LTBI)(2, 9). However, despite its cost-effectiveness(84), patient non-compliance and toxicity have resulted in limited success of this agent(2, 25, 59, 86).

In July of 1999, the American Thoracic Society (ATS) and CDC issued a joint official statement about LTBI treatment with the title "Targeted tuberculin testing and treatment of LTBI", which recommended using the term "treatment of LTBI", instead of "preventive therapy" or "chemoprophylaxis"(9). This official statement advocated the use of tuberculin tests to identify people at high risk of development of TB and who will benefit substantially from treatment of their LTBI. The report also endorses using a new rating system for treatment. Composed of two parts, this rating system first uses letters alphabetically to identify the strength of recommendation for a particular treatment regimen: A = "preferred"; B = "acceptable alternative"; C = "offer when A and B cannot be offered". The second part of this rating system numerically marks the strength of evidence supporting the chosen treatment: I= "randomized clinical trial"; II= "non-randomized clinical trial or trial conducted in other population"; III= "expert opinion". In brief, INH daily use for 9 months is the first choice with an A (II) recommendation, followed by INH twice weekly with a B (II) recommendation for both HIV positive and negative persons. Rifampin can be used for a shorter period of time: daily for 4 months with a recommendation rate of B (II) for HIV negative and B (III) for HIV positive. Rifampin-Pyrazinamide can be used daily for 2 months with a recommendation degree of B (II) for HIV negative and A (I) for HIV positive(9).

2.5 Treatment of Tuberculosis Disease

Treatment of tuberculosis by anti-tuberculosis drugs is a vital and highly effective step in the management and control of tuberculosis(22). Treatment of tuberculosis typically involves three objectives:

1) Kill as many of the active microorganism as possible;

2) Sterilize or eradicate the semi-dormant bacteria; and

3) Prevent the development of drug resistance.

There are two lines of therapy for the treatment of TB(2, 59). The first line of therapy includes the following medications: Isoniazid, Rifampin, Rifabutin, Pyrazinamide, and Ethambutol. The second line of treatment for resistant strains of TB includes Cycloserine, Ethionamide, Capreomycin, Kanamycin and Fluoroquinolones(59). INH, Rifampin, and Streptomycin are bactericidal in nature and as such are important agents for the elimination of active microorganisms. Rifampin and Pyrazinamide are sterilizing agents that eliminate bacteria in the dormant state (59). A key element in the success of treatment is patient compliance. The optimal way to achieve this compliance is by implementing Directly Observed Therapy (DOT). DOT is defined by the Canadian Tuberculosis Standards (5^{th} Edition) as: "The process whereby the ingestion of every dose of therapy for active disease is directly observed, by healthcare worker or pill dispenser, to have been ingested, also referred to as fully supervised therapy"(59, 76).

2.6 Contact Tracing Investigation

Individuals who come into contact with tuberculosis patients are at risk of contracting the disease. At the time of contact, the risk of infection for healthy individuals is 2% to 3% and increases to 5% to 12% during the next two years(76, 106). Public health officials usually conduct contact tracing to identify individuals who have knowingly or unknowingly come into contact with an infected individual and are themselves at risk(86). Contact tracing could also be useful for the identification of latent cases for preventative treatment, a procedure which in the past has typically relied on the implementation of tuberculin testing among high-risk groups(16, 98).

Deficiencies and limitations of contact tracing as a control strategy arise from the fact that a good number of tuberculosis cases do not undergo contact tracing investigations, and even among those cases that do go through contact tracing investigation, there are significant gaps in the information obtained. Consequently, large numbers of actual contacts do not go through testing or treatment(87, 98).

2.7 Restriction Fragment Length Polymorphism (RFLP) Typing

Restriction fragment length polymorphism typing (RFLP), or DNA fingerprinting, is a relatively new and very useful technique in tuberculosis strain identification and tracing during outbreaks of tuberculosis(24, 28). RFLP can also overcome some of the limitations of traditional typing methods(28, 120). The DNA insertion sequence 6110 (IS6110) is the key to RFLP typing(31), and is a 1361 base-pair (bp) fragment that belongs to the insertion sequence (IS3) family(49). These genomic components of the *Mycobacterium tuberculosis* chromosome exist in different copy numbers (10-15) that vary from strain to strain(30, 42, 57), and reports indicate that these insertion sequence elements are stable over 2 to 3 years (24, 42, 49). The basic RFLP technique involves DNA extraction from live bacterial cultures followed by Southern blotting and hybridization with specific probes. This, in turn, identifies the strain RFLP type or DNA fingerprint(30, 57, 120). In order to facilitate the comparison of results from various laboratories, the procedure itself has been standardized(120).

When compared with classical identification techniques, DNA fingerprinting has been found to be comparable and reliable(116). Consequently, it has been used in different epidemiological studies around the world to investigate the spread of tuberculosis in different communities(20, 23, 44, 47, 67, 119), including multi drug resistant strains(108),

and to assess the effectiveness of tuberculosis infection control(43). Research and mini epidemic (local outbreak) investigations are the main use of RFLP typing in Canada today(15, 32, 74, 75).

2.7.1 Limitations of RFLP

RFLP typing depends on IS6110, and though most TB strains have multiple copies of IS6110, some strains of TB have few or no copies of IS6110(10, 74, 103, 131). RFLP is based on the premise that when TB patients are infected with a TB strain of the same fingerprint, they are epidemiologically linked(74, 110, 131). This in turn depends on the presumption that the DNA genotype remains constant, and if any changes occur, they produce significant fingerprint diversity within the population(74, 131). It has been suggested that the fingerprint stays stable for at least two years(110).

If the DNA genotype changes occur quickly, then all the epidemiological links of transmission will be disguised(74, 110, 131). However, if the changes progress very slowly, the result will link cases which were not transmitted recently(74, 110, 131). Therefore, analysis and conclusions from RFLP results must be considered with caution(74). TB cases with similar fingerprints in one community can be explained in several ways(74): the source of the strain with one fingerprint can be from recent transmission, concurrent relapse of a latent TB infection acquired in the past, dominance of a local strain, or lab contamination. Using other typing methods beside RFLP will be very helpful in overcoming some of the RFLP limitations(10, 74, 110, 131). RFLP was the only method used in this study.

2.8 Tuberculosis Control Among Canadian First Nations

The persistent high incidence, plus the emergence of drug-resistant tuberculosis cases and HIV has forced several Canadian government departments to formulate a joint strategy for tuberculosis elimination among First Nations communities. The goal of the strategy is to eliminate tuberculosis by the year 2010, although government reports identify the difficulty in attaining such a goal in this time frame(26, 39, 60).

Regional policies used to control tuberculosis in First Nations communities include:

- Case finding and case holding;
- Contact tracing with Directly Observed Therapy (DOT) for LTBI;
- Surveillance at the community, regional, and national levels;
- BCG vaccinations;
- Health education and training; and
- Research.

Regardless of the policies implemented for the control of tuberculosis among the Canadian born with treaty status population, the ongoing involvement of community members and increased sensitivity towards the Aboriginal culture will undoubtedly be critical for the success of these tuberculosis control programs(26, 39, 60).

2.9 Obstacles and Challenges Facing TB Control Programs

Unfortunately, tuberculosis screening and even infection control programs may be compromised by lack of accurate diagnosis and treatment of the disease and patient non-compliance to treatment (84). As Canadian doctors become less experienced with tuberculosis disease, delays in proper diagnosis result in increased transmission, morbidity

and mortality, and could eventually lead to failure of tuberculosis control in Canada(22, 36). Iseman (63) has categorized the challenges facing tuberculosis control into major and minor obstacles.

Major obstacles include such issues as:

- Failure to deliver required medications;
- Inability to identify tuberculosis cases (case finding); and
- Failure to provide satisfactory (vaccine) protection.

Minor obstacles include:

- Drug resistance;
- Lack of prophylaxis; and
- Nosocomial spread(63).

Chapter 3: Social Network Analysis (SNA)

3.1 Basic Concepts in SNA

Social network analysis (SNA) is a set of techniques and measures used to understand the relationships among people and to evaluate the consequences of their interactions. The history of SNA dates back more than 100 years(71). Since its creation, SNA has been used for a wide range of purposes. In 1940, Burnet pioneered a similar process when he attempted to explain the spread of infection. Recently, researchers have used SNA in the study and control of infectious and sexually transmitted inflections(3, 5, 69-72, 107, 117, 122).

Social networks are composed of entities (called "nodes", "actors", or "vertices" in SNA terminology) connected to each other by relations, represented by a line. The nodes and lines together form a "network", and networks create a "system." Each system has structure made of nodes and their relations. The relational data explaining any network belongs to the system and not to the entities making up the system. In contrast, attribute data (age, gender, etc.) belong to the individuals they describe(107, 122). SNA is the analytical method used for relational data, while attribute data are appropriately analyzed by variable analysis(107, 122).

Networks studying can be carried out through either "socio-metric" or "ego-centric" approaches. In the socio-metric approach, which is also known as the whole network approach, all individuals from the targeted group or community will be included in the analysis. This approach is expensive and time-consuming, particularly if the target

group is large in number. Alternatively, the egocentric approach focuses only on the personal network of selected individuals from the group or community under investigation. The personal network, or "egocentric network", is composed of the index person or index node and the people or contact nodes with whom the index has a relationship. In SNA terminology, these people are known as alters.

SNA has the capacity to identify the most important nodes in the networks and can study the structure of networks(107, 122). To identify key nodes in the network, SNA uses special mathematical measures called "centrality measures" to rank nodes in accordance with their importance or prominence in the social network(1, 107, 122). One centrality measure is the "degree" of the node, also known as "local centrality" and defined as the number of nodes adjacent to the node. Other measures are used for determining the global centrality of a network, such as "closeness", which calculates the closeness of any network member to the remaining network members. "Betweenness" is a unique measure, because it can identify the network member with the power of being gatekeeper or network power broker. Betweenness can demonstrate to what extent any member lies between other members on the network paths(107, 122).

Social network analysis can describe the structure of a network, including its components, cliques, and whether there is a core group within the network where there is high prevalence of active infection(41, 71). Each network has a density, which depends on the extent or number of connections each node has. The greater the number of nodes connected to one another, the greater the network density. The maximum value of network density cannot exceed one; at this point, each node is connected to every other node in the network, i.e. density is the ratio between the present ties in the network and the

hypothetical maximum number of ties. The spread of infection in a network depends on the connectivity of the network or its density: the more connected the network, the faster a communicable disease can spread in the network(1, 3, 107, 122).

Social network analysis can investigate both microstructures and macrostructures or sub-structures of networks. The procedures start either from the bottom or the smallest parts of the network (nodes) up all the way to the largest parts of the network (components), in a "bottom-up" approach, or starting from the top of the network all the way down to the bottom, in a "top-down" approach. Social network analysis can identify any structural patterning of the network nodes connecting with each other. Substructures of the network (e.g. "triads") connect nodes with each other and provide shape to the network. Networks substructures have a wide spectrum of formations, from cliques to more complex subgroups like components. To understand network behaviors it is necessary to dissect and examine network structures(107, 122, 137, 138, 139)

3.2 Application of SNA in Population Health and Tuberculosis Research

SNA has been used in public health for different purposes: as a social field to assess the effect of social support on disease status(53, 54, 117); to monitor HIV transmission(69, 117); to investigate sexually transmitted infections (64, 65, 70, 130) and intravenous drug use (72); in family planning(118, 122); and as a tool to assess the performance and dissemination of information within a community(17, 122).

Ultimately, the goal of SNA in infectious disease control is two-folds: first, to understand how a social network can facilitate or impede propagation of disease; second, to develop models of the structure and dynamics of networks that permit prediction of

incidence. Relationships and social connections between people can channel not only information but infection as well(107, 122).

Using SNA in combination with molecular fingerprinting provides a unique opportunity to study and understand the dynamics of TB spread(41, 71). As we know, TB spread is closely related to socioeconomic factors(54, 60, 84) and, as such, its control involves the study and identification of those factors. SNA has the capacity to deal with such complex interacting variables.

3.3 Disadvantages of SNA

Disadvantages of using SNA range from its being a new method to several other obstacles, all of which are minor in nature compared to the significant benefits SNA can offer in return. Commonly noted disadvantages are listed below.

1. SNA is a new method in health care research. Any new method usually faces barriers and challenges in its application.

2. Sample size concerns. Extreme sample sizes like very large or small outbreaks usually create challenges for researchers, ranging from data collection difficulties to analysis problems.

3. Skills requirement. SNA requires highly skilled personnel who are particularly computer literate.

4. Generalization. SNA only deals with data describing the relations in the communities under investigation; caution must be used in applying results to other communities(107, 122).

A SNA researcher might face other issues in a small, remote community, including the physical, geographical accessibility of the community for the researcher and subjects alike.

Network boundaries and other SNA challenges:

Choosing network boundaries is very important for SNA research, and boundary selection should be based on clear rationale, as it is connected to the research outcome on more than simple or natural options like geographical boundaries. Unless the boundaries of the network are picked correctly, study objectives may not be met. For example, TB in small communities of the north is part of the general TB problem in Manitoba. To investigate TB in a particular community in the north or to confine the project to a particular geographical boundary may limit the researcher's ability to investigate the whole network (the entire picture of TB in Manitoba in this case) and not just part of it.

Solutions for network boundaries and other SNA challenges:

Selection of network boundaries should be based on clear rationale and should be tied to the research objectives. Such was the case in our case control study in which the study boundaries were tied with the study objective. The main objective of the case control study was to compare the tuberculosis network in Community 1 with the social network of the general population of Community 1. Therefore, the Community 1 geopolitical boundaries were used as the case control study network boundaries. Residence in Community 1 was the main inclusion criteria and living outside the community was one of the exclusion criteria.

3.4 Considerations for Research in Remote Aboriginal Communities

All research faces challenges, and conducting research in a remote Aboriginal community poses specific challenges. Aboriginal people may view people of European origin negatively, due to the history of Europe as a colonizing power in this region, and the perceived negative consequences for Aboriginal people. Therefore, research conducted within Aboriginal communities by non-residents is often regarded with suspicion and lack of trust by community members(45, 61, 135).

In addition, past Aboriginal experiences with researchers have not all been positive. These experiences include past research which started with a simple, straightforward idea about Aboriginal culture, and subsequently evolved to the investigation of, for example, Aboriginal body characteristics, reaching the conclusion that Aboriginal people are biologically inferior to "Canadian" people(18, 61). Frequently, Aboriginal communities have not received any briefing regarding research findings, and so have not had any opportunity to correct the misconceptions about them, their culture, and their communities. For Aboriginal people, these experiences are perceived as an extension of colonialization(18, 61).

The following is a list of problems facing any research in remote Aboriginal communities, and proposed solutions to these problems, in the context of the current research.

1. Social reality and community benefits:

Due to the past history of Aboriginal people with Europeans, there is mistrust of researchers. The community in general and the tribal government in particular inquires about the benefits to the community from the project, and for the community members, it

would be an unacceptable project if the community did not receive any benefits. Researchers must understand the political structure of the community and work in partnership with the local leadership(18).

Solutions for social reality and community benefits problems:

To maximize the relevance and feasibility of SNA research, researchers need to approach the First Nation leadership and request their permission and cooperation to conduct the research. If the leaders approve the project and express willingness to participate, researchers need to create an advisory board for the project from the community to help run the project, since the goal is to build partnerships with the communities involved(18). The project also should be conducted with the help of the TB Registry Office and the local health care providers. In general, if TB is a major problem in a community, its residents may realize there are benefits to their participation in a research effort and become consequently more understanding and helpful.

2. Interpreters and staff recruitment:

Scientific research requires highly qualified individuals and most of the remote Aboriginal communities do not possess such skills. However, local residents are essential for the study to proceed, especially if the researcher does not speak the local language. Caution must be exercised in selecting appropriately qualified local interpreters(18).

Solutions for interpreter and staff recruitment problems:

Researchers need to recruit interpreters and research assistants from the selected communities or neighboring areas. Study staff must spend some time to train the hired individual or individuals for the research in general and to explain the aim of the study and its benefit to the community. Interpreters must sign confidentiality agreements(18).

3. Subject recruitment:

In Aboriginal communities, it may be difficult to recruit subjects due the fact the researcher is from outside of the community; mistrust may affect the research progress and how the research affects the subjects of research. Further, researchers can be seen as extensions of the colonial past. In general, it is easier to recruit participants from urban settings because they offer a larger sampling pool(18).

Solutions for subject recruitment problems:

Meeting with the tribal and health authority in the selected communities plus creating a community-based advisory board will help advance the importance of the project and its goal. The significance of TB in selected communities, combined with the importance of the project, encourages potential subjects to participate in the research as a priority(18).

4. Confidentiality:

Small and remote communities can pose a real problems for researchers from the privacy point of view, whereas large urban communities do not have this challenge. In an urban community, people who work together are not necessarily friends or do not necessarily even know each other or the other subjects involved. In a small and remote Aboriginal community, almost everybody knows everybody else. Furthermore, problems arise from the fact that a large number of residents in small, remote communities do not speak or fully understand English and therefore require an interpreter. Subjects participating in the project might not trust the researcher or the team working with them. The problem becomes greater if the local interpreter breaks the confidentiality oath. Leaking information about participants in a project can create a very serious problem and therefore must be prevented(18).

Solutions for Confidentiality problems:

Privacy and confidentiality are important issues in any research. The best possible person must be recruited for the job of interpreter, and training the interpreter for research is vital. Emphasizing the confidentiality of research by demonstrating fictional roles or struggles between the interpreter's familial and professional duty, and guiding him or her accordingly, is a necessary step of training. The integrity of the interpreter's adherence to the oath of confidentiality is an essential component of the project(18).

5. Community control of research and publication:

Wanting to control misconceptions about their community in any publications that result from research, some communities demand that researchers obtain permission from the band authority (local leadership) to conduct research. Unfortunately, in addition, certain communities may refuse to grant researchers permission to publish their work unless they modify their findings to bring them in line with the community's desires. Researchers naturally protest such a request, express their dissatisfaction, and even consider this attitude as censorship curtailing scientific freedom(18).

Solutions for community control of research and publication problems:

Although the communities of interest in the project do not always initially express their concern about publication of data, this must be considered. Any valid concerns from the community or their leaders must be addressed. To address potential concerns about a bad or negative impression or stigmatization of the people or the communities due to the research results, the identity of the communities and participating individuals will be coded in the research data to keep the identities concealed(18).

6. Time requirements:

Recruiting communities for research, obtaining permission from the ethics or review board, and recruiting interpreters and subjects are time consuming, particularly when working in a remote and isolated community. Each of these tasks means extra time to complete the project(18).

Solutions for times requirements problems:

Some delays in finishing the project are not unexpected. Researchers must be prepared to accommodate any unforeseen delays(18).

7. Legal encounters:

Information collected through research is of a sensitive nature, and legal authorities may, in theory, subpoena the raw data. Researchers usually promise their subjects confidentiality and a breach of this pledge has serious consequences. First, it affects future recruitment. Second, subjects stop believing the researcher's promises and may become hostile toward the research in general(18).

Solutions for legal encounters:

Generally, to protect the research subjects from future subpoena, data collection should be made anonymous, not just confidential. It must be emphasized that participants have the right not to answer any questions they do not wish to answer(18).

SECTION III: METHODS

Chapter 4: Overview of Methods Used

The present study is composed of three parts. The first part involves a description of the epidemiology of TB in Manitoba between 1992 and 1999. The second part is a pilot study of the feasibility of using Social Network Analysis in tuberculosis epidemiology in Manitoba, by applying this methodology to eight high-incidence communities. The third and major part of the study is a case-control study in which SNA was used in conjunction with primary interview data and TB registry data from one northern Manitoba Aboriginal community with very high TB incidence.

4.1 Descriptive Epidemiology of Tuberculosis in Manitoba, 1992-1999

In the first part of the study, conventional and molecular epidemiology methods were to describe the problem of TB epidemiology in Manitoba during the study period. Complete data on all Manitoba's TB patients diagnosed from January 1, 1992 to December 31, 1999 were obtained from the Manitoba Central Tuberculosis Registry. These data include each patient's unique identifier number, date of birth, place of birth, gender, country of birth, residential postal code, origin (Canadian-born treaty, Canadian-born non-treaty, and foreign-born), date of diagnosis, case finding method, tuberculosis status (new active and relapse), BCG status, year of arrival for foreign-born patients, smear results, culture results, TB isolate DNA fingerprint type, drug resistance, chest x-ray results, cavity status, and contact tracing investigation results.

All collected diagnostic specimens were sent to the Clinical Microbiology Laboratory at the Health Sciences Centre in Winnipeg, Manitoba for isolation and identification of *M. tuberculosis.* The Accuprobe *M. tuberculosis* complex kit (Gen-Probe Incorporated, San Diego, CA, USA) was used according to the manufacturer's instructions to identify *M. tuberculosis* in the clinical samples. Samples were subsequently processed for DNA sub-typing according to the international standardized methodology of IS6110 restriction fragment length polymorphism (RFLP)(120). There was no evidence of culture cross contamination at the Health Science Centre, as the staff strictly adheres to the procedures outlined by Small *et al* to minimize that possibility(111).

Standard descriptive epidemiological methods were used to describe the incidence, trends and characteristics of TB cases in Manitoba over the study period. Discriminant analysis (using NCSS 2000 statistical software, NCSS, Kaysville, UT) was performed by dividing the cases into three categories (foreign-born, Canadian-born treaty, and Canadian-born non-treaty) and then determining whether there were significant differences between these categories using the following variables: age, gender, BCG status, TB molecular subtype, TB drug resistance, site of infection, and whether the infection was new or relapsed.

4.2 Social Network Analysis Methods

The approaches used in the SNA can be either whole network (sociometric) or personal network (egocentric) approaches. During the pilot project phase, data used for SNA were extracted from existing data of contact tracing investigations. The approach used in this stage was the sociometric network approach. For the case-control study in

Community 1, both sociometric and egocentric approaches were used. The data used for analysis in the case-control study were extracted from contact tracing records and primary interview data from Community 1 residents. Thus, the data used in the case-control study contained more information about the places index nodes frequently visited. However, for the case-control study we were not permitted by the university ethics board to contact the alters of index nodes due to privacy concerns.

For both the pilot study and the case-control study, UCINET (138) and PAJEK (137) social network analysis software programs were used to conduct the analyses.

4.2.1 SNA Approach Used in Pilot Project

Selected communities in Manitoba with high TB incidence were identified for the pilot project. The selected communities were Northern Manitoba Aboriginal communities (reserves) with the highest TB incidence in Northern Manitoba. One additional non-reserve community (Community 7) which had close contact relations with one community under investigation (Community 6) was added to the investigation, raising the total number of communities under investigation in the pilot project to nine. The study used data from contact tracing investigations of TB cases reported between 1992-1996 in the selected communities. SNA was used to plot and visualize the TB transmission networks in these communities, one community at a time, then in pairs of socially connected communities. Finally, the transmission network of one single tuberculosis strain which infected members of several communities was plotted. The selected strain was labeled as fingerprint type 5 (FP5). This TB strain (FP5) was the second most prevalent strain in Manitoba over the study period.

One particular community (Community 1) was then investigated further. Centrality measures were used to identify key nodes in the network. Network substructure was investigated and cliques in the networks were identified. *A clique is the maximally complete subgraph of at least three nodes(107, 122, 139).* Theoretically, clique members have close relations with each other. This close relationship makes cliques viable environments for TB transmission. Core structures of the network were extracted by the core collapse sequence technique using Seidman k-core (107, 122). In the core collapse sequence, the network is collapsed gradually by reducing the graph one degree in each step, until the network disintegrates. Members of each core collapse step or layer were then identified and analyzed further (107).

4.2.2 SNA Approach Used in Case Control Study of Community 1

In the case control study, both sociometric and egocentric approaches were used. First, contact tracing investigation data was used to plot the Community 1 TB outbreak. The plotted network was then used to identify the actors (nodes) to invite for participation in the case control study. Contact tracing investigation data were also used to plot the network of the Fingerprint 1 (FP1) TB strain in Manitoba, as this was the dominant TB strain in the Community 1 outbreak, as well as being the most prevalent strain in Manitoba. By plotting the network for FP1 for the whole province, the boundaries of the Community 1 TB outbreak were able to be identified. Thus, we were able to analyze the Community 1 TB outbreaks both locally and globally.

Chapter 5: Case Control Study Methods

5.1 Case Control Study Objectives

This case-control study research questions arose from the results of the pilot project. The specific objectives of the case-control study were to:

1. Explore the feasibility and utility of using primary (interview) data in conjunction with contact tracing data in applying SNA to TB control.
2. Explore whether or not a TB transmission network is a reflection of the affected community's social structure, by comparing the characteristics of the TB networks with the characteristics of:
    a. The social networks of the contacts of the TB cases; and
    b. The social networks of age- and sex-matched non-TB controls from the same community.

5.2 Community Recruitment and Participation

Initially, three high-incidence potential study communities were selected, and local authorities were contacted and invited to participate in the study. Field trips to the selected potential study communities were conducted. The goals at this point were to introduce the researchers to the involved communities, to gauge community interest in participation, and to establish rapport with the political leadership and local health care providers for a more productive relationship.

Community 1 was ultimately chosen as the study community, due to community members' strong interest in participating, as well as the fact that the community had the highest incidence of TB over the study period. An advisory committee of local individuals from Community 1 was formed to advise the research team about the project. As described above, a control group from the community, matched to TB cases by age and gender, was invited to participate.

5.3 Community 1

Community1 is about 1000 kilometers north of Winnipeg. According to regional population statistics from 1996, this First Nation has an on-reserve population of 1217, off-reserve population of 651, and 47 live on crown land. The total population is 1915. This community is not accessible by all-weather roads. There is a winter road usually constructed after freeze-up in late November that lasts until April. The road crosses lakes and is mainly used to bring freight. The community has an airport facility and is serviced by scheduled flights. The community has a nursing station. Hospitals in the area are located in neighboring communities(62).

5.4 Case Control Study Design

As stated above, Community 1 had the highest TB incidence of any community in Manitoba between the years 1992 and 1999. After securing the participation of the community and obtaining permission from relevant authorities (the Health Research Ethics Board at University of Manitoba, and the First Nation and Inuit Health Branch of Health

Canada), the TB network of the participating community was plotted using case and contact tracing investigation data available from the TB registry. Next, the network was pruned (collapsed) to degree 2; i.e. each node in the network had contact with at least two other nodes in the same network. The researchers, through the Director of the Manitoba TB Control Program, invited all the members of the collapsed network to participate in the study. Local leadership (through the local band office) was asked to help recruit a control group of an equal size as the TB case group, matched with the TB cases by age and gender. Age range required for the study was prepared and sent to the local leadership. The local leaderships were able to recruit pool of suitable candidates.

5.5 Inclusion and Exclusion Criteria

The inclusion and exclusion criteria were based on the design of the study as explained below.

5.5.1 Inclusion criteria:

Participants were included in the study according to the following criteria:

1. All pulmonary TB cases diagnosed in Community 1 residents between 1992 and 1999 (pulmonary cases were chosen specifically because of greater likelihood that these cases would have contact tracing records);
2. All included cases must have contact tracing investigation records available for social network analysis;
3. All reported contacts of TB cases in Community 1 during the study period.

4. All participating contacts must have known two or more members of the network of which they are a member (nodes with degree two);
5. Eligible controls included community residents without TB, of the same age and gender as the TB cases (n=29);
6. All participants must have been residing in Community 1 during the study period; and
7. All participants had to give informed consent to participate in the study.

5.5.2 Exclusion criteria:

Subjects were excluded in the study according to the following criteria:

1. Residence outside of Community 1 during the study period; and
2. Potential control subjects with any documented history of contact with the study TB cases (named by TB cases in contact tracing investigation).

5.6 The Interview

The case control study of Community 1 originally targeted 134 prospective individuals (29 TB cases, 76 contacts, and 29 controls), but field recruitment yielded 83 participants. The study subjects were interviewed to collect information on their personal (group) and network characteristics. The interview was 25 to 45 minutes long, and started first by asking participants for identifying data like name, sex, and date of birth. The second part contained questions regarding ethnic background, residence and length of time spent in the residence, living arrangements, condition of the house they live in, travel and mobility from Community 1 to other communities, and employment status. The questionnaire's

third part explored the medical history of the subjects, asking about tuberculosis, sexually transmitted infection and co-morbid conditions. The fourth and final part examined the network of the participants. The aim of this part of the questionnaire was to explore the types of places and the people attended by or interacting with the subjects of the study. The network includes contacts of people (alters) and places of social aggregation attended by the index nodes. From the responses of the participants we built the study variables.

5.6.1 Group characteristics

The personal characteristics studied were gender, age, education, residence, travel, living arrangements, work status, housing conditions, and co-morbid conditions (Tables 30, 31).

Residence and traveling

The subjects were asked about their places of residence in the past 5 years, with answers categorized into: present community, other reserve, other rural non-reserve community, Winnipeg, or other urban community.

Living arrangements

Participants were asked to identify their living arrangements. Living arrangements among the participants were diverse. To simplify the classification of living arrangements and consolidate the numbers of each stratum for statistical purposes, living arrangements were divided into three separate strata: single parent family (SPF), in which the family was composed of one parent (mother or father) and the children; a second stratum composed of the parents (mother and father) with the children (Family parent and children, FPC); and a third stratum composed of the remaining, or everybody else (an "Others" category). The others category consisted mainly of different forms of extended families.

Work status

Working status was categorized as working, not working, and student.

Housing condition and density

Study participants were asked about their houses, other occupants, and types of services available for the residents. The questions asked about the number of house occupants, the number of rooms in the house, the presence or absence of heat, hot water, tap water, bathroom, and toilet, whether the toilet is equipped with flush or not, and if the toilet is inside or outside the house. Dividing the number of occupants by the number of rooms in the house gives the density of the house.

Co-morbid illness

Participants were asked about their health and if they suffer from any other illness. To further quantify and compare the extent of co-morbid illness between the groups, a co-morbidity index was created in the analysis by assigning different weights to different diseases reflecting their potential importance in contribution to susceptibility to tuberculosis. Diabetes (DM), cancer and other illnesses affecting immunity in general were given 10 points. Asthma and other chronic respiratory conditions were given 5 points. All other chronic conditions were given 1 point.

5.6.2 Network characteristics

To understand the dynamics and structure of the networks, characteristics of these networks must be collected and investigated. The network characteristics studied were network size (number of people and places), contacts (number of people only) size, age of alters, closeness of the index nodes to the alters, frequency of index nodes' interaction with

alters, and place diversity (proportion of places visited by index nodes which are public places).

Network Size

Network size is the number of contacts (people and places) the index cases interact with or visit.

Contact Size

Contact size is the number of alters or contacts (people only).

Age of alters

Alters are the contacts of the index node or the members of egocentric network of the participants.

Closeness

Closeness is used to measure how close index nodes or the participants in the study are to their alters. Closeness of the nodes was measured by asking each index node about their relations with their contacts. The relationship was coded into one of three categories: strong relationship, not strong relationship (medium), or weak relationship. These relationships were defined by index case statements only and were not validated or confirmed by the alters. Scores were assigned according to strength of relationship. Strong relationships were given 3 point. Not strong (medium strength) relationships were given 2 points, and weak relationships were given 1 point.

Frequency of contacts

Frequency of contacts is the number of times index individuals interact with their alters per week. The frequency of visits ranges from zero to seven times per week.

Place Diversity

Public places are important in disease transmission, particularly airborne ones like TB. Therefore, identifying public places attended by TB index cases is vital for any control effort. In the Community 1 case control study, all index nodes were asked about places they regularly attended. The place diversity variable in the study is the percent of places visited regularly by index nodes which are public places. This variable was called place diversity to reflect node behavior. Place diversity was extracted for each node of the study by calculating the number of public places frequently visited by the index node and dividing this number by the total number of places visited by the same node (Table 31).

Network Type (Risky & Non-Risky)

The participants were asked to name their contacts and what kind of relations or activities they engaged in. The answer choices offered were casual contact, sexual contact, and alcohol or drug-using partner. In addition, participants were asked whether they knew if the contact had TB or not. If the study subjects admitting to have sex, alcohol, or drugs with one of their alters, or if they know that one of their alters had TB, then their network was classified as a risky network (or a network with high risk).

Behavior Types (Risky & Non-Risky)

The risky behaviors variable is similar to risky network variable; with the exception that history of TB in alters was not a criterion for the risky behavior variable. Any index node interacting with their alters through drinking alcohol, taking drugs, or sexual contacts were classified as having risky behaviors, and having non-risky behaviors if they were not.

5.7 Network Approaches Used in Case Control Study

The Community 1 TB outbreak and the FP1 TB outbreak in Manitoba were plotted as separate networks in a one-mode network. In one or single mode network each node in the people network under investigation can be related to any other node. Different SNA techniques were used to identify key nodes in the networks and explore the network characteristics.

5.7.1 Whole Network "Socio-Metric" Approach

The whole network was constructed from the TB nodes, contact nodes, control nodes, and alters of the three groups. This network was composed of 391 nodes. Next, a second Community 1 network was built consisting of people (index nodes, alters) and places frequently visited by index nodes. Our original study plan was to collect data from all members of egocentric networks for comparison with sociometric data from the contact tracing investigations, but due to restrictions imposed by the ethics board on the study (i.e. we were not permitted to contact alters), we were not able to collect enough data to build complete egocentric networks and therefore we were not able to do these comparisons.

The groups (TB, contact, and control) were divided into three separate networks. Each network was plotted separately, first as a network composed of people only (index nodes and alters), then as a second network composed of people and places (index nodes, alters, and places). Basic analysis was conducted on one mode network. In this analysis network density and centralization were calculated and centrality measures were used to identify key nodes in the network. Four centrality measures were used to identify the key individuals in the Community 1 network and the TB, contact and control networks. Centrality measures were used again to identify which of the public places were most

significant in the networks. Cliques and k-cores from the "bottom-up" approach and components from the "top-down" approach were used.

A core collapse sequence procedure was applied to the FP1 network in Manitoba to identify the significance of Community 1 in the spread or maintenance of the FP1 TB strain in Manitoba, as well as to demarcate the boundaries of the Community 1 TB outbreak in relation to other communities in the province.

Two-mode network analysis was used to study the impact of interaction between people and places regularly visited by index nodes on tuberculosis spread. Affiliation network is two mode network in which two different set of nodes are used (e.g. people, places) and the nodes are usually indirectly connected.

Affiliation networks of the study were composed of case control study people as one set of nodes, and places attended by the index nodes as second set of nodes. The people of the network were connected with each other through the places in the network or second set of nodes. Node1 (TB patient) for example was connected to node 2 (healthy person) through place1 indirectly. Through place 1, node 1 have unique opportunity to pass his TB disease to node 2 or other visitors to place 1.

The main reason for studying affiliation networks was to identify key places in the network. Places (particularly public places) create connections between nodes. Through visiting places of interest, TB patients have opportunities to connect with other community residents. TB may spread unnoticed in the community through these connections.

Finally, the cumulative degree distribution was investigated to determine whether or not the network is a scale-free network.

5.7.2 Personal Network "Egocentric" Approach

This approach focuses on the individual networks of index cases, so as to understand personal networks and the variables which influence the formation of the network. In this study, there were 83 personal networks in total. The variables under investigation can either describe the index cases and be called personal (group) characteristics, or be related to the individual network and be called network characteristics.

All 83 of the participants went through the same interview and answered the same questions; essentially, the questionnaire aimed to probe both individual and network characteristics. The egocentric approach focuses on two elements, which together give the answers to the questions raised earlier: first, whether TB group networks are similar to control group networks or not; and secondly, what are the differences between contact group and the TB and control groups. Whereas the control group was expected to reflect the social structure of Community 1, the contact group represented the pool from which most potential new TB cases develop.

Chapter 6: Statistical Approaches

6.1 Univariate Statistical Analyses

Descriptive and frequency tables of categorical and continuous variables of the case control study were created (Tables 30, 31) to describe the study population and groups.

Variables in the case control study were divided into group characteristics and network characteristics. Group characteristic variables included: gender, age, education, residence, travel, living arrangements, work status, housing conditions, and co-morbid conditions. The network characteristics of the participants included: network size (number of alters and places), contacts size (number of alters only), age of contacts (alters), closeness of the index nodes to the alters, frequency of index nodes' interactions with alters, and place diversity.

The variables are about describing individuals that are obtained from the survey or from network properties, and either of those can be examined at the level of the individual or can be aggregated.

The Chi-square test was used for categorical variable analysis (Table 17). These variables included: living arrangements, housing density, work status, behavior type, network type, and number of TB cases known. The living arrangement variable was coded into three categories: single parent family (SPF), family with parents and children only (FPC), and other, which included all other arrangements, the most frequent of which were various forms of extended family. The housing density variable was coded into two

categories: high-density houses were defined as those with more than the community mean of 1.56 people per room, and low-density houses were defined as those with fewer than 1.56 people per room. Two categories were defined to classify participants according to number of TB patients known: one category for those who knew no TB patients, and a second category for those knowing one or more TB patient. The network type variable was coded into two categories: risky and non-risky. If the study subjects admitting to have sex, alcohol, or drugs with one of their alters, or if they knew that one of their alters had TB then their network labeled as a risky network. The behavior type variable was also coded into risky and non-risky categories, based on the participant behaviors with their contacts. If the participant engaged in sex or drug use or alcohol drinking with their alters they were labeled as a subject with risky behaviors. Finally, the working status variable was coded into three categories: working, not working, and student.

The student's t-test was used to compare the means of continuous variables (both group characteristics and network characteristics) of two groups at a time: the TB group with the control group (Table 14), the contact group with the control group (Table 16), and contact group with the TB group (Table 15). Analysis of variance (ANOVA) was used to compare the means of the continuous variables of the three study groups simultaneously (Tables 12, 13). A correlation matrix (Table18) was created and used to identify related variables of the network characteristics before putting them through analysis of covariance (ANCOVA) (Tables 19, 20, and 21) and multivariate analysis. In the ANCOVA procedure, while controlling for personal and group characteristics, those network characteristics which were significant by the student's t-test, were entered into the ANCOVA one at a time.

6.2 Multivariate Statistical Analyses

Discriminant analysis was used twice in the study. In the first part of the study (description of TB in Manitoba 1992-1999), all TB cases were divided into three categories (foreign-born, Canadian-born treaty, and Canadian-born non-treaty), and then discriminant analysis (using NCSS 2000 statistical software, NCSS, Kaysville, UT) was used to determine whether there were significant differences between these categories using the following variables: age, gender, BCG status, TB molecular subtype, TB drug resistance, site of infection, and whether the infection was new or relapsed (Figure 2).

Secondly, discriminant analysis was used in the case-control study to compare the differences among study groups in the Community 1 network (TB, contact, and control). The outcome variable in this analysis was "type of network" (i.e. TB, contact, or control). The independent variables used were those found to be significant in univariate analyses, including: age, home density, number of TB patients known by index nodes, place diversity, co-morbid illness, and age of alters. The centroids of the groups were plotted on a graph (Figure 50). The contribution power of each variable on the axes is summarized in Table 32.

Logistic regression was used in analysis of the case-control study data. The objectives of using logistic regression analysis were:

1. To simultaneously compare multiple characteristics/factors associated with TB status (being a TB case, a TB contact, or a non affected control, respectively);
2. To determine which social network centrality measures best predict TB status (TB case versus TB contact versus control).

To accomplish the first objective, the outcome variable was defined as type of network, i.e. TB, contact or control network. The independent variables used in the analysis were age, highest schooling years, home density, work status, number of TB patients known by index nodes, place diversity, co-morbid illness, contact size, network size, frequency of contact, type of network (risky versus non risky), and age of alters. Variable selection was by backwards stepwise selection, whereby all variables were entered initially, and then removed from the model one variable at a time if not significant ($p<.05$). The final results included three final models, each representing a comparison between two groups (TB/control, TB/contact, and contact/control).

For the second logistic regression analysis, the outcome variable was TB status (TB case or not), to determine which of the centrality measures was most strongly associated with TB status. Independent variables were all calculated network centrality measures (degree, Betweenness, Bonacheis, and closeness) of the Community 1 network. Variable selection was again by the backward stepwise procedure.

6.3 Degree Distribution of Community 1 Network Estimation

To study and investigate Community 1 network degree distribution, the first step taken was to plot degree distribution (or frequency of degree in the network) and cumulative distribution on normal axes. Next, the curve was fitted on log-log axes. The curve visually appeared to have a long tail to the right on normal axes, and was almost a straight line on log-log axes, indicating that it followed a power law distribution. To confirm the power law distribution, further investigation of the curve was indicated.

Regression was used to estimate the curve and determine if it fit a linear, power or exponential distribution, with significance testing using the R^2 goodness of fit. Again, plotting on log-log axes was used to reconfirm the results. As the curve seemed to have two components with potentially different distributions, the curve was split into two parts (body and tail) at degree 8 and curve estimation for both parts was redone using regression. Finally, both parts (body and tail) were re-fitted to log-log curves to reconfirm the results.

SECTION IV: RESULTS

Chapter 7: Tuberculosis in Manitoba

7.1 Descriptive Epidemiology of Tuberculosis in Manitoba 1992-1999

Between January 1992 and December 1999, there were 855 reported TB cases in Manitoba (Table1). Females accounted for 459 (53.7%) cases and males accounted for 396 (46.3%) cases. Canadian born TB patients represented 604 cases (70.6%) while foreign-born (FB) people represented 251 (29.4%); on a year-by-year basis, the foreign-born TB case ranged from 19% to 43% of all TB cases in the time span between 1992 and 1999. Most of the foreign-born TB cases (96.8%) resided in Winnipeg, the major urban area in Manitoba, having immigrated to the city from 40 different countries in four continents: Asia, Africa, Europe and South America (Table 2). The majority of the foreign-born TB cases (196 cases, or 78.1%) came from Asia. Specifically, the Western Pacific Region (according to the WHO regional classification) was the most important source of foreign-born TB patients to Manitoba, with 171 (68.1%) cases. European immigrants represented 25 patients: only 10% of all foreign-born TB patients. On an individual country basis, immigrants from the Philippines were the largest contributors with 44.6% of the FB TB patient's pool, or 112 cases; Vietnamese immigrants accounted for 24 (9.6%) patients; India 18 (7.2%) patients; and China 13 (5.2%) patients. A breakdown of the remaining countries of origin is given in Table 2.

The bulk of these TB cases (88.4%) were clinical cases detected by health care facilities; only 3.6% and 7.2% of TB cases were diagnosed through immigration or contact tracing, respectively. We calculated the time lapse between the arrival of foreign-born individuals into Canada and the diagnosis of those individuals with TB. Most of the foreign-born TB cases were diagnosed within the first five years of arrival into Canada (43.8%), with a further 23.5% diagnosed 6 – 10 years after arrival. The proportion declined to 12% for the period 11 – 15 years after arrival, and to 4% for the period 16 –20 years after arrival. TB cases diagnosed 20 years or more after arrival represented 12.3% of FBTB.

Among the foreign-born TB cases, there were 156 (62.2%) pulmonary cases and 95 (37.8%) extra pulmonary cases. In contrast, Canadian-born TB patients consisted of 494 (81.8%) pulmonary cases and 110 (18.2%) extra-pulmonary cases ($p<0.001$). Drug resistance among FB TB cases was found among 30 (11.9%) out of 251 foreign-born TB cases. Single drug resistance to one of the five first line antituberculous drugs (Isoniazid, Ethambutol, Rifampin, Pyrazinamide, and Streptomycin) was found among 26 isolates. Four isolates were resistant to INH; three isolates were resistant to Streptomycin, and one to Ethambutol. One isolate was resistant to Isoniazid, Ethambutol, Rifampin, Pyrazinamide, and Streptomycin. In comparison, the Canadian-born population under study produced only 11 cases (1.8%) with drug resistance ($p< 0.001$). Eight of these isolates were mono resistant, one isolate was resistant to two of the first line drugs, and two isolates were resistant to three or more.

Canadian-born TB cases with treaty status formed 44.4% (380) of the total 855 cases, while Canadian-born individuals without treaty status formed 26.2% (224). The

treaty cases were concentrated on First Nation reserves (243 out of 380 cases, 64%). One-third of all Canadian-born with treaty status TB cases lived in Winnipeg (122 of 380). Most of the treaty cases living on reserves were living on 8 selected reserves (n=153, 63% of reserve cases). The total population of those reserves is 16,989 people, representing 17% of all Manitobans with treaty status and 1.5% of Manitoba's total population.

7.2 Molecular Typing of Manitoba Tuberculosis Cases 1992-1999

Culture isolates were available for 629 of the 855 Manitoba TB cases diagnosed between 1992 and 1999 (73.6%). Of the cases with isolates, 203 were unique (one TB case per fingerprint, i.e. previously uncharacterized), and 426 (46 fingerprints) were clustered or not unique. The clustered fingerprints infected a wide range in number of patients, ranging from as few as two patients to as many as 162. The predominant fingerprint type identified in Manitoba during this period was fingerprint type 1 (hereafter labeled FP1). FP1 was identified in 162 out 629 (25.8 percent) of all TB patients with isolates. Canadian-born people with treaty status composed 122 out of the 162 TB cases with the FP1 strain (75%), while 34 (21%) were Canadian-born people without treaty status. Winnipeg harbored 92 of 162 TB cases with FP1 (56.8%), while 64 (39.5%) of the cases with strain occurred on reserves.

The second most frequently encountered fingerprint identified was fingerprint type 2. In contrast to FP1, this fingerprint was identified in only 32 TB patients (5.1% of total). In all, 44% of all Manitoba TB cases were attributable to one of five fingerprint patterns.

7.3 Discriminant Analysis of TB Case Characteristics

The results of the discriminant analysis investigating differences between three subgroups of TB cases based on place of birth and treaty status (Canadian-born without treaty status, Canadian-born with treaty status, and foreign-born) are presented in Figure 2. The figure is composed of two axes: axis 1 is composed from the contribution of 5 variables: gender; whether the fingerprint type was clustered or unique; whether the case was new or relapsed; whether the site of infection was pulmonary or extra pulmonary; and whether the case was drug resistant or not. Axis 1 confirmed that the groups were significantly different on these variables; for example, Canadian-born TB cases with treaty status were more likely to be female and have clustered strain fingerprints than Canadian-born cases without treaty status or foreign-born cases. Canadian-born cases with treaty status were also more likely to be new (versus relapse) and pulmonary cases than Canadian-born people without treaty status or foreign-born people. Axis 2 is formed from 4 variables (age, drug resistance, BCG history, and site of infection. Significant differences between the groups were also confirmed on this axis; for example, Canadian-born cases with treaty status and foreign-born cases were younger than Canadian-born cases without treaty status. In summary, this discriminant analysis indicates that TB cases in Manitoba over the study period can be categorized into three subgroups with significantly different clinical, demographic and molecular characteristics: Canadian-born cases with treaty status, Canadian-born cases without treaty status, and foreign-born cases.

7.4 Summary of Descriptive Epidemiology of TB in Manitoba

Tuberculosis incidence in both Manitoba and Canada as a whole has declined markedly over the last century, but has recently plateaued. Looking more critically at the problem in Manitoba shows that TB is highly concentrated in certain sub-populations. TB incidence among Canadian-born people with treaty status (First Nation people) is much higher than in any other population subgroup. Tuberculosis incidence in selected First Nation communities surpasses the TB incidence in some developing countries. Most TB cases in reserves are concentrated in a handful of communities in Northern Manitoba. Discriminant function analysis confirms that Canadian-born treaty status cases, Canadian non-treaty status cases, and foreign-born cases have statistically different clinical, demographic and molecular characteristics.

Chapter 8: The Pilot Project

The contact investigation data for Community 1, which had the highest TB incidence in Manitoba over the study period, were further analysed in the pilot project. The objective of this analysis was to evaluate the relationship between exposure of contacts to TB patients and their subsequent risk of acquiring disease.

8.1 Relationship Between Level of Exposure and TB Status

In Manitoba, during contact tracing investigation exposure levels are divided into "high", "medium", and "low". High exposure means contacts had high daily exposure to tuberculosis patients – for example, living together with a TB patient. Medium exposure indicates that contacts had medium exposure (three times per week) to tuberculosis patients. Low exposure indicates that contacts had at the most a single exposure to TB patients.

TB incidence in Community 1 was 271 per 100,000 person-years between 1992 and 1996 (Figure14). Out of 26 TB cases in the community during this period, only 16 patients with active disease underwent contact tracing investigations (3 of them were repeated). For this study analysis, two extra levels of exposure were added in addition to the three levels of exposure described above. Creating two extra levels of exposure reflected the reality that some contacts in Community 1 were exposed to two or more TB patients, thereby putting them in a different class of exposure. The second added level of exposure was created to acknowledge the fact of missing contact data for 10% of the population under investigation (Table3). Consequently, levels of exposure were reorganized in this study to include the following levels:

Level 5: Contacts exposed to two or more tuberculosis patients (can be different levels of exposure).

Level 4: Contacts with daily exposure to one tuberculosis patient (living together, for instance).

Level 3: Contacts with medium exposure to tuberculosis patient (average 3 times per week).

Level 2: Contacts with low exposure to tuberculosis patient (single exposure).

Level 1: Lack of information (LOI), which was excluded from the analysis.

In total, 16 TB patients went through contact tracing investigation, resulting in 238 total contacts. The contact investigations found that 5.5% had active or inactive TB disease, 18.9% were Mantoux positive, 65.6% were Mantoux negative, and 10% had no follow-up data.

Figures 3, 4, and 5 illustrate the relationship between level of exposure to TB patients and TB incidence and Mantoux status for Community 1. Figure1 shows how tuberculosis incidence increased with the increasing levels of exposure to TB patients. This trend was statistically significant. Of note is the marked difference in incidence between exposure levels 4 and 5; i.e. exposure to more than one TB patient raised one's risk for TB significantly higher than daily exposure to only one TB patient. This observation is consistent with a similar relationship between level of exposure and risk of infection (as indicated by Mantoux positivity) depicted in Figure 4. Finally, Figure 5 illustrates the inverse relationship observed between level of exposure and Mantoux negativity.

8.2 Pilot Application of Social Network Analysis

The first step in the pilot application of SNA was to plot the network of TB cases and their contacts in several communities including Community 1 (Figures 6-9). The profiles of these communities' TB outbreaks and comparative incidence rates are presented in Tables 5 and 6, respectively.

Hundreds of pages of contact investigation results were able to be summarized in one page containing the plot (picture) of each community's TB outbreak network. Creation of clear figures enabling visualization of these TB outbreak networks guided further social network analysis. Examples of subsequent analyses (identification of cliques; core collapse sequence) of the Community 1 network are presented in Tables 4 and 8. The Community 1 TB outbreak network graph (Figure 6) reveals a compact network with an obvious and dense core (dense connections in the network). SNA centrality measures of degree, betweenness, and closeness were used to identify key individuals. Furthermore, the core structure of the network was extracted using core collapse sequences (k-core), and the people forming the core were identified (Figures 15-22).

Network structure was further investigated by identifying the clique composition of the network. UCINET software identified 38 different cliques (Table 4), ranging in size from 3 to 5. Out of that total number of cliques, 30 cliques had at least one node (member) who did not have a history of TB disease, while the remaining members did. Certain nodes were members in multiple cliques, which created overlap, important in network structure and disease spread. Community1 TB cliques were composed of both TB patients and people without TB disease (Mantoux negative or positive). For practical purposes, identifying disease free nodes is essential in TB control, particularly if the nodes are

connected to one or more TB cases in the network. These individuals should receive priority status for prophylaxis treatments.

The next stage in the pilot study was to plot TB networks for several other communities, one at a time, comparing the structural composition of the communities' TB networks with each other (Figures 6-9). We then plotted two communities' TB outbreaks together (Figures 10, 11). The final step was to plot a single strain's (FP5) network in Manitoba, across multiple communities (Figures 12, 13).

8.2.1 Observations from Pilot Application of SNA

Three important observations were made in this part of the study which raise important questions regarding the structure of the networks, the boundaries of the networks, and the missing links in the network structure which could shed more light on TB transmission dynamics.

First, the plots of the TB networks in the communities under investigation display the TB network structures of several First Nation communities (Figures 6 to 9). The structures of these networks varied in compactness. Some communities had compact networks, while others were not compact and may even have been composed of several components. One community (Community 1) had an evident core (Figure6).

The second observation concerned TB outbreak boundaries. Usually, TB outbreaks are defined by the name of the community in which they first are detected. The definition usually follows geopolitical boundaries or jurisdictional responsibility. However, the pilot SNA study confirmed that TB outbreaks usually extend beyond the geographical and political boundaries of communities. The networks of two communities together (Figures

10, 11) and of FP5 (Figures 12, 13) demonstrate that outbreaks frequently extend beyond the boundary of just one community.

The third important observation has to do with the TB network of strain 5 (FP5) (Figures 12, 13). Although all TB patients in this network were infected with the same strain, the structure of the network was fragmented into several components. This observation directs attention to the need to look for missing links in this network. Deficiencies in contact tracing procedures can result in missing information, which in turn can create an unconnected network. Ongoing TB transmission is related to the fact that TB is mainly transmitted through the air. The idea of "risk space" is important as a possible explanation for TB transmission between two strangers, as has been reported in the literature(73, 83, 85, 88, 93, 102, 121). The concept of risk space is based on the fact that active TB cases and non-infected individuals attend confined places for various reasons (e.g. social interactions in bars), and sharing that space has led to TB transmission. Other known or unknown factors in the location can help make such environments fertile ground for TB transmission.

8.2.2 Questions

The preceding observations raised valid and intriguing questions:

1. Does the TB network structure of First Nation communities under investigation reflect the social structure of the communities?
2. What kind of boundary definition should be used for outbreaks?
3. Are contact sites (places) important in TB outbreak networks or for TB transmission?

To answer these questions required further study. The observations and questions raised by the pilot project were instrumental in designing and conducting the case control study in Community 1.

Chapter 9: Case Control Study

9.1 Tuberculosis in Community 1

There were 51 TB cases in Community 1 between 1992 and 1999 (Table 7). Out of these 51 TB cases, 29 cases (56.9%) had been investigated further using contact tracing. The remaining 22 cases (43.1%) did not go through contact tracing investigations. Using the RFLP technique, all available isolates (n=42) from these TB cases were sub-typed. DNA fingerprint (FP) analysis of TB isolates from Community 1 revealed that:

- 42 cases had 6 different FP types:
    - FP1 (37 cases)
    - $FP77_b$ (1 case)
    - FP98 (1 case)
    - FP179 (1 case)
    - FP80a (1 case)
    - FP171 (1 case)

Annual TB incidence in Community 1 varied widely between 1992 and 1999, from 84 to1222 per 100,000 (Figure14, Table 7). Poisson regression was used to look for significant trends in incidence over this time period; no significant trends were detected.

9.1.1 Contact Tracing Data of Community 1 TB Outbreak

The TB network formed in Community 1 was based on the 29 TB cases for whom contact investigations had been conducted, and their contacts (alters), forming a total of 504 individuals (nodes). The Community 1 TB outbreak network had 68 individuals with

history of TB disease. Contact tracing investigation revealed that 109 individuals (21.6%) had Mantoux positive tests. The remaining 327 contacts (64.9%) were Mantoux negative.

Collapsing the network gradually using Seidman *K*-core in accordance with the degree of contact revealed an important finding, shown in Table 8 and Figures 15-22. The whole network disappears when the degree of contact is raised to 9. In other words, nobody in the network has connections with 9 or more other people. Further, key stations in the core collapse sequence occur at degree 4 and 7 (Table 8; Figures18, 21 and 22). In the case of the degree 4 result, the total number of nodes was 22; 15 nodes (68%) had TB disease, and 7 nodes (31%) were Mantoux positive (Table 8, Figure18). This means that any person from the network who had connections to four other people in the network was either a TB patient (68% of the degree 4 members) or infected by the bacteria (32% of the degree 4 members). The degree 7 result indicates that if a person from the network had a connection to 7 other people in the network, that person was definitely a TB patient (Table 8, Figure 21).

9.1.2 Contact Tracing Data of Fingerprint Type 1 (FP1) Network in Manitoba

FP1 was the dominant strain in Manitoba with 162 cases (25.8%) over the study period. The reason for the dominance of FP1 remains unclear. Contact tracing investigation done on 121 FP1 TB cases found a total of 2195 contacts. The average number of contacts per TB case was 20.3. These data showed that most of the TB cases and their contacts were relatives (46% of all contacts with no history of TB disease and from 57% to 86% of all contacts with a history of TB). Homes were the most common place for interaction among the network members. Thus, relations among the FP1 network members were very close, and degree of exposure to each other was high.

Network plots confirm that FP1 was the main strain in Manitoba (Figure 23) and Community 1 (Figure 15). The relationship between these two networks (FP1 TB network in Manitoba and FP1 TB network in Community 1) is significant. Looking at both networks separately and together (Figure 23) demonstrates how tightly both networks are interconnected. The graphs also point to the deficiency of the geopolitical definition of outbreaks usually used to delineate TB outbreaks. The Community 1 TB outbreak network shows that the boundaries of the network extend outside Community 1 to include most of Manitoba.

Pruning the FP1 network gradually according to degree of contact using the Seidman *K*-core process identifies the core composition of the network. This collapsing process helps to understand the FP1 network structure and expose the significance of the Community 1 nodes to the whole FP1 network. The results are shown in Figures 23 to 32 and are summarized in Table 9.

The core collapse sequence exercise further illustrates the inadequacy of a geopolitical definition of the Community 1 outbreak. Again, the true boundaries of the Community 1 outbreak were observed to extend well beyond the geopolitical boundaries of Community 1; the outbreak's sphere of influence extended to involve most of Manitoba. This finding is important for design of effective control strategies.

9.2 Community 1 Case Control Study

Field recruitment yielded 83 participants, which was 62% of the target 134 individuals. Twelve TB case-patients (41% of 29 TB cases targeted), 44 contacts (58% of 76 contacts targeted), and 27 controls (93% of 29 controls targeted) were recruited. Descriptive characteristics of the participants in the three groups are presented in Tables 30 and 31.

9.2.1 TB Group Network

The TB group was made of 12 index nodes, which in turn form a network of 88 nodes. Structurally, the TB network is composed of five components. The index nodes form 14.5% of the total number of TB network members. Degree centrality ranges from one to fifteen, and the mean value of degree centrality is 1.932. Degree centralization is 15%, while Betweenness centralization is 12%. Adding places attended by index nodes to the network increases the size of the network to 120 nodes, and the number of components in the TB network was reduced to one. Betweenness centralization increases to 50%, as shown in Table 11.

9.2.2 Contact Group Network

The contact group was the largest group, with 44 individuals, and the contact network size is 198 nodes, including 44 index nodes. The contact index nodes represent 22% of the network. The network is divided among 9 components. The degree centrality ranges from zero to twelve, and the mean value of degree centrality is 2.251. Degree centralization is 4.9%, while betweenness centralization is 29.1%. Adding places attended by index nodes to the network increases the size of the network to 289 nodes, while

reducing the number of components to 6. Betweenness centralization increases to 34.9% (Table 11).

9.2.3 Control Group Network

The control group was composed of 27 individuals matched to TB cases by age and gender. The network size is 170 nodes. Index cases represent 15.9% of the control network size. The network has seven components. Degree centrality ranges from one to twenty (2.071 mean). Degree centralization is 10.7%. Betweenness centralization is 25.4%. Adding the places attended by index cases to the network increases the size of the network to 217 nodes and reduces the number of components to one. The value of betweenness centralization increases from 25.4% to 49.0%.

Thus, the trend of observed changes with the addition of places to the TB group and control group networks was similar. The number of components decreased from five and seven for the TB and the control networks, respectively, to just one component in both cases. In other words, adding places to both these networks transformed them structurally from fragmented to compact and well connected (Table 11).

9.2.4 Community 1 Whole Network (Three groups combined)

Since all study group members (TB, contact, and control) lived in Community 1, combining them is more natural and will reflect the Community1 network more accurately than any individual group. The whole network is composed of 83 index nodes. Those 83 individuals (6.9% of Community 1 residents) with their contacts (alters) form a network of 391 nodes (32.6% of Community 1 residents) with 10 components. Index nodes represent 21.2% of the whole network. Degree centrality ranges from zero to 20 with a mean value of

2.47. Degree centralization is 4.5%, and betweenness centralization is 19.9%. Adding places to the network (145 places) has a major structural impact: network components are reduced to 4 from 10; betweenness centralization is increased significantly from 19.9% to 34.4%; and degree centralization is slightly increased from 4.5% to 5.6% (Table11).

9.3 Community 1 Case Control Study & Public Places

Places attended by index nodes can be divided into public and private places. For tuberculosis, public places are more important than private places from a public health point of view. Figure 47 and Table 10 show public places attended by index nodes of the study groups independently and collectively. Degree betweenness, and Bonacich eigenvector centrality measures were used to rank the importance of public places. The general store, school, and the nursing station were the top public places attended by index nodes in the Community 1 network (Table 10). Significance of the various public places varied between study groups. The nursing station, general store, and school, respectively, were the top public places attended by index nodes of the TB group. The top three public places attended by contact group index nodes were the school, general store, and nursing station. Finally, the control group network attended mainly the general store, school, and the community recreational hall.

Comparing the network graphs for the study groups (TB, contact, control, and whole network) with and without places illustrates the effect of adding places on network structure and formation. The TB network without places, for example, is fragmented, but is radically transformed with the addition of places. This addition converts it to a well-

connected and non-fragmented network (Figures 33, 37). Betweenness centralization significantly increases to 50% from 12% (Table 11).

In the two-mode network of all three groups, the average distance between reachable pairs of nodes in the node subset affiliation network is 2.176, and the average distance between reachable pairs of places in the places overlap network is 3.035. The average number of places each node visited was 29.9, and the average number of visits per place was 24.8 for public places, and 5.9 for private places (Figures 46, 47). Centrality measures (degree, betweenness) on the whole affiliation network in Community 1 identify the general store, school, and L store as the three most significant public places in the network of Community1 respectively.

9.4 Differences in Personal Characteristics Between Groups

The individual characteristics of the participants studied were gender, age, education, residence, travel, living arrangements, work status, housing conditions, and co-morbid conditions (Tables 30, 31).

9.4.1 Gender

Gender distribution among the study population of 83 was 39 females (47%) and 44 males (53%). There were no significant differences in gender distribution between groups (Table 30).

9.4.2 Age

The overall mean age was 34.6 years (SD 19.8) (Table 31). The TB and control groups were matched by age (within a 3 to 5 year range) and gender. Table 31 shows that

members of the TB and control individuals were all older than 11 years of age, whereas some members of the contact group were younger than 12.

Comparison of mean age for each group using paired t-tests and ANOVA (excluding children less than 12 from the contact group) found that there were no significant differences in group mean ages. Age was also found not to be significantly different between groups on multiple logistic regression analysis.

9.4.3 Education and Schooling

Education was measured by the highest number of school years attended. The overall mean was 7.1 years (Table 31). The only statistically significant difference in education between groups was between the TB group and contact group (Table 15). This finding can be explained by the fact that contact group members were often children or other family members of TB cases, and therefore likely to be younger in age. No significant differences in education were found between the TB and control group, or between the contact and control group (Tables 14 and 16).

9.4.4 Place of Residence

Forty-nine participants (59%) lived in Community 1 for all of the five years, while 34 (41%) lived some of the past 60 months outside of the community (Table 30). The average length of time lived outside the community was 11.5 months (range 1 to 36). No significant differences were found between groups in place of residence over the last 5 years (Tables 12, 14, 15).

9.4.5 Living arrangements

Living arrangements by group are summarized in Tables 17 and 30. There were no significant differences in this variable between groups (Table 17).

9.4.6 Work status

Results reflect the social reality of many remote northern Canadian communities, where unemployment is prevalent. Overall, 51% of study participants were unemployed (Table 30). The only significant difference between the TB, control and contact groups on this variable was that the contact group members were more likely to be students (Table 17, 23, 30), and this result was confirmed in the logistic regression analysis Table 36 (contacts 18.6 times more likely to be students than controls, $p<0.01$).

9.4.7 Housing conditions and density

Dividing the number of occupants by the number of rooms in the house gives the density of the house. Number of rooms per house and number of occupants per room are summarized in Tables 30 and 31. Results of tests of statistical differences between groups on this variable are found in Tables 12, 14-16, 22 and 23.

The mean number of persons per room (house density) in Community 1 was 1.56 (range 0.25 to 3.0). The TB group had the highest house density among the study groups (mean of 1.74), followed by the contact group (mean of 1.67). The control group has the lowest house density (mean of 1.3). The TB and contact groups had significantly higher mean house density than the control group.

The community mean house density (1.56 people per room) was used to divide the study group subjects into two groups: living in dense houses (house density above 1.56)

or living in non-dense houses (house density below 1.56). Bivariate comparisons between groups using this categorization also found that both the TB and contact groups had more dense housing in comparison to the control group. Finally, multivariate analysis using logistic regression confirmed house density as a significant predictor of TB status; the adjusted odds ratio is 12.1 (compared to controls) ($p < .05$, Table 34). In other words, while controlling for other covariates, when housing density increased by one person per room, the odds that the individual was a TB case increased by twelve times.

9.4.8 Number of TB patients known by index nodes

Asking the participants about the people they contact who are known to have history of TB revealed 31 individuals (39%) who reported knowing TB patients. The overall mean value was 3.6 TB patients known (Table 31).

TB group members were more likely to know other TB cases than their counterparts in the contact and control groups. The mean number of other TB cases known by the TB group was four times that of the contact group and six times that of the control group (Table 31). These differences are statistically significant (Tables 12, 14, 15). When number of TB patients known was coded into two categories (none versus one or more), chi-square testing found significant differences between groups on this variable as well (Tables 17, 22, 24). Finally, multivariate analysis using logistic regression confirmed that number of TB patients known was a significant predictor of belonging to the TB group with an adjusted odds ratio of 3.32 (versus the control group) ($p <. 05$, Table 34). As well, contact group members were more likely than controls to know other TB patients (adjusted odds ratio 2.71, $p <. 05$, Table 36).

9.4.9 Co-morbid illness

Half of the TB group members reported other chronic illnesses, compared with one-third of the control group, and 19% of the contact group (Table 30). The mean index of co-morbid conditions for each group is presented in Table 31. The values indicate the TB group had the highest point total, followed by the control group, with the contact group coming last. ANOVA detected significant inter-group differences in means (Table 12). Paired comparisons between group means confirmed a significant difference only for the TB: contact group comparison (Table 15).

9.5 Network characteristics

This study recruited 83 individuals, representing three different groups: TB, contact, and control. The study also represents 83 small networks. To understand the dynamics and structure of the networks, characteristics of these networks must be compiled and analyzed. The network characteristics analyzed were: network size (number of people and places), contacts (number of people only) size, age of alters, closeness of the index nodes to the alters, frequency of the index nodes' interaction with alters, and place diversity (ratio of public places visited by index cases to the private places).

9.5.1 Network Size

Network size is the number of contacts (people and places) the index cases interact with or visit. Network size is important in transmission of TB among community residents. Mean network size for all participants was 9.6 (range 0 to 25). Table 31 shows the mean values of network size of all participants collectively and for each individual group.

Individual group means were 11.4 for the TB group, 8.6 for the contact group, and 10.3 for the control group. Tests for differences between groups by t-test, ANOVA and ANCOVA revealed no significant differences (Tables 13-15, 19-21).

9.5.2 Contact Size

Contact size is the number of alters or contacts (people only). In this study, the size ranged from zero to 20, with a mean value of 6.0 (Table 31). The student t-test, ANOVA and ANCOVA tests for differences between the three study groups revealed no statistically significant differences (Tables 13-15, 19-21). However, multivariate analysis using logistic regression confirmed that contact size was a significant predictor of belonging to the TB group with adjusted odds ratio of 1.61 (versus the contact group) ($p<.05$, Table 35).

9.5.3 Age of alters

Alters are the contacts of the index node, i.e. the members of the egocentric network of the participants. The overall mean age of index node alters was 20 (SD 13), ranging from a mean of 16 for the contact group, 25 for the control group, and 27 for the TB group (Table 31). ANOVA revealed that there were significant between-group differences in the age of contacts (Table 13). The student t-test confirmed significant differences on this variable between the TB group and the contact group (Table 15), and between the contact and control groups (Table 16), but not between the TB and control group.

9.5.4 Closeness

The mean scores for this variable are presented in Table 31. The logistic regression model comparing contact and control groups showed an adjusted odds ratio of 8.12 ($p<.01$,

Table 36). No other significant between-group differences were found for this variable (Tables 13-16, 19-21).

9.5.5 Frequency of contacts

Frequency of contacts is the number of times index individuals interact with their alters per week. The frequency of contacts in this study ranged from zero to seven times per week. The mean value for all participants was 3.0 (Table 31). Only one significant between group difference was found on this variable: ANCOVA detected a significantly greater frequency of contacts in the control group than the TB group (Table 19). All other comparisons were non-significant (Tables 13-16, 20, 21). Logistic regression analysis confirmed that frequency of contact was significantly less for the TB group than the contact group (adjusted odds ratio 0.35, $p < .05$, Table 35).

9.5.6 Place Diversity

Public places can be important in disease transmission, particularly for airborne ones like TB. Therefore, identifying public places attended by TB index cases is vital for control efforts. In the case control study, all index nodes were asked about places they regularly attended. The most popular public places attended by the study participants in the single mode network are shown in Table 10, which illustrates that the control group goes most frequently to the general store, followed by the school (according to the betweenness measure) or the nursing station (according to the degree and Bonacich eigenvector measures), and thirdly, by the recreational hall (according to the betweenness measure) or the band office (according to the degree and Bonacich eigenvector measures). The contact group is composed mainly of family members of the TB group, and the school is the public

place they visit most frequently, followed by the general store and the nursing station. The TB group visited the nursing station most frequently, followed by the general store and the school. Places visited by the study groups may be a reflection of age, health and social status. Visiting the recreational hall may reflect the healthy status of the control group. The contact group is mainly composed of children of the TB group, therefore school is the most common place attended. Finally, the TB group's frequent visits to the nursing station may reflect co-morbidity and generally poorer health status.

Place diversity results are presented in Table 31. The only between-group significant difference in place diversity was between the contact group and control group, as tested by ANOVA (Table 13), student t-test (Table 16) and ANCOVA (Table 21). Multivariate analysis using logistic regression confirmed that the contact group visited significantly fewer public places than the control group with adjusted odds ratio 0.215 ($p<$ 05, Table 36).

9.5.7 Network Type (Risky & Non-Risky)

Among the 83 participant nodes there were 45 (54%) whose network was classified as risky, according to the criteria outlined in the Methods section. The proportion of the three groups' members whose networks were classified as risky differed significantly, with 75% of the TB group members, 67% of control group members, and 40% of contact group members being classified as being part of risky networks (Table 17, $p < .05$) Paired comparisons of this variable between groups confirmed significant differences on bivariate (Tables 23 and 24) and multivariate logistic regression analysis confirmed that the TB group had more risky network (almost 4 times) than the contact group with adjusted odds ratio 3.97 ($p< .05$, Table 35).

9.5.8 Behavior Types (Risky & Non-Risky)

As described in the Methods, index nodes interacting with their alters through partaking in alcohol or drugs together, or through sexual relations, were categorized as reporting high risk behaviours. Table 30 shows that 44 of the total group were categorized as having risky behaviors (53%). The TB group had 9 index cases with risky behaviors (75%). The control group had 16 index individuals labeled with risky behaviors (59%), while the contact group had 19 individuals with high-risk behaviors (43%). These differences in proportions of the three groups reporting high-risk behaviors were not statistically significant (Tables 17, 24).

9.6 Whole network "socio-metric" approach

The whole network "socio-metric" approach was very successful in answering some of the questions which arose from the pilot project, and was helpful in identifying important aspects of the Community1 outbreak locally and globally (Manitoba). The study found that the Community 1 TB outbreak network is very compact and the nodes are well connected (figures 15 and 22). Pruning of the network demonstrated that any person with connections to 4 or more nodes would be have a high likelihood of being a TB case (68%) or having a positive Mantoux test (32%). If person had connections to 7 other nodes from the same network, then there was 100% likelihood that that was a TB case (Figure 23, Table 8).

The FP1 outbreak network (Figures 23-32, Table 9) included most of the Community 1 outbreak. In reality, this figure demonstrates the real boundaries of the

Community 1 outbreak. The network extends to incorporate many Manitoba communities, and even extends beyond Manitoba into neighboring provinces.

The TB group network of people only (Figures 33, 43) was composed of 5 components. Adding places frequently attended by index cases to the network transformed this fragmented network into a well-connected network with only one component (Figure 37). This tightly connected network with only one component was fertile ground for TB transmission.

The Degree centralization value did not change much when places were added to the network (Table 11). In contrast, the Betweenness centralization value increased when places were added to the network. The TB group network Betweenness centralization value was 12.3% with people only. This value increased significantly to 49.9% in a network of both people and places (Table11). The change in the Betweenness centralization value demonstrates the significance of places in transforming outbreak networks to a more efficient environment for transmission of TB. In daily contact tracing practice, place information was either not collected or the collected information was never documented in the contact tracing data. In general, public places had more visitors than private places, and therefore require more consideration in the future. Table 10 summarizes the differences in public places attended by the TB group, contact group, control group, and all groups together respectively.

The Community 1 network is composed of 10 components (Tables 10, 25, and 26). The largest component consists of 353 nodes, or 90% of the Community 1 network (Tables 25, 26 and Figures 40, 43, 44, and 45). All of the TB cases and most of the contact and control nodes were in the main component; the remaining 9 components represent only 10% of the

Community 1 network members. The diameter of the Community 1 network and its main component was 13 (Table 26). The main component of Community 1 network represents the largest all connected part of Community 1 network (Figures 44, 45).

The use of selected SNA measures like degree and betweenness centrality measures to calculate and identify important nodes in the network structure (Figures 38, 39, 41, and 42) provided useful information and could potentially help direct scarce intervention resources. Figures 38 and 39 demonstrate significant nodes in the TB network (people and places) in single mode network, as does Figure 47. Figure 46 depicts the Community 1 network (people and places), in two mode network; places are blue in color and located in the center of the graph due to the fact that significance of places is high in the network structure.

Extracting cliques and studying clique overlaps was also found to provide important information (Figure 48, and Tables 4, 27). Some TB cases were found to be members of several cliques each. Members of the cliques who were free of TB disease require priority attention particularly if they have contacts with more than two TB patients; their chances of contracting the disease are very high. Theoretically, these people are strong candidates for TB prophylaxis. Table 4 shows the cliques in the TB network of Community 1, and the clique members' TB status.

9.7 Personal network "ego-centric" approach

The student's t-test was used as a first step to find which characteristics differed significantly between study groups. The student's t-test analysis results are displayed in Tables 14, 15, and 16 as follows:

- Table 14 compares the TB group and the control group.

- Table 15 table contrasts the TB group with the contact group.

- Table 16 compares the contact group with the control group.

TB versus control group

Comparing the TB group with the control group using the student's t-test yields three statistically significant variables: house density, number of contacts known to be TB patients, and frequency of contact. The house density results revealed that the TB group lived in more crowded houses than the control group. Second, the index nodes of the TB group had an average number of four contacts with the TB disease, while the control group nodes had an average of only 0.63 contacts with TB. Finally, the TB group index nodes interacted with their alters 1.9 times per week versus 3.4 interactions per week for the control group.

TB versus contact group

The contact group individuals were mainly family members of the TB group. Some contact group members were younger than 12 years of age, while among the TB group the youngest node was older than 12. Therefore, all individuals younger than 12 years of age among the contact group were excluded from the analysis. Table 15 shows the comparison between all continuous variables (education, house density, residence outside Community 1, co-morbid conditions, number of known TB patients, age of alters, network size, contact size, place diversity, and frequency of contact) of the TB and contact groups using the student t-test.

The average number of school years attended by the TB group was 8.9 years, while for the contact group it was 6.5 years. However, this difference may be explained by the fact that the contact group had more children in it than the TB group. Co-morbid conditions

were more prevalent among TB group members than among the contact group. People from the TB group tend to know more TB cases (4 on average) than the contact group (0.95 on average). The remaining variables did not differ significantly (Table 15).

Contact group versus control group

Comparison between the contact and control groups was important because the control group reflects the underlying social structure of Community 1, while the contact group members represent potential future TB cases.

The student t-test identified three statistically significant variables: house density, age of alters, and place diversity. Contact group members live in denser houses than the control group. The contact group interacts with younger alters (average age 15.7 years) than the control group's alters (average age 24.4 years). The last statistically significant variable was place diversity, which indicated that less than half (48%) of the places attended by contact group people were public places, while 79% of the places attended by control group members were public places (Table 16).

As described in Methods, the chi-square test was used to compare categorical variables of the study groups. Results for all groups are presented in Table 17, TB versus control group in Table 22, contact versus control group in Table 23, and TB versus contact group in Table 24. The variables analyzed included living arrangements, number of TB cases known, house density, work status, network type, and behavior type. No statistically significant differences in living arrangements were found between groups. Significantly fewer control group members (18%) than TB group (75%) or contact group (54%) members lived in high density houses, as defined in Methods (Tables 17, 22, 23, and 24).

The TB group members knew significantly more other TB patients than did either controls or contacts (Table 24). A significantly greater proportion of the contact group were students, in comparison to the other two groups (Tables 17 and 23). The TB and control groups had a higher percentage of risky network relationships than did that contact group (Table 17.)

ANOVA results show that house density ($F\ 3.630,\ p < 0.05$), co morbid illness ($F\ 4.631,\ p < 0.05$), and number of TB patient known by the index nodes ($F\ 11.090,\ p < 0.001$) were the only significant differences among group characteristic variables. ANOVA results of the network characteristics point to age of alters ($F\ 5.943,\ p < 0.01$) and place diversity ($F\ 8.824,\ p < 0.001$) as the only two variables which differed significantly (Tables 12, 13).

Table 18 is a correlation matrix for all of the individual and network characteristics analyzed. The first column contains variable names and an indication of whether the variable is significantly correlated with other variable(s). The second column contains the mean value of each variable. The third column contains standard deviation. The remaining columns contain the Pearson's correlation coefficients and the associated p-values.

The following variables were found to be significantly correlated to other variables:

1. The *age* variable was correlated with house density, co-morbid conditions, age of contacts, and place diversity.
2. The *age 12* variable was correlated with education, living outside Community 1, co-morbid conditions, and age of contacts.
3. The *education* variable was correlated with the number of TB patients known, network size, and contact size.

4. The *house density* variable was correlated with co-morbid conditions and age of contacts.
5. The *living outside Community 1* variable wss correlated with network size, contact size, and frequency of contacts.
6. The *co-morbid conditions* variable was correlated with age of contacts only.
7. The *number of TB patients known* variable was correlated with network size and contact size.
8. The *age of alters* variable was correlated with network size, contact size, place diversity, and frequency of contacts.
9. The *network size* variable was correlated with contact size and frequency.
10. The *contact size* variable was correlated with frequency of contacts only.

Analysis of covariance was the next step in the statistical analysis plan to answer the study questions. The ANCOVA step was conducted as a three-way comparison:

1. TB group compared with control group (Table 19).
2. TB group compared with contact group (Table 20).
3. Contact group compared with control group (Table 21).

1. TB group compared with control group

Comparing the TB group with the control group revealed no significant statistical differences in network characteristics except frequency of contact ($F\ 6.68,\ p<0.05$). The frequency variable results indicates that the TB group interacted with their contacts 1.5

times per week, while the control group interacted with their contacts 3.5 times per week. The remaining variables did not differ significantly (Table 19).

2. TB group compared with contact group

The ANCOVA results show no significant differences in the comparison between the TB group and the contact group (Table 20).

3. Contact group compared with control group

Most of the variables in this comparison were not statistically significant, except for one: place diversity ($F\ 9.68,\ p < 0.01$). Fifty-two percent of the places frequently visited by the contact group were public places, compared to 79% for the control group members (Table 21).

Discriminant analysis

Discriminant analysis results for the case control study are presented in Figure 50, and Table 32. The differences lie in two dimensions or functions. Both axis 1 and axis 2 were statistically significant (Figure 50). The highest score contributions to the axis 1 (or x-axis) were from the *number of TB patients known* and *mean age of alters* variables. For axis 2 (y-axis) the highest score contributions were from the *place diversity* and *mean age of alters* variables (Figure 50, Table 32). The results show that the TB group scored high on axis 1; with the highest number of TB patients known and highest mean age of alters, while both the contact and control groups both scored low on axis 1. The control group scored highly on axis 2, with the contact group scoring the lowest on this axis. The control group

had the greatest place diversity, while the contact group had the least place diversity (Figure 50, Table 32).

Logistic regression

Logistic regression was the last step in statistical analysis plan to answer the study questions. Three final models were obtained from backwards stepwise logistic regression modeling, as detailed in the methods:

1. TB group compared with control group.

The final model for the comparison between the TB group and the control group (Table 34, R^2 0.43, $p < 0.001$) contains two variables: house density and number of TB patients known. Thus, the TB group members lived in more crowded houses and knew more other TB patients than did members of the control group.

2. TB group compared with contact group.

The final model for the comparison of the TB and contact groups (Table 35, R^2 0.36, $p < 0.001$) contains three variables: contact size, network type, and frequency of contacts. Interestingly, these variables are all network characteristics.

3. Contact group compared with control group.

The third final model, for the contact group vs. control group comparison (Table 36, R^2 0.41, $p < 0.001$) contains four statistically significant variables, including two group characteristics and two network characteristics. The group characteristics are number of TB patients known and work status, while the network characteristics are place diversity and closeness.

Investigating the significance of SNA centrality measures:

Four centrality measures were used (degree, closeness, Betweenness, and Bonacich). In univariate analysis, degree, Betweenness, and Bonacich were statistically significant. Putting all the centrality measures (variables) together in a logistic regression model resulted in a final model where degree centrality is the only remaining significant measure (Table 37, R^2 0.055, $p < 0.001$). In other words, degree centrality was the best single predictor of TB cases among the network population when put all centrality measures results of the network population as independents variables against TB status as outcome variable in logistic regression model.

Degree Distribution in Community 1 Network

The main component of the Community 1 network represented 90% of the members of the whole network (Table 25). The frequency of degree distribution of the main component of the Community 1 network ranged from 1 to 20 (Table 33). Cumulative distribution figures (Table 33 and Figure 51) confirm that most of the network nodes were connected individually to a small number of other nodes, with few connected to a large number of nodes. Plotting degree distribution and cumulative distribution on normal axes (Figure 51) produced a curve similar to power law distribution (long tail to the right). Plotting the same values of degree and cumulative distributions on log-log axes produced a curve resembling a power distribution (Figure 52). Regression was used to estimate if the curve follows a linear, exponential, or power distribution. Regression results (model R^2 values) for curve estimation of Community 1 network cumulative and degree distribution were 0.59, 0.92, and 0.97 for linear, exponential, and power distribution respectively (Figure 53). Thus, linear distribution was the least compatible with the curve and was

excluded. The same values were plotted on log-log axes to estimate the curve compatibility with exponential and power distribution. Regression results for curve estimation were statistically significant, with R^2 of 0.97 for the power distribution and 0.92 for the exponential distribution (Figure 54).

The upper part of the curve (body) was closer to a power distribution than the lower part of the curve (tail), which fit an exponential distribution more closely (Figure 54). Diversion of the curve started from power to exponential distribution at degree 8. To investigate the curve further, we split the data at degree 8 and called the part that contained degree 8 or higher the body of the curve. We plotted it separately from the remaining part of the curve (Figure 55). The second part of the data, which represented by the remaining part of the curve, was called the tail (lower than degree 8, Figure 56).

The Community 1 network cumulative and degree distribution regression results for curve estimation of 8 degrees or more (body) shows an R^2 of 0.99 for the power distribution interpretation, and R^2 of 0.87 for the exponential distribution interpretation (Figure 55). Therefore, it is more appropriate to state that the body part of the curve follows the power law distribution. Regression results for curve estimation of 8 degree or less (tail) shows a curve estimation with an R^2 of 0.99 for the exponential distribution interpretation and an R^2 of 0.97 for the power distribution (Figure 56). Therefore, it was more appropriate to state that the tail part of the curve follows the exponential distribution the closest. In summary, the body part of the curve on log-log axes follows the power law distribution and the tail part of the curve on log-log axes follows an exponential distribution.

SECTION V: DISCUSSION AND CONCLUSIONS

Chapter 10: TB in Manitoba

Manitoba invests considerable resources to control TB among its residents, an effort that has been largely successful, as indicated by the status of TB among two of Manitoba's population subgroups. The TB incidence among Canadian-born people without treaty status is the first indication of success, with an incidence declining steadily over time to the current level of 3 per 100,000 per year. Likewise, the historical decline in TB incidence in Canadian-born people with treaty status is a second success story, despite the fact that TB incidence in this group remains much higher than their Canadian-born without treaty status counterparts. The fact remains that comparison of TB status among Canadian-born people with treaty status as they are now and as they were in the past demonstrates impressive improvement and tremendous progress in TB status, particularly when recognizing that the mortality rate among Aboriginal people due to TB not that far away in history and in neighboring provinces (Saskatchewan and Alberta) was as high as 9,000 per 100,000 (1909) (52, 125, 126).

In spite of these successes, there are some remaining concerns about TB incidence in Manitoba. Inequality of TB incidence between provincial and national levels is one of these concerns. The difference in TB incidence between Manitoba and the overall Canadian level were described previously. Of further concern, and perhaps more alarming, is the difference in TB incidence among different population groups of Manitoba's residents. There is enormous difference in incidence between population subgroups as

defined by sociocultural identity, a difference that is not acceptable and requires revisiting of the TB problem on the part of TB control and policy making people in Manitoba.

The difference in TB incidence between Canadian-born people with and without treaty status is not a unique problem that only affects Manitoba's residents. This problem is extensively reported in the literature of most of Canadian communities that have suffered from this disease in the past and present(11, 13, 21, 22, 34, 39, 40, 46, 51, 52, 59, 60, 77, 90, 126, 132, 133, 135).

The TB problem in Manitoba parallels the national trends, but Manitoba's TB incidence remains higher than Canada's TB incidence. This study of TB cases in Manitoba between 1992 and 1999 shows that TB is a multifaceted problem, in practical terms, when identifying the cases in terms of ethnicity or birthplace. These findings have several implications. First and foremost, Aboriginal people require more TB control attention. Demographically, Aboriginal people are over represented in TB cases as an ethnic group. Second, to control TB among Aboriginal people and succeed in doing so means potentially preventing almost half of all TB cases in Manitoba. Third, Winnipeg and eight reserves harbor most of the Aboriginal TB cases; therefore these places function as the main reservoirs of the disease in Manitoba. Residents of these reserves represent a tiny portion of Manitoba's population (1.5%), and for that matter, a fraction of the total Aboriginal population of Manitoba (17%). By focusing more on these communities instead of on all of Manitoba's reserves or on all the Aboriginal people in Manitoba, the total workload could be reduced without compromising the effectiveness of the control effort. Such a move would help in organizing priorities in control efforts – particularly in light of the scarcity of

and huge demand for health service resources. Focusing on those eight reserves would likely result in a more effective control outcome.

In Manitoba, the strain of TB called FP1 is most prevalent, and currently infects mainly Canadian-born people with treaty status. This dominance of FP1 is an important finding and is extremely significant for control efforts. The reason for the dominance of FP1 remains unclear. However, it has been speculated that FP1 may have been the first strain introduced into Manitoba. Routine BCG vaccine administered among treaty status infants may have further fostered the natural selection of FP1 by increasing immunological protection against other susceptible strains through selective immunity. Host genetics, strain virulence (including expression, regulation and mutation of bacterial virulence factors), and host immune evasion mechanisms may have also played a role in the success of this strain(66, 114). However, the role of these factors in the dominance of FP1 remains to be elucidated(66).

The dominance of Canadian-born with treaty status cases representing the largest single sub population group of TB cases in Manitoba applies equally to the patients with the FP1 TB strain. In other words, Canadian-born with treaty status patients were over-represented in FP1 TB cases in Manitoba. The majority of FP1 TB cases were Canadian born with treaty status living in Winnipeg, followed by Canadian born with treaty status living on reserves.

10.1 Main Findings of the Study

Contact tracing investigation is one of the pillars used in TB control practice to identify potential TB patients or latent infection. Limitations of the contact tracing investigation procedure have been frequently cited in the literature(48, 78). The limitations

basically are twofold: first, not all TB patients go through contact tracing investigation; only pulmonary TB patients do. Secondly, missing data from contact tracing investigations is very common, for several possible reasons, including lack of compliance from participants, logistical difficulties facing patients and health care providers, and both the procedure and patient behavior together (16, 19, 48, 78, 87, 98, 105, 110, 119, 121). However, limitations do not reduce the importance of contact tracing in TB control. The steady reduction in TB incidence among the mainstream population (Canadian-born people without treaty status) indicates that contact tracing has worked effectively in the past in most contexts. However, the transformation of the TB landscape in Manitoba and Canada presents new challenges to health authorities, which dictate a new reality, and therefore require new or novel approaches to improve contact tracing practice to accommodate these challenges.

This study of tuberculosis in Community 1 was valuable in identifying important facts about TB in Aboriginal communities. The study shed more light on TB groups as individuals and as group. Implementing social network analysis in the study, and applying whole network (socio-metric) and personal network (ego-centric) analytic approaches was found to provide unique and useful insights. These findings and their implications are highlighted below.

The whole network analysis approach is important because it gives researchers the opportunity to visualize the whole network and understand its structure. A network fragmented into several components is not as efficient in TB transmission as a network with only one component. Our study found that the use of single mode network analysis in conjunction with two-mode network analysis was complementary. We found that the

addition of centrality measures and centralization of the network or cohesion of the network nodes around focal nodes was helpful in identifying the significance of public places and their contribution to cohesive networks conducive for TB transmission in the community. Extracting the core composition of the network by using the core collapse sequence identified the core nodes and further elaborated the network compactness, structure and implications for TB transmission. An important finding of the study was the observation that as *k* increases in the Seidman *k*-core collapse sequence, the probability of active or latent tuberculosis infection increases. Those in the network who were exposed to 7 or more TB cases had a 100% probability of being TB cases themselves. The whole network approach was also able to identify the boundaries of the outbreaks of the Community 1, FP1, and FP5 networks. As shown in the FP1 TB network (Figure 23), TB outbreak boundaries in Community1 extend beyond the community boundaries. Hence, boundaries of TB outbreaks in Community1 were not bounded by geopolitical boundaries. It is understandable that individual health authorities prefer using geopolitical boundaries in defining outbreaks for their logistic and administrative purposes, but using geopolitical boundaries gives a false impression about the real or natural boundaries of any outbreak. For an effective control program, natural boundaries of outbreaks should be investigated and identified first.

Our analysis of TB and Mantoux status in the Community1 TB outbreak network and the degree of exposure of the same network members revealed that high exposure to TB cases leads to a high likelihood of infection or disease. We found that connecting with more than one TB patient results in a marked increase in risk of TB infection and disease. The study found that comparing graphs of the case control study groups with and without places

demonstrated that places attended by index nodes are vital in creating a well-connected network. Public places are more significant and attended by higher number of people than private places, and naturally more connected networks tend to spread TB more effectively. Comparison of the whole network parameters of the TB group network with the control group network revealed that both networks had a similar structure and behaves similarly. These findings indicate that the TB network was a reflection of the social structure of the community, in this case, of Community1.

The study found that if community members had social contacts with more than two TB patients, their chances of contracting the disease were extremely high. This is an important finding, as, theoretically, these people would be strong candidates for TB prophylaxis.

From the data collected in the study, we were able to build a profile for each study group (TB, contact, and control groups). The TB group members live mainly in overcrowded houses. They have more social connections to other TB cases, thus are part of a high-risk network for TB disease development. Most members of the TB group are unemployed. In contrast, the control group members, by and large, do not live in overcrowded houses. They have few social connections to TB cases. The control group members were similar to the TB group in one way, however, in that most were also unemployed. Finally, the contact group members tended to live in overcrowded houses, putting them at risk for TB. A majority of the contact group were either students or unemployed. Contact group members were not found to be members of risky networks, and were found to have less social interactions with TB patients than the TB group members.

Multivariate analysis using logistic regression confirmed two variables to remain significant in testing for differences between the TB and control groups: house density and number of TB patients known. There were no significant differences in network characteristics between these groups. These results reconfirm a well-known fact about TB: living in overcrowded places is conducive to TB transmission. It is important that these independent risk factors were observed *within* a context where socioeconomic conditions are generally poor, that of a northern Canadian Aboriginal community.

The study found that the number of other TB cases that the TB group members had social relations with was 4.4 times more, on average, than the contact group. The TB group was also found to have more risky behaviors (alcohol, sex, and drugs) than the contact group (borderline statistically significant). Finally, TB group people usually have riskier networks (risky behaviors among index nodes and their contacts, plus some contacts are TB cases) than the networks of the contact group. The logistic regression model comparing the TB and contact groups identified three significant variables: contact size, network type, and frequency of contacts. In summary, the TB group members in comparison with the contact group had more alters but saw them less frequently, and their networks were riskier in nature than those of the contact group.

Previously published literature has pointed to the significance of overcrowded housing in TB transmission among First Nation people(27, 39). Clark *et al*(27), estimated average house density at 0.4 persons per room in Canada, 0.7 persons per room for First Nations people on reserves (all over Canada), and 0.8 persons per room for First Nations people on reserves in Manitoba.

In our study, we found the mean house density in Community 1 to be 1.6 persons per room, which is double the overall rate for Manitoban First Nations people on reserve. Three quarters (75%) of the TB group population, 54% of the contact group, and 19% of the control group live in dense houses where density is above 1.6 persons per room. The Community 1 house density is extremely high; therefore, it is not surprising that TB incidence is high in this community. This study showed that 75% of the TB group population had a connection with other TB patients, which would amplify the impact of overcrowded houses in connection to TB transmission. Overall, the TB and control groups were found to have similar network characteristics, with one exception: TB group members see their contacts less frequently (1.6 times per week), while the control group sees their contacts more frequently (3.5 times per week). This result suggests that TB patients are more isolated than other individuals of similar age in Community 1. It is unknown whether this isolation was a precursor to or a result of tuberculosis.

Overall, contact and control networks were found to be different in place diversity. The contact group visits fewer public places than the control group. The place diversity of the control group members reflects their higher social status in the community. Visiting more pubic places demonstrates their wider sphere of influences, in contrast with the contact group. The contact group members, generally, had a lower social status than the control group. Members of the contact group are important people in TB transmission, owing to the fact that they represent the first cycle around TB cases with a direct contact. Therefore, their chance of contracting the disease is higher than somebody else living in the same community. In addition, they live in overcrowded houses. Most of the contact group

people are family members of TB cases (children and spouses). Therefore, future TB cases most probably will come from the contact group.

10.2 Social Network Analysis and Tuberculosis

The knowledge gained from executing SNA in TB outbreaks in the course of two phases, the pilot project and the case control study, proved to be useful. In the pilot project, SNA helped initially to visualize TB outbreaks in several communities, identify the outbreak networks and their structures, and delineate the boundaries of the outbreaks. It was clear from the generated graphs of TB outbreaks that TB networks found in Aboriginal communities have different degrees of compactness. Some networks were simple and not compact, while others were very compact with a complex structure. Differences in the network structure compactness raise some valid questions. The first question was why some communities had a complex networks while others had simple networks, and whether the TB network structure is a reflection of the social structure of the community it represents or not. The second question concerned the real boundaries of the outbreaks, for example, the FP1, FP5, and Community 1 outbreaks. For logistical reasons, most outbreaks tend to be reported and defined by geopolitical boundaries – understandable from a jurisdictional responsibility standpoint, but inefficient from a control standpoint, and not reflecting the reality that outbreak boundaries often extend far beyond local geopolitical boundaries, as was demonstrated by this study. The third apparent question concerns the fragmentation of some TB networks, the reasons for such fragmentation, and whether this phenomenon is reflects the dynamic nature of TB transmission, or if it results from a failure

of contact tracing, systematically or locally. A good example of fragmentation is found in the network of the FP5 TB strain; fragmentation is also obvious in the networks of FP1. The answers to the question of fragmentation are not simple, though one aspect of the answer is obvious because the contact tracing results used in this study resulted from the cumulative effort of large numbers of health care providers throughout different communities with a timeline extending over eight years.

Another important point in TB transmission is consideration of the places frequently attended by index cases. These places are usually missed in contact tracing investigations, as TB transmission occurs through the sharing of air space, or "risk space", and does not necessarily require close relationships. Contact tracing investigations are usually conducted without asking index cases about places attended regularly in a timeline fashion. Index cases are usually asked only about their contact names, relations, and place of contacts.

Certainly, understanding the community and the transmission network structure is the first step in the right direction toward TB outbreak control. SNA accomplished this for the outbreaks analyzed in this study, and made it easy to visualize the outbreaks and help identify key nodes in the networks. The magnitude of any node and its influence is derived from the type and strength of its relations to the remaining members of the network. Centrality measures in SNA are one tool for that purpose; they can identify the significance of the nodes in the network and rank them accordingly. In this study, we used measures of degree, closeness, betweenness and Bonacich as epidemiological tools for screening and assessment of the nodes' significance from the disease status magnitude in the community. In future follow up, health care providers could feasibly choose to visit key nodes in

outbreak networks for assessing disease status, and could quantify the degree of success achieved in disease control due to the intervention measures applied in the community.

Other techniques were used in this study to investigate the TB network structures: "Components", "Cliques", and "Core Collapse Sequence", for example. Cliques in Community 1 were recognized and nodes requiring further attention were identified. Identifying cliques helps to discover nodes that equate together or have strong relation. Recognizing disease status for each clique member helps tailor priority and proper management.

Compact or dense networks in general allow infection to spread more efficiently and faster than sparse or non-compact networks. Having prior knowledge about communities' network density would alert communities to the level of risk facing them in a TB outbreak. In a dense network, it would not take long before the disease spread widely and deeply in the community to the point where it would be difficult to control. Therefore, the density of any network is important and could be used as a parameter to predict the pace of TB transmission and the effort required for outbreak control.

Our thinking following the pilot project observations was that if TB outbreak networks are similar to the social structure of the communities they occur in, then some communities are more vulnerable to TB infection due to the nature of their social networks. If this is the case, then such communities should be made aware of their vulnerability. If, however, TB outbreak network structures differ from the underlying community social network structure, then the implications would be different. The distinct characteristics of the TB network could be identified and used for prevention and control of TB. The results of the present study support the former hypothesis; i.e. that the TB outbreak network

characteristics were similar to those of the underlying community social network. The key differences that were identified between the TB group and the control group were that the TB group members lived in more crowded housing and associated with more other people with a history of TB. Thus, we can conclude that all members of Community 1 are at increased risk of TB due to its tight and compact underlying social network structure, and community members living in crowded houses and/or knowing one or more people with a history of active TB are at particular risk for infection or active disease from tuberculosis.

To further understand the dynamics of TB transmission in Community 1, two additional angles were examined in this study. Community 1 is a small and remotely located community in Northern Manitoba with few public places. These public places are important for community activities and are deeply integrated in the community's social makeup. Secondly, our identification of the Community 1 TB network as a scale free network shed more light on the TB network structure and its transmission dynamics. The scale free network means that a few selected nodes in the network are super nodes, i.e. nodes connected to a large number of contacts (alters) in comparison with the rest of the network. The study demonstrated that together, these "social maverick" nodes and the public places in Community 1 are the main factors that keep the Community 1 TB network tightly connected, producing fertile ground for efficient TB infection and transmission. Housing shortages and overcrowded environments make this situation even more conducive to TB transmission. Therefore, effective TB control program in Community 1 should address the housing shortages and focus first on the super nodes and on the nodes with multiple connections to TB cases.

10.3 Advantages of Using SNA in TB Epidemiology and Control

The advantages of using SNA in outbreak investigations are considerable. These advantages range from visualizing to analyzing outbreaks and their structures. SNA is more than just an epidemiological tool; it has a wide range of applications and benefits.

Visualization: SNA graphing is very useful in understanding TB outbreaks. It can depict the outbreaks as a graphic showing characteristics and illustrating the density and level of connectedness of the network. Several studies using SNA have been able to plot the networks of outbreaks(41, 69-72, 130). In the present study, SNA was able to convert complex data generated by contact tracing investigation and display it graphically in a format easy to follow and understand. It allowed for plotting of networks of outbreaks in several settings: in one, two or more communities, individually or together. Network graphs generated by SNA have immense potential to be used in public health, particularly in outbreak investigation, and because of its ability to handle compound data and display it in a way which is easy to follow, it can even be used to create an epidemic curve portraying a complex combination of data. Our study also demonstrated how SNA can be used to study a TB network's structure, sub-structure and components. The structure of the networks is the key to understanding the dynamics of TB transmission and can point toward the best control approaches (64, 65, 94, 95, 130).

Epidemiological Tool: SNA can be used as an epidemiological tool for surveillance, core characterization, identification of groups and/or individuals at particular risk of infection, delineation of outbreak boundaries, and designing and evaluating disease control interventions. SNA can assist TB surveillance activities by identifying high risk communities, social networks, individuals and places, all of which could be subject to more

intense surveillance to facilitate TB control. Detection of the core of TB networks using SNA gives health care providers an opportunity to investigate the characteristics of the core group(64, 65, 100, 107, 130).Use of the SNA technique can identify people at particularly high risk of contracting TB. For example, extracting a clique's population from the network – particularly when more than one member of the clique has TB and the remaining do not – can pinpoint the people with very high risk of acquiring TB due to their exposure to more than one TB patient. Beside its ability to visualize the network of any TB outbreak, SNA can help in identifying an outbreak's boundaries. In general, outbreak boundaries extend beyond one community's geopolitical boundaries, and SNA has the ability to shed more light on an outbreak's natural boundaries (64, 65, 130). Finally, applying SNA to contact tracing investigations is a promising way to improve the results of these investigations. Since SNA can identify key people and structures as well as core groups in TB outbreaks, these key findings could help guide rapid and effective intervention for TB management.

10.4 Study Limitations

The present study is subject to several limitations. The first limitation is related to TB as a disease and the nature of its transmission. SNA has been widely used in studying sexually transmitted infections (STI), which are directly transmitted between people. STI require intimate relations, while for TB merely sharing air space is enough to transmit the disease to new vulnerable subjects even without knowing or having any sort of relation with the index cases. This fact could have introduced a source of information bias into a study of TB using SNA, since people may have shared airspace without having social relations that they are aware of. The use of TB contact tracing investigation data in SNA

provided an opportunity to assess its compatibility and usefulness in SNA. Overall, we did find that TB contact tracing data could successfully be used for social network analysis in the context of community TB outbreaks.

A second limitation is recall bias, an important potential source of bias in all case-control studies. All TB index cases in the study were diagnosed between 1992 and 1999, and the original contact-tracing investigations were performed at the time of diagnosis. A varying number of years lapsed between the time of diagnosis and the time of interview for the present study. There is some indication of recall bias in the number of contacts named in the study interview compared with the number of contacts named in the original contact tracing investigations. The number of contacts per patient from contact tracing investigation records was 16 individuals on average per case. The study, on the other hand, showed that for 12 TB cases interviewed, there were 76 contacts (6 per case).

Selection bias is a third limitation of the study. Whereas there were 29 total cases of TB reported in Community 1 members during the study period, twelve were available and consenting for interview for the study. These cases did not differ demographically from the non-participating case; however, there may be other characteristics of the study cases which differ from the non-participating cases.

Sample size is another limitation of the study, with implications for statistical power. The small group sizes in the case-control study (especially the TB group) may have limited the power of the study to detect small differences between groups.

A final potential limitation of this study is its generalizability: because the pilot and case-control studies were conducted in selected communities and sub-populations of Manitoba, the findings may not all be generalizable to other populations. However, the

characteristics of the populations studied are in many ways similar to other populations at-risk for tuberculosis, both in Canada and other countries.

10.5 Policy Implications

TB in Manitoba is a multifaceted problem. Canadian-born people with treaty status form the largest single group of all TB cases, while representing only a small fraction of Manitoba's population. Over-representation of this segment of Manitoba's population is an important reason to be concerned about TB control in this population subgroup. Eight selected reserves harbor the majority of Aboriginal TB cases, and along with the capital city of Winnipeg, these are the reservoirs of TB in Manitoba. Boundaries of Manitoba TB outbreaks extend far beyond the geopolitical boundaries of any single reserve or community. A single TB strain, FP1, dominates all TB cases with isolates in Manitoba. Canadian-born people with treaty status represent the majority of all FP1 cases. These findings dictate that new TB control measures are urgently required to target this population subgroup(11, 12). TB control efforts should focus first on the selected reserves with high TB incidence. The benefits from focusing control efforts on hot spots in Manitoba with a global plan certainly will not reduce control efforts, but will look at the problem in more strategic way. Control efforts should be expanded systematically in consideration of the natural boundaries of TB outbreaks and targeting all population subgroups affected. A priority list of communities and sub-populations can be created with the help of SNA. SNA can also help with evaluation of the successes of these TB control interventions.

Contact tracing is an important component of TB control (80, 101, 124, 127). The goal of any TB control program is to identify potential TB cases for proper treatment (36,

52, 77, 79, 86, 105). Contact tracing investigation was designed to follow up TB cases (index cases) and people who come in contact with TB index cases at work, school, home, and during many other activities (86). Control programs successes are measured by their ability to reduce disease incidence. TB control efforts in Canada have largely been successful in reducing TB incidence among all population subgroups(36, 52, 77, 86). One out of every 5 Canadians had TB at the turn of 19th century (77). Overall TB incidence in Canada decreased subsequently to 5.9 per 100,000 by the end of the 20th century (21, 22, 55, 104). TB in Canada has concentrated among high risk groups and in defined geographic areas(77, 86, 105). Therefore, it is reasonable to conclude that TB control efforts including contact tracing were vital and successful in controlling TB among the main segment of the population in Manitoba. By the same token, control efforts and contact tracing were not as efficient or successful in reducing TB incidence among Aboriginal people in comparison with Canadian born non-Aboriginal people.

Improved contact tracing result among Aboriginal populations is the first step to success in controlling TB among Aboriginal communities. Limitations of contact tracing have been documented in the literature(87, 98, 101). Not all TB patients receive contact tracing investigation, and gaps exist in the data collected from contact tracing (87, 98). Delays in TB management is another deficiency in TB control programs (87, 105). Delay in any management steps can have negative ripple effects on subsequent steps (87, 105). Rothenberg et al. (2003) noted that contact tracing procedures in TB are not properly documented and vary by location, and the stigma of TB as a disease does not makes the process difficult (101). Expert opinion rather than evidence remains the basis of most guidelines and recommendations issued for contact tracing (79).

Despite the existence of contact tracing limitations, it is unclear to what extent these limitations have contributed to the persistent relatively high incidence of TB in Canada's Aboriginal population, and what role other factors, such as poverty, marginalization, and delays or restrictions in access to health care may play. The period between becoming ill and asking for help from health care providers is important, because during this time patients may be contagious and most at risk of passing the disease to contacts(59, 86, 105). With the dwindling number of TB cases in the general population, identification of TB patients switched from being proactive to a rather passive procedure. Proactive case finding and enhanced contact tracing are required in order to improve control of TB in Canada's Aboriginal population.

The role of public places in tuberculosis transmission and control deserves further attention. Anecdotally, a recent TB outbreak in one northern Manitoba's community was associated with a social gathering place for community teenagers. Another TB outbreak in Minneapolis was reported among regular patrons of a neighborhood bar (68). To improve contact tracing and control efforts an effort should be made to collect the names of public places attended by TB cases regularly, in addition to their contacts. Collection of this information on an ongoing basis will lead to an improved understanding of those places most frequently associated with disease transmission. Further investigation of these places may aid control activity by pointing to or revealing the reasons for association of these places with disease transmission.

This study found that housing density is the most significant variable of community related factors associated with tuberculosis networks, and requires urgent attention. The association between overcrowded houses and tuberculosis has been known for a long time.

Community1 (and many other Aboriginal communities) is suffering from a chronic housing shortage. Attention to the housing problem is necessary for several problems in the community, including TB. A reasonable goal would be to reduce housing density in Community 1 (and other Aboriginal communities) to Manitoba's overall mean housing density. This reduction will help in reducing TB incidence and reducing expensive disease control costs. Estimation of the cost of TB control in Community1 is very complex. During the TB outbreak in Community 1 between 1992 and 1999, the TB network was composed of 504 individuals, almost one third of the entire population. All network members went through some form of investigation. TB control efforts ranged from mass screenings to individualized treatment. Some cases required transportation to secondary or tertiary hospitals in Winnipeg or other neighboring communities. All these expenses added together give a crude estimate of the direct financial costs of an extensive TB outbreak. The huge financial burden to local residents and different levels of governments (local, provincial, and federal) indicate that improved control of TB in Community 1 could potentially be cost-saving in the long run, in addition to reducing the morbidity and suffering of community members.

The ultimate test of success for SNA application to TB control is the implementation of the method in daily TB control efforts. Feedback from front line health care providers and public health staff is extremely important. Therefore, initial policy changes should include implementation of SNA in day-to-day TB control efforts. SNA enables visualization of TB networks(71, 80), identifies key individuals in the TB spread networks, and recognizes network structures of TB outbreaks. SNA is a useful epidemiological tool: it can improve contact tracing performance and potentially assist to

improve control interventions and management(41, 80, 98). SNA can offer more comprehensive and systematic in-depth analysis of the information collected in contact tracing investigation than is currently performed. This wide range of contributions SNA can make to control efforts, make it a suitable and indispensable tool for outbreak investigation and management practices. Finally, the use of SNA to detect transmission paths, outbreak boundaries, and network structures can make significant contributions to improve TB control. Transmission paths are the routes through which TB enters communities, spreads among the residents, and eventually spreads to other communities in a most efficient way. Knowing the boundaries of the outbreaks helps to delineate and coordinate control efforts across public health jurisdictions. Network structures are the blue prints, which can help public health practitioners design effective intervention programs.

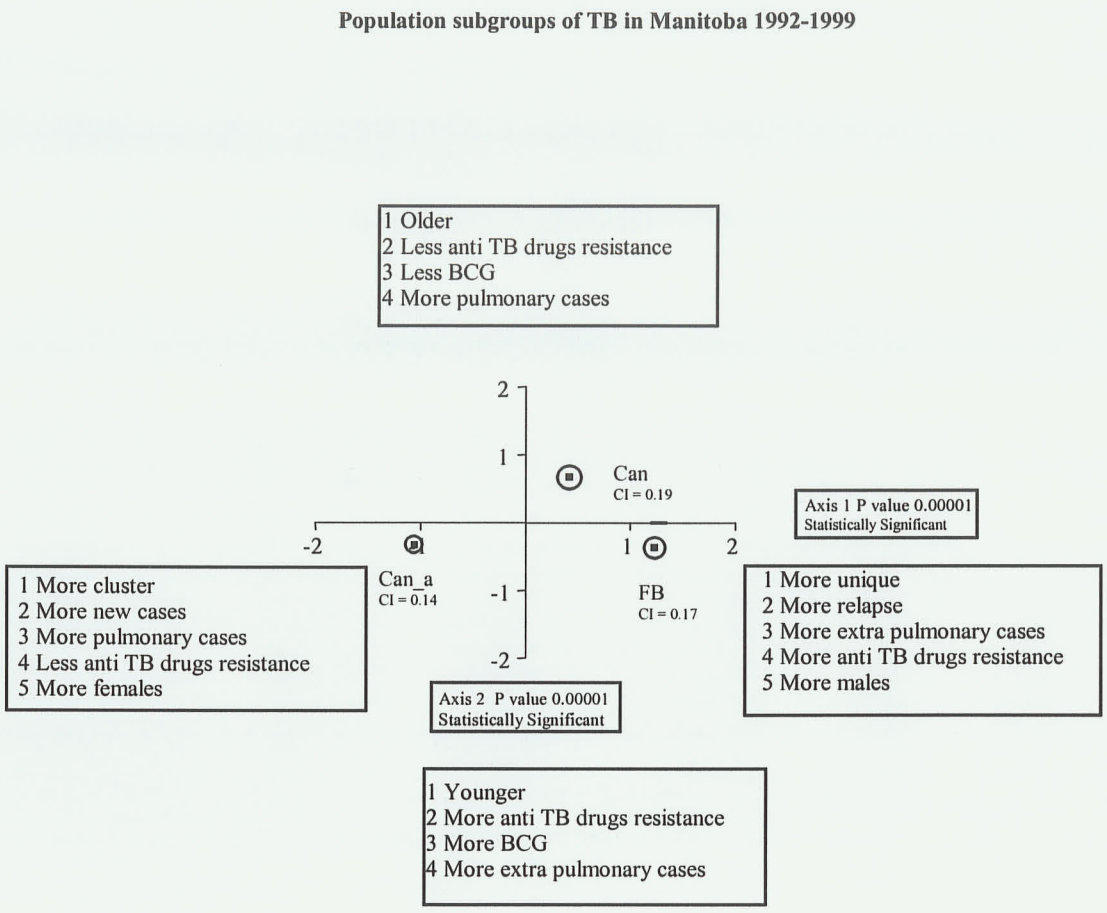
In summary, this study has demonstrated that the methods of social network analysis can be successfully applied to deepen our understanding of tuberculosis transmission in Manitoba's Aboriginal community, thereby potentially enhancing public health initiatives to control this important problem. The time has come for these methods to be applied more widely to tuberculosis epidemiology and control.

TABLES AND FIGURES

Figure 1: Reported TB incidence in Canada and Manitoba, 1985-1998

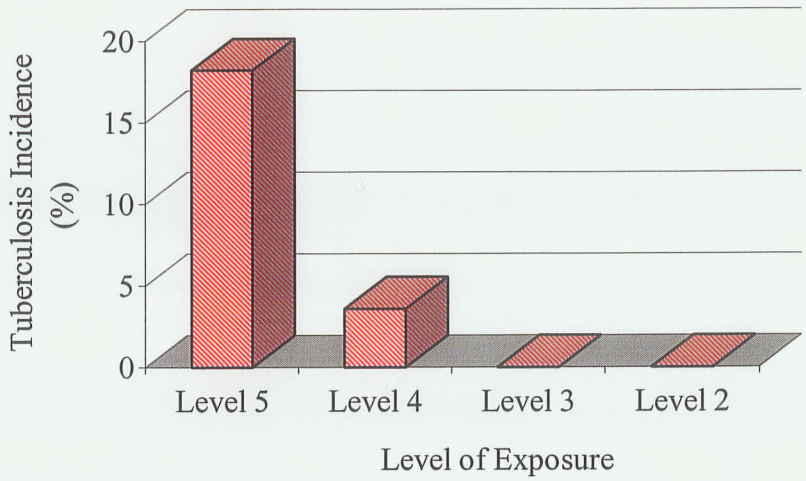
(90.91)

Figure 2: Discriminant analysis results and population subgroups of TB cases in MB



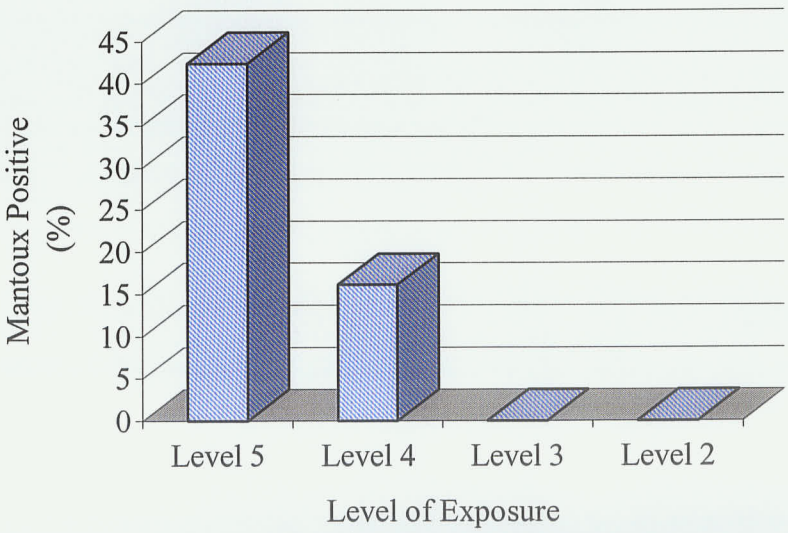
- Candian non treaty (Can)
  Candian treaty (Can_a)
  Forign Born (FB)

Figure 3: Trend of TB incidence by degree of exposure in Community 1



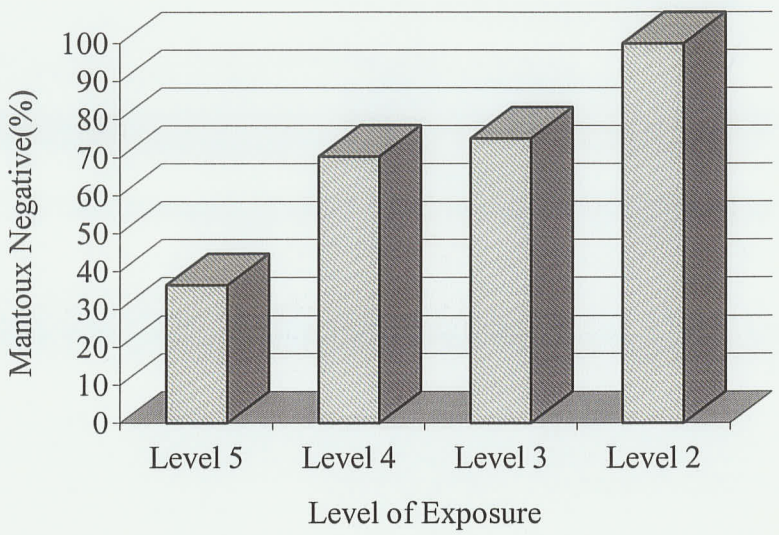
Trend Test: ($P < 0.01$)

Figure 4: Trend of Mantoux positivity (M+v) by degree of exposure in Community 1



Trend Test: ($P < 0.001$)

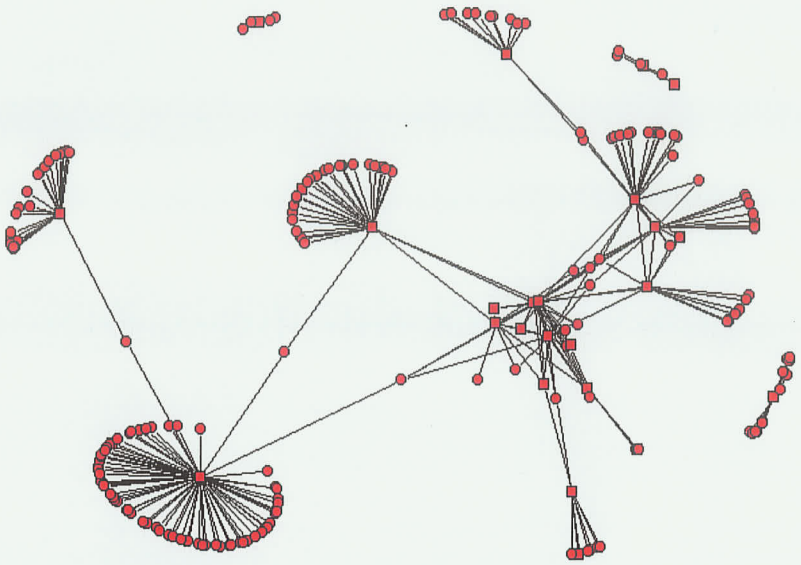
Figure 5: Trend of Mantoux negativity (M-v) by degree of exposure in Community 1



Trend Test: ($P < 0.001$)

Community 1: Number of patients = 16. Number of contacts = 238

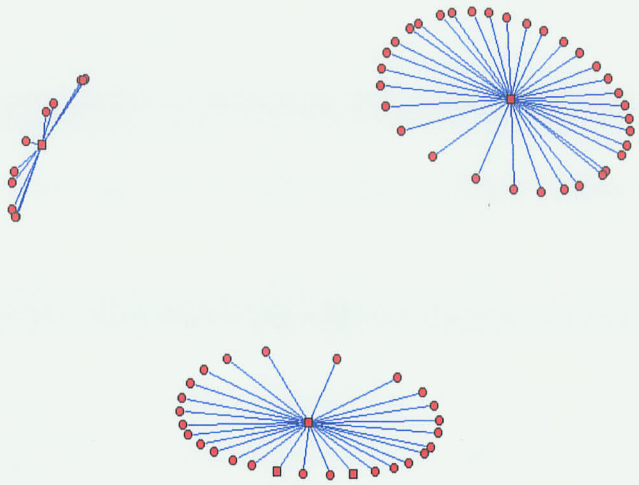
Figure 6: Community 1 TB outbreak network



Notes on Figure:

- Community 1 TB outbreak network 1992-1996 from contact tracing investigation data. The network is composed of TB index cases and their contacts (TB cases, Mantoux positive, and Mantoux negative).
- Boxes are TB cases and nodes with history of TB disease.
- Circles are contacts of TB index cases. The circles represent contacts with Mantoux skin test results (positive and negative).
- Lines represent relations or contacts between the nodes.
- The network is very compact with a clear core.

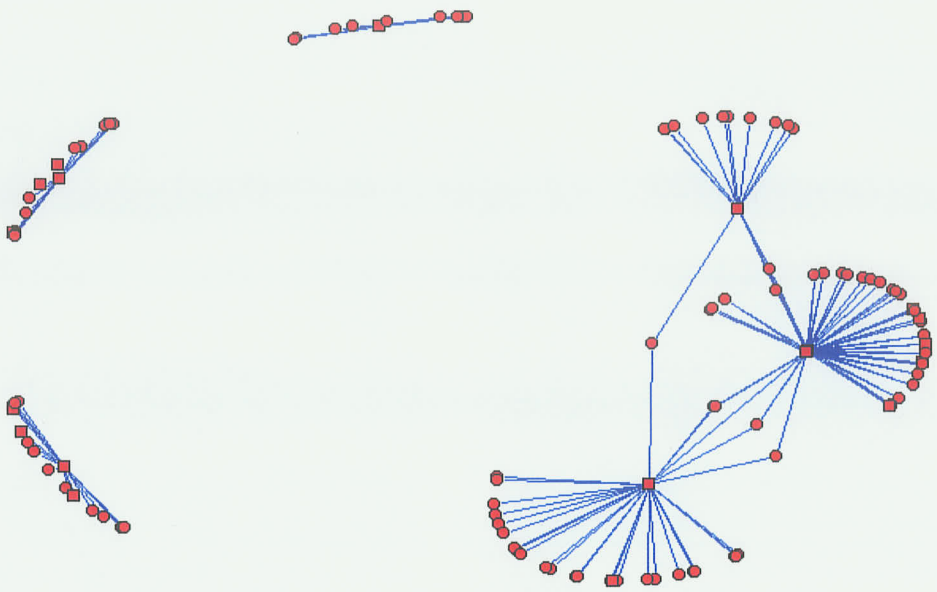
Figure 7: Community 2 TB outbreak network



Notes on Figure:

- Community 2 TB outbreak 1992-1996 from contact tracing investigation data. The network is composed of TB index cases and their contacts (TB cases, Mantoux positive, and Mantoux negative).
- Boxes are TB cases and nodes with history of TB disease.
- Circles are contacts of TB index cases. The circles represent contacts with Mantoux skin test results (positive and negative).
- Lines represent relations or contacts between the nodes.
- A network with three scattered components is observed.

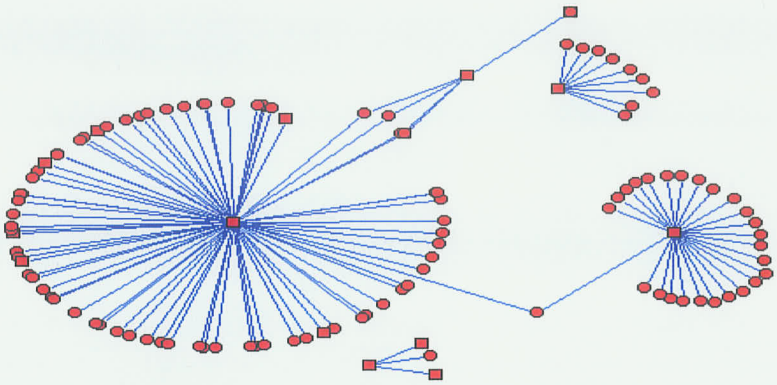
Figure 8: Community 6 TB outbreak network



Notes on Figure:

- Community 6 TB outbreak 1992-1996 from contact tracing investigation data. The network is composed of TB index cases and their contacts (TB cases, Mantoux positive, and Mantoux negative).
- Boxes are TB cases and nodes with history of TB disease.
- Circles are contacts of TB index cases. The circles represent contacts with Mantoux skin test results (positive and negative).
- Lines represent relations or contacts between the nodes.
- The network is composed of 4 components. One component is complex and larger than the remaining simple and scattered components.

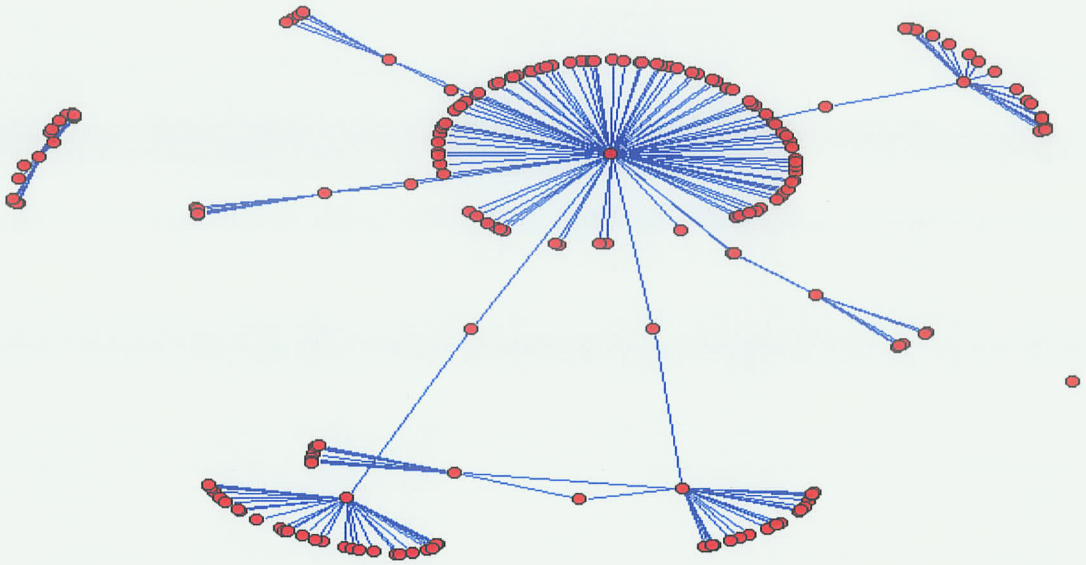
Figure 9: Community 9 TB outbreak network



Notes on Figure:

- Community 9 TB outbreak 1992-1996 from contact tracing investigation data. The network is composed of TB index cases and their contacts (TB cases, Mantoux positive, and Mantoux negative).
- Boxes are TB cases and nodes with history of TB disease.
- Circles are contacts of TB index cases. The circles represent contacts with Mantoux skin test results (positive and negative).
- Lines represent relations or contacts between the nodes.
- The network is composed of 3 components. One component is complex and larger than the remaining simple and scattered components.

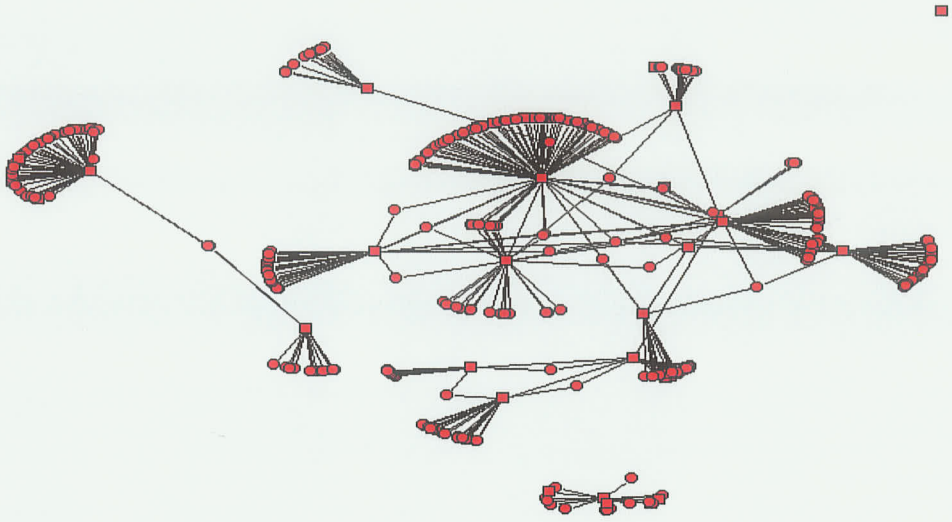
Figure 10: TB outbreak network of Communities 3 and 4 together



Notes on Figure:

- TB outbreak of Communities 3 and 4 together, 1992-1996, from contact tracing investigation data. The network is composed of TB index cases and their contacts (TB cases, Mantoux positive, and Mantoux negative).
- Boxes are TB cases and nodes with history of TB disease.
- Circles are contacts of TB index cases. The circles represent contacts with Mantoux skin test results (positive and negative).
- Lines represent relations or contacts between the nodes.

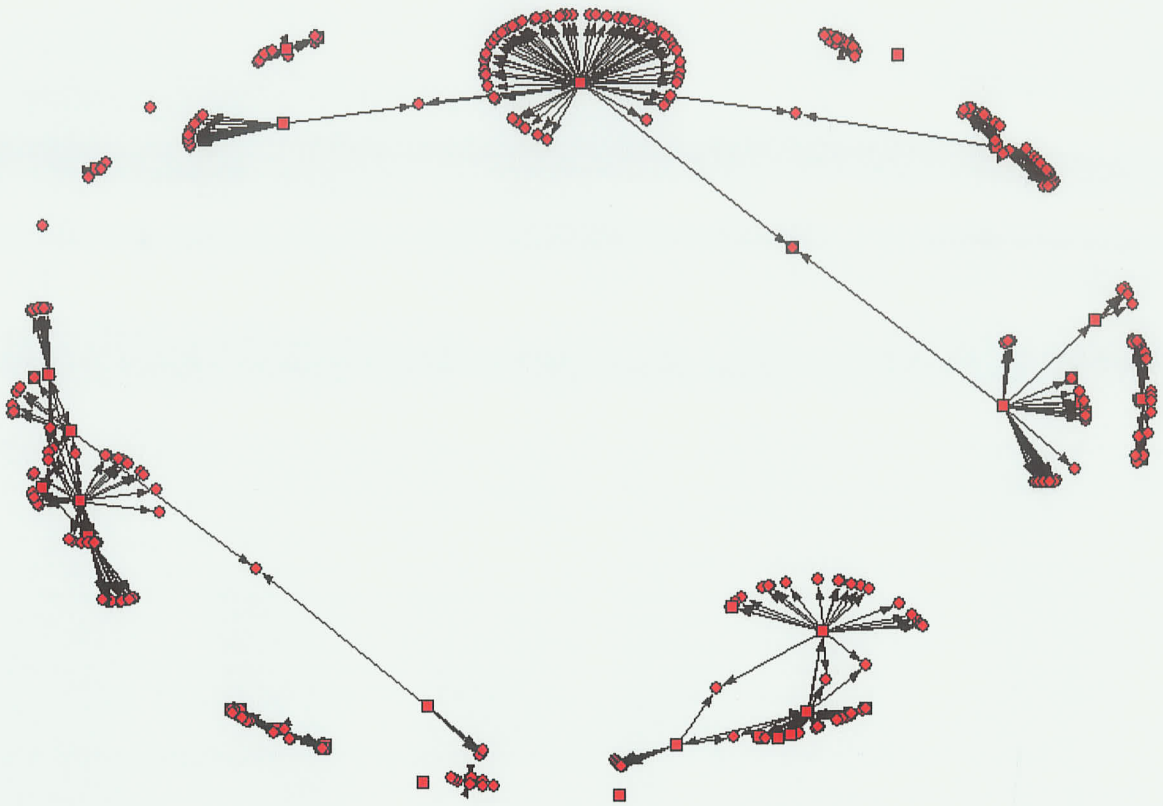
Figure 11: TB outbreak network of Communities 6 and 7 together



Notes on Figure:

- TB outbreak of Community 6 and 7 together, 1992-1996, from contact tracing investigation data. The network is composed of TB index cases and their contacts (TB cases, Mantoux positive, and Mantoux negative).
- Boxes are TB cases and nodes with history of TB disease.
- Circles are contacts of TB index cases. The circles represent contacts with Mantoux skin test results (positive and negative).
- Lines represent relations or contacts between the nodes.
- Communities 6 and 7 are 100 KM apart.

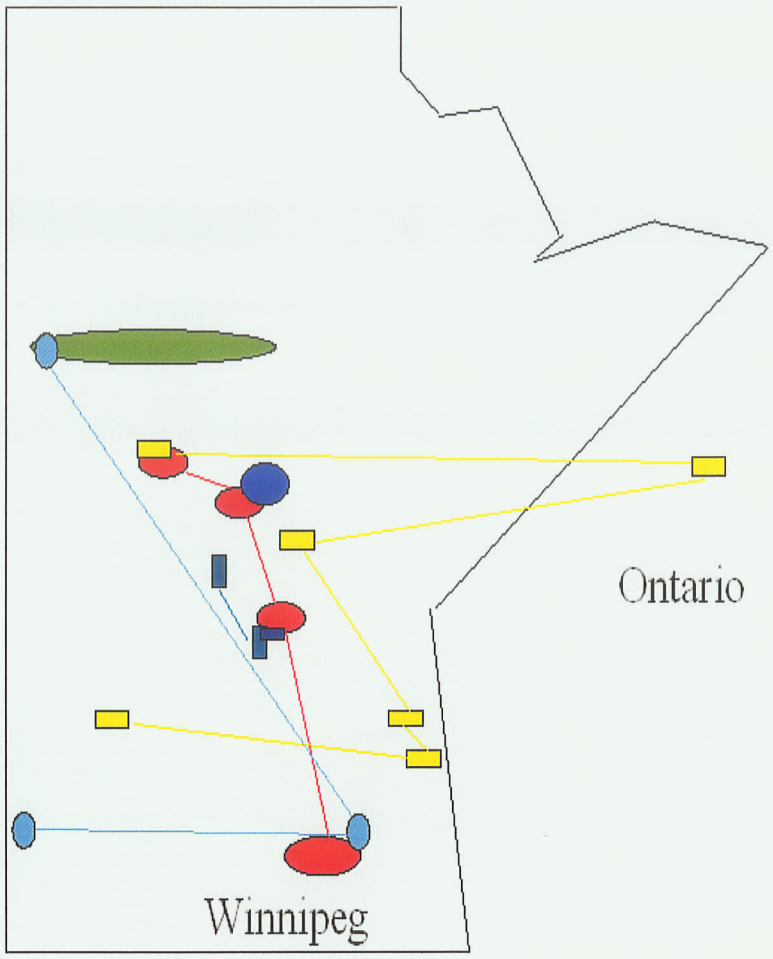
Figure 12: Network of FP5 TB strain in Manitoba's communities



Notes on Figure:

- Network of FP5 TB strain in Manitoba, 1992-1996, from contact tracing investigation data. The network is composed of TB index cases and their contacts (TB cases, Mantoux positive, and Mantoux negative).
- Boxes are TB cases and nodes with history of TB disease.
- Circles are contacts of TB index cases. The circles represent contacts with Mantoux skin test results (positive and negative).
- Lines represent relations or contacts between the nodes.
- The network is fragmented and extends to include scattered communities all over Manitoba and some just over the Manitoba-Ontario border

Figure 13: Schematic map FP5 TB strain in Manitoba's communities



Notes on Figure:

- Manitoba map identifying fingerprint type 5 (FP5) TB strain network components distribution 1992-1996.
- Each color and shape represents one connected component of FP5 TB strain network.
- Figure 13 is a schematic simplification of Figure 12. Communities' locations, component size, and nodes connections were taken in consideration.

Figure 14: TB incidence in Community 1, 1992-1999 (Cases per 100,000 by year)

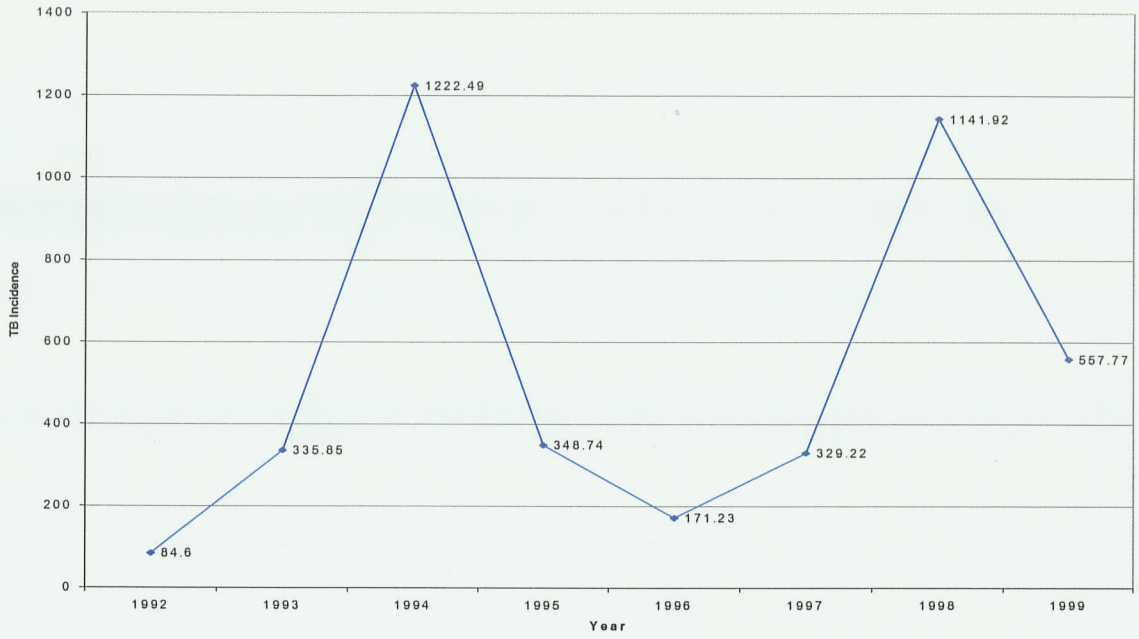
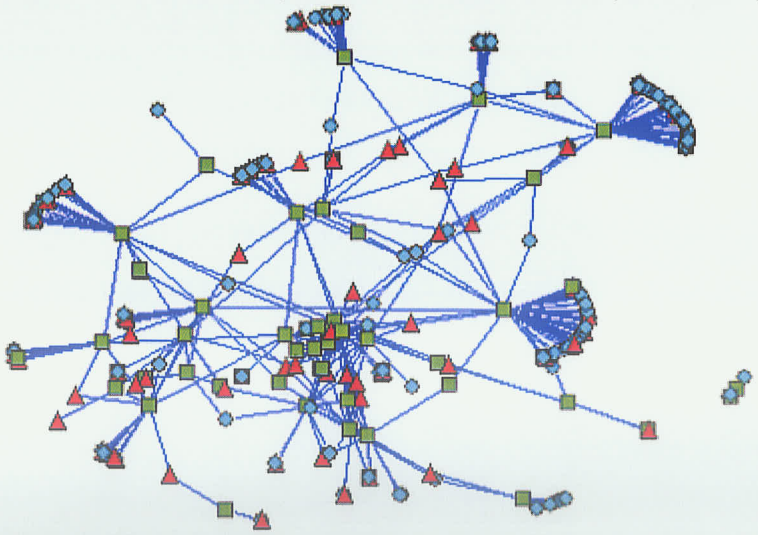


Figure 15: Community 1 TB outbreak network (1992-1999)

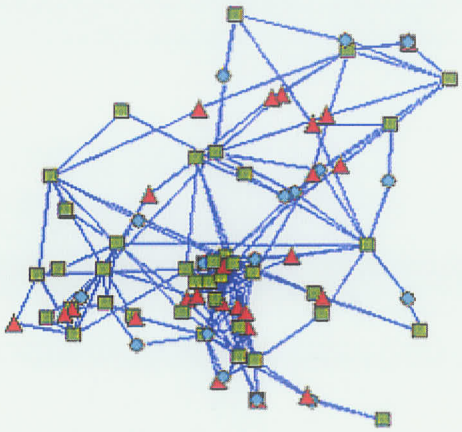


Notes on Figure:

Total number of nodes is 504. The network is composed of TB cases and contacts.

Green boxes are TB cases, red triangles are Mantoux positive contacts, and blue circles represent Mantoux negative contacts.

Figure 16: Community 1 TB outbreak degree two network (1992-1999)



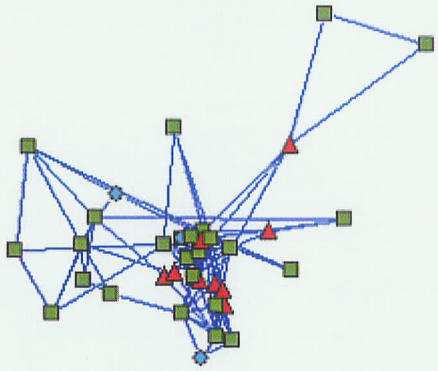
Notes on Figure:

Seidman *K*-core was used to collapse the network to degree 2.

Total number of nodes is 122. The network is composed of TB cases and contacts.

Green boxes are TB cases, red triangles are Mantoux positive contacts, and blue circles represent Mantoux negative contacts.

Figure 17: Community 1 TB outbreak degree three network (1992-1999)



Notes on Figure:

Seidman *K*-core was used to collapse the network to degree 3.

Total number of nodes is 45.

Green boxes are TB cases, red triangles are Mantoux positive contacts, and blue circles represent Mantoux negative contacts.

Figure 18: Community 1 TB outbreak degree four network (1992-1999)

Notes on Figure:

Seidman *K*-core was used to collapse the network to degree 4.

Total number of nodes is 22.

Green boxes are TB cases; red triangles are Mantoux positive contacts. There are no Mantoux negative contacts.

Figure 19: Community 1 TB outbreak degree five network (1992-1999)

Notes on Figure:

Seidman *K*-core was used to collapse the network to degree 5.

Total number of nodes is 18.

Green boxes are TB cases; red triangles are Mantoux positive contacts. There are no Mantoux negative contacts.

Figure 20: Community 1 TB outbreak degree six network (1992-1999)

Notes on Figure:

Seidman *K*-core was used to collapse the network to degree 6.

Total number of nodes is 10.

Green boxes are TB cases; red triangles are Mantoux positive contacts. There are no Mantoux negative contacts.

Figure 21: Community 1 TB outbreak degree seven network (1992-1999)

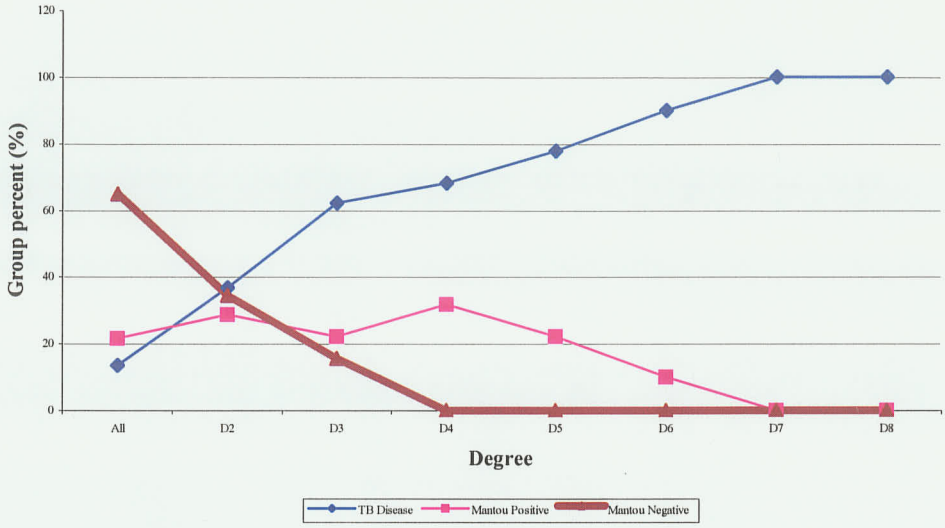
Notes on Figure:

Seidman *K*-core was used to collapse the network to degree 7.

Total number of nodes is 7.

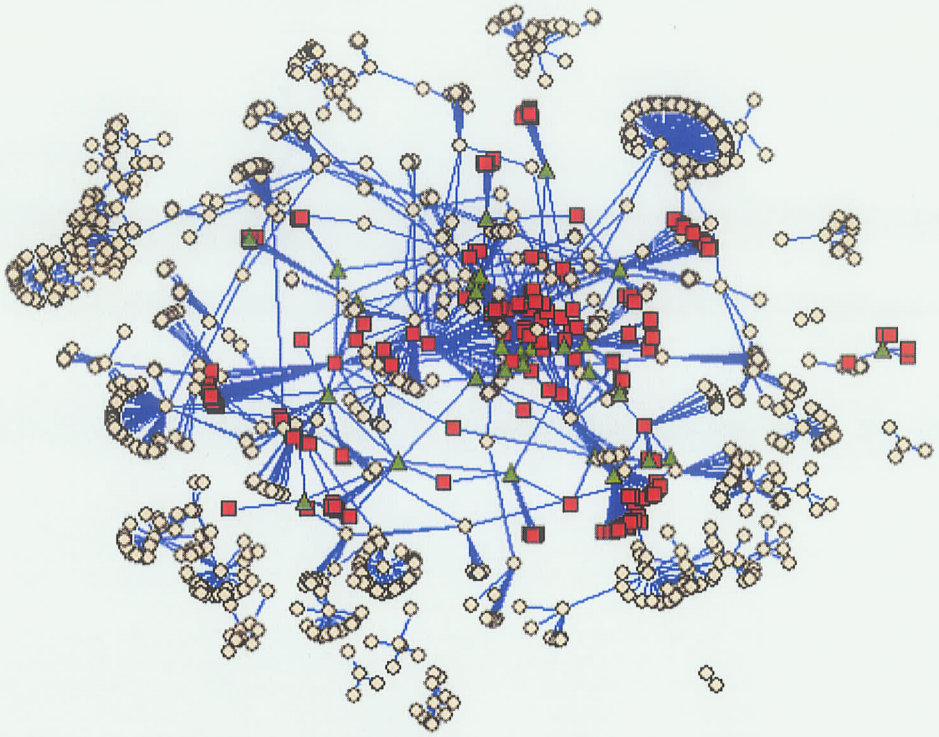
Green boxes are TB cases. There are no TB contacts.

Figure 22: Community 1 TB outbreak network core collapse sequence graph



Notes on Figure: Illustrates change in network composition with core collapse sequence of Community1 TB outbreak network until disintegration. The collapse is consistent with the degree of contact between TB cases and their contacts.

Figure 23: FP1 TB strain network in Manitoba



Notes on Figure:

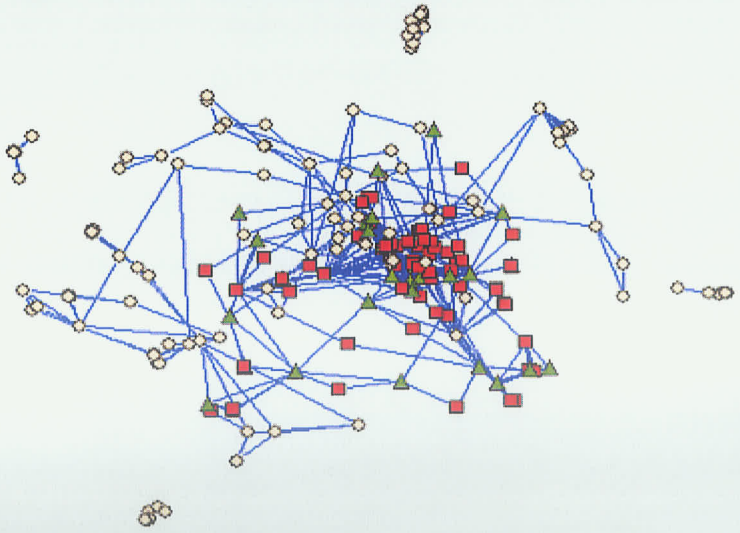
FP1 network includes 1795 total nodes, and includes Community 1 members.

Green triangles are TB cases with FP1 from Community 1.

Red boxes are contacts of Community 1 TB cases.

Pink diamonds are the remaining members of FP1 TB network.

Figure 24: FP1 TB strain network in Manitoba, degree two



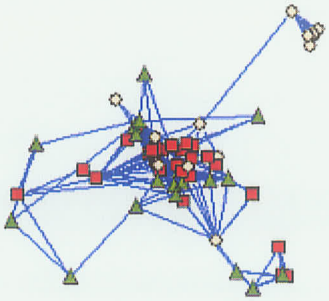
Notes on Figure:

Seidman *K*-core was used to collapse the network to degree 2

Total number of nodes is 303.

Green triangles are TB cases with FP1 from Community 1; red boxes are contacts of Community 1 TB cases; pink diamonds are the remaining members of the network.

Figure 25: FP1 TB strain network in Manitoba, degree three



Notes on Figure:

Seidman *K*-core was used to collapse the network to degree 3.

Total number of nodes is 79.

Green triangles are TB cases with FP1 from Community 1; red boxes are contacts of Community 1 TB cases; pink diamonds are the remaining members of the network.

Figure 26: FP1 TB strain network in Manitoba, degree four

Notes on Figure:

Seidman *K*-core was used to collapse the network to degree 4.

Total number of nodes is 43.

Green triangles are TB cases with FP1 from Community 1; red boxes are contacts of Community 1 TB cases; pink diamonds are the remaining members of the network.

Figure 27: FP1 TB strain network in Manitoba, degree five

Notes on Figure:

Seidman *K*-core was used to collapse the network to degree 5.

Total number of nodes is 28.

Green triangles are TB cases with FP1 from Community 1; red boxes are contacts of Community 1 TB cases; pink diamonds are the remaining members of the network.

Figure 28: FP1 TB strain network in Manitoba, degree six

Notes on Figure:

Seidman *K*-core was used to collapse the network to degree 6.

Total number of nodes is 22.

Green triangles are TB cases with FP1 from Community 1; red boxes are contacts of Community 1 TB cases.

Figure 29: FP1 TB strain in Manitoba, degree seven

Notes on Figure:

Seidman *K*-core was used to collapse the network to degree 7.

Green triangles are TB cases with FP1 from Community 1; red boxes are contacts of Community 1 TB cases.

Figure 30: FP1 TB strain network in Manitoba, degree eight

Notes on Figure:

Seidman *K*-core was used to collapse the network to degree 8.

Green triangles are TB cases with FP1 from Community 1; red boxes are contacts of Community 1 TB cases.

Figure 31: FP1 TB strain network in Manitoba, degree nine

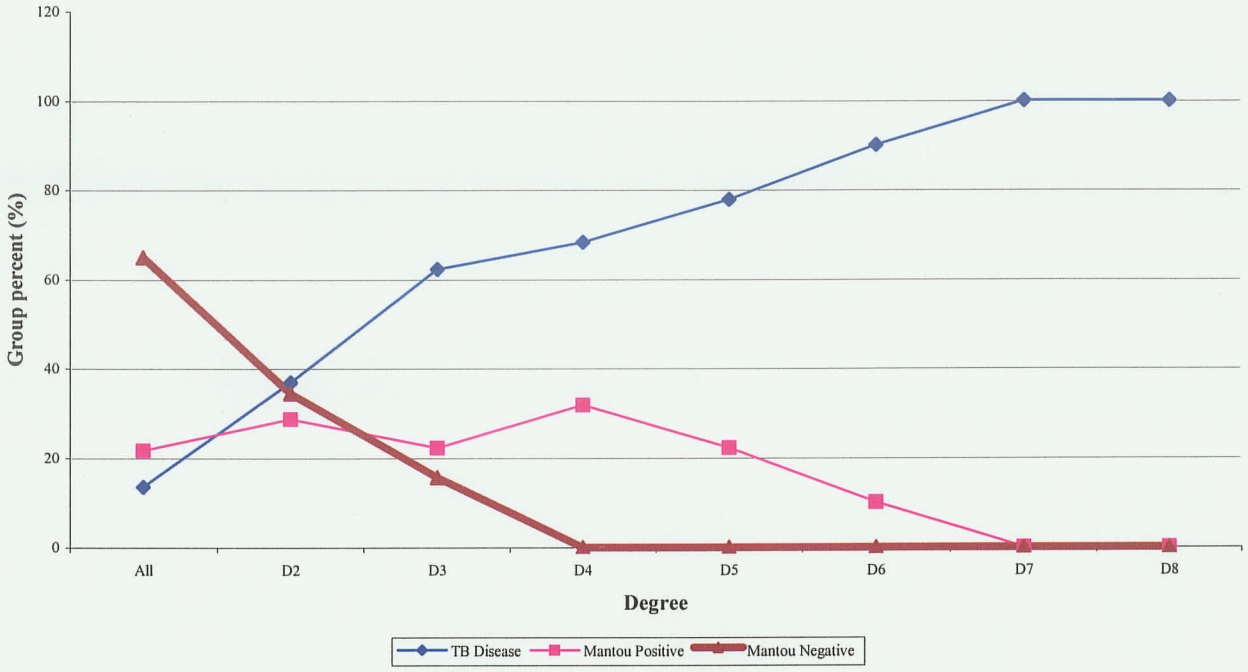
Notes on Figure:

Seidman *K*-core was used to collapse the network to degree 9.

Total number of nodes is 3. All nodes are TB cases.

Green triangles are TB cases with FP1 from Community 1.

Figure 32: Core collapse sequence graph of FP1 TB strain network in Manitoba

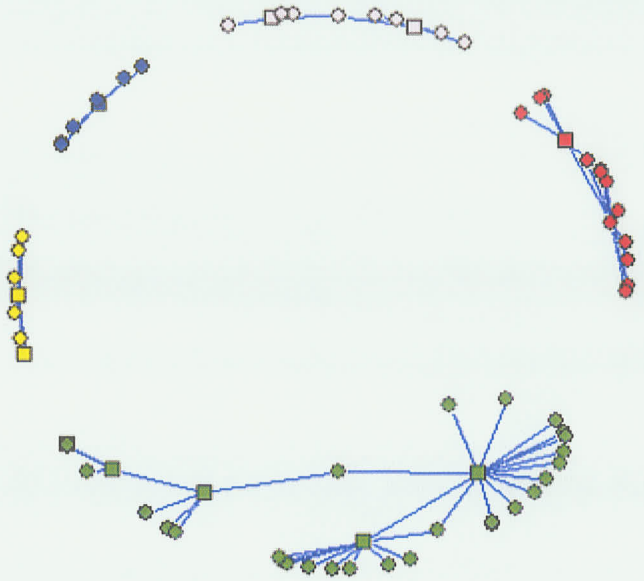


Notes on Figure:

Graph illustrates change in network composition with use of Seidman *K*-core to collapse the network until its disintegration.

The collapse is consistent with the degree of contact between TB cases and their contacts.

Figure 33: Community 1 TB group network components (people only)

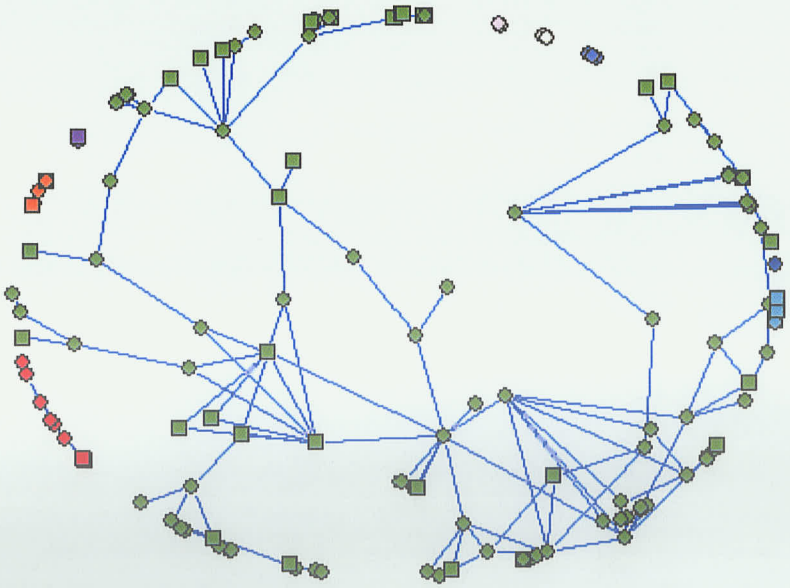


Notes on Figure:

The network is composed of TB cases and alters. TB cases or index nodes are boxes; contacts are circles. There are 88 nodes in total.

Five components are noted in network. Each component has a different color.

Figure 34: Community 1 contact group network components (people only)

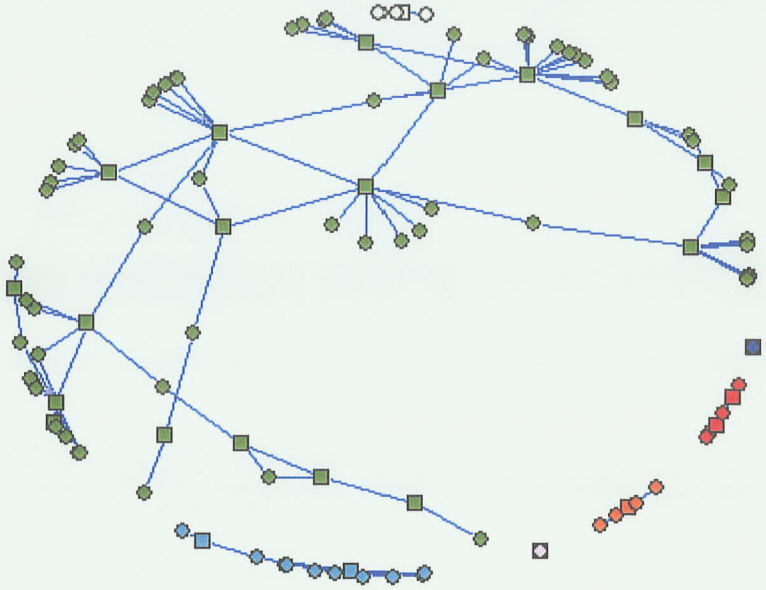


Notes on Figure:

The network is composed of contacts and alters. TB contacts are depicted as boxes; their contacts are circles. There are 198 nodes in total.

Nine components are observed. Each component has different color.

Figure 35: Community 1 control group network components (people only)

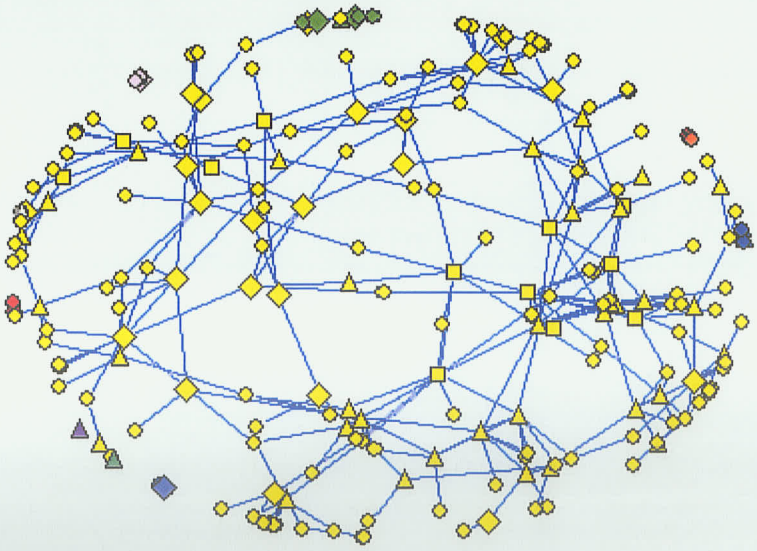


Notes on Figure:

The network is composed of control nodes and alters. Control nodes are depicted as boxes, and their contacts (alters) as circles. There are 170 nodes in all.

Seven network components are observed. Each component has different color.

Figure 36: Community1 network components (TB, contact, and control groups & alters)



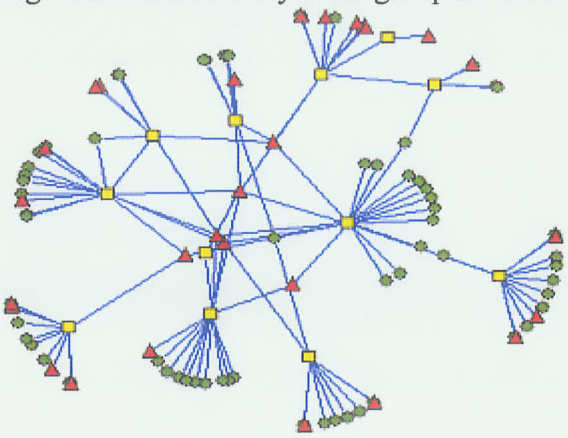
Notes on Figure:

The network is composed of index nodes and alters for all 3 groups:

- TB group (boxes)
- Control group (diamonds)
- Contact group (triangles)
- Alters of all groups (circles)

There are 391 nodes in all. Ten network components are observed. Each component has a different color.

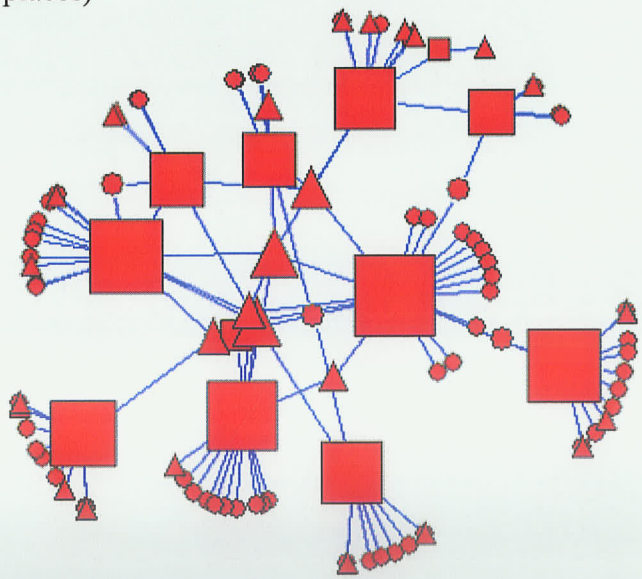
Figure 37: Community 1 TB group network (people and places)



Notes on Figure:

TB cases are depicted as boxes, contacts of the index cases as circles, and the places most frequented by the index cases are triangles.

Figure 38: Degree Centrality Graph of Community 1 TB group network (people and places)

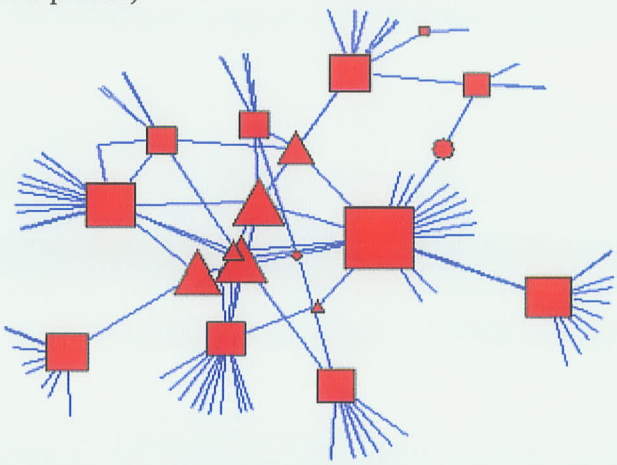


Notes on Figure:

Boxes are people and triangles are places.

Degree centrality is reflected by the size of the nodes, being directly proportional to the nodes' importance in the network.

Figure 39: Betweenness Centrality Graph of Community 1 TB group network (people and places)

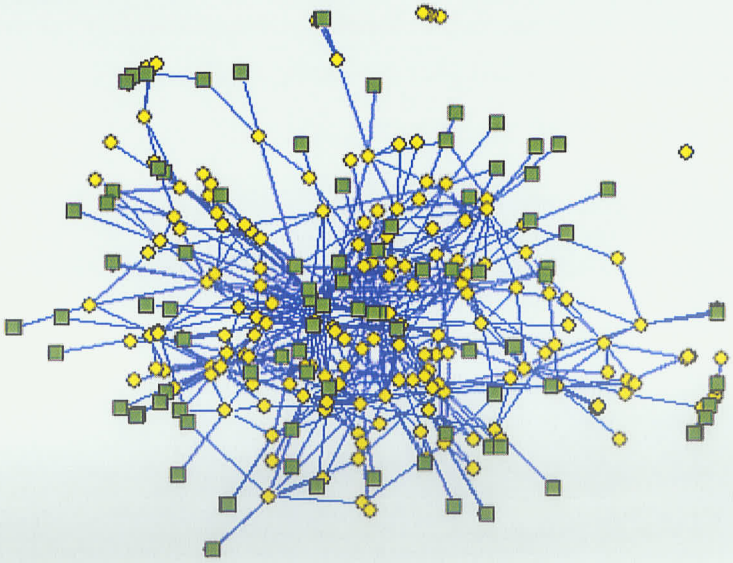


Notes on Figure:

Boxes are people and triangles are places.

Betweenness centrality is reflected by the size of nodes, which is directly proportional to the nodes' importance in the network.

Figure 40: Community 1 network: people (TB cases, contacts, controls, and alters) and places

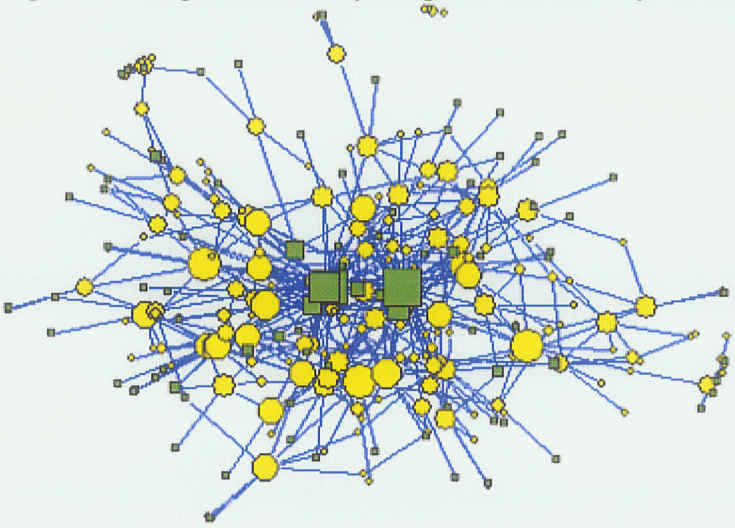


Notes on Figure:

Boxes (green) are places and circles (yellow) are people.

People include TB cases, contacts, controls and alters together.

Figure 41: Degree Centrality Graph of Community1 network (people and places)

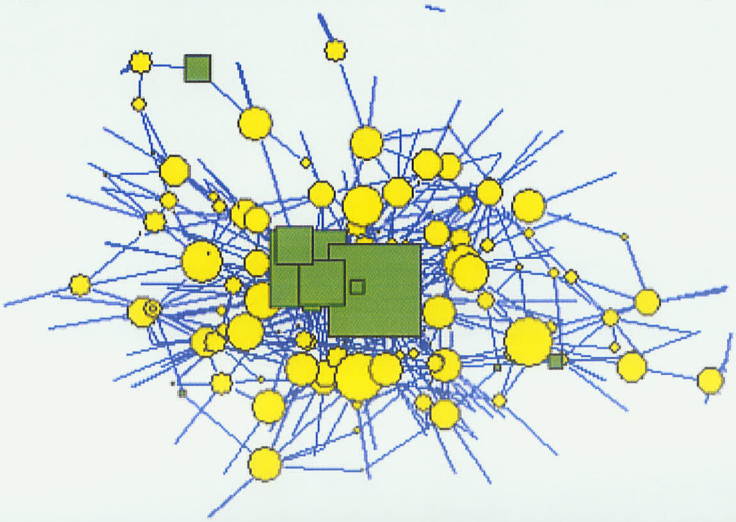


Notes on Figure:

Boxes are places (green) and circles (yellow) are people.

Degree centrality is reflected by the size of nodes, which is directly proportional to the nodes' importance in the network.

Figure 42: Betweenness Centrality Graph of Community 1 network (people and places)

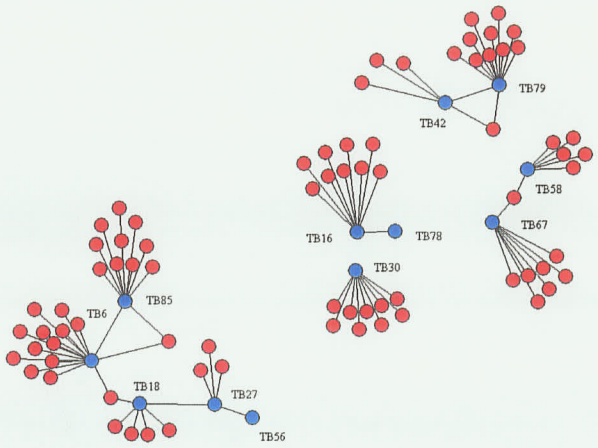


Notes on Figure:

Boxes are places (green), and circles (yellow) are people.

Betweenness centrality is reflected by the size of the nodes, which is directly proportional to the nodes' importance in the network.

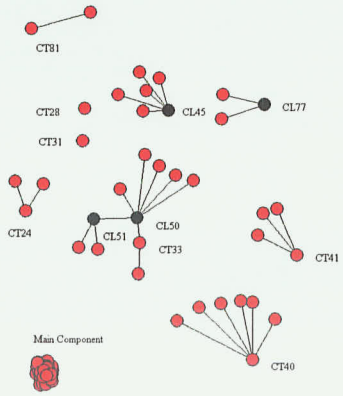
Figure 43: Community 1 TB group network, showing alters exposed to two TB cases



Note on Figure:

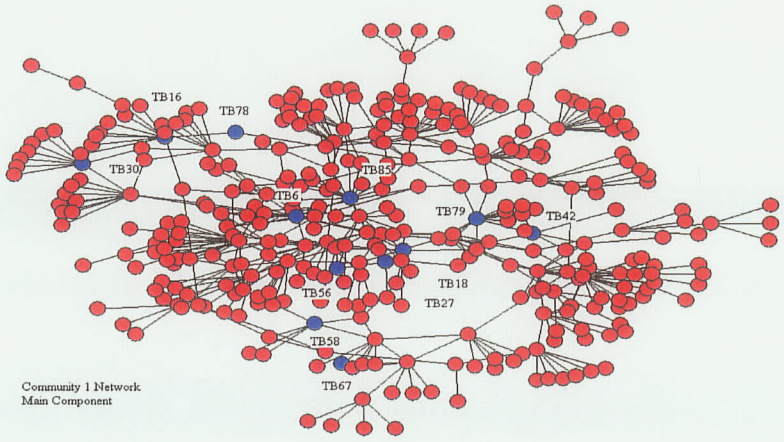
TB group network showing TB cases (blue) and altars (red). Some alters have been exposed to two TB cases; such individuals are high risk for TB infection.

Figure 44: Community 1 Network Components



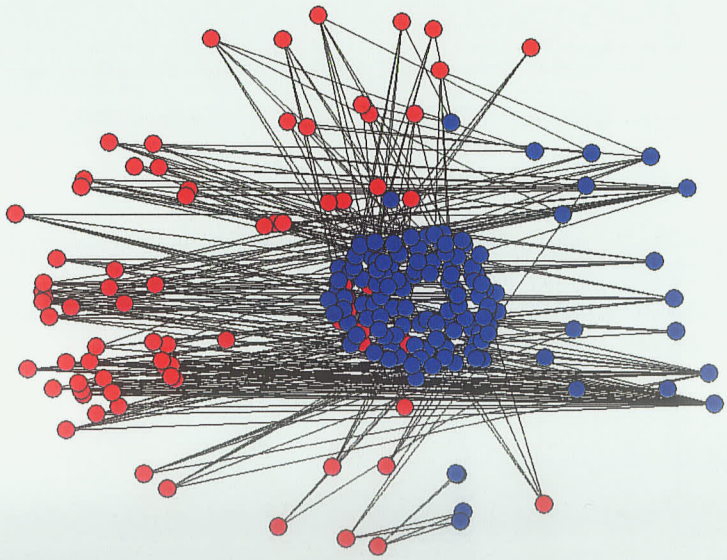
Notes on Figure:
The Community 1 network has 10 components. The main component (in lump) is composed of 353 nodes. The remaining 9 components are depicted in the Figure. CT represents contact group members, CL represents control group members, and non-identified nodes are alters.

Figure 45: Main component of Community 1 network.



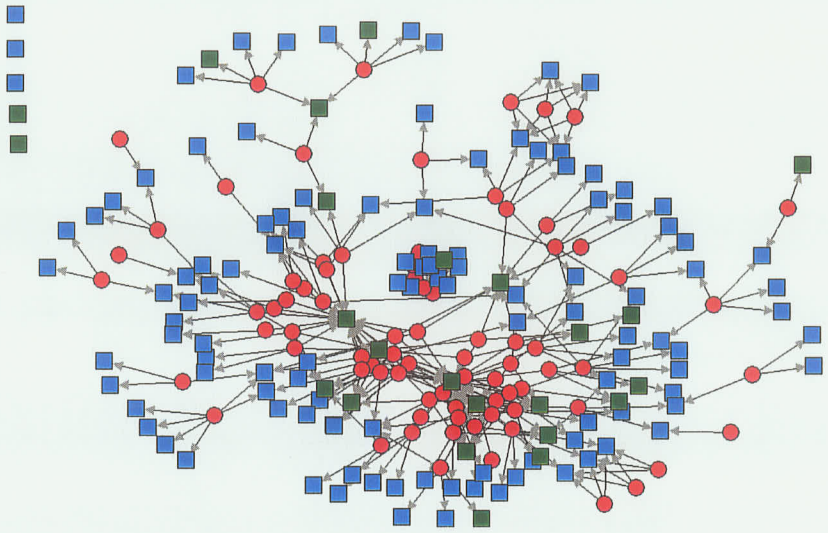
Notes on Figure:
The main component of the Community 1 network consists of 353 nodes. TB index cases are in blue, and the remaining nodes are red.

Figure 46: Visual representation of Community 1 in two-mode network using MSD of Geodesic distance.



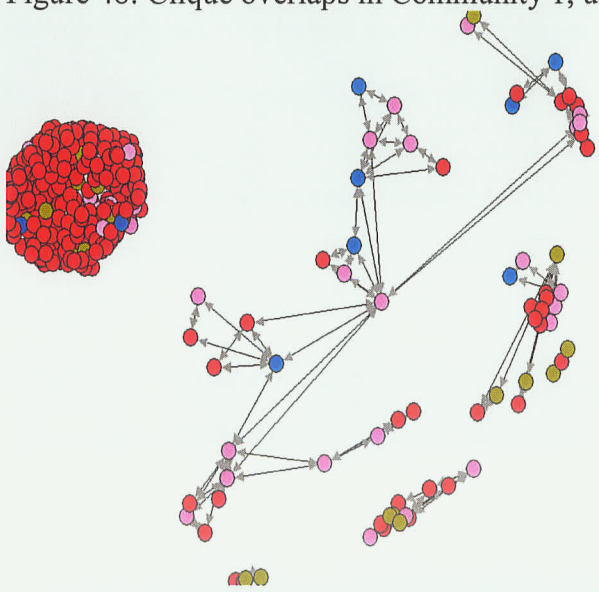
Notes on Figure:
Red circles represent people and blue circles represent places frequently attended by index nodes in Community 1 network.

Figure 47: Community 1 network plotted in two-mode graph (people and places)



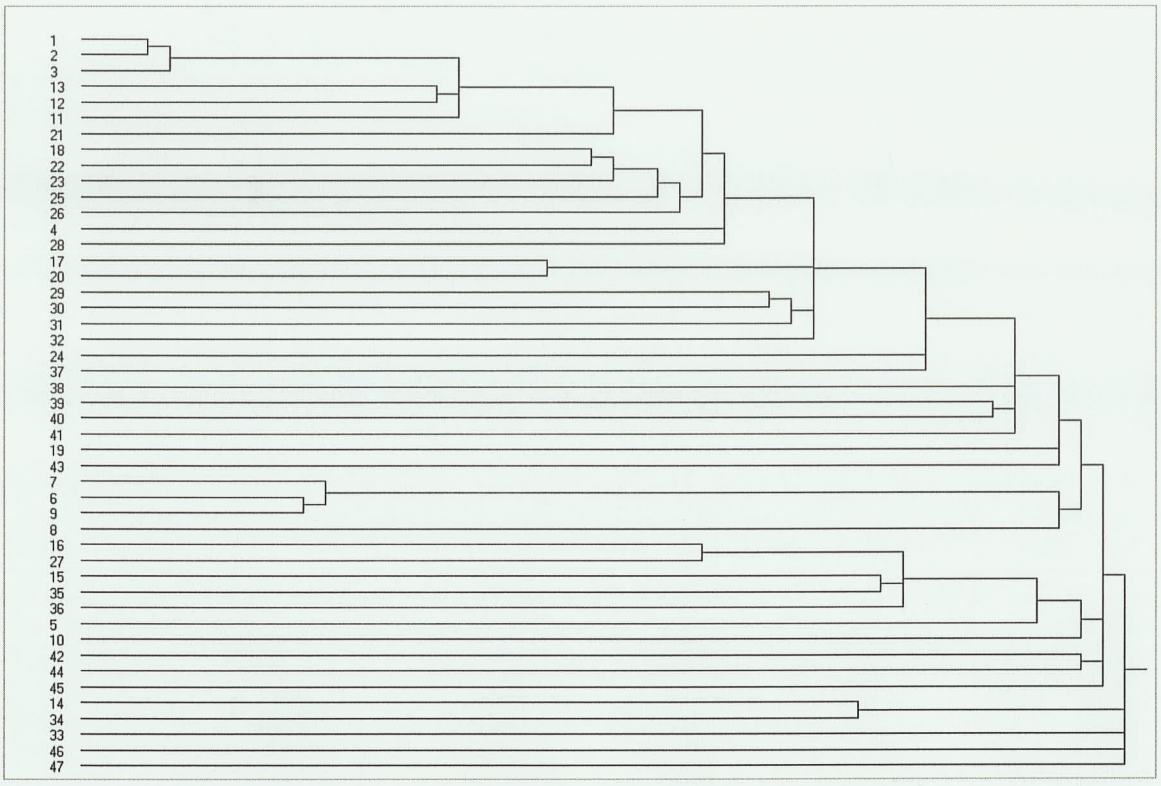
Notes on Figure:
Blue boxes represent private places, green boxes represent public places, and red circles represent people.

Figure 48: Clique overlaps in Community 1, using MSD



Notes on Figure:
Blue circles represent TB cases, pink circles contacts, green circles controls, and red circles alters.
TB6 and TB18 are very active nodes; they participate in several cliques each.
TB6 (found at cliques: 5, 42, 44, 45) and TB18 (found at cliques: 11, 12, 15, 16) were members in four different cliques each.

Figure 49: Dendrogram of hierarchical clustering of the clique overlaps in the Community 1 network



Notes on Figure:
Clique overlaps in Community 1 network. Seventeen cliques have at least one member from the TB group. Four cliques (15, 37, 42 and 45) have two members from the TB group.

Figure 50: Discriminant analysis of Community 1 Case Control Study Groups

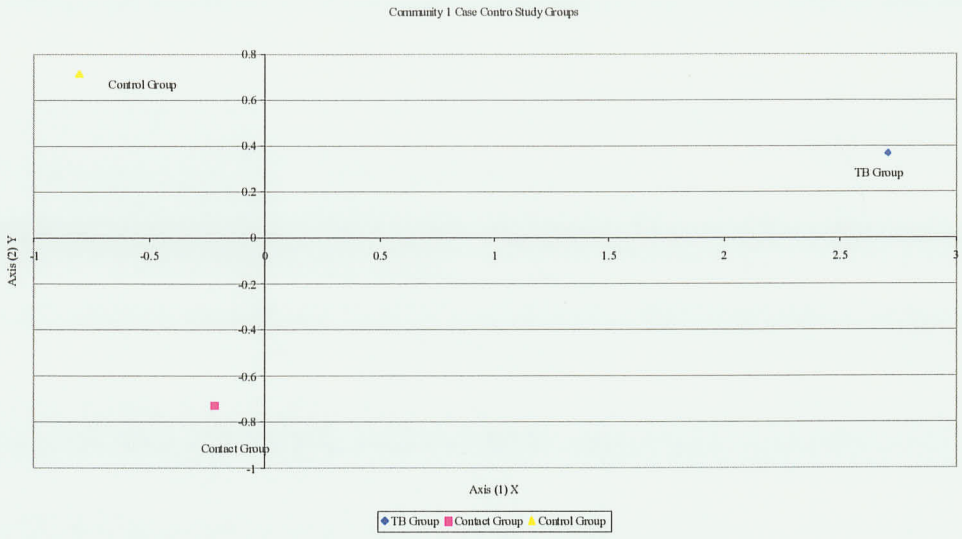


Figure 51: Cumulative and degree distribution of the Community 1 network main component on normal axes

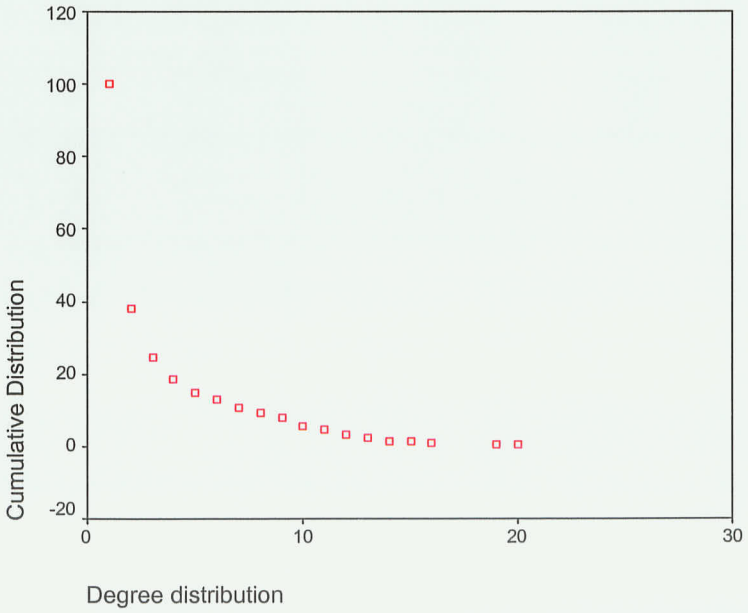


Figure 52: Cumulative and degree distribution of the Community 1 network main component on log -log axes

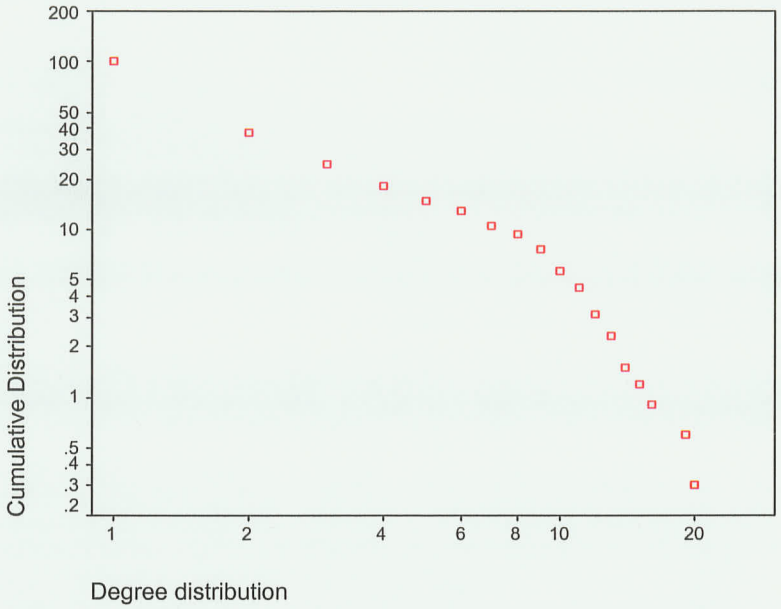


Figure 53: Community1 network main component, cumulative and degree distribution regression results of curve estimation with linear, power, and exponential distribution

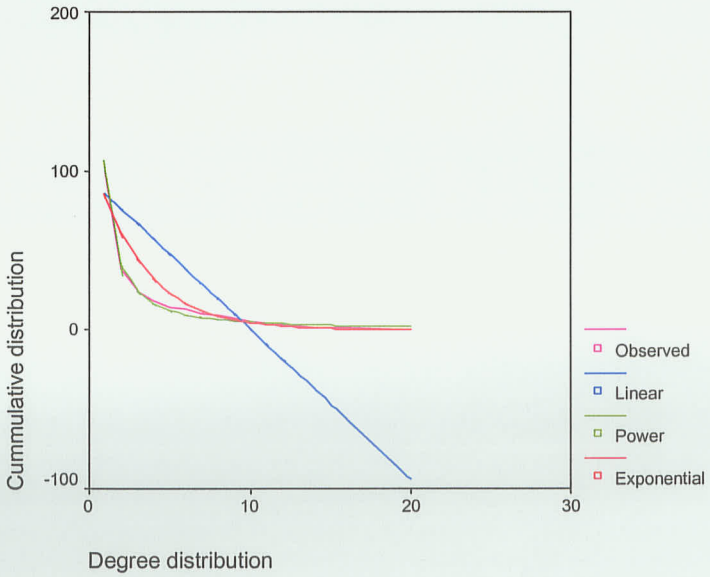


Figure 54: Community 1 network main component, cumulative and degree distribution regression results of curve estimation with power and exponential distribution on log-log axes

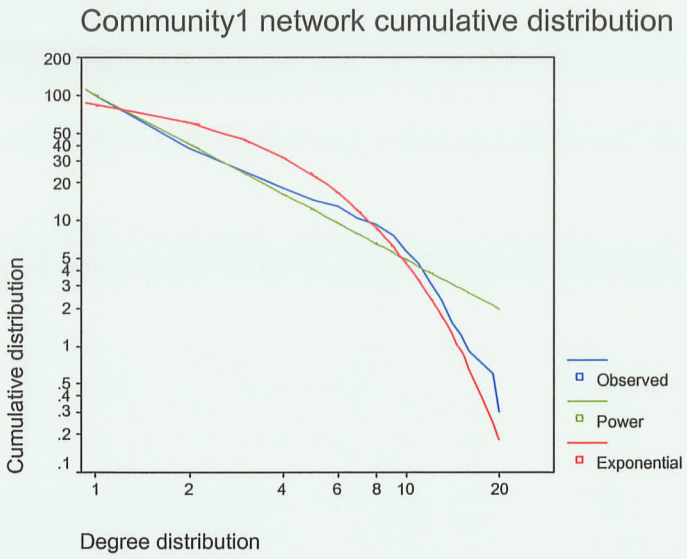


Figure 55: Community 1 network main component, cumulative and degree distribution regression results of curve estimation (body of the curve) with power and exponential distribution on log-log axes

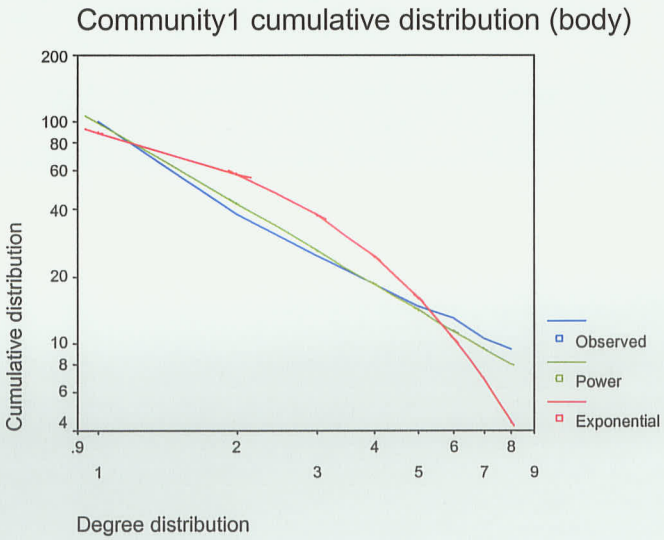


Figure 56: Community 1 network main component, cumulative and degree distribution regression results of curve estimation (tail of the curve) with power and exponential distribution on log-log axes

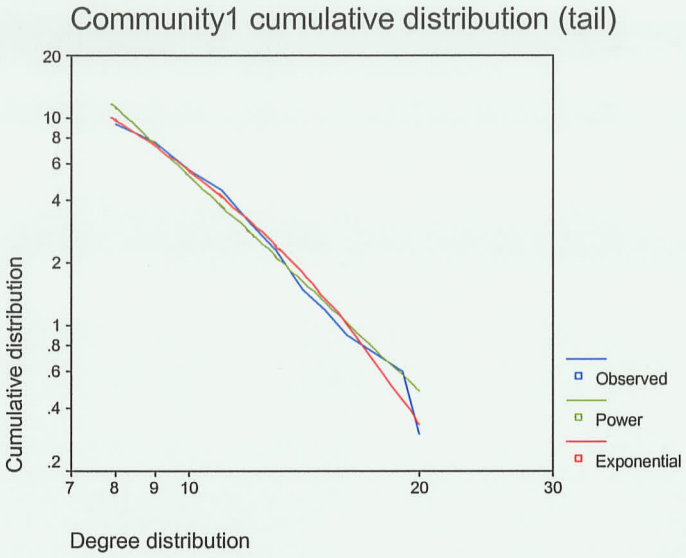


Table 1: Proportion of Canadian born and Foreign-born (FB) TB cases in Manitoba 1992-1999

Year	Canadian[1] (n)	(%)	Foreign[2] (n)	(%)	Total (n)
1992	58	67.4	28	32.6	86
1993	87	80.6	21	19.4	108
1994	85	73.3	31	26.7	116
1995	80	74.1	28	25.9	108
1996	64	65.98	33	34.02	97
1997	66	68.8	30	31.2	96
1998	71	62.3	43	37.1	114
1999	94	71.8	37	28.2	131
Total	604	70.6	251	29.4	855

[1] Canadian-born [2] Foreign-born (FB)

Table 2: Country of origin of FB TB cases in Manitoba between 1992-1999

Country	(n)	(%)	Country	(n)	(%)	Country	(n)	(%)
Philippine	112	44.6	**Cambodia**	2	0.8	**Indonesia**	1	0.4
Vietnam	24	9.6	**Italy**	2	0.8	**Korea**	1	0.4
India	18	7.2	**Jamaica**	2	0.8	**Malaysia**	1	0.4
China	13	5.2	**Romania**	2	0.8	**Malta**	1	0.4
Laos	9	3.6	**Saudi Arabia**	2	0.8	**Mozambique**	1	0.4
Hong Kong	8	3.2	**Trinidad**	2	0.8	**Nicaragua**	1	0.4
Ethiopia	7	2.8	**Zambia**	2	0.8	**Somalia**	1	0.4
Poland	6	2.4	**Afghanistan**	1	0.4	**Sudan**	1	0.4
Africa	4	1.6	**Austria**	1	0.4	**Sweden**	1	0.4
Germany	4	1.6	**Bosnia**	1	0.4	**Taiwan**	1	0.4
Kenya	3	1.2	**El Salvador**	1	0.4	**Yugoslavia**	1	0.4
Pakistan	3	1.2	**Chile**	1	0.4	**Zaire**	1	0.4
Russia	3	1.2	**Guatemala**	1	0.4	**Unknown**	1	0.4
Ukraine	3	1.2	**Haiti**	1	0.4	**Total**	251	100

Table 3: TB status by level of exposure, Community 1, 1992-1996

	Level 5		Level 4		Level 3		Level 2		Level 1		Total
	(n)	(%)	(n)	(%)	(n)	(%)	(n)	(%)	(n)	(%)	(n)
Tuberculosis	6	18.2	6	3.6	0	0	0	0	1	4	13
Mantoux (positive)	14	42.4	27	16.1	0	0	0	0	5	20	46
Mantoux (Negative)	12	36.4	118	70.2	6	75	4	100	16	64	156
Unknown	1	3	17	10.1	2	25	0	0	3	12	23
Total	33	100	168	100	8	100	4	100	25	100	238

Definitions: Level 5: exposed to 2 TB cases or more; Level 4: high level of exposure to one case (live together);Level3: medium level of exposure (3 times/week); Level 2: low level of exposure (once/week). Level 1: lack of information

Table 4 Cliques in Community 1 TB outbreaks

Cliques	Clique Members	Cliques	Clique Members	Cliques	Clique Members
1	2, 4, 8, 9, *22**	***2***	2, 4, 9, *84**	***3***	2, 4, 9, *87**
4	2, 4, 9, *90**	***5***	2, 4, 9, *93**	***6***	2, 4, *24**
7	2, 4, *29**	***8***	2, 4, 8, *30**	***9***	2, 4, 8, *36**
10	2, 4, *52**	***11***	2, 4, *12**	***12***	2, 4, *58**
13	2, 4, *94**	***14***	2, 6, 8, 9, *22**	***15***	2, 6, 8, 17
16	2, 6, 8, *30**	***17***	2, 6, 8, *36**	***18***	2, 6, 8, *37**
19	2,6,8,11	***20***	2, 4, *88**	***21***	4, 9, *252**
22	6,7,8,11,14	***23***	6,7,8,11, *18**	***24***	6,7,8,11,20
25	6,7,8,11,32	***26***	6,7,8,11, *33**	***27***	6, 7, 8, 9
28	6, 7, 8, *21**	***29***	6, 7, 8, *27**	***30***	6, 7, 8, *30**
31	6, 7, 8, *36**	***32***	6, 7, 8, *37**	***33***	6, 7, 8, 14, *38**
34	6, 8, 11, 17	***35***	6, 8, *39**	***36***	6,8,14,49
37	7, 8, *45**	***38***	7, 11, *19**		

(*) Tuberculosis disease-free clique members

Table 5. TB status profile of nine selected rural communities of Manitoba

Community	TB Cases	Age (Mean)	FP Types	No Isolates	Treaty	Not Treaty	FB	Incidence Cases/ 100,000 person-year
Community1	51	34	6	9	50	1	0	334.2
Community2	8	54	6	2	6	2	0	19.9
Community3	10	36	2	1	10	0	0	117.3
Community4	9	43	3	2	9	0	0	223.1
Community5	15	25	3	7	15	0	0	47.6
Community6	14	26	3	4	14	0	0	67.1
Community7	6	44	2	0	6	0	0	0
Community8	14	28	2	8	14	0	0	182.9
Community9	32	20	3	20	32	0	0	496.3

Notes and definitions for Table:

*TB cases = number of TB cases in the community between 1992-1999, e.g. community1 had 51 TB cases between 1992-1999.

*FP Types = number of different molecular RFLP TB strains or fingerprint types, e.g. Community 1 TB cases infected by 9 different TB strains.

*No isolates = number of TB cases without isolate, e.g. among Community 1 TB cases, there were 9 cases which did not have isolates for culture and molecular typing.

*Treaty = number of TB cases with treaty status, e.g. Community 1 had 50 out of 51 TB cases with treaty status.

*Not Treaty = number of TB cases without treaty status, e.g. Community 1 had one out of 51 TB cases without treaty status.

*FB = number of TB cases with Foreign born status, e.g. Community 1did not have any TB cases with foreign-born status.

*Incidence (case per 100,000 person-year)= TB incidence at the community e.g. TB incidence in Community 1 between 1992-1999 was 334 per 100,000 person-year.

Table 6: Comparison of TB incidence in 8 selected rural communities of Manitoba with overall TB incidence of Treaty, Manitoba, and Canadian people.

Community (Reserves)	Incidence per/100,000	Ratio of Community Incidence To Treaty's Incidence[1]	Ratio of Community Incidence To Manitoba's Incidence[2]	Ratio of Community Incidence To Canada's Incidence[3]
Community1	334.2	6.9	36.3	56.6
Community2	19.9	0.4	2.2	3.4
Community3	117.3	2.4	12.8	19.9
Community4	223.1	4.6	24.3	37.8
Community5	47.6	0.9	5.2	8.1
Community6	67.1	1.4	7.3	11.4
Community8	182.9	3.8	19.9	31.0
Community9	496.3	10.3	53.9	84.1

[1] Treaty people TB incidence = 48.4 per 100,000 person-year. [2] Manitoba TB incidence= 9.2 per 100,000

Table 7: Number of TB cases, population and TB incidence, Community 11992-1999

Year	TB cases	Resident	TB Incidence per 100,000
1992	1	1182	84.6
1993	4	1191	335.9
1994	15	1227	1222.5
1995	4	1147	348.7
1996	2	1168	171.2
1997	4	1215	329.2
1998	14	1226	1141.9
1999	7	1255	557.8
1992 – 1999	51 **(Total)**	1201 **(Average)**	523.9* **(Average)**

* Incidence in the community was calculated using only community residents and excluded off-reserve residents.

Table 8: Core collapse sequence (to degree 8) of Community 1 TB network, with Mantoux positivity at each stage of collapse.

Network Degree	TB		Mantoux Positive		Mantoux Negative		Total
	(N)	(%)	(N)	(%)	(N)	(%)	
All	68	13.5	109	21.6	327	64.9	504
2	45	36.8	35	28.7	42	34.4	122
3	28	62.2	10	22.2	7	15.6	45
4	15	68.2	7	31.8	0	0	22
5	14	77.8	4	22.2	0	0	18
6	9	90	1	10	0	0	10
7	7	100	0	0	0	0	7
8	7	100	0	0	0	0	7

Table Notes:
Community 1 TB network was composed of 504 nodes: 68 (13.5%) were TB cases, 109 (21.6%) were Mantoux positive, and 327 (64.9%) were Mantoux negative.

Degree 2: the network was collapsed to degree2 (every node in the network should know at least 2 more nodes.) The network was composed of 122 nodes: 45 (36.8%) were TB cases, 35 (28.7%) were Mantoux positive, and 42 (34.4%) were Mantoux negative.

Degree 3: the network was collapsed to degree3 (every node in the network should know at least 3 more nodes.) The network was composed of 45 nodes 28 (62.2%) were TB cases, 10 (22.2%) were Mantoux positive, and 7 (15.6%) were Mantoux negative.

Table 9: Manitoba FP1 network collapse sequence, identifying Community 1 proportion in the network composition

	Network size	TB cases in Community 1	Contacts of TB cases In Community 1	Remaining members Of FP1 network in MB
FP1 all network	1795 (100%)	64 (3.6%)	369 (20.5%)	1362 (75.9%)
FP1 degree 2	303 (100%)	48 (15.8%)	84 (27.7%)	171 (56.4%)
FP1 degree 3	79 (100%)	31 (39.2%)	26 (32.9%)	22 (27.9%)
FP1 degree 4	43 (100%)	21 (48.8%)	12 (27.9%)	10 (23.3%)
FP1 degree 5	28 (100%)	17 (60.7%)	10 (35.7%)	1 (3.6%)
FP1 degree 6	22 (100%)	17 (77.3%)	5 (22.7%)	0
FP1 degree 7	11 (100%)	10 (90.9%)	1 (9.1%)	0
FP1 degree 8	9 (100%)	8 (88.9%)	1 (11.1%)	0
FP1 degree 9	3 (100%)	3 (100%)	0	0

Table 10: Public places attended by different population groups in Community 1 and their ranking in significance in accordance with Degree, Bonacich and Betweenness centrality measures

Public Places	TB Network Centrality			Contact Network Centrality			Control Network Centrality			Community 1 Network Centrality		
	D	E	B	D	E	B	D	E	B	D	E	B
General Store	12	5	5	21	5	2	5	1	1	1	1	1
School	14	10	6	1	1	1	28	23	5	2	3	2
Nursing Station	11	3	4	31	8	9	23	8	26	7	2	5
Recreational Hall				48	33	34	32	29	10	61	44	8
Band Office	13	9	11	201	87		26	11	28	39	11	50
L Store	16	8	16	85	65	66	31	15	39	45	12	58
Restaurant				207	92		27	12	24	74	34	74
Arena							35	46	23	111	118	75
T Inn	119	24					212	74		163	112	101
Airport				200	81		37	18	41	82	49	105
Pentecostal							34	24	40	94	91	107
S Store	17	12	19							125	74	118
IUN	115	20					206	51		170	85	124
E Store							48	47	48	131	125	149

D= Degree centrality ranking. E= Bonacich Eigenvector centrality ranking

B= Betweenness centrality ranking.

TB Network Degree Centrality range = 1 to 21

Contact Network Degree Centrality range = 0 to 18

Control Network Degree Centrality range = 1 to 25

Table 11: Characteristics and network measures for the three Community 1 study groups

Network Measures	TB Network		Contact Network		Control Network		All groups Network	
	People*K*	*People & placesσ*	People*K*	*People & placesσ*	People*K*	*People & placesσ*	People*K*	*People & placesσ*
Number of index people	12	*12*	44	*44*	27	*27*	83	*83*
Number of places		*32*		*90*		*47*		*145*
Network total number	88	*120*	198	*289*	170	*217*	391	*536*
Density	*0.02*	*0.019*	*0.01*	*0.009*	*0.013*	*0.01*	*0.007*	*0.006*
Degree								
Mean	1.932	*2.250*	2.251	*2.588*	2.071	*2.516*	2.471	*2.91*
Range	1 – 15	*1 – 21*	0 – 12	*0 – 18*	1 – 20	*1 – 225*	0 – 20	*0 – 33*
Degree centralization	15.4%	*16.0%*	4.9%	*5.4%*	10.7%	*10.5%*	4.5%	*5.7%*
Betweenness								
Mean	0.546	*2.510*	1.617	*0.894*	1.137	*1.561*	1.084	*0.617*
Range	0.0–12.6	*0.0–51.9*	0.0-30.5	*0.0-35.7*	0.0-26.4	*0.0-30.3*	0.0-20.9	*0.0-34.9*
Betweenness centralization	12.3%	*49.9%*	29.1%	*34.9%*	25.4%	*49.0%*	19.9%	*34.4%*
Components	5	*1*	9	*6*	7	*1*	10	*4*
Closeness centralization		*30.4%*		-		*34.4%*		-

Table Notes:
Network types include network with people only or (K) (index nodes and alters) in normal fonts, and network of people and places (σ) (index nodes and alters including places attended by index nodes counted as member of alters) in italic fonts.

Table 12: Means, SD, ANOVA F Values and *P* values for group characteristics.

	Groups			Comparison of Means	
	TB	**Contact**	**Control**	**F**	***P***
Age	41.83 (13.7)	28.05 (19.9)	42.02 (18.4)	5.724	0.005
Age (12 over)	41.83 (13.7)	34.42 (19.1)	42.04 (18.4)	1.554	0.219
Education	8.92 (1.93)	6.52 (2.97)	7.19 (3.52)	2.935	0.059
House Density	1.743 (0.637)	1.667 (0.638)	1.304 (0.541)	3.630	0.031
Living out of Community1	0.054 (0.123)	0.089 (0.147)	0.084 (0.140)	0.293	0.747
Co morbid illness	3.67 (4.62)	0.62 (1.89)	1.89 (3.93)	4.631	0.013
Number of known TB patients	4 (4.02)	0.95 (1.92)	0.63 (1.28)	11.090	0.000

Table 13: Means, SD, ANOVA F and *P* values for study group network characteristics

	Groups			Comparison of Means	
	TB	**Contact**	**Control**	**F**	***P***
Age of Alters	27 (14)	16 (11)	25 (13)	5.943	0.004
Network Size[1]	11.42 (4.48)	8.61 (3.95)	10.33 (5.84)	2.001	0.142
Contact Size	7.08 (4.48)	5.20 (3.29)	6.78 (5.15)	1.695	0.190
Place Diversity	0.63 (0.31)	0.49 (0.26)	0.80 (0.28)	8.824	0.001
Closeness	2.75 (0.39)	2.75 (0.43)	2.50 (0.71)	1.806	0.172
Frequency of Contacts with alters	1.98 (1.61)	2.89 (1.83)	3.45 (1.87)	2.437	0.094

1=: network members include people and places attended or interaction with the index individuals.

Table 14: Student's t-test results comparing TB group versus control group

	TB vs. control	Mean	Std. Deviation	t	*P*
Age	TB cases Control	41.83 42.04	13.717 18.444	-0.034	0.973
Education	TB cases Control	8.92 7.19	1.929 3.520	1.593	0.120
House Density	TB cases Control	1.74306 1.30350	.637088 .540914	2.218	0.033
Living out of Community1	TB cases Control	.05417 .08395	.123322 .140472	-0.633	0.531
Co morbid illness	TB cases Control	3.67 1.89	4.619 3.926	1.237	0.224
Number of known TB patients	TB cases Control	4.00 .63	4.023 1.275	2.840	0.015
Age of alters	TB cases Control	27.3900 24.4642	13.96604 13.05552	0.594	0.556
Network size (people + places)	TB cases Control	11.42 10.33	5.854 5.844	0.534	0.597
Contacts size	TB cases Control	7.08 6.78	4.481 5.147	0.178	0.860
Place diversity	TB cases Control	.62857 .79565	.312606 .283966	-1.508	0.142
Closeness to alters	TB cases Control	2.7516 2.5066	.38615 .71227	1.027	0.311
Frequency of contact with alters	TB cases Control	1.9830 3.4481	1.60570 1.86748	-2.194	0.035

Table 15: Student's t-test results comparing TB group versus contact group

	TB vs. contact	Mean	Std. Deviation	T	*P*
Age (12 over)	TB cases Contacts	41.83 34.42	13.717 19.133	1.228	0.226
Education	TB cases Contacts	8.92 6.52	1.929 2.969	3.351	0.002
House Density	TB cases Contacts	1.74306 1.66705	.637088 .637800	0.366	0.716
Living out of Community1	TB cases Contacts	.05417 .08939	.123322 .146849	-0.760	0.451
Co morbid illness	TB cases Contacts	3.67 .62	4.619 1.899	2.232	0.045
Number of known TB patients	TB cases Contacts	4.00 .95	4.023 1.916	2.545	0.025
Age of alters	TB cases Contacts	27.3900 15.7285	13.96604 11.02383	2.814	0.007
Network size (people + places)	TB cases Contacts	11.42 8.61	5.854 3.954	1.952	0.056
Contacts size	TB cases Contacts	7.08 5.20	4.481 3.289	1.618	0.111
Place diversity	TB cases Contacts	.62857 .48548	.312606 .258476	1.447	0.156
Closeness to alters	TB cases Contacts	2.7516 2.7714	.38615 .45569	-0.126	0.900
Frequency of contact with alters	TB cases Contacts	1.9830 2.8976	1.60570 1.83220	-1.446	0.154

Table 16: Student's t-test results comparing contact group versus control group

	Contact vs. Control	Mean	Std. Deviation	T	*P*
Age (12 over)	Contact Control	34.42 42.04	19.133 18.444	-1.558	0.125
Education	Contact Control	6.52 7.19	2.969 3.520	-0.850	0.398
House Density	Contact Control	1.66705 1.30350	.637800 .540914	2.565	0.013
Living out of Community1	Contact Control	.08939 .08395	.146849 .140472	0.154	0.878
Co morbid illness	Contact Control	.62 1.89	1.899 3.926	-1.567	0.126
Number of known TB patients	Contact Control	.95 .63	1.916 1.275	0.780	0.438
Age of alters	Contact Control	15.7285 24.4642	11.02383 13.05552	-2.915	0.005
Network size (people + places)	Contact Control	8.61 10.33	3.954 5.844	-1.479	0.144
Contacts size	Contact Control	5.20 6.78	3.289 5.147	-1.420	0.163
Place diversity	Contact Control	.48548 .79565	.258476 .283966	-4.181	0.000
Closeness to alters	Contact Control	2.7714 2.5066	.45569 .71227	1.858	0.068
Frequency of contact with alters	Contact Control	2.8976 3.4481	1.83220 1.86748	-1.203	0.233

Table 17: Chi Square test results comparing categorical variables of Community 1

	TB* n (%)	Contact* n (%)	Control* n (%)	χ^2	P Value
Living arrangement				2.515	0.642
SPF[1]	2 (16.7)	9 (20.5)	4 (15.0)		
FPC[2]	6 (50.0)	14 (31.8)	13 (48.0)		
Others	4 (33.3)	21 (47.7)	10 (37.0)		
Density				13.57	0.001
Density ≥ 1.56	9 (75)	24 (54.5)	5 (18.5)		
Density ≤ 1.56	3 (25)	20 (45.5)	22 (81.5)		
Number of TB patients they know				7.784	0.020
Known 0	3 (25)	26 (65)	19 (70.4)		
Known 1or more	9 (75)	14 (35)	8 (29.6)		
Work status				14.802	0.005
Working	2 (16.7)	4 (9.1)	9 (33.3)		
Not working	8 (66.6)	19 (43.2)	15 (55.6)		
Student	2 (16.7)	21 (47.7)	3 (11.1)		
Network type				6.913	0.032
Risky	9 (75)	18 (41)	18 (67)		
Not risky	3 (25)	26 (59)	9 (33)		
Behavior type				4.459	0.108
Risky	9 (75)	19 (43)	16 (59)		
Not risky	3 (25)	25 (57)	11 (41)		

SPF[1]: single parent family. FPC[2]: family parent and children. *=Network

Table 18: Correlation matrix of people and network characteristics variables

	M	SD	1	2	3	4	5	6	7	8	9	10	11	12
People (group) Characteristics														
1-Age Sig. (2 tail)	34.95	19.77												
2-Age 12 Sig. (2 tail)	38.51	18.25												
3-Education Sig. (2 tail)	7.08	3.12	.019 .866	-.231 .004										
4- House Density Sig. (2 tail)	1.56	.63	-.336 .002	-.231 .051	-.094 .399									
5-Living out of community1 Sig. (2 tail)	.083	.14	-.205 .063	.288 .014	.162 .143	-.028 .802								
6-Co morbid conditions Sig. (2 tail)	1.49	3.31	.389 .000	.352 .003	.180 .107	-.288 .009	-.213 .056							
7-Number of known TB Patients Sig. (2 tail)	3.57	2.87	.088 .427	-.022 .856	.364 .001	-.060 .589	.125 .262	.034 .762						
Network Characteristics														
8-Age of Alters Sig. (2 tail)	20	13	.592 .000	.487 .000	.201 .084	-.234 .043	-.011 .928	.325 .005	.095 .419					
9-Network Size Sig. (2 tail)	9.58	4.98	-.107 .334	-.207 .081	.319 .003	-.086 .439	.279 .011	.111 .325	.294 .007	.303 .008				
10-Contact Size Sig. (2 tail)	5.99	4.18	-.103 .353	-.200 .092	.367 .001	-.072 .519	.315 .004	.063 .574	.280 .010	.319 .005	.950 .000			
11-Place Diversity Sig. (2 tail)	0.62	0.31	.267 .033	.155 .253	.209 .097	-.129 .310	-.076 .550	.133 .300	.074 .563	.364 .004	.142 .264	.235 .061		
12-Closeness to alters Sig. (2 tail)	2.67	0.55	-.008 .947	.111 .373	-.173 .133	.201 .080	-.184 .109	.026 .826	-.101 .381	.020 .863	-.156 .177	-.177 .123	-.096 .463	
13-Frequency of contact with alters Sig. (2 tail)	2.97	1.85	-.212 .062	-.212 .085	.049 .670	.157 .170	.256 .024	-.142 .220	-.191 .094	.262 .024	.358 .001	.405 .000	.244 .058	.168 .146

Table 19: ANCOVA results. Comparison between TB group and control group

	Mean		Adjusted Mean		Comparison of Means	
	TB	**Control**	**TB**	**Control**	**F**	**P**
Age of Alters	27.3900	24.4642	27.461	24.438	0.257	0.615
Network Size	11.42	10.33	9.672	11.109	0.438	0.512
Contact Size	7.08	6.78	5.557	7.456	1.079	0.306
Place Diversity	.62857	.79565	0.603	0.807	2.443	0.129
Closeness to alters	2.7516	2.5066	2.858	2.467	2.027	0.164
Frequency of contact with alters	1.9830	3.4481	1.589	3.594	6.680	0.014

Age was controlled by study design. Density and Number of known TB patients variables have been controlled. These variables were significant in t test.

Table 20: ANCOVA results. Comparison between TB group and contact group

	Mean		Adjusted Mean		Comparison of Means	
	TB	**Contact**	**TB**	**Contact**	**F**	P
Age of Alters	27.3900	15.7285	25.043	20.174	1.737	0.197
Network Size	11.42	8.61	11.614	8.746	2.950	0.093
Contact Size	7.08	5.20	7.062	5.371	1.727	0.196
Place Diversity	.62857	.48548	0.574	0.543	0.071	0.791
Closeness to alters	2.7516	2.7714	2.709	2.723	0.007	0.936
Frequency of contact with alters	1.9830	2.8976	1.949	2.765	1.386	0.247

Age (over12) and Density variables have been controlled.

Table 21: ANCOVA results. Comparison between contact group and control group

	Mean		Adjusted Mean		Comparison of Means	
	Contact	**Control**	**Contact**	**Control**	**F**	**P**
Age of Alters	15.7285	24.4642	20.946	22.823	0.387	0.537
Network Size	8.61	10.33	8.734	10.436	1.609	0.210
Contact Size	5.20	6.78	5.278	6.883	1.954	0.168
Place Diversity	.48548	.79565	0.522	0.793	9.684	0.003
Closeness to alters	2.7714	2.5066	2.711	2.504	1.577	0.215
Frequency of contact with alters	2.8976	3.4481	2.638	3.576	3.407	0.071

Age (over12) and Density variables have been controlled. These variables were significant in t test or to control for the difference in particular among control group.

Table 22: Chi Square test results. Comparison of categorical variables of TB group and control group in Community 1

	TB n (%)	Control n (%)	χ^2	P Value
House Density			11.52	0.001
Density ≥ 1.56	9 (75)	5 (18.5)		
Density ≤ 1.56	3 (25)	22 (81.5)		
Number of known TB patients			6.96	0.008
Known 0	3 (25)	19 (70.4)		
Known 1or more	9 (75)	8 (29.6)		

Table 23: Chi Square test results. Comparison of categorical variables in contact group and control group in Community 1

	Contact n (%)	Control n (%)	χ^2	P Value
House Density			8.99	0.003
Density ≥ 1.56	24 (54.5)	5 (18.5)		
Density ≤ 1.56	20 (45.5)	22 (81.5)		
Work status			12.54	0.002
Working	4 (9.1)	9 (33.3)		
Not working	19 (43.2)	15 (55.6)		
Student	21 (47.7)	3 (11.1)		
Network type			4.44	0.035
Risky	18 (41)	18 (67)		
Not risky	26 (59)	9 (33)		

Table 24: Chi Square test results. Comparison of categorical variables of TB group and contact group in Community 1

	TB n (%)	Contact n (%)	χ^2	P Value
Number of known TB patients			5.99	0.014
Known 0	3 (25)	26 (65)		
Known 1or more	9 (75)	14 (35)		
Network type			4.39	0.036
Risky	9 (75)	18 (41)		
Not risky	3 (25)	26 (59)		
Behavior type			3.82	0.051
Risky	9 (75)	19 (43)		
Not risky	3 (25)	25 (57)		

Table 25: Community1 network components size distribution

Components	TB (N)	Contact (N)	Control (N)	Alter (N)	Total (N)(%)
1	12	37	23	281	353(90%)
2	0	1	2	8	11(2.8%)
3	0	1	0	6	7(1.8%)
4	0	0	1	5	6(1.5%)
5	0	1	0	3	4(1%)
6	0	1	0	2	3(0.8%)
7	0	0	1	2	3(0.8%)
8	0	1	0	1	2(0.5%)
9	0	1	0	0	1(0.3%)
10	0	1	0	0	1(0.3%)
Totals	12	44	27	308	391(100%)

Table 26: Network Characteristics of Community 1 Network and its Main Component

Network Measures	Community 1 Network	Community 1 Network Main Component
Density	0.00652	0.0073
Degree (Mean) **Range**	2.47 0 – 20	2.57 1-20
Degree Centralization	4.52%	4.98%
Components	10	1
Diameter	13	13
Cliques (N=3)	45	47
2-Cliques (N=3)	145	137
3-Cliques (N=3)	277	270
4-Cliques (N=3)	1834	1828
5-Cliques (N=3)	3988	3982
6-Cliques (N=3)	6437	6431
7-Cliques (N=3)	2732	2726
8-Cliques (N=4)	296	292
2-Plex (N=3)	2217	2162
2-Plex (N=4)	106	106

Table 27: Clique structure of Community1 network

Cliques	Cliques Members	Cliques	Cliques Members	Cliques	Cliques Members
1	77, CL20, CL19	*2*	197, CL20, CL19	*3*	CL21, CL20, CL19
4	3, CL8, CL15	*5*	CT86, 205, TB6	*6*	CT47, 22, CT48, CT46
7	CT47, CT46, 342	*8*	CT48, CT22, CT46	*9*	CT48, 26, CT46
10	CT48, CT46, CT66	*11*	CT75, TB18, CT60, CT59	*12*	CT75, TB18, 169
13	TB56, CT60, CT59	*14*	59, CT5, CT2	*15*	TB18, TB27, CT66
16	TB18, CT59, CT66	*17*	CL68, 113, CL65	*18*	86, CT84, CT87
19	99, CT22, CT83	*20*	104, CL65, CL4	*21*	CT17, 108, CL1
22	106, CT84, CT87	*23*	108, CT84, CT87	*24*	CT84, 120, CT87, TB79
25	CT84, CT41, CT87	*26*	CT84, 117, CT87	*27*	CT84, CT87, CT66
28	118, CL8, CL15	*29*	CT64, CT39, CL65	*30*	CT64, CL65, 258
31	CT64, CL65, 267	*32*	CT64, CL65, 275	*33*	135, CT3, CT2
34	136, CT5, CT2	*35*	TB27, 185, CT66	*36*	TB27, CT43, CT66
37	TB42, TB79, 268	*38*	TB58, CT69, CT76	*39*	TB67, 181, CT76
40	TB67, 182, CT76	*41*	TB67, CT73, CT76	*42*	189, TB6, TB85
43	CT22, CT83, 272	*44*	CT46, CT66, TB6	*45*	CT66, TB6, TB85
46	237, CT3, CT2	*47*	CL70, 366, CL71		

(TB) Clique Members from tuberculosis group, (CT) Clique Members from contact group,
(CL) Clique Members from control group, (Number only) Clique Members from *ego's alters*,

Table 28: Community 1 network core collapse sequence

	TB Group N (%)	Contact group N (%)	Control Group N (%)	All groups alters N (%)	Total N
Community1 Network	12 (3)	44 (11)	27 (7)	308 (79)	391
2 Degree	11 (9)	33 (27)	20 (16)	60 (48)	124
3 Degree	5 (19)	14 (53)	1 (4)	6 (23)	26
4 Degree	0	3 (100)	0	0	3

Table 29: Community 1 network main component core collapse sequence

	TB Group N (%)	Contact group N (%)	Control Group N (%)	All groups alters N (%)	Total N
Community1	12 (3.4)	37 (10.5)	23 (6.5)	281(79.6)	353
2 Degree	12 (9.7)	32 (25.8)	20 (16.1)	60 (48.4)	124
3 Degree	5 (22.7)	12 (54.6)	0	5 (22.7)	22
4 Degree	0	0	0	0	0

Table 30: Case control study categorical variable frequencies

		TB	Contact	Control	All
Gender	Male	8	20	16	44
	Female	4	24	11	39
Residence1	All time in Community 1	9	26	14	49
	In & outside Community 1	3	18	13	34
Residence2	Lived in Winnipeg	2	10	6	18
	Did not live in Winnipeg	10	34	21	65
Residence3	Lived in Thompson	2	13	9	24
	Did not live in Thompson	10	31	18	59
Living Arrangement	Single parent family	2	9	4	15
	Parent and children family	6	14	13	33
	Others	4	21	10	35
Work Status	Working	2	4	9	15
	Not Working	8	19	15	42
	Students	2	21	3	26
House Density	Density > 1.56	9	24	5	38
	Density < 1.56	3	20	22	45
Number of known TB patients	Know 1 or more TB patients	9	14	8	31
	Did not know any TB patient	3	26	19	48
Comorbid Illness	Present	6	8	9	23
	Not present	6	34	18	58
Type of Network	High Risk	9	18	18	45
	Low Risk	3	26	9	38
Type of Behaviors	High Risk	9	19	16	44
	Low Risk	3	25	11	39

Table 31: Case control study continuous variable frequencies

	TB				Contact				Control				All			
	Mean	SD	Mn	Mx	Mean	SD	Mn	Mx	Mean	SD	Mn	Mx	Mean	SD	Mn	Mx
People (group) Characteristics																
Age (all)	41.8	13.7	24	65	28.1	19.9	5	79	42	18	12	82	35	19.8	5	82
Age (>12)	41.8	13.7	24	65	34.4	19.1	12	79	42	18	12	82	38.5	18.3	12	82
Education	8.9	1.9	6	12	6.5	3.0	1	12	7.2	3.5	0	12	7.1	3.1	0	12
House Density	1.7	0.6	0.5	2.3	1.7	0.6	0.5	2.7	1.3	0.5	0.3	3	1.6	0.6	0.3	3
Comorbid Illness	3.7	4.6	0	11	0.6	1.9	0	11	1.9	3.9	0	12	1.5	3.3	0	12
Number of known TB patients	4	4	0	10	0.95	1.9	0	10	0.6	1.3	0	5	3.6	2.9	0	10
Network Characteristics																
Age of Alters	27	14	11	59	16	11	3	39	25	13	5	61	20	13	3	61
Network Size	11.4	5.9	1	21	8.6	4	0	16	10.3	5.8	3	25	9.6	5	0	25
Contact Size	7.1	4.5	0	15	5.2	3.3	0	11	6.8	5.2	1	20	6	4.2	0	20
Place Diversity	0.6	0.3	0.2	1	0.5	0.3	0.2	1	0.8	0.3	0.3	1	0.6	0.3	0.2	1
Closeness to alters	2.8	0.4	1	3	2.8	0.4	1	3	2.5	0.7	0	3	2.7	0.6	0	3
Frequency of Contact with alters	2	1.6	0	4.8	2.9	1.8	0	7	3.5	1.9	0.3	7	3	1.9	0	7

SD (Standard Deviation), Mn (Minimum), Mx (Maximum).

Table 32: Discriminant analysis of study group characteristics - Standardized canonical discriminant function coefficients

	Axes	
	X (1)	**Y (2)**
Age	-0.871	0.157
House Density	0.430	-0.159
Number of TB patients known by index nodes	1.026	0.154
Place Diversity	-0.286	0.703
Co-morbid illness	0.551	-0.067
Age of alters	0.803	0.403

Wilks' Lambda of discriminant analysis axes

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1 through 2	.273	69.486	12	.000
2	.673	21.202	5	.001

Table 33: Cumulative and degree distribution for main component of Community 1 network

Degree distribution	**Cumulative Distribution**
1	100.00
2	37.80
3	24.50
4	18.30
5	14.60
6	12.90
7	10.40
8	9.30
9	7.60
10	5.60
11	4.50
12	3.10
13	2.30
14	1.50
15	1.20
16	.90
19	.60
20	.30

Table 34: Final logistic regression model comparing TB and Control groups in Community 1 network

Name/Variables	Regression Coefficient	Standard Error	Probability	Adjusted Odds Ratio
Intercept	- 7.255075	2.848612	0.010869	
House Density	2.492945	1.211088	0.039549	12.09
Number of known TB patients	1.200299	0.5124936	0.019177	3.32

Model R^2	Model D.F	Model Chi-Square	Model Probability
0.426281	2	21.55	0.000021

Table 35: Final logistic regression model comparing TB and Contact groups in Community 1 network

Name/Variables	Regression Coefficient	Standard Error	Probability	Adjusted Odds Ratio
Intercept	-3.512008	1.890186	0.063166	
Contacts size	0.4778132	0.2388469	0.045446	1.61
Network type	1.378924	0.5909692	0.019631	3.97
Frequency of contact	-1.050794	0.4626637	0.023136	0.35

Model R^2	Model D.F	Model Chi-Square	Model Probability
0.362235	3	17.61	0.000530

Table 36: Final logistic regression model comparing Contact and Control groups in Community 1 network

Name/Variables	Regression Coefficient	Standard Error	Probability	Adjusted Odds Ratio
Intercept	- 4.820419	2.760206	0.080742	
Number of known TB patients	0.9966365	0.4352204	0.022024	2.71
Place diversity	- 3.836263	1.582327	0.015332	0.0215
Work status (Student)	2.923687	1.08502	0.007047	18.6
Closeness	2.094671	0.9163269	0.022258	8.12

Model R^2	Model D.F	Model Chi-Square	Model Probability
0.413319	4	31.00	0.000003

Table 37: Final logistic regression model comparing centrality measures in Community 1 network

Name/Variables	Regression Coefficient	Standard Error	Probability	Adjusted Odds Ratio
Intercept	-4.22418	0.5063903	0.000000	
Degree	1.10343	0.2259653	0.000001	3.014

Model R^2	Model D.F	Model Chi-Square	Model Probability
0.055976	1	23.07	0.000002

REFERENCES

1. Centrality in Social Networks Conceptual clarifications. Social Networks 1978;1:215-39.

2. Treatment of tuberculosis and tuberculosis infection in adults and children. American Thoracic Society
   24. Monaldi Arch.Chest Dis 1994;49:327-45.

3. Chossing a centrality measures: Epidemiology correlates in Colorado Spring study of social networks. Social Networks 1995;17:273-97.

4. Screening for tuberculosis and tuberculosis infection in high-risk populations. Recommendations of the Advisory Council for the Elimination of Tuberculosis 4. MMWR Recomm.Rep. 1995;44:19-34.

5. Social network approach to corroborating the number of AIDS/HIV+ victims in the US. Social Networks 1995;17:167-87.

6. Proceedings of the national consensus conference on tuberculosis, Dec 3-5, 1997. Canada Communicable Disease Report. 1998;Vol 24S2.

7. Recommendations for prevention and control of tuberculosis among foreign-born persons. Report of the Working Group on Tuberculosis among Foreign-Born Persons. Centers for Disease Control and Prevention. MMWR Recomm.Rep. 1998;47:1-29.

8. Diagnostic Standards and Classification of Tuberculosis in Adults and Children. 3. Am.J Respir.Crit Care Med. 2000;161:1376-95.

9. Targeted tuberculin testing and treatment of latent tuberculosis infection. American Thoracic Society. MMWR Recomm.Rep. 2000;49:1-51.

10. Agasino CB et al. Epidemiology of Mycobacterium tuberculosis strains in San Francisco that do not contain IS6110 1. Int.J Tuberc.Lung Dis 1998;2:518-20.

11. Al-Azem A et al. Rural and localized tuberculosis outbreaks in Manitoba (1992-1999). ATS 2000 CDC-NCHSTP-DTBE.Public Health Poster Session . 2000.

12. Al-Azem A et al. The impact of foreign born tuberculosis in Manitoba (1992-1999). ATS 2000 CDC-NCHSTP-DTBE.Public Health Poster Session . 2000.

13. Al-Azem A et al. Tuberculosis in Manitoba: Conventional and moleculor epidemiology (1992-1999). ATS 2000 CDC-NCHSTP-DTBE.Public Health Poster Session . 2000.

14. Al-Mouaiad Al-Azem A. Epidemiology of Tuberculosis in Manitoba (1992-1997). 7-9-1999.

15. Arvanitakis Z et al. M. tuberculosis molecular variation in CNS infection: evidence for strain-dependent neurovirulence 3. Neurology 1998;50:1827-32.

16. Bailey WC et al. Predictive model to identify positive tuberculosis skin test results during contact investigations 2. JAMA 2002;287:996-1002.

17. Barnes JA. Social networks

1. Reading, Mass: Addison-Wesley Pub. Co., 1972.

18. Bass MJ. Conducting research in the practice setting. Newbury Park, Calif: Sage Publications, 1993.

19. Behr MA et al. Predictive value of contact investigation for identifying recent transmission of Mycobacterium tuberculosis. Am.J.Respir.Crit Care Med. 1998;158:465-9.

20. Borgdorff MW et al. Transmission of tuberculosis between people of different ages in The Netherlands: an analysis using DNA fingerprinting

10. Int.J Tuberc.Lung Dis 1999;3:202-6.

21. Brancker A et al. A statistical chronicle of tuberculosis in Canada: Part I. From the era of sanatorium treatment to the present 3. Health Reports 1992;4:103-23.

22. Brancker A et al. A statistical chronicle of tuberculosis in Canada: Part II. Risk today and control 2. Health Reports 1992;4:277-92.

23. Casper C et al. The transcontinental transmission of tuberculosis: A molecular epidemiological assessment 2. Am.J Public Health 1996;86:551-3.

24. Cave MD et al. Stability of DNA fingerprint pattern produced with IS6110 in strains of Mycobacterium tuberculosis 9. J Clin.Microbiol. 1994;32:262-6.

25. CDC. Essential Components of a Tuberculosis Prevention and Control Program. Screening for Tuberculosis and Tuberculosis Infection in High-Risk Populations. RR-11, 1-32. 9-8-1995. MMWR.

26. Clark M, Riben P. Tuberculosis in First Nations Communities, 1999. 1-66. 1999.

27. Clark M, Riben P, Nowgesic E. The association of housing density, isolation and tuberculosis in Canadian First Nations communities. Int.J Epidemiol. 2002;31:940-5.

28. Cohn DL, O'Brien RJ. The use of restriction fragment length polymorphism (RFLP) analysis for epidemiological studies of tuberculosis in developing countries 5. Int.J Tuberc.Lung Dis 1998;2:16-26.

29. Cowie RL, Sharpe JW. Tuberculosis among immigrants: interval from arrival in Canada to diagnosis. A 5-year study in southern Alberta
1. CMAJ. 1998;158:599-602.

30. Drobniewski FA et al. Molecular biology in the diagnosis and epidemiology of tuberculosis 6. J Hosp.Infect 1994;28:249-63.

31. Drobniewski FA, Pozniak AL. Molecular diagnosis, detection of drug resistance and epidemiology of tuberculosis 2. Br.J Hosp.Med. 1996;56:204-8.

32. Elliot LJ. An evaluation of the role of tuberculosis DNA fingerprinting in delineation of Manitoba outbreak. 1-60. 1-12-1995.

33. Enarson DA. Tuberculosis in Aboriginals in Canada
1. International Journal of Tuberculosis and Lung Disease 1998;2:S16-S22.

34. Enarson DA, Grzybowski S. Incidence of active tuberculosis in the native population of Canada 1. CMAJ. 1986;134:1149-52.

35. Enarson DA, Wang JS, Dirks JM. The incidence of active tuberculosis in a large urban area 1. Am.J Epidemiol. 1989;129:1268-76.

36. Fanning A. Tuberculosis: 1. Introduction 2. CMAJ. 1999;160:837-9.

37. Fanning EA. Globalization of tuberculosis 21. CMAJ. 1998;158:611-2.

38. FitzGerald JM. Tuberculosis in Canadian schools 21. Can J Public Health 1993;84:294-6.

39. FitzGerald JM, Wang L, Elwood RK. Tuberculosis: 13. Control of the disease among aboriginal people in Canada. CMAJ. 2000;162:351-5.

40. Fitzgerald JM, Fanning A, Gangadharam PRj. Tuberculosis among the disadvantaged. International Journal of Tuberculosis and Lung Disease 1998;2:S1-163.

41. Fitzpatrick LK et al. A preventable outbreak of tuberculosis investigated through an intricate social network 1. Clin.Infect Dis 2001;33:1801-6.

42. Fomukong NG et al. Insertion sequence typing of Mycobacterium tuberculosis: characterization of a widespread subtype with a single copy of IS6110
1. Tuber.Lung Dis 1994;75:435-40.

43. French AL et al. Use of DNA fingerprinting to assess tuberculosis infection control 1. Annals of Internal Medicine 1998;129:856-61.

44. Frieden TR et al. The molecular epidemiology of tuberculosis in New York City: the importance of nosocomial transmission and laboratory error 3. Tuber.Lung Dis 1996;77:407-13.

45. Garro LC. Individual or societal responsibility? Explanations of diabetes in an Anishinaabe (Ojibway) community 10. Social Science and Medicine 1995;40:37-46.

46. Gaudette LA, Ellis E. Tuberculosis in Canada: a focal disease requiring distinct control strategies for different risk groups 1. Tubercle and Lung Disease 1993;74:244-53.

47. Genewein A et al. Molecular approach to identifying route of transmission of tuberculosis in the community 1. Lancet 1993;342:841-4.

48. Gerald LB et al. Standardizing contact investigation protocols. International Journal of Tuberculosis and Lung Disease 2003;7:S369-S374.

49. Gicquel B. RFLP typing of the Mycobacterium tuberculosis bacilli 3. Tuber.Lung Dis 1993;74:223-4.

50. Grzybowski S. Epidemiology of tuberculosis and the role of BCG. Clinics in Chest Medicine 1980;1:175-87.

51. Grzybowski S. Natural history of tuberculosis. Epidemiology 14. Bull.Int.Union Tuberc.Lung Dis 1991;66:193-4.

52. Grzybowski S, Allen EA. Tuberculosis: 2. History of the disease in Canada 1. CMAJ. 1999;160:1025-8.

53. Hall GB, Nelson G. Social networks, social support, personal empowerment, and the adaptation of psychiatric consumers/survivors: path analytic models 1. Social Science and Medicine 1996;43:1743-54.

54. Hammer M. 'Core' and 'extended' social networks in relation to health and illness 1. Social Science and Medicine 1983;17:405-11.

55. Health Canada. Canada Communicable Disease Report. 1998. Supplement. Proceeding of the national consensus conference on tuberculosis, Dec 3-5, 1997. Vol 24S2. 1998.

56. Health Canada. Tuberculosis in Canada 2000. Melissa Phypers. (29) 2, 9-16. 1-15-2003. Health Canada. CCDR.

57. Hermans PW et al. Insertion element IS986 from Mycobacterium tuberculosis: a useful tool for diagnosis and epidemiology of tuberculosis 2. J Clin.Microbiol. 1990;28:2051-8.

58. Hershfield E. Tuberculosis Still a major health problem 20. Can J Infect Dis 1991;2:131-3.

59. Hershfield E. Tuberculosis: 9. Treatment 13. CMAJ. 1999;161:405-11.

60. Hoeppner VH, Marciniuk DD. Tuberculosis in aboriginal Canadians. Canadian Respiratory Journal 2000;7:141-6.

61. Holmes W et al. Researching Aboriginal health: experience from a study of urban young people's health and well-being 1. Social Science and Medicine 2002;54:1267-79.

62. Indian and Northern Affairs Canada. First Nations Community Profiles 1997-1998 Manitoba region. QS-N038-000-EE-A1, 1-167. 1997. Ottawa, Indian and Northen Affairs Canada. QS-N038-000-EE-A1.

63. Iseman MD. TB elimination in the 21st century, a quixotic dream? 1. International Journal of Tuberculosis and Lung Disease 2000;4:S109-S110.

64. Jolly AM et al. Sexual networks and sexually transmitted infections: a tale of two cities 1. J Urban.Health 2001;78:433-45.

65. Jolly AM, Wylie JL. Sampling individuals with large sexual networks: an evaluation of four approaches 2. Sex Transm.Dis 2001;28:200-7.

66. Kato-Maeda M et al. The nature and consequence of genetic variability within Mycobacterium tuberculosis 1. J Clin.Invest 2001;107:533-7.

67. Kimerling ME et al. Restriction fragment length polymorphism screening of Mycobacterium tuberculosis isolates: population surveillance for targeting disease transmission in a community 3. Int.J Tuberc.Lung Dis 1998;2:655-62.

68. Kline SE, Hedemark LL, Davies SF. Outbreak of tuberculosis among regular patrons of a neighborhood bar. N.Engl.J.Med. 1995;333:222-7.

69. Klovdahl AS. Social networks and the spread of infectious diseases: the AIDS example 3. Social Science and Medicine 1985;21:1203-16.

70. Klovdahl AS. Networks and pathogens 1. Sex Transm.Dis 2001;28:25-8.

71. Klovdahl AS et al. Networks and tuberculosis: an undetected community outbreak involving public places 1. Social Science and Medicine 2001;52:681-94.

72. Klovdahl AS et al. Social networks and infectious disease: the Colorado Springs Study 2. Social Science and Medicine 1994;38:79-88.

73. Ko G, Thompson KM, Nardell EA. Estimation of tuberculosis risk on a commercial airliner. Risk Anal. 2004;24:379-88.

74. Kulaga S, Behr MA, Schwartzman K. Genetic fingerprinting in the study of tuberculosis transmission 2. CMAJ. 1999;161:1165-9.

75. Kunimoto D et al. Investigation of tuberculosis transmission in Canadian Arctic Inuit communities using DNA fingerprinting 1. Int.J Tuberc.Lung Dis 2001;5:642-7.

76. Long R. Canadian Tuberculosis Standards. Goverment of Canada, 2000.

77. Long R, Njoo H, Hershfield E. Tuberculosis: 3. Epidemiology of the disease in Canada 2. CMAJ. 1999;160:1185-90.

78. Macintyre CR, Plant AJ. Impact of policy and practice on the effectiveness of contact screening for tuberculosis. Prev.Med. 1998;27:830-7.

79. Marks SM et al. Outcomes of contact investigations of infectious tuberculosis patients. Am.J.Respir.Crit Care Med. 2000;162:2033-8.

80. McElroy PD et al. A network-informed approach to investigating a tuberculosis outbreak: implications for enhancing contact investigations. International Journal of Tuberculosis and Lung Disease 2003;7:S486-S493.

81. McKenna MT, McCray E, Onorato I. The epidemiology of tuberculosis among foreign-born persons in the United States, 1986 to 1993 29.

82. Menzies D. Tuberculosis crosses borders. The International Journal of Tuberculosis and Lung Disease 2000;4:S153-S159.

83. Menzies D et al. Efficacy of environmental measures in reducing potentially infectious bioaerosols during sputum induction. Infect.Control Hosp.Epidemiol. 2003;24:483-9.

84. Menzies D, Adhikari N, Tannenbaum T. Patient characteristics associated with failure of tuberculosis prevention 1. Tubercle and Lung Disease 1996;77:308-14.

85. Menzies D et al. Hospital ventilation and risk for tuberculous infection in canadian health care workers. Canadian Collaborative Group in Nosocomial Transmission of TB. Annals of Internal Medicine 2000;133:779-89.

86. Menzies D, Tannenbaum TN, FitzGerald JM. Tuberculosis: 10. Prevention 11. CMAJ. 1999;161:717-24.

87. Mohle-Boetani JC, Flood J. Contact investigations and the continued commitment to control tuberculosis 1. JAMA 2002;287:1040-2.

88. Nardell EA. Catching droplet nuclei: toward a better understanding of tuberculosis transmission. Am.J.Respir.Crit Care Med. 2004;169:553-4.

89. Njoo H. tuberculosis- A re-emerging public health threat in Canada. Can J Infect Dis 1998;9:273-5.

90. Njoo H, Nault P. Tuberculosis in Canada 1998. 1-51. 2001.

91. Njoo H, Nault P, Farzad E. Tuberculosis in Canada 1996. 1-56. 1998.

92. Pilheu JA. Tuberculosis 2000: problems and solutions 1. Int.J Tuberc.Lung Dis 1998;2:696-703.

93. Ponticiello A et al. Demographic risk factors and lymphocyte populations in patients with tuberculosis and their healthy contacts. International Journal of Tuberculosis and Lung Disease 2001;5:1148-55.

94. Potterat JJ et al. Sexual network structure as an indicator of epidemic phase. Sex Transm.Infect. 2002;78 Suppl 1:i152-i158.

95. Potterat JJ et al. Risk network structure in the early epidemic phase of HIV transmission in Colorado Springs. Sex Transm.Infect. 2002;78 Suppl 1:i159-i163.

96. Raviglione MC, Snider DE, Jr., Kochi A. Global epidemiology of tuberculosis. Morbidity and mortality of a worldwide epidemic 1. JAMA 1995;273:220-6.

97. Raviglione MC et al. Secular trends of tuberculosis in western Europe 1. Bull.World Health Organ 1993;71:297-306.

98. Reichler MR et al. Evaluation of investigations conducted to detect and prevent transmission of tuberculosis 3. JAMA 2002;287:991-5.

99. Rivest P, Tannenbaum T, Bedard L. Epidemiology of tuberculosis in Montreal 22. CMAJ. 1998;158:605-9.

100. Rothenberg R et al. Social network methods for endemic foci of syphilis: a pilot project. Sex Transm.Dis. 2000;27:12-8.

101. Rothenberg RB et al. Contact tracing: comparing the approaches for sexually transmitted diseases and tuberculosis. International Journal of Tuberculosis and Lung Disease 2003;7:S342-S348.

102. Roy CJ, Milton DK. Airborne transmission of communicable infection--the elusive pathway. N.Engl.J.Med. 2004;350:1710-2.

103. Sahadevan R et al. Restriction fragment length polymorphism typing of clinical isolates of Mycobacterium tuberculosis from patients with pulmonary tuberculosis in Madras, India, by use of direct-repeat probe 2. J Clin.Microbiol. 1995;33:3037-9.

104. Schmitt N. J. Douglas Galbraith fights TB among Indians and Inuit 10. CMAJ. 1986;134:1184-5.

105. Schwartzman K, Menzies D. Tuberculosis: 11. Nosocomial disease 10. CMAJ. 1999;161:1271-7.

106. Schwartzman K, Menzies D. Tuberculosis screening of immigrants to low-prevalence countries. A cost-effectiveness analysis 2. Am.J Respir.Crit Care Med. 2000;161:780-9.

107. Scott J. Social network analysis a handbook 16. London: SAGE Publications, 1991.

108. Shafer RW et al. Temporal trends and transmission patterns during the emergence of multidrug-resistant tuberculosis in New York City: a molecular epidemiologic assessment 2. J Infect Dis 1995;171:170-6.

109. Sharnprapai S et al. Genotyping analyses of tuberculosis cases in U.S.- and foreign-born Massachusetts residents 1. Emerg.Infect Dis 2002;8:1239-45.

110. Small PM et al. The epidemiology of tuberculosis in San Francisco. A population-based study using conventional and molecular methods 5. N.Engl.J Med. 1994;330:1703-9.

111. Small PM et al. Molecular strain typing of Mycobacterium tuberculosis to confirm cross- contamination in the mycobacteriology laboratory and

modification of procedures to minimize occurrence of false-positive cultures 1. J Clin.Microbiol. 1993;31:1677-82.

112. Smeja C, Brassard P. Tuberculosis infection in an Aboriginal (First Nations) population of Canada 1. International Journal of Tuberculosis and Lung Disease 2000;4:925-30.

113. Statistics Canada. Dimensions Series: Ethnocultural and social characteristics of the Canadian population. 94F001XCB. 1999. Otawa. Dimensions Series.

114. Stead WW. Variation in vulnerability to tuberculosis in America today: random, or legacies of different ancestral epidemics? 2. Int.J Tuberc.Lung Dis 2001;5:807-14.

115. Stehr-Green JK. Tuberculosis in New Zealand, 1985-90 1. N Z.Med J 1992;105:301-3.

116. Thierry D et al. Characterization of a Mycobacterium tuberculosis insertion sequence, IS6110, and its application in diagnosis 2. J Clin.Microbiol. 1990;28:2668-73.

117. Valente TW et al. Social network associations with contraceptive use among Cameroonian women in voluntary associations 2. Social Science and Medicine 1997;45:677-87.

118. Valente TW. Network models of the diffusion of innovations 1. Cresskill, N.J: Hampton Press, 1995.

119. van Deutekom H et al. A molecular epidemiological approach to studying the transmission of tuberculosis in Amsterdam 1. Clin.Infect Dis 1997;25:1071-7.

120. van Embden JD et al. Strain identification of Mycobacterium tuberculosis by DNA fingerprinting: recommendations for a standardized methodology 7. J Clin.Microbiol. 1993;31:406-9.

121. Wallace CE et al. Approaches to contact investigations in Texas. International Journal of Tuberculosis and Lung Disease 2003;7:S358-S362.

122. Wasserman S, Faust K. Social network analysis methods and applications 1. Cambridge: Cambridge University Press, 1994.

123. Watkins RE, Plant AJ, Gushulak BD. Tuberculosis rates among migrants in Australia and Canada 1. Int.J Tuberc.Lung Dis 2002;6:641-4.

124. Webb RM, Holcombe M, Pearson MM. Tuberculosis contact investigation in a rural state. International Journal of Tuberculosis and Lung Disease 2003;7:S353-S357.

125. Wherrett GJ. Tuberculosis in Canada: Royal Commission on Health Services. 1964. Ottawa.

126. Wherrett GJ. The miracle of the empty beds a history of tuberculosis in Canada 1. Toronto: University of Toronto Press, 1977.

127. Wilce M et al. Tuberculosis contact investigation policies, practices, and challenges in 11 U.S. communities. J.Public Health Manag.Pract. 2002;8:69-78.

128. Wilkins Kathryn. Tuberculosis, 1994. Health Reports Vol8, No. 1, 33-39. 1996. Statistics Canada. 1996.

129. Wobeser WL et al. Expanding the epidemiologic profile: risk factors for active tuberculosis in people immigrating to Ontario 1. CMAJ. 2000;163:823-8.

130. Wylie JL, Jolly A. Patterns of chlamydia and gonorrhea infection in sexual networks in Manitoba, Canada 3. Sex Transm.Dis 2001;28:14-24.

131. Yeh RW et al. Stability of Mycobacterium tuberculosis DNA genotypes 1. J Infect Dis 1998;177:1107-11.

132. Young TK. Aboriginal health: not just tuberculosis. International Journal of Tuberculosis and Lung Disease 1998;2:S23-S25.

133. Young TK, Casson RI. The decline and persistence of tuberculosis in a Canadian Indian population: implications for control. Can.J.Public Health 1988;79:302-6.

134. Young TK, Hershfield ES. A case-control study to evaluate the effectiveness of mass neonatal BCG vaccination among Canadian Indians 5. American Journal of Public Health 1986;76:783-6.

135.Young TK. Measuring the health status of Canada's aboriginal population a statistical review and methodological commentary 10. Winnipeg, Man: Northern Health Research Unit, Dept. of Community Health Sciences, Faculty of Medicine, University of Manitoba, 1994.

136. Waldram JB, Herring DA, Young TK. Aboriginal health in Canada. University of Toronto press, 1995.

137. Vladimir Batagelj and Andrej Mrvar. Pajek 1.03. November 1996 - January 2005.

138. Borgatti, S.P., M.G. Everett, and L.C. Freeman. 2002. *Ucinet 6 for Windows: Software for Social Network Analysis.* Harvard: Analytic Technologies.

139. Wouter de Nooy, Andrej Mrvar, and Vladimir Batagelj, *Exploratory Social Network Analysis with Pajek.* Cambridge University Press 2005.

140. National Advisory Committee on Immunization. *Canadian immunization guide.* 4th ed., Ottawa, Ont.: Health Canada, 1993:29-33. (Supply and Services Canada, Cat. No. H49-8/1993E.)

141. Advisory Committee on Immunization Practices. *The role of BCG vaccine in the prevention and control of tuberculosis in the United States.* MMWR 1996;45:1-18

Glossary

A

Actor: *Vertex, Node refers to person or organization that is involved in social relation(139).*

Alter: *The nodes that connected to the index node or ego in ego-centered network.*

Affiliation: *Affiliation networks are two mode networks; consist of two subsets of nodes (e.g. people and places). The connections among members of one of the modes are based on linkages established through the second mode(122).*

B

Betweenness centrality: *Betweenness centrality of a node is the proportion of all geodesics between pairs of other nodes that include this node(139).*

Betweenness centralization: *Betweenness centralization is the variation in Betweenness centrality of nodes divided by the maximum variation in between centrality scores possible in a network of the same size(139).*

C

Clique: *"A clique is a maximally complete subgraph of at least three nodes"(107, 122, 139).*

Closeness centrality: *Closeness centrality of a node is the number of other nodes divided by the sum of all distances between the node and all others(139).*

Closeness centralization: *Closeness centralization is the variation in the closeness centrality of nodes divided by the maximum variation in closeness centrality scores possible in a network of the same size(139).*

Component *"Component is a maximal connected subgraph"*

D

Degree: *The degree of a node is the number of lines incidents with it(139).*

Degree centrality: *The degree centrality of node is its degree(139).*

Degree centralization: *Degree centralization of a network is the variation in the degree of nodes divided by the maximum degree variation that is possible in a network of the same size(139).*

Density *"Density of a graph is the ratio of the number of the lines present to the maximum possible"*

Diameter of a graph *"Diameter of a graph considers the largest geodesic distance between any pair of nodes in a graph"*

E

Edge: *An edge is undirected line(139).*

Ego-network: *Ego-network consist of focal node (index node), termed ego, and set of alters (122).*

G

Graph *"Graph is a visual representation of a data, graph is a model for social network with undirected dichotomous relation"*

Geodesic *"Geodesic is the shortest path between two nodes"*

Geodesic distance *"Geodesic distance or simply distance between two nodes is the length of a geodesic between them"*

K

***K*-Core:** *K*-Core is a maximal subnetwork in which each node has at least degree k within the subnetwork*(139).*

N

Node: *Actor, Vertex* refers to person or organization that is involved in social relation*(139).*

Network: *A network consists of a graph and additional information on the nodes or the lines of the graph(139).*

O

One-mode network: *In a one-mode network, each node can be related to each other node(139).*

T

Triads *"Is a subgraph consisting of three nodes"*

Two-mode network: *In a two-mode network, nodes are divided into two sets and nodes can be related only to nodes in the other set(139).*

V

Vertex (*Pl.* Vertices): *Node, Actor* refers to person or organization that is involved in social relation.

Abbreviations

BCG: Bacillus Calmette Guerin

C1: Community 1

DOT: Directly Observed Therapy

FP1 Fingerprint 1

FB: Foreign born

INH: Isoniazid

IS: Insertion sequence

LTBI: Latent tuberculosis infections

MDR: Multi-drug resistant strains

RFLP: Restriction Fragment Length Polymorphism

SNA: Social Network Analysis

TB: Tuberculosis