Risk Factors in the Progression from Tuberculosis Infection to Disease

A Thesis Submitted to the College of Medicine

In Partial Fulfillment of the Requirements for the Degree of

Masters of Science

University of Saskatchewan

Saskatoon, Saskatchewan

Canada

Heather A. Ward

Permission to Use

In presenting this thesis as partial fulfillment of the requirements for a Masters degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying of publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or make other use of material in this thesis in whole or part should be addressed to:

Department Head, Community Health and Epidemiology

University of Saskatchewan

Saskatoon, Saskatchewan, CANADA

S7N 0W8

Abstract:

Tuberculosis (TB) is a two-stage disease, acquisition of infection and progression to disease. A complex interaction exists between the individual and their environment that determines who acquires infection and who progresses to disease. According to TB literature, 10% of individuals with infection will develop TB disease (1;2). Tuberculosis has been described a disease of poverty, but other factors may be important. The contribution of both individual measures, such as ethnic origin, gender and age and area-level measures, or socio-economic factors, to this two-stage process is not well understood. Understanding tuberculosis epidemiology and identifying those at risk for developing TB is important for effectively controlling the disease.

The objective of this study was to determine the individual (age, gender, ethnic origin, geographic location) and area-level measures (income, home ownership, housing density, education, and employment) that contribute to the progression from tuberculosis infection to disease. Data from all Canadian-born Caucasians, Status Indians, and non-Status Indians and Metis, with an initial positive tuberculin skin test (TST) documented in the Saskatchewan TB Control database from January 1, 1986 to January 31, 2002 was analyzed. Exclusion criteria included any previous BCG vaccination, treatment for latent TB infection, or missing data. Individual data was obtained from the TB Control database. Area-level measures were obtained by matching individual postal codes with Canada census data to obtain information from enumeration areas. Outcome was time to TB disease at $\geq$ 1 month following a documented positive tuberculin skin test. Analysis was completed using Cox regression proportional hazards model.

7588 individuals with a positive tuberculin skin test were included in the study and of these 338 (4.5%) developed TB disease. Thirty-four out of 4140 (0.8%) of Caucasians, 183 out of 2649 (6.9%) of Status Indians and 121 out of 799 (15.1%) non-Status Indians and Metis developed TB. The rate of progression to TB was 5.6/1000 person years for the entire study population. The incidence for Caucasians was 0.9/1000 person years, 7.7/1000 person years for Status Indians and 16.0/1000 person years for non-Status Indians and Metis. In the Cox regression model, including individual and area-level measures, the risk factors association with the progression to TB was age and ethnic origin (< 19 years of age HR 3.7, 95% CI 2.8 – 4.8 compared to $\geq$ 19 years and ethnic origin HR 5.1, 95% CI 3.0 – 8.6 for Status Indians and HR 7.4, 95% CI 4.1-13.3 for non-Status Indians and Metis both compared to Caucasians). No socio-economic factor was consistently associated with progression to disease.

We have found that age and ethnic origin are associated with an increased risk of TB infection progressing to disease. The differences in TB rates between Saskatchewan Caucasians and Status Indians, non-Status Indians and Metis can be explained by Grigg's natural history curve of TB epidemiology within a population (3). The Aboriginal population of Saskatchewan is much earlier in its epidemic resulting in higher disease rates compared to the Caucasian population.

Identifying those at risk of developing TB and understanding the determinants of TB epidemiology are important for establishing successful TB control programs.

Acknowledgements

Thank you to my thesis committee, Dr. Bruce Reeder, Dr. Nazeem Muharajine, Dr. Punam Pahwa, Dr. Darcy Marciniuk, Dr. Vern Hoeppner, for their encouragement and patience during the learning process involved in completing this thesis. A special thank you to Dr. Punam Pahwa who repeatedly provided helpful advice on the biostatistical analysis of this project.

Thank you to Rob Anderson, GIS Analyst, Corporate Information and Technology, Saskatchewan Health, who obtained the Canada Census data and completed the linkage with the individual-level data.

Without the assistance of my classmates Karen Hay, Maureen Jackson, and Shardelle Brown, this course would never have been completed. Thank you.

A special thank you to Joshua Lawson, a very talented and kind colleague would answered a multitude of statistical questions with very competent and comprehensive answers. Thank you to Carol Lynn Nagle for her encouragement.

Thank you to my medicine colleagues Sharon Card and Anne PausJenssen who provided the encouragement and practical advice necessary to complete this thesis

Thank you to my family members, Joyce Heaton, for her interest, encouragement and careful editing and to my mother, Ida Ward, for her support.
A very special thank you to my friends who supplied the chocolate necessary to make this project possible.

In memorium of Jack Ward, September 1940 to August, 2004.

Table of Contents

List of Tables

List of Figures

1.0 Introduction

Tuberculosis (TB) is an infectious disease transmitted between humans. Historically, it has had implications on the course of society and has been listed as one of history's most lethal killers (4). Evidence of tuberculosis disease in humans was documented in Egyptian civilization beginning in 3400 BCE (2). Babylonian writings dated 1948 to 1905 BCE document a disease that was most likely tuberculosis (1). Hippocrates (460-375 BCE) first referred to tuberculosis as phthisis, or the wasting disease (2). The estimated mortality rate from tuberculosis epidemic occurring in London, England during the Industrial Revolution was 900/100 000 in 1740 (3), while the highest documented mortality rate in the early epidemic of Fort Qu'Appelle First Nations was 13 700/100 000 in 1890 (5).Tuberculosis continues to have a role in one of society's more recent epidemics as tuberculosis is the number one cause of death in individuals with HIV/AIDS (6). Tuberculosis is a pervasive disease whose impact on society has never abated. Identifying the cause and developing an effective treatment has been an ongoing process of scientific discovery since Greek civilization (2).

The disease process and its epidemiology are not completely understood. There is a complex interaction between the individual and their environment that determines who acquires infection and who progresses to TB disease. Control of tuberculosis occurs by stopping transmission (identifying and treating infectious disease), preventing infection (BCG vaccination), and preventing disease among infected high risk individuals (preventive treatment). Identifying those at increased risk of developing TB disease is one priority of TB Control. Understanding the epidemiology of tuberculosis will enhance TB Control programs.

1.1 Objective

To determine the individual (age, gender, ethnic origin, geographic location) and area-level measures (mean level of income, home ownership, crowding, education, and employment) that contribute to the progression from TB infection to disease.

2.0 Literature Review

2.1 Tuberculosis Pathogenesis

Tuberculosis is a two-step process. The first step is acquisition of infection and the second is progression to disease. Infected individuals are not ill and not infectious, but may progress to active tuberculosis. Individuals with TB may be able to transmit infection and require treatment for cure.

2.1.1 Infection

Laennec initially proposed the unitary theory of tuberculosis when he stated that early progression to disease or delayed reactivation were different stages of the same process begun at the initial infection (7). Tuberculosis infection is acquired by inhaling respiratory droplet nuclei aerosolized by coughing from an individual with pulmonary tuberculosis. Respiratory droplets less than 5 microns in size must reach the lung alveoli where the *Mycobacterium tuberculosis* begin to multiply (2). Primary tuberculosis usually becomes established in the lower portion of the lung where ventilation is greatest. At initial exposure in individuals with no previous contact to *Mycobacterium*, minimal immune resistance to growth is encountered (7). During this period the *Mycobacteria* spread hematogenously, or circulate through the blood stream, and establish metastatic

foci of infection at other sites. This is thought to most commonly involve the lung apices, but may include other sites such as the kidneys, distal end of long bones, meninges and brain.

The replicating *Mycobacteria* stimulate a specific immune response. If the immune response is adequate, the primary and metastatic sites of infection regress and heal. If the response is not adequate, the initial infection may progress to primary disease. Individuals newly infected with tuberculosis have the greatest risk of progressing to disease within the first two years following infection (2;7). During this time, the situation is unstable and the infection may heal or progress to disease in the lung or other organs if the immune defence is not able to promote and sustain healing. Immunity is usually adequate to cause a regression and healing of the initial infection although the *Mycobacteria* may persist in a latent or dormant state. Reactivation of the original infection may occur after many years of latency.

Reactivation tuberculosis most commonly affects the lung apex as healing at primary foci is less likely to occur. High oxygen tension may favour persistence of dormant *Mycobacteria* as healed primary apical lesions contain culturable TB bacteria more commonly than at other sites (7). Although *Mycobacteria* become dormant in the lung apex, the potential for reactivation remains.

TB infection is documented by a positive tuberculin skin test (TST) (8). A positive TST is a delayed type hypersensitivity (Th1) immune response to purified protein derivative of *M. tuberculosis* injected into the dermis. False positive TST results can occur as a result of sensitization from other species of *Mycobacteria* and from BCG

vaccine. Presently there is no other screening or clinical diagnostic method used for identifying TB infection.

The initial infection may progress to disease, resolve, or become latent with the potential to reactivate in the future. Reactivation is not evidence of re-exposure, as the initial infection protects against re-infection, but part of the unitary theory of tuberculosis proposed by Laennec (7). Primary infection with *Mycobacterium tuberculosis* has a long-term clinical perspective in primary progression to disease, reactivation or latency.

Individuals with TB infection are not ill or infectious. It is estimated that one third of the world population is infected with TB (6). Ninety percent of individuals with TB infection are reported never to develop TB disease (2;8). Prophylactic treatment to prevent the progression to disease can be provided to those with infection identified at increased risk of progression (8).

As TB is spread by respiratory droplets, concentration of airborne bacilli and duration of exposure are two key factors in transmission of TB infection. Living conditions such as overcrowded housing with increased droplet concentration will increase risk of TB transmission (9;10).

2.1.2 Disease

Individuals with TB disease have actively multiplying *Mycobacterium tuberculosis*. Clinical symptoms are most commonly cough and fever (11). Effective treatment with antibiotics is available. Treatment is required to improve clinical symptoms and halt infectiousness. On average, 85 percent of TB cases are pulmonary; however the remaining 15% can involve any area that the TB bacteria seeded during

infection at the initial hematogenous spread (7;12). Individuals may be infectious if they have pulmonary TB and aerosolize *Mycobacterium* with coughing.

Ten percent of individuals with infection progress to disease – an estimated 5% develop primary and the remaining 5% reactivation TB (2;8). Delayed type hypersensitivity involving T lymphocytes and their messengers is responsible for preventing TB disease. A few high risk groups in the progression to TB have been identified. Individuals with impaired T lymphocyte immunity, those with HIV/AIDS or immunosuppression following organ transplantation, for example, are at increased risk of progressing to TB. Children are at increased risk of developing TB, particularly meningeal and disseminated disease (2). In the majority of individuals with TB disease, however, the etiology for progression from infection to disease is not well understood. The epidemiology of high risk groups for progression to TB is still being examined.

In U.S. population groups, such as First Nations, Hispanics, and African Americans, TB disease rates are higher than in the Caucasian population (2;13). This may result from either increased risk of infection and/or increased risk of progression to disease. Tuberculosis is often described as a disease of poverty and socio-economic determinants have been identified as the explanation for increased TB rates among different ethnic groups (2;14). The exact mechanism of action of socio-economic factors and their effect on TB rates is not well defined in the literature. Tuberculosis as a two-step process and the effect of each socio-economic factor on these two stages must be evaluated to further define TB epidemiology and the disease process.

2.2 Socio-economic Factors and Tuberculosis

2.2.1 Defining Socio-economic Factors

Social and economic factors have long been recognized as determinants of health. In the 14th century, Paracelsus identified high rates of medical conditions in miners (15). By the 19th century, the effects of living and working conditions of the Industrial Revolution on the health of the poor were identified by several individuals including Virchow and Engles (15). Differences in health outcomes among socio-economic groups is a consistent finding in epidemiological research (15). Higher socio-economic status results in better health outcomes (15). Skills, knowledge and accessible resources that accompany social position are important for determining health (15). Quantifying this association has been difficult because of the limitations associated with measuring socio-economic status.

Capturing definitions of socio-economic factors is challenging. Different terminology, such as socio-economic status, socio-economic position, or social class, refers to unique concepts and theory. This results in variable definitions utilized for researching the relationship between social and economic factors and health (16). Socio-economic factors include not only measures of income, education and occupation, but the historical, cultural, political and institutional factors that interact with individual factors to determine the association between status and health outcomes (15). Many of these factors are interdependent, dynamic or possibly even unmeasurable, making standard measures of socio-economic factors difficult to identify.

Socio-economic position, social inequality and social status, are aggregate concepts of socio-economic factors that include both resource and prestige-based

measures (15). Resourced-based measures refer to material and social resources, including income, wealth, and education. Prestige is defined by occupation, income and education. Prestige-based measures refer to rank or status in the social hierarchy and often determine access to resources (17). The interaction of resource and prestige-based factors will vary between communities and countries depending on historical, cultural, social, and political contexts. Measuring individual-level resources such as occupation, income and education may not capture the aggregate concepts included in socio-economic position or social status.

Clearly defining socio-economic factors and their measure is important for researching association with health outcomes. An awareness of the complexity of socio-economic measurement is necessary. Traditionally, individual measures of resource-based socio-economic factors, education, income, and occupation, have been used to define socio-economic factors. These individual measures may not account for contextual effects of the socio-economic environment. Individual-level indicators are derived from the larger social and economic interaction occurring in the area and their measurement alone may result in unmeasured aspects of socio-economic factors (15). Using individual-level measures may provide a limited assessment of the factors that determine health.

An alternative to individual-level measures are area-level measures of socio-economic factors. Although area-level measures are sometimes considered a proxy for individual-level data, they may also provide a measure of the socio-economic environment. Area-level measures may encompass the interaction of individual-level measures. Measurement of socio-economic factors needs to incorporate both individual-

level variables with the social and economic context to identify the association with health outcomes. The use of individual and area-level measures to encompass the concepts of socio-economic factors and identify the social and economic context is debated in the literature. Various established methods, using both individual and area-levels of measuring socio-economic factors have been utilized.

2.2.2 Quantifying Socio-economic Factors – Individual and Area-level Measures

Controversy exists in the literature around the various data collection methods used to measure socio-economic factors. Two methods of data collection exist – individual and area-level measures. Individual-level data is collected directly from individuals. Area-level data is aggregate data from defined geographic areas of varying size. This can include data from census tracts or enumeration areas, combined with a geographic locator such as the forward sorting area from a postal code. Although individual-level data provides the most direct information for each socio-economic factor, there is no measure of the context or neighbourhood effects of this data. Area-level measures can be an indirect measure or proxy for individual-level measures or may directly measure socio-economic context or environment. Individual-level data can be difficult and costly to obtain. Area-level measures can be more readily available in pre-existing data bases, such as census data, and less costly to obtain. The controversy reviewed in the literature surrounds the applicability of individual-level compared to area-level data for measuring the association between socio-economic factors and health.

Although individual-level measures of socio-economic characteristics have been associated with health, there are limitations to using individual-level data. Data may not be readily available when required, it may be costly to obtain, and it may not be

complete. In addition, individual data are considered to be too narrow a measure to solely provide the richness or variability of data that is required to explore the association between socio-economic context and health (18;19).

Defining variables for socio-economic factors is often limited by what is available. Area-level or aggregate data are often available from pre-existing data sources. For example, census data, grouped by enumeration area or postal codes, is available from previous years. Because individual data are often difficult to obtain, group level data has been used to examine the association between socio-economic status and health in two ways. The first is as a proxy or indirect measure for unavailable individual level data. The second is as a direct measure of the socio-economic context or environment as means to characterize qualities of the area itself. Area-level measures may provide information about socio-economic context beyond what individual level data are able to provide.

Socio-economic position is defined by interdependent individual characteristics (18). Socio-economic position or factors represent a complex of living and working conditions, assets, and attitudes (18). The environment or context created by the interaction of the various individual characteristics creates a setting that influences behaviours, access to resources, and attitudes (18;20). Looking at individual-level data alone may not capture this neighbourhood effect. Group-level data, defining neighbourhood socio-economic factors, may provide a context for health outcomes in a given area not measured by individual-level data (15;19;21). Using individual-level data may detract from the context or neighbourhood effect provided by area-level measures. Area-level measures may define an association between socio-economic factors and health not defined by individual-level data (20). Individual variability is lost but

information on area or socio-economic context may be defined. Although area-level measures may underestimate the association with socio-economic factors, individual-level data is often not available for research purposes (17). Identifying a neighbourhood effect may provide useful information when examining the association between community social and economic factors and health. Studies examining area-level measures as a proxy for individual measures and as a direct measure for association with health outcomes have been completed.

An Alberta study compared income as a measure of poverty to determine the accuracy of area-level measures as a proxy for individual measures (22). Individuals with taxable incomes below a threshold to be exempted from Alberta Health Insurance premiums (<$12, 620 per year) were identified. This data set was linked to the lowest income quintile identified by Canada census data using the first 3 digits (forward sorting area) of postal codes. The area-level measures in urban areas had a sensitivity of 26% and specificity of 83% to identify those with low incomes. In non-metropolitan areas, the sensitivity was 39% and specificity 69% (22). The authors concluded that there was only a weak association between personal poverty and postal code markers of poverty and that using only postal codes as a definition of personal poverty would result in 'substantial misclassification' (22). In non-urban areas the weaker association suggested heterogeneity of incomes in postal code areas. No individual data was collected and income was the only socio-economic factor of poverty identified. As a proxy for presumed individual data, area-level measures were able to identify individuals with low incomes with only limited association between area and presumed individual-level data.

Geronimus and Bound also found that micro-level data for income, education and occupation corresponded weakly with area-level data (23). The authors concluded that although correlations were weak, area-level data should not necessarily be used as an equivalent measure for individual-level data. The authors theorized that aggregate data measured different socio-economic factors than individual data. Aggregate data was a 'proxy for a broader construct than micro–level data' and may measure contextual effects, not individual variables (23).

Krieger compared individual data obtained from an insuring organization in California with census-based area-level data for education and social class and health outcomes (hypertension, height, smoking, reproductive history) (24). Individual and census-based socio-economic factors were similarly associated with health outcomes (24). Individual and area-level measures of socio-economic status were not directly compared; however, health outcomes using each of these measures were similar. Although area-level data may not measure the same factors as individual data, the association with health outcomes was similar.

An evaluation of household income compared to area-based measures as an indicator of oral health was completed in Ontario (20). The area-based measures were used to define a neighbourhood group or category. The individual measure of income was a better predictor of inequalities in oral health status and access to dental care than the area-level measures; however, both indicated a consistent gradient of socio-economic status and dental health. Area-level measures had more complete data compared to the individual measure of household income (98% complete vs 75%). Area-level measures demonstrated that neighbourhood data provided information independent of individual-

level data. The area in which people lived had an influence on health and health behaviour. Subjects from low income households living in high status areas had better oral health and used dental services more frequently than those living in low status areas. Area-level data in this study was used to provide information on neighbourhood context or setting. The area-level measures were thought to be more valuable in directing resources to areas to improve dental health (20). Individual and area-level measures had different roles in examining socio-economic determinants of dental health in Ontario neighbourhoods.

A second study by the same authors again used area-based measures of socio-economic factors to examine the correlation between mean enumeration area income and an individual's self-reported income and use of dental services (25). Despite a weak correlation, the two measures of income were comparable in their ability to identify variations in dental health and service use. With fewer income categories, correlation between individual and group-level data improved. After controlling for self-reported socio-economic factors, association between dental health and area-level measures of socio-economic factors persisted suggesting that the socio-economic status of the area in which individuals lived had an effect on dental health beyond that identified by self-reported individual characteristics alone.

Although the association between area and individual-level measures of income has not been strongly correlated in several studies, the gradient between socio-economic factors and health outcome exists for both individual and area-level measures (15;20;22;23;25). In addition, area-level measures seem to have an independent role, separate from measurements of individual-level variables, in identifying the relationship

between health and socio-economic environment. "Extra-individual socio-economic factors closely related to the physical and social infrastructure of the communities are thought to affect health above and beyond individual aspects" (15). The socio-economic characteristics of an area in which individuals live influences their health in addition to the effect exerted by personal characteristics (15).

A study completed in Alameda County, California, found an independent role for area-level measures on mortality rate and poverty (21). Mortality rates were high in areas of poverty independent of individual socio-economic characteristics including income. "The socio-physical environment may be an important contributor to the association between low socio-economic status and excess mortality. This contribution is independent of individual behaviour" (19). Variation in health between areas may not be completely identified by individual characteristics. "These apparent contextual or area-level effects may be due to individual characteristics, imprecise measurements or other statistical artefacts, or they may be real and need to be considered as part of the relationship between health and socio-economic environments" (19).

One possible limitation to using group data as a proxy for individual data is ecologic fallacy. Ecologic fallacy makes an incorrect causal inference from group data to individual behaviours or incorrectly assumes that correlations that apply to groups will also apply to individuals. Individual and group-level data may measure different components of socio-economic factors. It is important to state at the outset of any study the measures for socio-economic factors that will be used and the intended purpose for these measures.

Socio-economic factors can also be evaluated as a composite index using both individual and/or area-level measures. Commonly used indices, Jarman, Townsend, and Carstairs Indices, are measures of social deprivation (9;10;18;26-29). The Townsend Index, for example, is a set of characteristics with a subjectively defined threshold indicating deprivation for each factor (9;10;19;26-29). A socio-economic index can be used to represent individual-level measures or as a measure of the composite social and economic environment. As with all socio-economic measurements, it is important to know what socio-economic characteristics are included in the index and the threshold of each definition of deprivation in order to understand what is being measured (19). Knowing what the index measures and how it is applied is necessary to define the association between the index and health.

The challenge for all measures of socio-economic factors is to recognize the presence and importance of socio-economic context. Although a composite index may include information on individual and area-level characteristics, an index has limitations. A threshold definition of social deprivation that impacts on health may be difficult to define and justify (25). These indices are often used as a proxy for individual-level variables but may actually be representing neighbourhood effects of the combined characteristics (18;19). By using a composite measure, the opportunity to examine the relationship between individual measures of socio-economic factors and health is lost. Measures of both individual socio-economic characteristics and their context are important for identifying an association between health and socio-economic factors (18). Using an index may provide a composite assessment of the relationship between health

and social deprivation, but understanding what is being measured is necessary for interpreting study results (19).

Recognizing advantages and limitations of the data being utilized is necessary when researching the association between socio-economic status and health. Individual-level data may provide an association between each socio-economic factor and health. Area-level measures may provide a different method of examining the association between socio-economic factors and health by measuring a context or neighbourhood effect not identified by individual-level measures. Using more easily accessible area-level data as a proxy for individual-level data provides only a weak correlation of data and runs the risk of ecologic fallacy. When comparing measures for individual and area-level data, income for example, only a weak correlation between individual and area-level data was found (22-24). The gradient between socio-economic status and health seems to persist for area-level measures. Even though the correlation was weak, the associated gradient between socio-economic factors and health persisted (23;24). Area-level measures may provide a socio-economic context associated with health outcomes, unique from individual-level measures. Heterogeneity of group level data may result in a weaker association with health outcomes (20). Both individual and area-level measures have a role in defining plausible explanations for the causal link between socio-economic factors and health.

The socio-economic environment or context, as identified by area-level measures, is important in examining the relationship between socio-economic factors and progression from infection to disease. In exploring tuberculosis epidemiology and the association with socio-economic factors, it is important to identify which socio-economic

factors are being measured and how they are being used. This will assist in identifying how biological and social processes interact to affect susceptibility to disease (19;30).

2.2.3 Area-level Measures and Tuberculosis

Tuberculosis rates differ between population groups. In Saskatchewan, TB rates in Status Indians are higher than Canadian-born Caucasians or foreign-born (12). It is uncertain as to whether intrinsic genetic differences defining immune response or social and economic factors determine the varied TB rates between different populations. Identifying those at risk of developing TB and understanding the determinants of TB epidemiology is important for establishing successful TB control programs.

"Tuberculosis is a disease of poverty" is a frequently cited generalization that focuses the epidemiology of tuberculosis on impoverished groups (2;28). This generalization underestimates the complexity of the relationship between the TB bacillus, host immune response, age, ethnic origin, and socio-economic characteristics, including individual and neighbourhood effects. Both the method of measuring socio-economic factors, individual vs area-level data, and the pathogenesis of the two-stage disease process are important considerations in examining the association between socio-economic characteristics and tuberculosis.

Some studies examining the association between tuberculosis and socio-economic factors have used composite indices while others have used separate socio-economic variables measured by area or individual-level measures. Each method resulted in different associations between TB and socio-economic factors. The majority of studies identified the presence of disease as the outcome measure without separately examining the two stages of TB disease. Only one study has examined the association between

socio-economic status, ethnic origin and acquisition of infection and no studies have examined socio-economic status and the progression from infection to disease (31).

Bhatti et al examined the increase in TB rates in the United Kingdom from 1988 to 1992 (9). The documented rise in TB rates affected only the socio-economically poorest areas as measured by the Jarman Index. TB was therefore associated with socio-economic deprivation, according to the authors. The Jarman Index, using census data, was comprised of ten factors including overcrowding, poor housing and proportion of ethnic minority residents. The strongest correlation between TB and measures of socio-economic deprivation was with overcrowding. Although TB rates differed between ethnic groups, the authors felt that the association was not clinically significant. TB rates were increased among ethnic minorities, but these increases were partly attributed to the increased number of recent immigrants from TB endemic regions. Distinguishing ethnic origin and socio-economic factors and their respective contribution was not discernable because of the study method. The authors concluded that socio-economic factors affecting individuals of all ethnic origins were responsible for the observed increase in TB rates in the United Kingdom. Socio-economic deprivation, as defined by ten factors in the Jarman Index, was considered to be more important than ethnic group as the cause for increased TB rates.

Spence et al utilized both the Jarman and Townsend Indices as well as two independent socio-economic factors, proportion of council housing and free school meals, as measures of social deprivation (28). Census data was used to determine the Jarman and Townsend Indices and local council data was used for the independent factors. Information on TB rates were obtained from a data base in Liverpool, England

(28). Foreign sounding names of all notified TB cases were used to identify ethnic origin. The association between TB and socio-economic deprivation was strongest with both the Jarman and Townsend Indices. The association was less strong with the two independent factors of council housing and free school meals. As no association with individual socio-economic factors could be determined by the Jarman or Townsend Indices, it was uncertain if one socio-economic factor accounted for the identified association between the composite indices and tuberculosis. The decreased association with the two individual-level measures, suggests that not all socio-economic factors in the composite indices had the same level of association with the progression to TB. The correlation between socio-economic deprivation and TB was analyzed with and without defined ethnic origins as part of the Jarman and Townsend indices. Because poverty continued to be correlated with TB despite the exclusion of ethnic origins, ethnicity was not considered to be associated with an increased risk of TB.

Using indices with multiple variables makes it difficult to identify how socio-economic factors can be used to explain the increased TB rates. When multi-factorial socio-economic indices were used to measure the correlation between TB and poverty, a relationship has been identified, but the means by which poverty was associated with TB was not described (9;28). Composite definitions of socio-economic status do not establish if separate factors have a unique association with tuberculosis. The Jarman and Townsend Indices both include overcrowding, for example, as a measure of socio-economic deprivation. Overcrowding may increase TB infection rates and may be a possible cause for the relationship identified between TB and the indices; however, overcrowding may be the only associated socio-economic factor, but this cannot be

identified by a composite index. Examining socio-economic factors separately may provide a clearer explanation of the relationship between TB and poverty.

An ecologic study using United Kingdom census data examined the relationship between rates of change in unemployment, overcrowding, social class, and change in tuberculosis rates in several London boroughs (10). Components of the Townsend or Carstairs index were analyzed as separate variables. A correlation between tuberculosis and overcrowding, but no association with unemployment or social class, was identified. Changes in TB rates were also correlated with the proportion of immigrants in the study population. Separating ethnic origin and socio-economic status was identified as being difficult as recent immigrants may reside in areas of socio-economic deprivation. Potential confounding between overcrowding and ethnic origin was not examined further in this study.

Causality is difficult to determine with ecologic studies. An ecologic study using individual-level data for ethnic origin and census data for socio-economic status was undertaken in Birmingham, England (27). When ethnic origin and socio-economic measures were analyzed together, overcrowding was the only variable independently associated with tuberculosis. When the analysis was separated by ethnic group, overcrowding was only associated with TB in the Caucasian but not the Asian population. No socio-economic measures were correlated with TB in the Asian population. The relationship between TB and socio-economic status may not be generalizable to different ethnic groups. This suggests that ethnic origin may contribute independently to tuberculosis.

A study of children and adolescents in Leeds, United Kingdom, found that ethnicity explained a high proportion of disease independent of a socio-economic deprivation index (32). Although TB in this population was associated with socio-economic deprivation, ethnicity was an independent risk factor suggesting that the relationship between TB and ethnicity was not explained by socio-economic factors or population density. When risks for TB in South Asian children were analyzed separately, socio-economic deprivation and population density were not risk factors for TB. This suggested unidentified risk factors for the development of TB in this population. Different ethnic origins were hypothesized to have different risk factors.

A study using census data by Tocque et al, examined the relationship between immigrants, socio-economic status and tuberculosis in the United Kingdom (29). The authors demonstrated that areas with more immigrants also had higher levels of poverty. Immigration and the Jarman Index had significant correlation with tuberculosis rates. Both decreased socio-economic status and immigration were felt to contribute to increased TB rates although the stronger statistical correlation was with immigration.

Identifying the proportion of immigrants in an ecologic study does not establish if the observed TB rates were due to recent immigration or ethnic origin. Immigrants from countries with high rates of TB are more likely to have been exposed to and infected with TB. Because of increased infection prevalence in individuals from countries with high rates of TB, more individuals have the potential to progress to TB disease. A study completed in the United States compared the effect of poverty and ethnic origin between groups born in the United States (13). Cantwell, et al tried to dissect out the interaction between poverty, ethnic minorities and TB rates (13). TB cases were matched by zip

code to obtain census data for all demographic and socio-economic variables. No individual- level data was utilized. The authors concluded that socio-economic status impacted on the incidence of TB. The strongest association was with overcrowding. A gradient between socio-economic factors and tuberculosis was identified. The authors concluded that socio-economic factors accounted for much of the increased TB risk previously associated with ethnicity in the United States. However, only half of the increased risk of TB associated with ethnicity among US born African Americans, Hispanics and First Nations was accounted for by socio-economic factors. In addition there were different relative risks of TB among these U.S.-born ethnic groups. Socio-economic status did not entirely explain the varying TB rates among different ethnic groups in the United States as differences remained in the risk of developing TB independent of socio-economic status.

Tuberculosis is a two-stage disease – acquisition of infection followed by progression to disease. Variables determining susceptibility to infection may be independent of those that determine progression from infection to disease. All the previously reviewed studies did not distinguish between infection and disease. Overcrowding is hypothesized to be associated with increased transmission of infection, but it may not have a role in the progression to disease. It is not possible to understand the disease process and describe TB epidemiology without examining both the risk factors for infection and progression to disease.

Stead examined racial differences in susceptibility to infection by *Mycobacterium tuberculosis* (31). He compared African Americans and Caucasians with no previously documented TB infection living in racially mixed nursing homes (31). The nursing

homes provided a similar socio-economic environment for both ethnic groups thus controlling for socio-economic factors. Stead found that African Americans were infected with TB more frequently than Caucasians regardless of the ethnic origin of the source case. No difference was identified in the rate of progression from infection to disease in the African American population compared to Caucasians.

Ethnic origin and susceptibility to tuberculosis are discussed in medical textbooks and the literature, however, no studies specifically examine the progression from infection to disease among different ethnic groups. Studies that are the basis for determining increased risk for TB among different ethnic groups do not separate the process of infection from progression to disease and are confounded by socio-economic status. A paper that reviewed the studies supporting variations in ethnic origin and TB rates identified that the majority of studies were designed for other purposes – INH prophylaxis trials, BCG trails, and studies to follow progression from TB infection to disease (33). As these studies were not designed to compare ethnic origin and the progression to TB, the authors stated that results must be carefully considered (33).

In summary, all of the epidemiological risk factors for tuberculosis have not been clearly defined. Conflicting evidence suggests that either ethnic origin or socio-economic factors are associated with increased TB rates. Many of these papers are ecologic studies and as result, identifying interacting risk factors, such as socio-economic factors and ethnic origin has been difficult. Ecologic studies reduce the ability to separately examine socio-economic factors and ethnic origin. It is important to identify if ethnic origin interacts with socio-economic factors or is an independent risk factor for the development of TB among those with a decreased socio-economic status. Although socio-economic

factors can be measured by both individual and area-level measures, area-level measures provide a context or neighbourhood effect in the disease process. TB is a complex disease with a two-stage process each involving different factors. An understanding of TB epidemiology must include a description of the relationship between ethnic origin and socio-economic factors in both stages – acquisition of infection and progression to disease.

2.3 Tuberculosis Epidemiology in Saskatchewan

Tuberculosis rates vary among the ethnic groups in this study. Ethnic groups whose rates of TB are monitored in Saskatchewan include Canadian-born Caucasians, status Indians, and non-Status Indians and Metis. TB rates among the Caucasian population in Saskatchewan have been steadily declining in the latter half of this century. This trend mirrors the Canadian TB rates this century (Figure 2.1). From 1991 to 2001 the average rate of TB among Caucasians in Saskatchewan was 1.4/100 000 (12). Over the past 10 years, TB rates among the Status Indian population has been steadily declining, but remains at a rate higher than the Caucasian population. The average rate from 1991 to 2001 was 99/100 000 (12). Among the non-Status Indian and Metis population, the average TB rate from 1991 to 2001 was 93/100 000 (12). There is significant difference in the rates between these three groups in the province of Saskatchewan. (Figure 2.2)

Figure 2.1 TB Rates in the Saskatchewan Caucasian Population

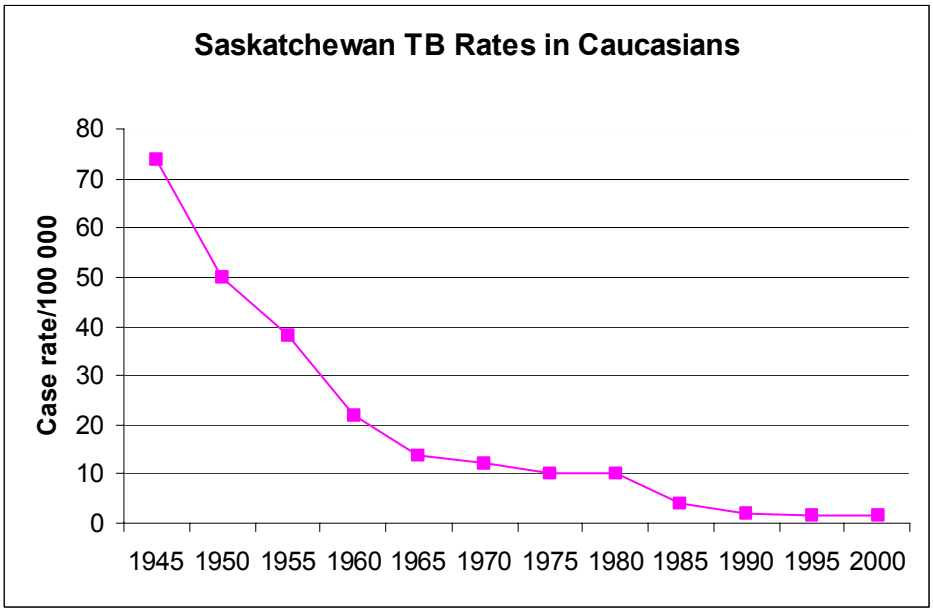
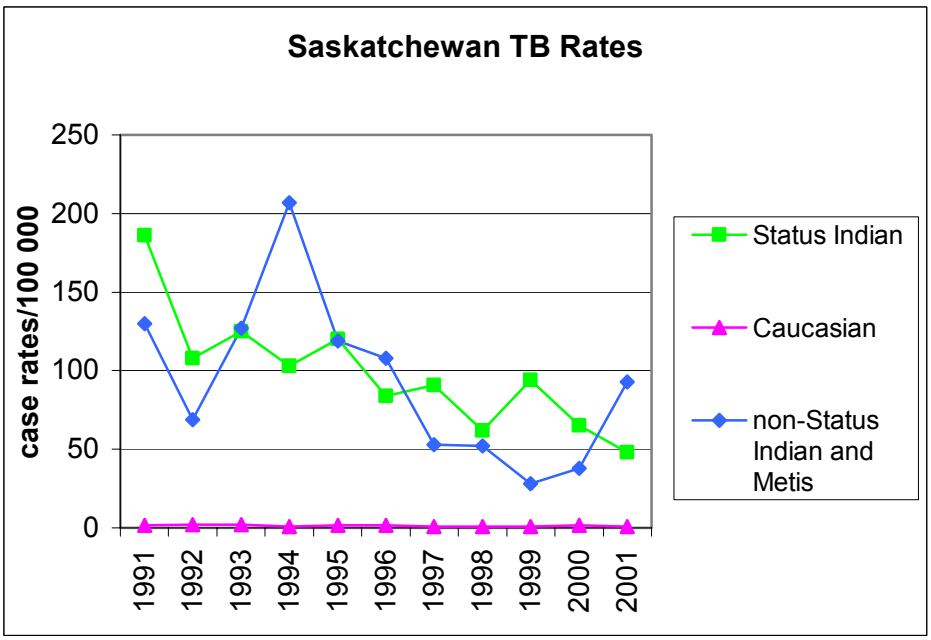


Figure 2.2 – TB Rates in Saskatchewan 1991-2001



TB rates differ among populations in the same geographic region. Various hypothesis, socio-economic status, immigration from high incidence areas, ethnic origin, are all considered to be possible explanations although the epidemiology and biology of the infection and disease process is not well understood. Further examination of the association between ethnic origin and socio-economic status and the progress from TB infection to disease may provide further insight into epidemiology of tuberculosis.

2.4 Saskatchewan TB Control Program

Saskatchewan had a population of 963 155 in 2001 (34). The Saskatchewan TB Control provides a program for controlling TB for the entire province. Elements of tuberculosis control include treatment of TB cases, early case finding by contact tracing and screening known high risk groups, and prevention by treatment of latent TB infection. Control of tuberculosis occurs by stopping transmission (identifying and treating infectious disease), preventing infection (BCG vaccination), and preventing disease among infected high risk individuals (preventive treatment). Identifying those at increased risk of developing TB disease is one priority of TB Control. All disease treatment is community-based directly observed therapy (DOT) with the objective of curing disease, preventing relapse and stopping disease transmission. Early case finding as identified by contact tracing to infectious cases and source tracing of primary cases are essential elements of TB Control. The screening program is undertaken in groups defined as high risk of acquiring TB infection and progressing to disease including those with HIV/AIDS, organ transplantation, and other medical causes of immunosuppression. Tuberculosis screening is performed by completing a tuberculin skin test (TST). Directly

observed preventative therapy is provided for these groups. It is important to identify groups at risk of progressing from infection to disease in order to prevent TB disease.

3.0 Methods

3.1 Saskatchewan TB Control Program Data Base

A computerized record of individuals with a documented tuberculin skin test result was maintained by Saskatchewan TB Control. TST date and results were recorded in the data base. Demographic data, including ethnic origin, date of birth, gender, and home address at the time of tuberculin skin testing, were provided by the individual at the time of testing. For ethnic origin a category, Canadian-born Caucasian, Status Indian, non-Status Indian or Metis, Asian or other (including all other ethnic origins) was selected by the individual. Previous BCG vaccine was recorded in the data base. A unique file number was assigned to individuals with a positive tuberculin skin test who were assessed by a TB Control physician. A hard copy file was maintained in addition to the computer data base. If an individual with a previously positive TST developed TB disease, information was updated in the existing file.

TST was provided for contact tracing and screening of high risk individuals as recommended by Saskatchewan TB Control (35). Contact tracing was provided to all known contacts with more than 10 hours of exposure in the previous 30 days to an infectious TB case and to potential source cases of a primary TB case. Recommended screening of high risk populations included First Nations children on reserve at age 2, school entry, and biannually in kindergarten through Grade 6. Occupations, including health care, day care and corrections employees as well as recent immigrants received

screening by TST. Status Indians were screened for contact tracing, preschool and school on reserve screening. Caucasians and non-Status Indians and Metis received TST screening predominantly as part of contact tracing and occupational health.

Tuberculin skin testing was completed using 5 TU of purified protein derivative (35). A positive tuberculin skin test was defined as $\geq$ 5 mm. for contact to a known infectious TB case or individuals with HIV/AIDS and $\geq$ 10 mm for all other individuals(8).

Once an individual had a positive TST, they were assessed by TB Control physicians to determine if disease was present or if preventative therapy was indicated. If no disease was present, individuals who were 15 years of age and under or significantly immunocompromised (HIV/AIDS, organ transplantation) were offered directly observed prophylactic therapy (DOP). Follow-up, at the request of the individual, local nurse or physicians, was arranged for those not meeting the criteria for DOP. (Appendix 2)

Tuberculosis disease was diagnosed by the presence of cultures positive for *Mycobacterium tuberculosis* from sputum, pleura or any other site. In the absence of a positive culture, a clinical diagnosis was made based on history of contact and/or symptoms, a TST and chest radiograph. The American Thoracic Society classification was used to identify TB cases (36).

3.2 Study Design

All Caucasians, Indians, and non-Status Indian and Metis in the province of Saskatchewan with an initial documented positive tuberculin skin test between January 1, 1986 and January 31, 2002 were identified in the TB control database. A positive tuberculin skin test was defined as $\geq$5mm. Five mm was chosen as a cut off as this was a

positive TST for individuals in close contact with an infectious case, Status Indian children who received on reserve screening, and for immunocompromised individuals.

3.2.1 Exclusion Criteria

Any individual who had a positive tuberculin skin test or TB disease prior to January 1, 1986 was not included in the study. Any individual who was reported to have received BCG vaccine or who had received prophylactic therapy following a positive tuberculin skin test were also not included. If TB disease was diagnosed less than 1 month after an initial positive skin test, this individual was excluded as it was most likely that the individual had TB disease at the time of the initial TST. Individuals with a positive TST skin test who did not show for clinic appointments with TB Control had incomplete documentation and were not included in the study. All TB cases in the province are managed by TB Control. Even if individuals did not show for their initial assessment, if they developed TB disease during the study period, they would be included in the database.

Anyone with incomplete data on ethnic origin, gender, or address and postal code, was not included in the study. As much missing demographic information as possible was completed from the database. For gender, names were reviewed in the computer database to assign gender wherever possible. Addresses were reviewed in the hard copy file if they were unavailable in the computer database. If an address was available, the postal code was identified through the Canada post website (37). If an address or postal code was unavailable, the entry was deleted from the study database. All data from the TB Control data base was de-identified for use in this study.

3.2.2 Canada Census Data

Census data was collected in Canada in 1991 and 1996. A list of available census data was obtained and from this area-level socio-economic factors similar to those frequently utilized in the literature to assess the association between TB and socio-economic factors were chosen (means levels of crowding, education, income, employment, home ownership and lone parent families).

Canada Census data was obtained by using the forward sorting areas, the first three letter/number combination of postal codes. The variables at the enumeration area-level were obtained from Canada Census and linked to the postal code file. Census data from both 1991 and 1996 was obtained and averaged together. Census data from 1991 and 1996 were utilized because this encompassed socio-economic factors for most of the study period duration. Averaging this data together would provide a context of area-level measures for socio-economic factors. The Canada Census dataset was obtained by Rob Anderson, GIS Analyst, Corporate Information and Technology, Saskatchewan Health. A de-identified file containing a subject number and postal code was used for the record linkage to the Canada Census dataset.

3.3 Study Database

3.3.1 Individual-level Data

Once exclusion criteria were met, a final study database was established. Individual-level data available for the study were, as follows; gender, ethnic origin, address at time of TST, BCG status, TST date and result (in mm). Geographic location was coded as a categorical variable for rural, mid-sized urban ≥ 5000 to $< 100\,000$

population and urban ≥100 000 population. Mid-sized urban areas included Prince Albert, Moose Jaw, Swift Current, Yorkton, Weyburn, Estevan, North Battleford, Battleford, Melfort and Melville. The two urban centres ≥ 100 000 were Saskatoon and Regina. Age at time of tuberculin skin testing was categorized into < 19.0 and ≥19.0 years of age (children and adolescents vs adults)

Date of positive tuberculin skin test and date of TB diagnosis were compared and all individuals who developed TB ≤ 1month after a positive TST were excluded. End point of the study was diagnosis of disease or end of study period. We were not able to identify if individuals died from other causes or moved from the province during the study. If TB disease was identified, geographic location and postal code remained the same as the initial positive tuberculin skin test.

3.3.2 Area-Level Measures

Enumeration area data obtained from Canada census included mean family income, mean number of persons per room, participation rate in the labour force and percent distribution of lone parent families, home ownership, highest level of education attained less than grade 9, and completion of a university degree. Total population for all ages and for those greater than age 15 were available for each enumeration area. Persons per room was obtained by dividing the average number of persons per household in the enumeration area by the average number of household rooms. Values for home ownership and lone parent families were expressed as percentage of families in the enumeration who owned their own home or were lone parent families. Grade 9 education was defined as individuals greater than age 15 years whose highest level of completed education was less than grade 9. Percentage of individuals greater than age 15 years who

had completed a university degree in an enumeration area defined the variable of university degree. Once the Canada Census data set was obtained, the de-identified data from the TB control database was matched to the Canada Census data by the primary investigator.

3.4 Analysis

Incidence rate of TB was calculated for the entire study population and for each ethnic group. The statistical analysis was completed for individuals who developed TB disease $\geq$ 1 month after an initial positive TST. Descriptive statistics for both individual and area-level data were completed for the group of individuals who developed TB. Categorical individual data included ethnic origin, gender, age categories of < 19 years and $\geq$ 19 years, and residence. Time to TB disease in months was the only individual continuous variable. Frequencies describing cross tabulations of those with and without TB disease were calculated for the individual categorical variables and statistical comparison using chi-square was completed. The individual-level data was then divided by ethnic origin in the categories of Caucasian, Status Indian, and non-Status Indian and Metis. Descriptive statistics were repeated, including time to TB. Mean, median, range and standard deviation were calculated for the continuous variables. Chi-square was used to compare categorical variables, independent t test for continuous variables with a normal distribution, and non-parametric tests for continuous variables without a normal distribution.

An average of 1991 and 1996 census data for each variable was used for the area-level measures. This included crowding (average number of persons per room), average family income, percentage of home ownership and lone parent families, participation rate

in the labour force and percentage of individuals greater than age 15 who had less than grade 9 or university education as their highest level of completed education. Descriptives, mean $\pm$ SD, for area-level measures were calculated. Presence and absence of TB disease were compared using independent samples t test. Census data was also divided by ethnic origin for those with and without TB and descriptive statistics were completed.

All individual and area-level variables with a significance of $p \leq 0.25$ were included in a Cox proportional hazards and Kaplan-Meier analysis. The Cox proportional hazards model examined the length of time to tuberculosis disease. Data was censored by presence of TB disease or end of study. Time zero was the date of the initial recorded positive tuberculin skin test after January 1, 1986. Censoring date was January 31, 2002. Kaplan-Meier analysis was completed to determine length of time to TB disease as reported by average and median survivals and to enable a graphic representation of the outcome. The analysis was completed for TB disease occurring anytime $\geq$ 1 month after the initial positive TST until then end of the study period.

Assumptions of the Cox proportional hazards analysis included an identifiable starting point at the date of an initially positive TST, and the study end point was defined as presence of disease or end of the study period. Information on loss to follow-up or TB identified outside the province of Saskatchewan during the study period were not available to be included in the study data base. Cox proportional hazards is a nonparametric model. The proportional hazards assumption requires the use of time independent variables (38;39). Covariates in this study were independent of time.

Univariate Cox proportional hazards analysis for individual level variables was completed first followed by analysis of the group-level data. The group-level data was stratified into six equal categories (hexatiles) for the univariate Cox proportional hazards analysis. Hexatiles were chosen in order to examine the effect of stratified socio-economic factors on the progression to TB disease. The association within each area-level measure and progression to TB was examined. Kaplan-Meier survival curves depicting the relationship within for each area-level measure were completed.

Multivariate Cox proportional hazards analysis using main effects, statistically significant variables from the univariate analysis of individual data, was completed. Presence of interaction between individual variables was examined by the use of product terms and adjusted hazard ratios were computed. A multivariate Cox proportional hazards model was completed for the area-level measures that were significant in the univariate analysis. Collinearity was examined for between area-level measures by assessing each variable with the other area-level measures. Collinearity statistics for the area-level measures using linear regression with time to TB as the dependent variable were calculated.

A multivariate Cox regression analysis using both individual and area-level measures that were significant in the univariate analysis was conducted. This multivariate Cox proportional hazards model was calculated using computer generated backward LR model. The –2 log likelihood ratio was calculated for the computer model to confirm the most parsimonious model.

The TB Control database, study database and linkage to Canada census data were completed using Excel (Microsoft Office 2000). All analysis was completed using SPSS version 11.0.

4.0 Results

There were 15 040 individuals, identified as Canadian-born Caucasian, Status Indian, or non-Status Indian and Metis who had an initial positive TST between January 1, 1986 and January 31, 2002 who selected from the Saskatchewan TB Control database. 2123 were excluded because they received prophylactic treatment. 1818 had a positive tuberculin skin test identified within one month of beginning treatment for TB disease. By excluding individuals with BCG, 1842 (31%) of the Caucasian population, 1408 (36%) of the Status Indian population and 729 (16%) of the non-Status Indian and Metis population were excluded. 160 were excluded because of unknown gender, 372 because of missing postal codes or incomplete Canada census data. 7588 individuals met the inclusion criteria and were included in the final study database.

4.1 Descriptive Statistics for Tuberculosis - Individual Data

There were 338 cases of tuberculosis; 34 among the Caucasian population, 183 among the Status Indian population and 121 among the non-Status Indian and Metis. The characteristics of individuals who developed TB disease are described in Tables 4.1 and 4.2. Among all study subjects, the prevalence of TB disease was 4.4% (388/7588). Only 0.8% Caucasians (34/4140) developed TB whereas 6.9% of Status Indians (183/2649) and 15.1% of non-Status Indians and Metis (121/799) progressed to TB.

Incidence density is the total number of new cases during the study period divided by the total person-time of observation. The incidence density for the total study population was 5.6/1000 person years. The incidence density for each of the ethnic groups was, as follows; Caucasians 0.9/1000 person years, Status Indians 7.7/1000 person years and 16.0/1000 person years for non-Status Indian and Metis. The incidence density by gender was 7.2/1000 person years for males and 3.6/1000 person years for females. Children and adolescents less than 19 years of age had an incidence density of 23.3/1000 person years whereas those $\geq$ 19 years of age had an incidence density of 2.6/1000 person years. Among the three geographic locations, incidence density rates were, as follows; rural 8.3/1000 person years, mid sized urban (<100 000) 2.1/100 000 and urban ($\geq$ 100 000) 1.2/1000 person years.

The mean age for those with TB disease was 23.6 years (SD $\pm$ 21.3), median age 17.6 years and age range 0.3 to 90.0 years. For those without TB the mean age was 37.2 ($\pm$ 14.6), median 36.3 and age range 0.0 to 100.8 years. There was a statistically significant difference in the median ages between those with and those without TB disease (Mann Whitney U test $p<0.00001$).

There was a statistically significant difference in gender, ethnic origin, age, and geographic location between individuals with TB and those without. Age and location were documented at the time of positive TST, not at the diagnosis of TB disease (Table 4.1).

The mean age for Caucasians with TB was 49.7 (SD $\pm$ 24.3) (median age 53.3, range 1.5 to 90.0 years). Among those without TB the mean age was 39.8 years (SD$\pm$ 12.9), median age 38.0 and range was 0.0 to 100.8 years. Thirty-four had tuberculosis.

Among Status Indians with TB, the mean age was 20.8 (±19.1), median age was 16.9 and range was 0.3 to 79.2 years. The mean age for those without TB was 34.3 years (SD±16.6), median age was 32.2 and range was 0.2 to 88.8 years. There were 183 cases of TB.

Among non-Status Indians and Metis with TB, the mean age was 20.5 (SD ± 18.5) years, median age was 16.2 and age range was 0.8 to 86.4 years. Among those without TB, the mean age was 32.1 (±16.1), median age was 28.2 and range was 0.3 to 91.7 years. 121 had tuberculosis. Of the non-Status Indian and Metis population (n=799) 47% of the study population were from two rural communities, as identified by postal codes. One community accounted for 32% and the second for 15% of the study group.

More males developed TB disease than females. Non-Status Indians and Metis and Status Indians progressed to TB more frequently than Caucasians. Individuals less than 19 years of age had more TB than those greater than 19 years of age (Table 4.1). There was a higher percentage of TB in rural areas compared to mid-sized urban and urban areas (Table 4.1).

The median time to TB was 2.5 months for Caucasians, 3.0 months for Status Indians and 3.4 months for non-Status Indians and Metis (Table 4.2). The distribution of time was not symmetrically distributed. For all ethnic groups combined, the distribution for the 25^{th} percentile was 1.6 months, 50^{th} percentile was 3.0 months and the 75^{th} percentile was 30.3 months. There was no significant difference in time to TB between the ethnic groups. As Cox proportional hazards analysis is a non-parametric statistical analysis, time to TB could be analyzed as a non-symmetrically distributed outcome.

The different use of TST among each population is reflected in table 4.2. Among Caucasians, more females than males received a TST. This likely reflects the predominance of females in the health care profession. Very few Caucasian children were screened for tuberculosis unless they were a contact to an infectious case. On reserve Status Indian children received regular screening during the preschool and school years. Among non-Status Indian and Metis TST screening would mostly likely occur for contact tracing and occupational screening. Higher percentages of TB were observed in males, rural areas, and children for all ethnic groups (Table 4.2). There was a statistically significant occurrence of TB between Status Indians and non-Status Indians and Metis ($p<0.00001$).

Table 4.1 Descriptive Statistics for TB Disease– Individual Data

	TB Disease N= 338(%)	**No TB Disease N= 7250**	**Total N=7588**	**P**
Gender				
Male	171 (6.5)	2451 (93.5)	2622	
Female	167 (3.4)	4799 (96.6)	4966	<0.0001
Ethnic Origin				
Caucasian	34 (0.8)	4106 (99.2)	4140	
Status Indian	183 (6.9)	2466 (93.1)	2649	
Non-Status Indian and Metis	121 (15.1)	678 (84.9)	799	<0.0001
Location				
Rural	289 (7.4)	3624 (92.6)	3913	
Urban < 100 000	20 (2.0)	996 (98.0)	1016	
Urban > 100 000	29 (1.1)	2630 (98.9)	2659	<0.0001
Age TST (years)				
<19.0	178 (19.8)	719 (80.2)	897	
≥19.1	160 (2.4)	6531 (97.6)	6691	<0.0001

Table 4.2 Descriptive Statistics by Ethnic Origin and TB Disease – Individual Data

	Caucasian N = 4140 (%)			**Status Indian N = 2649 (%)**			**Non-Status Indian and Metis N = 799 (%)**		
	TB 34(0.8)	**No TB** 4106 (92.2)	**p**	**TB** 183 (6.9)	**No TB** 2466	**p**	**TB** 121 (15.1)	**No TB** 678	**p**
Gender									
Male	18(0.4)	1055(25.5)		93(3.5)	1089(41.1)		60(7.5)	307(38.4)	
Female	16(0.4)	3051(73.7)	<0.001	90(3.4)	1377(52.0)	ns	61(7.6)	371(46.4)	ns
Location									
Rural	24(0.6)	1333(32.1)		156(5.9)	1808(68.2)		109(13.6)	483(60.5)	
Urban < 100 000	3(0.07)	657(15.9)		10(0.4)	237(8.9)		7(0.9)	102(12.7)	
Urban > 100 000	7 (0.2)	2116(51.1)	<0.001	17(0.6)	421(15.9)	<0.01	5(0.6)	93(11.6)	<0.01
Age TST									
<19.0	6(0.1)	92(2.2)		100(3.8)	478(18.0)		72(9.0)	142(17.8)	
≥ 19.1	28(0.7)	4014(97.0)	<0.001	83(3.1)	1988(75.0)	<0.001	49(6.1)	529(66.2)	<0.001
Time to TB (months)									
Median	2.5 ± 2.3			3.0 ± 2.8			3.4 ± 3.7		
Range	1.0 to 65.6			1.0 to 188.9			1.0 to 138.9		

4.2 Descriptive Statistics for Tuberculosis – Area-level Measures

The mean total population in each enumeration area (EA) was 725 ± 317. There was a statistically significant difference in the mean value for all socio-economic factors when comparing those with and without TB disease (Table 4.3). Differences were observed in each of the area-level measures among each of the ethnic groups (Table 4.4).

Table 4.3 Descriptive Statistics for TB Disease –Area-level Measures

	No TB Disease N= 7250 Mean ± SE	**TB Disease N=338 Mean ± SE**	**Total N=7588 Mean ± SE**	**p**
Persons per room	0.520 ± 0.003	0.82 ±0.04	0.53 ±0.003	<0.0001
Average Family Income ($)	35 923.29 ± 232.81	22 954.80 ± 795.19	35 384.21 ± 227.58	<0.0001
Own Home (%)	20.6 ± 0.1	18.0 ± 0.6	20.2 ± 0.1	<0.0001
Lone Parent Family (%)	4.25± 0.03	5.8± 0.2	4.3 ± 0.3	<0.0001
<Grade 9 (> age 15) (%)	20.0 ± 0.2	35.9 ± 0.8	20.6 ± 0.2	0.1
University Degree (>age 15) (%)	8.71 ± 0.09	5.8 ± 0.3	8.6 ± 0.09	<0.0001
Labour Force Participation Rate(%)	59.7 ± 0.2	46.0 ±0.7	59.1 ± 0.2	<0.0001

Table 4.4 Descriptive Statistics by Ethnic Origin and TB Disease –Area-level measures

	Caucasian Mean ± SE			**Status Indian Mean ± SE**			**Non-Status Indian and Metis Mean ± SE**		
	TB 34 (0.8)	**No TB** 4106	**p**	**TB** 183 (6.9)	**No TB** 2466	**p**	**TB** 121 (15.1)	**No TB** 678	**p**
Person per room	0.41±0.01	0.83±0.44	ns	0.65 ± 0.01	0.82 ±0.02	<0.0001	0.72 ± 0.01	0.82 ± 0.02	<0.0001
Family Income ($)	44231 ±303	40597 ±4665.43	ns	25775 ±255	23543 ± 875	0 .03	22524 ± 544	18020.29 ± 1052	<0.001
Home Ownership (%)	25.5 ± 0.1	26.0 ±0.02	0 .05	14.9 ± 0.2	10.0 ± 0.8	<0.0001	11.8 ± 0.3	8.4 ± 0.7	<0.0001
Lone Parent Family (%)	3.7 ±0.04	2.9 ±0.4	<0.001	4.47 ± 0.05	4.9 ± 0.2	ns	6.7 ± 0.1	7.9 ± 0.3	ns
<Grade 9 (%)	10.1 ± 0.1	12.7 ± 1.1	ns	18.0 ± 0.2	24.1 ± 0.5	<0.001	17.6 ± 0.3	21.9 ± 0.6	<0.01
University with Degree (%)	8.7 ± 0.1	8.8 ± 2.0	0.08	3.8 ± 0.1	3.2 ± 0.2	ns	3.8 ± 0.1	3.5 ± 0.2	0.02
Labour Force Participation Rate (%)	66.3 ± 0.2	60.3 ± 1.9	<0.01	51.3 ± 0.3	45.2 ± 0.9	<0.0001	50.3± 0.5	43.9 ± 1.0	<0.0001

4.3 Univariate Cox Proportional Hazards Model for Risk of Tuberculosis – Individual Data

Males were more likely to progress to TB disease than females (HR 2.0 95% CI 1.6 – 2.4). Status Indians of all ages were 8.6 times (95% CI 6.9 – 12.5) more likely and Non-Status Indian and Metis 19.2 times (95% CI 13.1 – 28.0) more likely to progress to TB disease than Caucasians. Living in a mid-sized urban area (<100 000) compared to an urban area with a population ≥ 100 000 increased the risk of progressing to TB disease by 7.0 times (95% CI 4.8 – 10.3) whereas living in a rural compared to an urban area increased the risk 1.8 times (95% CI 1.0 – 3.2). Individuals less than 19 years of age were more likely to develop TB than those greater than 19 years of age, across all ethnic groups (HR 9.1 95% CI 7.4 – 11.3) (Table 4.5). This was a univariate analysis with the objective of identifying significant risk factors for the multivariate model. Reference categories are in bold print.

Table 4.5 Univariate Cox Regression for Risk of Tuberculosis – Individual Data

	β	SE(β)	P value	HR	95% CI
Gender					
Female	-	-	-	1.0	-
Male	0.7	0.1	<0.0001	2.0	1.6 – 2.4
Ethnic					
Caucasian	-	-	-	1.0	-
Status Indian	2.2	0.2	<0.0001	8.6	6.0 – 12.5
Non-Status Indian and Metis	2.9	0.2	<0.0001	19.2	13.1 – 28.0
Location					
Urban ≥ 100 000	-	-	-	1.0	-
Urban < 100 000	1.9	0.2	<0.0001	7.0	4.8 – 10.3
Rural	0.6	0.3	0.04	1.8	1.0 – 3.2
Age TST (years)					
≥19.0	-	-	-	1.0	-
< 19.0	2.2	0.1	<0.0001	9.1	7.4 – 11.3

4.4 Univariate Cox Proportional Hazards Model for Risk of Tuberculosis – Area-level measures

An analysis was completed for each of the socio-economic factors and its association with tuberculosis. Each socio-economic factor was divided into six categories of equal size (hexatile). This was done to identify any gradient between socio-economic status and tuberculosis disease. The socio-economic factors were crowding (persons per room), average family income, percent home ownership and lone parent families, percent with highest completed education less than grade 9 and percent completion of a university degree in among those greater than age 15. The lowest hexatile for each variable was the reference category.

4.4.1 Risk of Tuberculosis by Crowding (Persons per room)

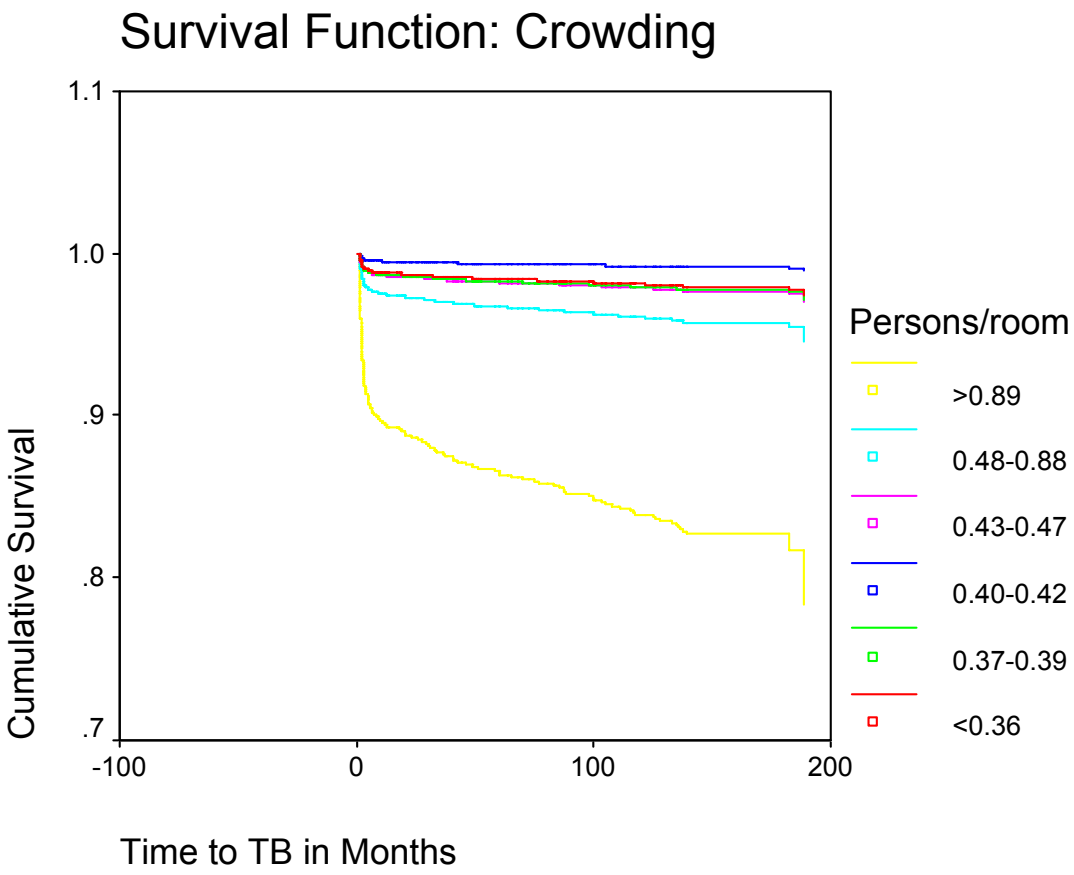
As crowding increased, the risk of developing TB disease appeared to increase. This was seen in the two categories with the highest number of persons per room. (table 4.6). The reference category, <0.36 represented the enumeration areas with the lowest number of people per household. In Figure 4.1, the Kaplan Meier curve, hexatiles 4 and 5, those enumeration areas with the highest number of persons per room, were associated with a higher risk of progressing to TB disease. Those living in enumeration areas identified by the hexatile with the second highest number of persons per room (0.48 to 0.88) had an increased risk of 2.1 compared to those in the hexatile with the fewest number of persons per room. Enumeration areas with the highest number of persons per room had an increased risk of 9.2 of progressing to disease compared to enumeration areas with the lowest numbers of persons per room. In figure 4.1, the reference category, enumeration areas with the lowest number of persons per room was identified by the red line whereas the enumeration areas with the largest number of persons per room was

demonstrated by the yellow line. There is a doubling of the number of individuals per room in the fourth row (0.48 – 0.88 persons per room). It was in this hexatile and the hexatile representing the enumeration area with the largest average number of persons per room that there is an association with an increased risk in the progression to TB.

Table 4.6 Univariate Cox Regression for Risk of Tuberculosis by Hexatiles of Crowding (Persons Per room) (n=7560)

	Persons per room	β	SE (β)	P value	HR	95% CI
0	<0.36					
1	0.37 – 0.39	0.09	0.29	0.8	1.1	0.6 – 1.9
2	0.40 – 0.42	-0.90	0.39	0.02	0.4	0.2 – 0.9
3	0.43 – 0.47	0.12	0.29	0.7	1.1	0.6 – 2.0
4	0.48 – 0.88	0.74	0.26	0.004	2.1	1.3 – 3.5
5	>0.89	2.21	0.22	<0.0001	9.2	5.9 – 14.1

Figure 4.1 Risk of Tuberculosis by Hexatiles of Crowding (Persons Per Room)



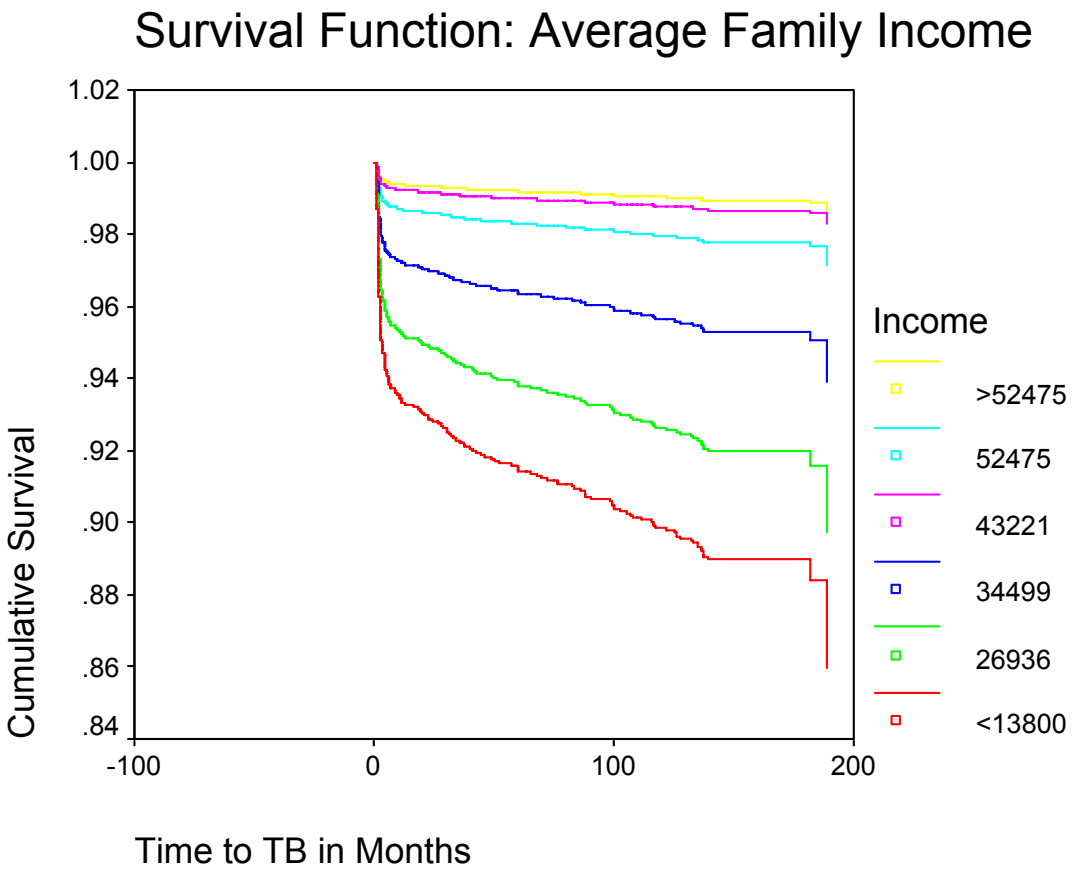
4.4.2 Risk of Tuberculosis by Average Annual Family Income

Enumeration areas with lower average family incomes were more likely to progress from infection to TB disease. Enumeration areas with higher average family incomes provided a protective effect as the risk of TB decreases with higher family income (Table 4.7 and Figure 4.2). As income increased by hexatile, the Kaplan Meier curves demonstrate that the risk of progressing to TB decreased. This decrease was most prominent in the first and second hexatile compared to the reference hexatile. There was less of a gradient in the risk of TB in the highest three income hexatiles.

Table 4.7 Univariate Cox Regression for Risk of Tuberculosis by Hexatiles of Average Annual Family Income (n=7575)

	Family Income ($)	β	SE (β)	P value	HR	95% CI
0	< 13 800.50					
1	13 800.51 – 26 936.50	-0.34	0.14	0.02	0.7	0.5 – 0.9
2	26 936.51 – 34 499.00	-0.88	0.16	<0.0001	0.4	0.3 – 0.6
3	34 399.01 – 43 221.50	-2.16	0.27	<0.0001	0.1	0.07 – 0.2
4	43 221.51 – 52 475.00	-1.65	0.22	<0.0001	0.2	0.1 – 0.3
5	> 52 475.01	-2.39	0.30	<0.0001	0.09	0.05 – 0.17

Figure 4.2 Risk of Tuberculosis by Hexatiles of Average Annual Family Income



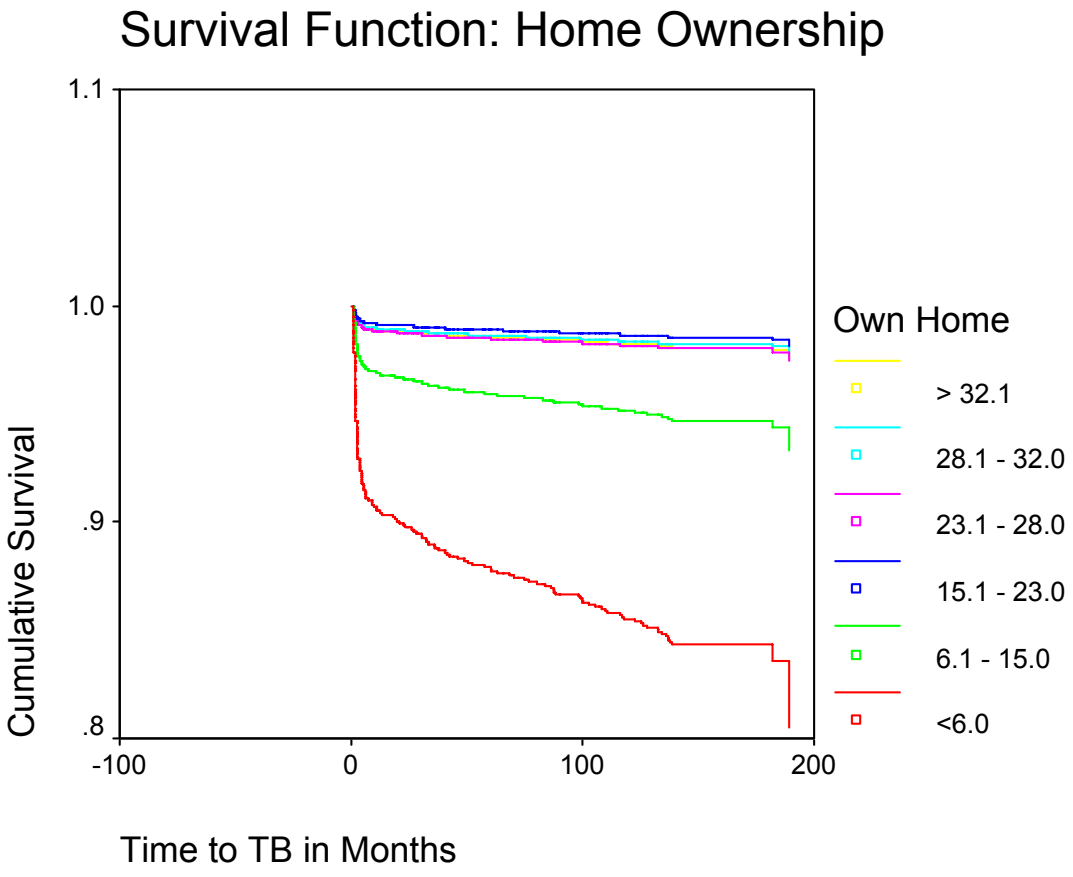
4.4.3 Risk of Tuberculosis by Percentage of Home Ownership

Home ownership decreased the risk of progressing from TB infection to disease (Table 4.8 and Figure 4.3). In enumeration areas with higher percentages of home ownership, there was a further reduction in risk of progressing to TB disease. Enumeration areas with the lowest percentages of home ownership, hexatiles 0 and 1 of the Kaplan Meier survival curves, have the highest risk of progression to TB disease. Among hexatiles 3, 4, and 5, there is no discernable difference in risk of tuberculosis.

Table 4.8 Univariate Cox Regression for Risk of Tuberculosis by Hexatiles of Percentage of Home Ownership (n=7568)

	% Home Ownership	β	SE (β)	P value	HR	95% CI
0	<6.0					
1	6.1 – 15.0	-1.14	0.15	<0.0001	0.32	0.24 – 0.43
2	15.1 – 23.0	-2.44	0.26	<0.0001	0.09	0.05 – 0.15
3	23.1 – 28.0	-2.13	0.23	<0.0001	0.12	0.08 – 0.19
4	28.1 – 32.0	-2.24	0.24	<0.0001	0.11	0.07 – 0.17
5	>32.1	-0.23	0.23	<0.0001	0.11	0.07 – 0.18

Figure 4.3 Risk of Tuberculosis by Hexatiles of Percentage of Home Ownership



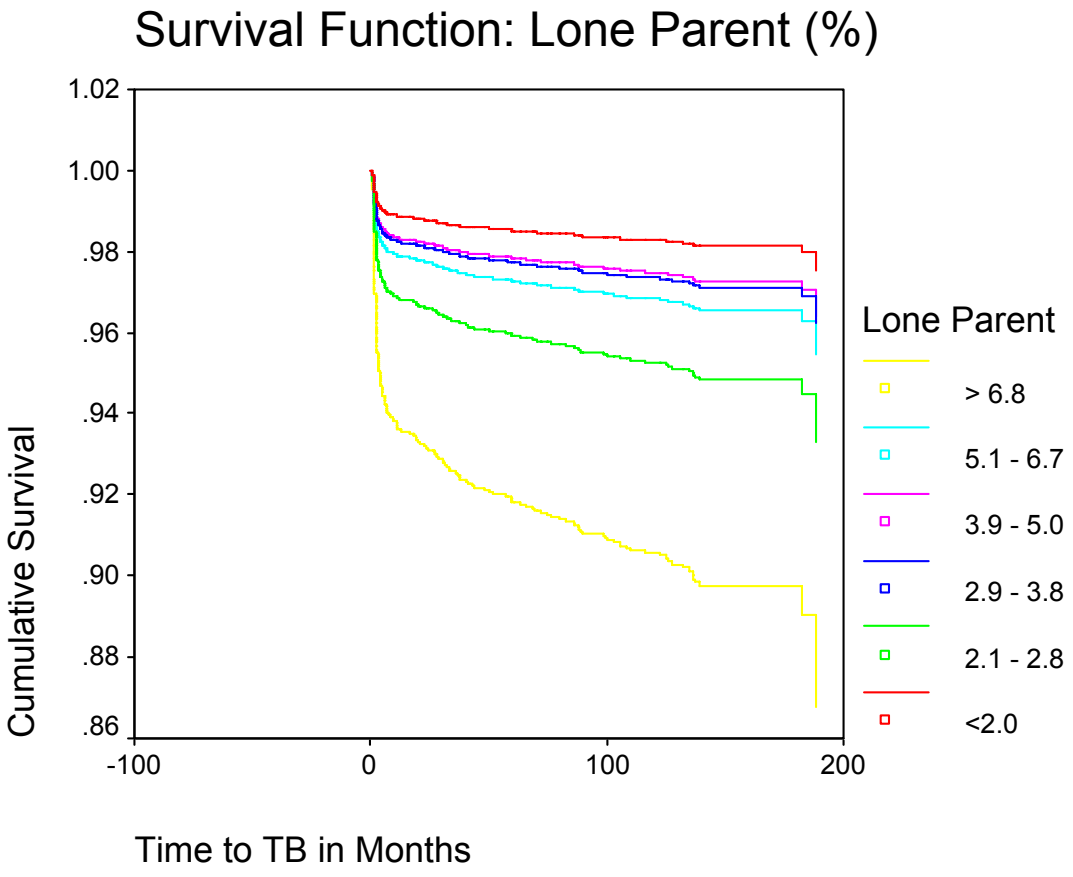
4.4.4 Risk of Tuberculosis by Percentage of Lone Parent Families

The relationship between lone parent families and progression to TB disease was and highest in the first and last hexatiles (Table 4.9 and Figure 4.4). The enumeration areas with highest percentage of lone parent families were more likely to have increased TB rates. No increase in risk was seen in the e2umeration areas with the middle categories of lone parent families (hexatiles 1,2, and 3).

Table 4.9 Univariate Cox Regression for Risk of Tuberculosis by Hexatiles of Percentage of Lone Parent Families (n=7574)

	% Lone Parent Families	β	SE (β)	P value	HR	95% CI
0	<2					
1	2.1 –2.8	1.03	0.27	<0.0001	2.79	1.63 – 4.77
2	2.9 – 3.8	0.44	0.29	0.13	1.55	0.89 – 2.70
3	3.9 –5.0	0.39	0.29	0.19	1.47	0.83 – 2.60
4	5.1 – 6.7	0.62	0.27	0.02	1.86	1.09 – 3.17
5	>6.8	1.74	0.24	<0.0001	5.72	3.60 – 9.09

Figure 4.4 Risk of Tuberculosis by Hexatiles of Percentage of Lone Parent Families



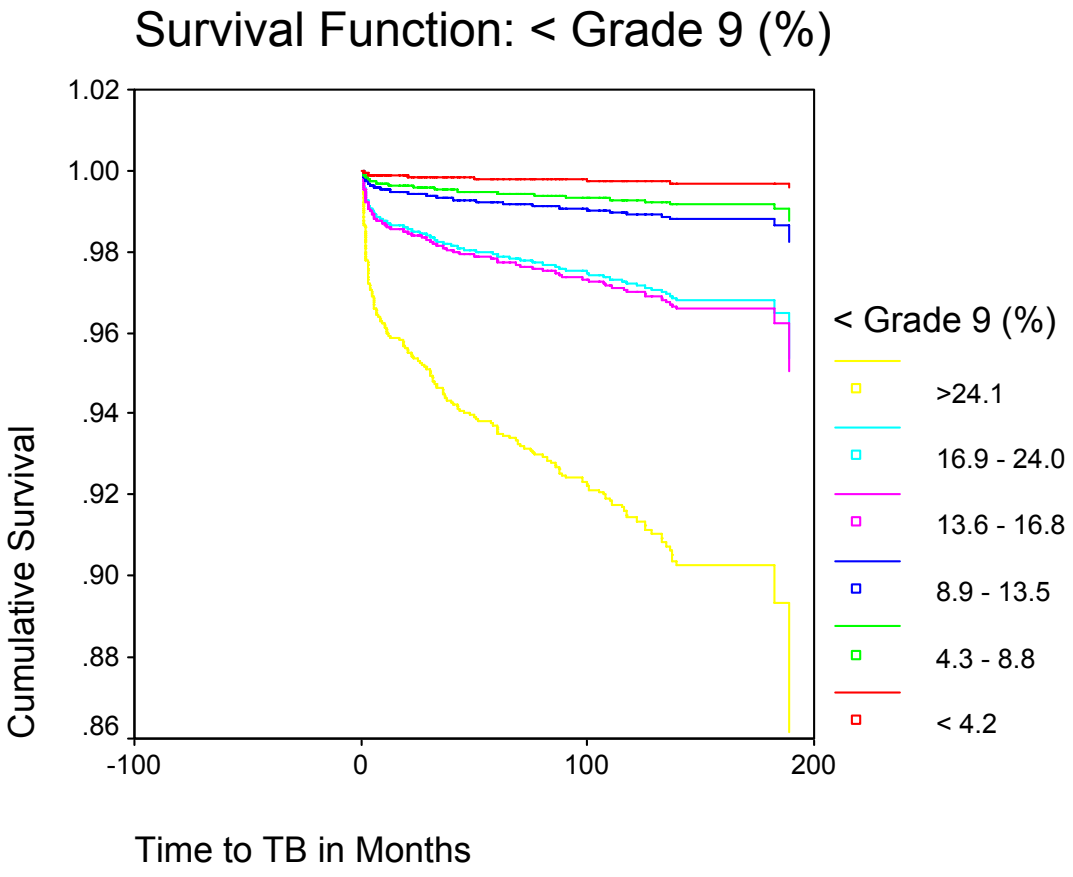
4.5 Risk of Tuberculosis by Hexatiles of Percent Highest Completed Education < Grade 9

In the enumeration areas with the two lowest hexatiles of less than grade 9 as the highest education attained, there was no statistically significant association with the risk of progression to TB (Table 4.10 and Figure 4.5). As the percentage of those who completed less than grade 9 in an enumeration area increased, the risk of progressing from TB infection to disease also increased. The highest risk of progressing to TB occurred with the 5th hexatile, or the highest percentage of individuals > age 15 who had completed less than a grade 9 education.

Table 4.10 Univariate Cox Regression for Risk of Tuberculosis by Hexatiles of Percent Completion of Less Than Grade 9 Education (n=7568)

	% < Grade 9	β	SE (β)	P value	HR	95% CI
0	<4.2					
1	4.3 – 8.8	1.02	0.68	0.13	2.79	0.74 – 10.50
2	8.9 – 13.5	1.39	0.65	0.03	4.01	1.13 – 14.22
3	13.6 – 16.8	2.43	0.60	<0.0001	11.37	3.49 – 37.08
4	16.9 – 24.0	2.37	0.61	<0.0001	10.67	3.26 – 34.90
5	>24.1	3.52	0.59	<0.0001	33.61	10.66 – 105.99

Figure 4.5 Risk of Tuberculosis by Hexatiles of Percent Completion of Less Than Grade 9 Education



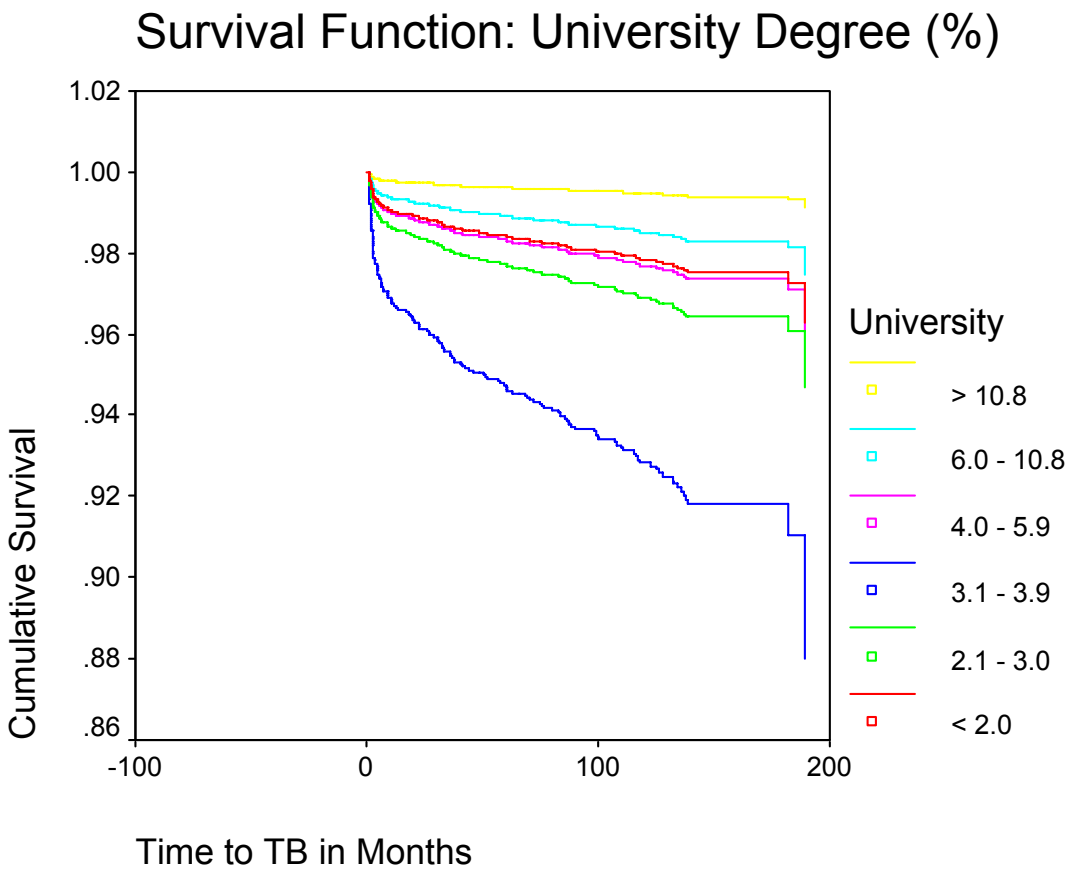
4.4.6 Risk of Tuberculosis by Percentage of Completion of a University Degree

An increased risk of progressing from infection to disease was observed only in the enumeration area with the second lowest percentage (2^{nd} hexatile) of those who had completed a university degree (Table 4.11 and Figure 4.6). The risk of progressing to TB disease was decreased in the enumeration areas with the highest percentage of university degree completion (5^{th} hexatile). No clear pattern of association or gradient was identified between completing university degree and progressing to TB disease.

Table 4.11 Univariate Cox Regression for Risk of Tuberculosis by Hexatiles of Percent Completion of a University Degree (n=7571)

	% University Degree	β	SE (β)	P value	HR	95% CI
0	<2.0					
1	2.1 – 3.0	0.37	0.28	0.18	1.45	0.84 – 2.50
2	3.1 – 3.9	1.23	0.24	<0.0001	3.41	2.13 – 5.45
3	4.0 – 5.9	0.06	0.28	0.82	1.07	0.61 – 1.86
4	6.0 –10.8	-0.38	0.33	0.24	0.68	0.36 – 1.29
5	>10.8	-1.43	0.46	0.002	0.24	0.099 – 0.590

Figure 4.6 Risk of Tuberculosis by Hexatiles of Percent of University Degree Completion



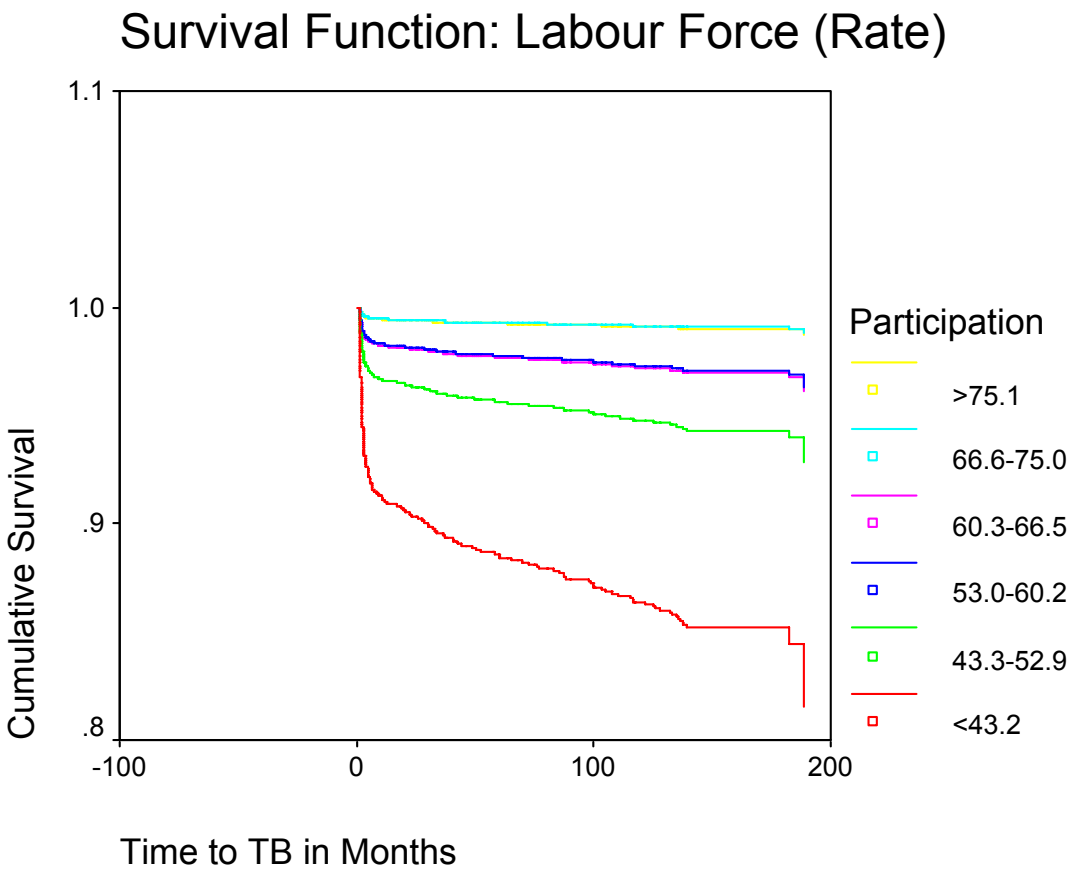
4.4.7 Risk of Tuberculosis by Participation Rate in the Labour Force

Participating in the work force resulted in a decreased risk of progressing from infection to TB for all hexatiles (Table 4.12 and figure 4.7). The reference category, or hexatile representing the enumeration areas with the lowest participation rate in the labour force had diminished the risk the least compared to the other hexatiles. There was minimal difference in the decreased risk of progression to TB for hexatiles 3, 4, and 5.

Table 4.12 Univariate Cox Regression for Risk of Tuberculosis by Hexatiles of Participation Rate in the Labour Force (n=7575)

	Labour Force Participation(%)	β	SE (β)	P value	HR	95% CI
0	<43.2					
1	43.3 – 52.9	-1.01	0.15	<0.0001	0.37	0.28 – 0.49
2	53.0 – 60.2	-1.69	0.19	<0.0001	0.19	0.13 – 0.27
3	60.3 – 66.5	-1.65	0.19	<0.0001	0.19	0.13 – 0.28
4	66.6 – 75.0	-2.83	0.33	<0.0001	0.06	0.03 - 0.11
5	>75.1	-2.79	0.31	<0.0001	0.06	0.03 – 0.11

Figure 4.7 Risk of Tuberculosis by Hexatiles of Participation Rate in the Labour Force



4.5 Multivariate Cox Proportional Hazards Model for Risk of Tuberculosis – Individual and Area-level Measures

In the multivariate Cox proportional hazards model, an analysis with individual level variables was completed first and interaction was assessed. A multivariate Cox proportional hazards analysis was completed using the area-level measures and collinearity was assessed. In the final multivariate model, individual-level variables and socio-economic factors were assessed together. Reference categories used to compare rate of progression to TB (for example, Status Indians compared to Caucasians) are underlined and in bold.

4.5.1 Multivariate Analysis of Individual-level Data and Risk of Tuberculosis

In the multivariate model for individual-level variables, there was no difference between males and females in progression to TB disease (Table 4.13). This is different from both the descriptive statistics (Table 4.1) and univariate analysis (Table 4.5) of gender and progression to disease where males had a statistically significant increased rate of tuberculosis. Status Indians and Non-Status Indian and Metis were more likely to progress to TB disease than Caucasians. This association has persisted through all the analysis. Individuals living in mid-sized urban communities of < 100 000 were more likely to progress to TB compared to individuals in urban centres $\geq$ 100 000. This differs from the descriptive statistics (Table 4.1) in which individuals in rural areas had the highest rates of tuberculosis. Individuals less than 19 years of age continued to be more likely to develop TB than those greater than 19 years of age.

There are only minimal differences in the hazard ratio between the univariate and multivariate Cox proportional hazards analysis with the individual data. As the change was minimal, no interaction was occurring. Interaction terms for individual-level variables were calculated (Table 4.14). Interaction occurred between age of TST and gender, and between ethnic origin and gender. As there was no interaction between age and ethnic origin, this interaction term was not included in the final model. There was no second level interaction between age and gender and ethnic origin combined. When interaction was sought by including interaction terms in the analysis (for example, age and gender), males again had an increased risk of progressing to TB disease (Table 4.14). This suggests that interaction was occurring between gender and age and gender and ethnic origin. Rural and mid-sized urban areas had increased rates of TB compared to urban areas, but unlike the descriptive statistics (Table 4.1), individuals in mid-sized urban areas were more likely to progress to TB than those in rural areas compared to the urban centres. Although no interaction was identified in the analysis, other factors have affected geographic location or residence at the time of positive TST and risk of TB when comparing rural and urban locations.

Table 4.13 Multivariate Cox Regression for Risk of Tuberculosis - Individual Level Variables

	β	SE	Sig	HR	95% CI
Gender					
Female	-	-	-	1.0	-
Male	0.06	0.07	0.5	1.1	0.9 – 1.2
Ethnic Origin					
Caucasian	-	-	-	1.0	-
Status Indian	1.3	0.2	<0.0001	3.7	2.5 – 5.4
Non-Status Indian and Metis	2.0	0.2	<0.0001	7.5	5.9 – 11.4
Location					
Urban ≥ 100 000	-	-	-	1.0	-
Urban <100 000	0.9	0.2	<0.0001	2.3	1.7 – 3.3
Rural	0.1	0.3	0.8	1.1	0.7 – 1.8
Age TST					
≥19.0	-	-	-	1.0	-
<19.0	1.6	0.1	<0.0001	4.8	4.1 – 5.7

Table 4.14 Multivariate Cox Regression for Risk of Tuberculosis – Individual Interaction Terms

	β	SE	Sig	HR	95% CI
Gender					
Female	-	-	-	1.0	0
Male	1.2	0.3	0.004	3.4	1.7 – 6.6
Ethnic Origin					
Caucasian	-	-	-	1.0	-
Status Indian	1.6	0.3	<0.00001	5.1	2.9– 9.0
Non-Status Indian and Metis	2.3	0.3	<0.00001	9.8	5.4 – 17.7
Location					
Urban ≥ 100 000	-	-	-	1.0	-
Urban < 100 000	0.9	0.2	<0.0001	2.5	1.7 – 3.8
Rural	0.2	0.3	0.6	1.8	0.7 – 2.1
Age TST					
≥19.0	-	-	-	1.0	-
<19.0	1.8	0.2	<0.00001	5.9	4.3 – 8.1
Age TST x Gender	-0.6	0.3	0.01	0.6	0.4 – 0.9
Ethnic Origin x Gender					
Male	-0.9	0.4	0.03	0.4	0.2 – 0.9
Female	-0.8	0.4	0.06	0.5	0.2 – 1.0

Table 4.15 Adjusted Hazard Ratios for Risk of Tuberculosis by Individual Data

Age x gender (reference categories: age ≥19.0 years, gender female, Caucasian)

Gender	**Age**	
	<19	≥19
Male	6.1 (4.3 – 8.4)	3.3 (1.7 – 6.5)
Female	3.4 (2.5 – 4.6)	1.9 (0.9 – 4.1)

Ethnic Origin x gender

Ethnic Origin	**Gender**	
	Male	Female
Caucasian	3.3 (1.7 – 6.6)	1.4 (1.0 – 2.0)
Status Indian	5.0 (2.8 – 8.7)	2.1 (1.2 – 3.4)

Ethnic Origin	**Gender**	
	Male	Female
Caucasian	3.3 (1.7 – 6.6)	1.6 (1.0 – 2.5)
Non-Status Indian and Metis	10.0 (5.5 – 18.0)	4.7 (2.7 – 8.3)

Screening protocols used by TB control result in a selection of different age groups and ethnic origins for tuberculin skin testing. Status Indian children receive TST through the on reserve preschool and school screening programs. Status Indians ≥ 19 years of age receive TST through contact tracing to infectious cases and occupational screens. Caucasians and non-Status Indians and Metis would receive screening for occupational screening and as a contact to an infectious case. Selection by age would occur among Status Indians compared to Caucasians and non-Status Indians and Metis. Selection by gender would occur among all those ≥ age 19 years as health care workers are the main indication for occupational screening and the majority of these are female.

As interaction is present between age and gender and gender and ethnic origin, adjusted hazards ratios were calculated (40). The hazard ratio of the main effect for gender was adjusted by the covariate age (Table 4.15). Although females are less likely to progress to TB than males, females < 19 years of age are at increased risk of progressing to disease than those ≥ 19 years. The increased risk of progressing to TB disease is greater in Status Indians and the highest in non-Status Indians and Metis compared to Caucasians for both males and females. The risk of progression by ethnic origin is higher in males than females. Despite varying use of TST, the increased rate of progression from infection to disease persists among ethnic group and age after adjusting for interaction among the hazard ratios.

4.5.2 Multivariate Analysis for Area-level Measures and Risk of Tuberculosis

The risk of progression from infection to disease had the most consistent association with level of education. As the percentage of individuals in enumeration areas whose highest completed education was grade 9or less increased, the risk of progressing to TB disease also increased. This gradient plateaued in the three hexatiles with the highest percentage of individuals whose highest academic completion was grade 9 or less with no increase in the risk of progression to TB. This differed from the univariate analysis where there continued to be a gradient with enumeration areas with increasing percentages of individuals whose highest level of education was completion of grade 9.and progression to disease (Table 4.10). Progression to TB disease was associated with obtaining a university degree but there was no gradient throughout the different enumeration area categories. In the univariate analysis there was an inconsistent association between university degree and developing TB disease.

In the enumeration areas with the two lowest participation rates in the work force, there was decreased risk of progression to TB disease. In the univariate analysis, enumeration areas with higher participation rates had a decreased risk of progression to TB disease. This differed from the univariate analysis. Other variables, both individual and area-level measures included in the analysis, may have resulted in this difference.

A greater number of persons per room was associated with an increased risk of developing TB disease only for the two highest hexatiles. A similar association was seen in the univariate analysis. A greater proportion of home ownership consistently decreased the risk of progressing to TB disease. A greater proportion of lone parent families were associated with an increased risk of TB disease in the univariate analysis. Differences in association between the hexatiles of the area-level measures may be occurring with the addition of all the area-level measures and the individual measures to the Cox regression analysis.

The multivariate Cox regression was the computer generated model developed by using backward likelihood ratio. In calculating – 2 log likelihood ratio, this was the most parsimonious model or model with the fewest number of variables without a change in statistical significance of the final model. Only family income was removed from the model. Collinearity was assessed in the model and none was identified.

Table 4.16 Multivariate Cox Regression for Risk of Tuberculosis by Area-level Measures (Main Effects)

	β	SE	Sig	HR	95% CI
Persons per room					
0.37 – 0.39	-0.14	0.31	0.65	0.87	0.47 – 1.61
0.40 – 0.42	-0.74	0.40	0.07	0.48	0.22 – 1.05
0.43 – 0.47	0.41	0.32	0.20	1.51	0.81 – 2.83
0.48 – 0.88	0.59	0.28	0.04	1.81	1.04 – 3.15
>0.89	1.35	0.33	<0.0001	3.86	2.02 - 7.36
< Grade 9 (%)					
4.3 – 8.8	0.68	0.57	0.23	1.97	0.65 – 6.03
8.9 –13.5	1.56	0.67	0.02	4.78	1.30 – 17.57
13.6 – 16.8	2.19	0.71	0.002	8.97	2.23 – 36.05
16.9 – 24.0	2.22	0.73	0.002	9.20	2.21 – 38.28
>24.1	2.08	0.75	0.006	7.98	1.82 – 34.89
University with degree (%)	0.81	0.25	0.001	2.26	1.38 – 3.71
2.1 – 3.0	1.45	0.24	<0.0001	4.24	2.65 – 6.81
3.1 – 3.9	1.05	0.29	<0.001	2.85	1.62 – 5.20
4.0 – 5.9	1.08	0.29	<0.002	2.93	1.66 – 5.20
6.0 –10.8	1.18	0.49	0.016	3.24	1.25 – 8.44
>10.8					
Participation rate					
43.18 – 52.90	-0.80	0.18	<0.0001	0.45	0.31 – 0.64
52.91 – 60.15	-0.69	0.30	0.02	0.50	0.28 – 0.91
60.16 – 66.50	-0.25	0.30	0.39	0.78	0.43 – 1.39
66.51 – 75.00	-0.73	0.47	0.12	0.48	0.19 – 1.21
>75.01	-0.24	0.59	0.69	0.79	0.25 – 2.49
Home Ownership					
7 – 15	0.04	0.33	0.91	1.04	0.55 – 1.96
16 – 23	-0.84	0.44	0.06	0.43	0.18 – 1.04
24 – 28	0.24	0.48	0.62	1.27	0.49 – 3.28
29 – 32	0.24	0.49	0.61	1.29	0.49 – 3.38
>33	-0.26	0.50	0.60	0.77	0.29 – 2.05
Lone Parent Family					
2.1 – 2.8	0.19	0.34	0.57	1.21	0.62 – 2.35
2.9 – 3.8	-0.56	0.39	0.16	0.57	0.27 – 1.23
3.9 – 5.0	0.19	0.35	0.58	1.21	0.61 – 2.41
5.1 – 6.7	0.45	0.35	0.20	1.56	0.79 – 3.09
>6.8	0.56	0.36	0.12	1.75	0.86 – 3.56

4.5.3 Multivariate Cox Regression for Risk of Tuberculosis by Individual Data and Area-level Measures (Main Effects)

Ethnic origin was statistically significant and strongly associated with the progression from infection to TB disease (Table 4.16). The risk of progressing to TB disease did not meaningfully change from the full individual model (HR 5.6 95% CI 3.1 – 10.2) to the full model with individual and area-level data (HR 5.1 95% CI 3.0 -8.6). The risk of progressing to TB disease decreased in the Non-Status Indian and Metis population from the multivariate individual-level model (HR 7.5 95% CI 5.9 – 11.4) to the full multivariate model combining individual and enumeration area data (HR 7.4 95% CI 4.1 – 13.3). No interaction was identified. Being less than 19 years of age continued to be associated with the increased risk of progression to TB disease. Gender was not a significant contributor to developing TB disease.

Among the socio-economic variables, persons per room, family income, home ownership and lone parent families did not have a consistent association with progression to TB disease. Only one category in the persons per room variable was significantly associated with decreased risk of TB disease. These factors seemed to be less significant once individual variables were included in the model. In regards to education, achieving less than grade 9 as the highest level of education was no longer identified as significant in the computer-generated model once age was added to the model. Holding a university degree was the only area-level measure that was consistently associated with an increased risk of progression from TB infection to disease in all enumeration areas except that with the highest percentage of degree holders. Participation rate had a statistically significant association with a decreased risk of TB disease.

4.17 Multivariate Cox Regression for Risk of Tuberculosis by Individual and Area-level Measures (Main Effects)

	β	SE	Sig	HR	95% CI
Gender	-0.26	0.13	0.052	0.77	0.60 – 1.00
Ethnic Origin					
Caucasian	-	-	-	1.0	-
Status Indian	1.62	0.27	<0.00001	5.1	3.0 – 8.6
Non-Status Indian and Metis	2.00	0.30	<0.00001	7.4	4.1 – 13.3
Age Category					
≥19.0	-	-	-	1.0	-
<19.0	1.3	0.1	<0.00001	3.7	2.8 – 4.8
Person per room					
0.37 – 0.39	-0.59	0.37	0.11	0.56	0.27 – 1.14
0.40 – 0.42	-1.47	0.48	0.002	0.23	0.09 – 0.59
0.43 – 0.47	-0.36	0.43	0.40	0.70	0.30 – 1.63
0.48 – 0.88	-0.71	0.44	0.11	0.49	0.21 – 1.17
>0.89	-0.69	0.49	0.16	0.50	0.19 – 1.31
Family Income					
13800.51 – 26936.50	0.02	0.29	0.95	1.02	0.58 – 1.79
26936.51 – 34499.00	-0.04	0.28	0.89	0.96	0.09 – 0.59
34399.01 – 43221.50	-0.82	0.43	0.06	0.44	0.30 – 1.63
43222.51 – 52475.00	0.29	0.44	0.52	1.33	0.21 – 1.17
>52475.01	-0.17	0.56	0.76	0.85	0.19 – 1.31
Home Ownership (%)					
7 – 15	0.06	0.34	0.85	1.06	0.55 – 2.06
16 – 23	-0.63	0.45	0.16	0.54	0.22 – 1.29
24 – 28	0.24	0.52	0.64	1.28	0.47 – 3.50
29 – 32	0.73	0.53	0.16	2.08	0.74 – 5.82
>33	-0.12	0.53	0.82	0.89	0.32 – 2.50
Lone Parent (%)					
2.1 – 2.8	0.06	0.38	0.87	1.06	0.51 – 2.20
2.9 – 3.8	-0.60	0.40	0.14	0.55	0.25 – 1.21
3.9 – 5.0	0.20	0.37	0.59	1.22	0.60 – 2.49
5.1 – 6.7	0.40	0.34	0.25	1.48	0.76 – 2.89
>6.8	0.46	0.37	0.22	1.58	0.32 – 2.50
University Degree (%)					
2.1 – 3.0	0.91	0.33	0.006	2.48	1.30 – 4.72
3.1 – 3.9	0.85	0.31	0.007	2.34	1.26 – 4.33
4.0 – 5.9	0.98	0.33	0.003	2.65	1.39 – 5.07
6.0 –10.8	0.83	0.35	0.02	2.30	1.15 – 4.61
>10.8	0.86	0.53	0.10	2.35	0.84 – 6.58
Participation Rate					
43.18 – 52.90	-0.95	0.27	0.0004	0.39	0.23 – 0.66
52.91 – 60.15	-0.84	0.32	0.01	0.43	0.23 – 0.82
60.16 – 66.50	-0.70	0.33	0.03	0.50	0.26 – 0.95
66.51 – 75.00	-1.73	0.47	0.0003	0.18	0.07 – 0.45
>75.01	-1.40	0.49	0.004	0.25	0.10 – 0.64

Individual-level variables age and ethnic origin have the strongest association with the progression from TB infection to disease. None of the hexatiles for area-level measures demonstrated a gradient in association with progression to TB disease. Possessing a university degree had a statistically significant association with progression to TB disease, but no gradient was identified. Participation in the labour force decreased the risk of TB. Individual-level factors are more significant than socio-economic factors, as measured by area-level measures, in predicting who is at risk of developing TB disease.

As all individual and area-level measures were statistically significant in the univariate analysis, all were included in the final multivariate Cox proportional hazards regression analysis. Time to TB and presence of TB were outcome measures. By including all of the individual and area-level measures, their effects on each other are included in the analysis. Despite differences in area-level measures between each ethnic group identified in the descriptive statistics (Tables 4.2 and 4.4), ethnic origin and age continue to contribute to the risk of progression to TB disease despite differences in individual and area-level measures for these two groups. The Cox regression main effects analysis identifies Status Indians and non-Status Indians and Metis and all individuals less than 19 years of age, considering the variation in socio-economic factors as represented by area-level measures, as having an increased risk of progression to TB disease.

5.0 Discussion

Of the 7588 individuals with a positive tuberculin skin test 338 (4.5%) developed TB. There were 34 cases of TB in the Caucasian population, 183 in the Status Indian population and 121 in the non-Status Indians and Metis populations. According to TB literature, 10% of all individuals will develop TB (1;2). In this study, the progression to TB varied between ethnic origins. Thirty-four out of 4140 (0.8%) of Caucasians developed TB disease, 183 out of 2649 (6.9%) of Status Indians and 121 out of 799 (15.1%) non-Status Indians and Metis. The rate of progression to TB reported in the literature did not apply uniformly to the ethnic groups.

Progression to TB was 5.6/1000 person years for the entire study population. The incidence rate for Caucasians was 0.9/1000 person years, 7.7/1000 person years for Status Indians and 16.0/1000 person years for non-Status Indians and Metis. Status Indian and non-Status Indian and Metis populations have an increased risk compared to the Caucasian population of progressing to disease. Both age < 19 years and Aboriginal ethnic origin resulted in a shorter time to progress from TB infection to disease.

The rate of progression to TB disease for other individual level variables, including gender, age and geographic location also differed. Males developed TB disease at a rate of 7.2/1000 person years and females at 3.6/1000 person years. Individuals < 19 years of age developed TB at a rate of 23.3/1000 person years while individuals $\geq$ age 19 developed TB at a rate of 2.6/1000 person years. The rate of TB disease was 8.3/1000 person years in rural areas, 2.1/1000 person years in mid-sized urban areas and 1.2/1000 person years in urban areas.

There was no consistent gradient of association between socio-economic factors and progression to TB disease. In the multivariate Cox proportional hazards model, each socio-economic factor had varying associations with progression to TB disease $\geq$ 1 month following a positive TST, but no consistent gradient was identified across the hexatiles of area-level measures. In the Cox proportional hazards model, including individual and area-level measures, the strongest association was with ethnic origin (HR 5.1 95% CI 3.0 – 8.6 for Status Indians and HR 7.4 95% CI 4.1-13.3 for non-Status Indians and Metis both compared to Caucasians). Completing a university degree increased this risk of progression to TB. Participation in the labour force was associated with a statistically significant decrease in the risk of progression to TB. Interaction between ethnic origin and socio-economic status was investigated in the multivariable Cox proportional hazards analysis and it was not significantly associated with the progression to TB disease. No collinearity was identified among the area-level measures.

Variations occurred in the association between individual and area-level measures and progression to TB in both the univariate and multivariate Cox regression analysis. Gender was significant in the univariate analysis and in the multivariate analysis gender modified the effect of other individual variables. Gender was not significant ($p=0.052$) in the final multivariate analysis that included both individual and area-level measures. In the adjusted hazard ratios, males were at increased risk of TB compared to females (Table 4.15). Individuals from rural areas had increased rates of TB identified by incidence rates, but individuals from mid-sized urban communities had an increased risk of progressing to TB compared to individuals from rural areas, although both had increased risks compared to urban areas. In the final Cox regression analysis, geographic location

was not statistically significant. Although no interaction was identified, the addition of other variables to the final model may have confounded geographic location changing the identified risk of association in rural compared to mid-sized urban areas making it not a statistically significant variable in the final analysis. The area-level measure of percent of individuals in the enumeration areas whose highest level of academic completion of less than grade 9 was significant in the Cox regression analysis of area-level measures, but was not significant in the final analysis that included both individual and area-level measures. Interaction with other area-level measures and individual variables, although not significant, may have resulted in this change of statistical significance.

Tuberculosis literature identifies that the life time risk of progressing from TB infection to disease is 10% (2). In our study 0.8% of Caucasians, 6.9% of Status Indians and 15% of Non-Status Indian and Metis developed TB. The risk of progressing from infection to disease varied across the three ethnic groups. This difference in susceptibility to tuberculosis by ethnic origin remained in the multivariate model including group level socio-economic variables. The descriptive data identify that there was a difference in socio-economic factors between Caucasians, Status Indians and non-Status Indians and Metis. Despite this difference, no consistent gradient was associated with the hexatiles of socio-economic factors and the strongest association remained with ethnic origin. This suggests that ethnic origin is an important contributing factor in the progression from TB infection to disease distinct from socio-economic factors. Socio-economic factors, as measured by area-level measures, do not provide a consistent gradient in association with progression to TB disease.

There are limitations to this study. Because area-level measures were used, ecologic fallacy is of concern. Ecologic fallacy occurs when only area-level data are used in an analysis as results ascribed to group members may not apply to individuals. This may be particularly true in rural areas where one postal code applies to an entire town as well as the surrounding rural area. Ecologic fallacy makes it difficult to delineate the association between TB and its environment. Clustering of disease may be occurring in geographic areas either by community or family groups, resulting in an overestimation of the risk factors for progression to TB (3). Determining how socio-economic factors impact the progression to TB disease, as individual measures or a neighbourhood effect, will continue to provide a challenge in defining the factors that determine TB epidemiology. As previously discussed, area-level measures of socio-economic factors can be used to provide a limited proxy for individual measures. Area-level measures can also be used to identify a neighbourhood effect or define a socio-economic environment that may impact on the progression from infection to disease. Identifying which measures, individual or area-level measures, and how they interact in the acquisition of infection and progression to disease will always be challenging because of their close association.

The Canada Census data was averaged for the years 1991 and 1996. Although 2001 study data was not included, the average of 1991 and 1996 census data most likely encompassed the main duration of time in the progression from infection to disease. A positive TST does not identify when TB infection initially. A positive TST only identified that TB infection has occurred since the last negative TST. If this was a subject's first TST, for example, infection could have occurred at any point in time in that

subject's life prior to the skin test. Identifying the time period between infection and disease is difficult. Potentially past census data, as found in 1991 and 1996 may be more representative of socio-economic environment than present socio-economic data as time of infection is generally unknown. An alternate approach to this study could have been to match the available census data to the time period between the onset of infection and progression to disease. This would have been difficult to accurately do as the TST result is documented when it is initially positive, not necessarily when infection first occurs. The linkage between onset of infection and progression to disease would not have been accurate. Average the 1991 and 1996 Canada census data provided the area-level measures for the socio-economic environment during the study period.

There are other limitations to this study. One limitation is follow-up. We could not determine if individuals in the study died or moved from the province before the end of the study period. Not being able to identify subjects lost to follow-up could make the denominator larger than it truly is, underestimating the rate of TB. Anyone who moved out of province and then developed TB would not have been included in this study. This may underestimate the numerator of the rate of progression to TB disease. Missing individual data, postal code, gender, and ethnic origin resulted in a loss of some individuals in the study. Excluding individuals because of missing data would also decrease the size of the denominator and overestimating the rate of TB. This underestimation of TB rates was presumed to be minimal.

A positive tuberculin skin test indicating infection was defined as 5 mm for this study. More than one cut-point for a positive TST is generally used (8). Five mm is considered positive for contacts to an infectious case and immunocompromised

individuals. In Saskatchewan, children participating in screening programs were considered to have a positive tuberculin skin test at 5mm until 1999. After 1999, 10 mm was considered a positive TST for children enrolled in a screening program. For all immunocompetent individuals with no known contact to infectious tuberculosis, 10 mm is a positive tuberculin skin test. By using a cut-point of 5mm, all those individuals at increased risk of progressing to disease, recent contact, children, and immunocompromised individuals, are included in the study. These individuals may increase the observed rate of progression from infection to disease. Some individuals without tuberculosis infection would be included in this study by using a cut-point of 5mm. Individuals without infection would not progress to disease and this would decrease the observed rate of progression to TB disease. By using a cut-point of 5 mm, individuals at increased risk and no risk of progressing to disease are both included in the study. The direction of bias is difficult to determine.

Individuals who had previously received BCG vaccination were not included in the study for two reasons. The first is BCG vaccination could have provided a false positive TST. Theoretically, this is unlikely as tuberculin skin test response to BCG wanes over time (2). A positive tuberculin skin test is more likely to be due to infection if the BCG vaccination is remote. The second reason those who received BCG vaccination were excluded is that BCG may influence the progression from infection to disease. Many studies have been completed on the efficacy of BCG vaccination and the general conclusion is that BCG vaccination does not prevent progression to TB, but decreases the risk of the more severe manifestations of TB, meningitis and disseminated TB(1;2;8). By excluding individuals with a previous BCG, 31% of the Caucasian, 36%

of the Status Indian and 16% of the non-Status Indian and Metis population were not included in the study population. Use of BCG vaccination was similar in both the Caucasian and Status Indian study population. An analysis was done for individual-level variables, including an independent variable for BCG vaccination. A history of BCG vaccination was not statistically significant in the progression from infection to disease. Excluding individuals who had received BCG vaccination did not change the study outcome.

Cox regression analysis in this thesis was utilized because of the time dependent outcome, time to TB. Cox regression models analyze both individual and area-level measures; however, area-level measures are analyzed as fixed effects i.e. the same characteristic is attributed to all individuals in the area-level measures. For example, all individuals in an enumeration area would be attributed as having the same average family income. An alternative method for analyzing area-level measures involves assessing variability within an area, such as an enumeration area. In this analysis, the area-level measures are analyzed as random and not fixed effects. Frailty models, which are an extension of Cox Regression models, analyze a time to event outcome utilizing area-level measures as random effects (41). This was attempted by using SAS by a member of the thesis committee, but no outcome was available as the models did not converge. Reanalysis of the data will be reviewed as software for completing frailty models and analyzing random effects for area-level measures with a time dependent outcome is more readily available.

Tuberculosis is a two-stage disease. Which factors determine whether or not TB infection is established after exposure and why only a proportion of individuals progress

to disease has never been clearly defined. The T lymphocyte immune response is known to be an important entity in the outcome in each of these stages (2) . It is estimated that one-third of the world's population is infected with TB bacteria yet only a small percentage progress to TB disease (6). Some factors that advance the progression to TB disease have been identified and include HIV/AIDS and immunosuppressive medications for organ transplantation. Each of these selectively suppress the T lymphocyte component of the immune system. The majority of individuals with TB disease in Saskatchewan do not have HIV/AIDS nor have they received immunosuppressive medications (12). Other disease processes such as diabetes mellitus, chronic renal disease, and silicosis, have been reported to contribute to the development of TB; however, this would still contribute to only a small portion of individuals who progress to TB (2). The variations in TB rates among different ethnic groups and the lack of understanding as to why the majority of individuals progress to TB disease may make socio-economic factors an appealing hypothesis. If socio-economic factors are important in TB disease, an explanation of their biologic role in each of the two stages of TB would need to be identified. Other hypotheses, including a potential role for ethnic origin and age, also need to be considered. One theory proposed by Grigg suggests that a TB epidemic occurs in immune naive populations to tuberculosis. As the immune response of a population evolves to become more effective, the TB epidemic begins to decline (3). This epidemic takes several centuries to occur within a population. Although socio-economic factors, ethnic origin and susceptibility to tuberculosis are commented on in medical textbooks and the literature, the close relationship between ethnic origin and

socio-economic environment and contribution of each to the two stage process of tuberculosis are not well defined and other hypothesis need to be considered (33).

Ecologic studies in the United Kingdom have identified that overcrowding is the only single socio-economic factor that is repeatedly associated with TB (10;27). Ecologic studies using composite indices that include overcrowding as a measure make it difficult to ascertain if socio-economic factors other than overcrowding are associated with progression to TB disease (9;28). Ethnic origin has been associated with tuberculosis but confounding with socio-economic factors could not be discerned by the study (32). No distinction was made between acquisition of infection and progression to disease in any of these ecologic studies. In our study the hexatile with the enumeration areas that had the largest number of people per room (>0.89) was strongly associated with increased progression from infection to disease in the univariate analysis (HR 9.2, 95% CI 5.9 – 14.1). In the multivariate analysis of TB $\geq$ 1 month following a documented positive TST, the hexatile with the enumeration areas having the largest number of persons per room remained significant although no gradient was identified (HR 3.9 95% CI 2.0 – 7.4). The effect of overcrowding, *i.e.* recent transmission, may still have been seen in those who developed TB disease during the study period. Overcrowding presumably increases disease transmission, but it does not have a clear role in the progression from infection to disease.

One study has focused on the relationship between socio-economic status and ethnic origin as an explanation for increased TB rates among different ethnic populations. Cantwell *et al* observed the increased rates of TB among ethnic minorities in the United States and hypothesised that this increased risk was due to confounding by socio-

economic factors (13). Individual TB cases were matched by zip code for age group, gender, and ethnicity. Socio-economic factors of crowding, income, poverty, public assistance, unemployment and education were also matched by zip code. Outcome was the presence of tuberculosis disease and prior existence of TB infection was not documented. The only individual-level data was zip code. A health gradient was identified for each of the socio-economic factors for all ethnic groups combined; however this gradient was not consistently statistically significant and the clinical importance seemed minimal as OR ranged from 0.9 to 1.5. Confidence intervals were not recorded. The lowest quartile representing the most crowding had the strongest association with TB among all ethnic groups (OR 1.7). Absolute case rates were not significantly different for any of the socio-economic indicators other than crowding. Overcrowding was the socio-economic risk factor that had the strongest association with TB disease in all ethnic groups. In the multivariate analysis adjustment for all of the socio-economic factors decreased the risk of developing TB only by half among the different ethnic groups. Although the authors concluded that socio-economic factors "accounted for much of the increased risk of TB previously associated with race/ethnicity", they did acknowledge that these factors accounted for half of the increased risk (13). The increased rates of TB disease could not completely be accounted for by socio-economic factors, particularly crowding. Crowding most likely contributes to the transmission of infection. This study did not separately examine the contribution of socio-economic factors to the two-stage disease process.

Only one study has tried to separate out the first stage of tuberculosis disease, acquisition of infection, and the role of ethnic origin. Stead *et al* examined the

relationship between acquisition of infection and ethnic origin while adjusting for socio-economic factors (31). Stead *et al* noted a difference in tuberculosis rates between Caucasians and African Americans in nursing homes in the southern United States. This difference had been attributed to social factors, but Stead *et al* theorized that the difference in tuberculosis rates was due to variation in ethnic susceptibility to tuberculosis infection. A database was maintained as part of the tuberculosis control program for nursing homes in Arkansas. New admissions to nursing homes were screened with a two-step tuberculin test. All individuals with TB infection acquired 60 days after admission to a nursing home, as defined by a conversion to a positive TST following a negative two-step tuberculin test were included in the study. Socio-economic factors were controlled for as all subjects lived in racially integrated nursing homes with the same social environment. The relative risk of African Americans acquiring TB infection compared to Caucasians was 1.9 (95% CI 1.7 – 2.1 n=25 398). They found that African Americans were more readily infected with *Mycobacterium tuberculosis* regardless of the ethnic origin of the source case.

Among the nursing home residents not treated for TB infection, there was no association between ethnic origin and progression to TB disease identified in this study. Although no differences were identified in progression to disease, the duration of follow-up in this elderly population was not defined. Differences in ethnic susceptibility to disease was hypothesized by the authors to be due to differences in the initial immune response to tuberculosis (31). The authors suggested that the immune response to the disease process was different for infection than it was for progression to disease (31). This may occur as different aspects of the immune system may be responsible for the

immune response at each of the two stages of TB disease. Macrophage function, as well as, to a lesser extent cell-mediated immunity, are important at initial exposure and acquisition of infection. Cell-mediated immunity is important for initiating and maintaining control of infection once it is established to prevent TB disease. The immune system is complex and all of the intricate features involved in the immune response to *M. tuberculosis* are not known. Variations in this complex immune response among ethnic groups could determine the epidemiology of tuberculosis.

There are biologic theories as to why African Americans, for example, are more likely to acquire TB infection than Caucasians. Some of these biologic theories involve macrophage function and Vitamin D levels. Both macrophage function and cell-mediated immunity may play a role in these biologic theories. One theory is that Vitamin D, which is required for immune and macrophage function, is at lower levels in African Americans compared to Caucasians (31). Another theory suggests that there is considerable variability in the macrophage ability to ingest *Mycobacterium*. Genetic difference in the ability of unstimulated macrophages to lyse TB bacilli and threshold of bacterial counts to stimulate macrophages may contribute to the difference in infection rates between two different populations (31). Susceptibility in mice to tuberculosis has been associated with the NRAMP 1 allele, an inherited variation in the immune response to tuberculosis. The average immune response of a population to *M. tuberculosis* and the observed epidemiology of TB within that population may relate to the immune response within a defined group. These biologic theories may apply to differences in TB rates observed between other ethnic groups.

The T helper cells are important in an effective immune response to *Mycobacterium tuberculosis.* Alveolar macrophages are the first line of defence against tuberculosis (1;2). *M. tuberculosis* is transmitted from individuals with infectious pulmonary TB as aerosolized droplets. Individuals in contact with infectious tuberculosis inhale these aerosolized droplets into their lung alveoli. Macrophages recognize the TB bacilli as foreign and initiate the immune defence cascade. TB bacilli are engulfed and destruction within the macrophage is attempted. Signals, or cytokines, are released from the macrophages to call for further assistance from the immune system. If the macrophages are successful, infection with *M. tuberculosis* is not established. If the macrophages are not successful, T lymphocytes are called to continue the immune response. Various T lymphocytes, including CD4 lymphocytes are involved in the ongoing immune defence against TB bacilli. Cytokines released by T lymphocytes again play an important role in signalling for assistance and coordinating the immune response. TB infection is established if the initial TB bacilli exposure is not contained by macrophages. Once T lymphocytes are involved in the immune response an immune memory is formed. This is identified by a positive tuberculin skin test. If the macrophages are successful in preventing infection after exposure, T lymphocyte involvement is minimal, no immune memory is formed and the tuberculin skin test remains negative despite TB exposure.

Once TB infection occurs, as documented by a positive tuberculin test result, the TB bacteria can remain dormant or suppressed for the lifetime of the host or TB disease, either primary progressive or reactivation, can develop. Primary progressive TB occurs if the *Mycobacterium* are not initially controlled by the immune system. This is defined as

occurring within 24 months after infection. Reactivation TB occurs if the infection is initially contained in a dormant or latent phase, but the *Mycobacterium* reactivate or begin to grow again resulting in symptoms and disease. Without an intact and co-ordinated response from T lymphocytes and cytokines, the body is not able to defend itself against TB. The T lymphocyte response will determine if an individual progresses from infection to disease and the length of time between infection and disease.

Macrophages have been shown to vary in their ability to defend against *M. tuberculosis* (31). Stead et al hypothesized that the immune defence before infection was different than the defence after infection and that variation in the initial immune response to TB varied between ethnic groups (31). Macrophages from African Americans permit significantly more replication of *M. tuberculosis* than Caucasians (31). This may result in an increased risk of developing an established infection in African Americans compared to Caucasians following exposure. Innate resistance to tuberculosis was thought not to occur from acquired immunity, but by the inherited ability of the host's macrophages to engulf the *M. tuberculosis* and prevent TB infection (31). This would account for a difference in establishing infection but not in the progression from infection to disease as this second step is dependent upon the immune response of the T lymphocytes and cytokines.

Vitamin D has been identified as an important mediator in macrophage activation (42). It is possible that active disease may occur with lower levels of vitamin D. A study of an Asian population in west London identified that those with serum vitamin D levels too low for detection had almost a tenfold increase risk in developing active TB disease (42). Vitamin D levels (25-hydroxycholecalciferol) levels were recorded in patients with

TB disease and healthy contacts with TB infection. Vitamin D deficiency was associated with a 2.9 (95% CI 1.3 to 6.5) increased risk of developing TB disease (42). Variation in vitamin D levels may play a role in the different TB rates noted between different population groups. Prior to chemotherapy, during the sanatorium era, patients were treated with Vitamin D through the use of 'sunshine balconies' and cod liver oil.

Animal studies have been undertaken to see if genetic factors controlling the occurrence of TB disease could be identified. A gene in mice that provides resistance to tuberculosis has been identified. This mouse gene is identified as NRAMP 1 (natural resistance associated macrophage protein 1). Human alleles of the NRAMP gene have been identified and may have an effect on the susceptibility of the African population to tuberculosis (43). Variations (polymorphisms) of the NRAMP1 gene have been identified with an increased incidence of TB disease compared to West Africans without these variations (43). West Africans, who were HIV negative, with smear positive pulmonary tuberculosis, were matched by age and ethnicity to individuals without TB disease. Individuals with tuberculosis disease had an increased frequency of two polymorphisms associated with TB disease (OR 4.07 95%CI 1.86 – 9.12). It was not possible to determine from this study if there was an increased risk of TB infection or progression to TB disease among individuals with these variants. As the frequency of these two variants were higher in the population with TB disease than the general population, this suggests that variations of the NRAMP1 gene were associated with progression to TB disease. One NRAMP1 variant associated with TB disease is very uncommon in Europeans, but was present in approximately one quarter of West Africans. This may explain in part why African Americans are more at risk of TB than Caucasian

Americans (43). The exact role of the NRAMP1 gene in the immune response has not been identified, however, macrophages from mice with distinct NRAMP phenotypes have an impaired capacity to restrict the growth of *M. tuberculosis* (43). No studies in other ethnic origins have been completed.

Other genes that have been identified that affect immune response to TB include allotypes of the Km1 light chain immunglobulin and haptoglobin. It is uncertain whether these proteins are directly involved in the immune response or merely reflect gene linkage (42). T helper cells have different patterns of maturation, T helper 1 (Th1) and T helper 2 (Th2). Th1 generate gamma interferon and Th2 generates interleukin. Differences in maturation of both Th1 and Th2 can alter host immune response, including granuloma formation and tissue destruction by tuberculosis. Although a relationship between ethnic origin and disease presentation has not been described, the maturation of T helper cells may be one more difference in the immune response that contributes to the variation seen in TB rates among ethnic groups and families (42).

Identical twin studies indicate that inherited factors are important in determining resistance to TB disease following infection (42). This suggests that genetic make-up may determine immune response and susceptibility to either TB infection, TB disease or both of these stages. Separating socio-economic or environmental and genetically determined factors in the immune response to TB is difficult (42). The immune response to tuberculosis is important in determining who acquires TB infection and who progresses from to disease. The various immune responses, including macrophage function, vitamin D levels, NRAMP 1 and other genes, and maturation of T helper cells, determine an average level of immune response within an ethnic group and play a role in

the TB rates within that population. One theory for the varying TB rates among different populations or ethnic groups is that each is in a different stage in the development of an effective immune response to the tuberculosis epidemic resulting in varying tuberculosis epidemiology in each group (3;42).

The Yanomami Indians of the Amazon river basin had their first prolonged contact with individuals of European descent in the mid-1960's (44). The first reported case of tuberculosis occurred in 1965 and by the 1980's TB was defined as an epidemic. The epidemic prevalence of TB disease was 6.4% in the population studied and smear positive cases were identified among 4.5% of the population. Mortality was unknown, but was estimated at 1280/100 000 annual rate. A study to observe the immune response to tuberculosis newly introduced into a population was undertaken (44).

Profound differences in response to BCG and immunoglobulin levels compared to other ethnic groups of the region suggested differences in the immune response to tuberculosis. Fifty-eight percent of Yanomami Indians who had received BCG vaccination had a negative TST whereas a control population of military recruits who received BCG vaccination had a positive TST rate of 73%. Among the bacteriologically confirmed cases 82% had received BCG vaccination. This suggested that BCG vaccination provided very little protection to TB disease. Antibody titres, predominantly IgM antibodies, were higher among Yanomami Indians (>70% above normal range) compared to control subjects of Brazilian or European descent (14%) with TB. Although the authors were not able to complete an analysis of cell-mediated immunity because of remoteness of study population, they hypothesized that there was a diminished cell-mediated response and an increased antibody response (IgM) to TB in this population.

The authors suggested that initial introduction of tuberculosis into a population results in a lower level of cell-mediated immune response and an ineffective antibody response compared to populations with a longer history of tuberculosis exposure. This suggests an immunologic naivety before introduction of TB resulting in poor protective immune responses. It has been theorized that a high antigen exposure induces a Th 2 response and a low antigen exposure induces a Th 1 or protective response (45). In immune naive populations any exposure to TB is likely to be high dose eliciting a non-protective (Th2) immune response. The result may be a selective pressure within the population to establish an effective immune response to tuberculosis infection and disease.

Grigg, in his paper "The Arcana of Tuberculosis", describes an asymmetrical wave or curve that defines the natural history of TB epidemiology within a population (figure 5.1) (3). This wave identifies TB morbidity and mortality rates following the introduction of *M. tuberculosis* into a population. The ascending limb is the rapid increase in mortality and morbidity in a previously non-exposed or immune naive population. The prolonged decline in mortality and morbidity occurs, according to Grigg, because of the gradual increase in the average immunity within the population. This results in a natural selection of individuals with increased immunity that prevents TB mortality then morbidity (3). Dubos and Dubos also described a 'rhythmic' course of epidemics that changed with the susceptibility of the host population (14).

Figure 5.1 The Natural History of TB with a Population(3)

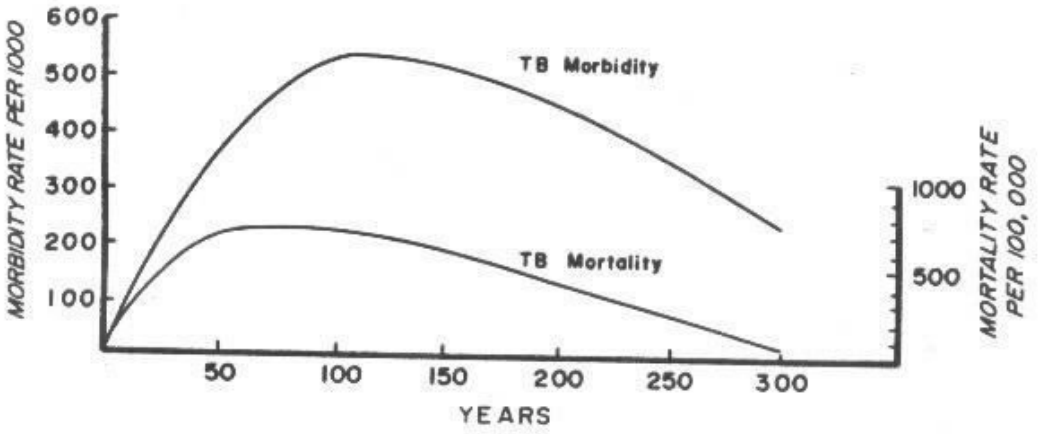


Figure 1. A theoretic plot of TB morbidity and mortality after the disease is first introduced into a population having no prior experience with TB. *(Adapted from* Grigg ER: The arcana of tuberculosis. American Review of Tuberculosis and Pulmonary Disease 78:160, 1958, with permission.)

The theoretical epidemiology curve described by Grigg has occurred within several populations throughout history (3) . Although the data are incomplete, the TB epidemic in London, England and Stockholm followed this projected asymmetrical curve from 1750 to 1950 (3). The initial TB mortality rates peaked early after the onset of the TB epidemic, approximately 900/100 000 in the late 1700's in London. These mortality rates gradually but persistently decreased over the following 150 to 200 years. Similar shaped epidemiology curves have been identified by Grigg in New York City in the early 1800's, the African American population following the Civil War, and the First Nations population in the southwest United States in the early 1900s (3).

The similar shape of the TB natural history curve suggests a similar disease pattern among each of the populations reviewed (3). As average individual immunity improves, susceptibility of the population to TB declines. This pattern was repeated among different populations or ethnic groups, suggesting that a TB epidemic elicits a

similar response in increasing average immunity over time. The epidemiology of TB is modified by the immune potential development of the host population (3;14). The Yanomami Indians of South America were early in their TB epidemic (44). The high mortality rates may have been due to the ineffective antibody (Th2) immune response in the population at the beginning of the epidemic. As predicted by Grigg, as a protective delayed type (Th1) immune response begins to occur within the population, the mortality curves may begin to decline.

Although the asymmetrical shape of the epidemiology curves recorded by Grigg are similar, differences between population groups exist (3). These differences include the start date of the epidemic and peak mortality rates. Peak mortality rates occurred in Europe in the late 1700's and early 1800's (3;14). The TB epidemic was thought to have begun in Europe with the Industrial Revolution as people moved from dispersed rural to crowded urban settings (3;14). Different onsets in time of the epidemic curves are documented by Grigg across the continental United States (3). With the arrival of European immigrants to the eastern seaboard of the United States, at the same time as the mortality from the TB epidemic in Europe was peaking, the first wave of the TB epidemic in the United States began in the Caucasian population of the Atlantic coast cities. Following the Civil War, as African Americans moved into urban areas, the TB rates peaked in approximately 1880, at rates higher than those documented in the Caucasian population (3). The TB rates were particularly high in urban centres that were sites for the underground railway. Tuberculosis had been documented among the First Nations population prior to the arrival of immigrants but with a change in the traditional lifestyle and crowding onto reserves, the mortality from TB accelerated rapidly at the

beginning of the epidemic (1;3;46). TB rates among First Nations in the south west United States and in Saskatchewan peaked in the early 1900's (3;5;14). The mortality rate among Fort Qu'Appelle First Nations decreased from 4000/100 000 in 1882 to 13 700/100 000 in 1890 (5).

As in the European population, the TB rate among African American and First Nations populations have steadily been declining (3;12). Because the epidemic, or asymmetrical wave of TB, began at different times in different populations, each group is at a different point in its epidemic. The TB epidemic began in Caucasians approximately 500 years ago. This group is now far along the descending limb of the epidemic curve and TB rates are low in this population (12). It is not unexpected that the Aboriginal population in Saskatchewan has a higher rate of TB than the Caucasian population. Within the Saskatchewan First Nations population, the TB epidemic began in the late 1890's (5;46).

The Aboriginal population passed the peak of their mortality rates from TB 100 years ago and began the descent as the average population immunity to TB theoretically adapted (3;46). There is a trend of TB rates among First Nations to increase from the coastal region (east then west) to the centre of Canada and from the south to northern areas of Canada (47). The difference in rates of progression from infection to disease in Caucasians (0.8%; 0.9/1000 person years) compared to Status Indians (6.9%; 7.7/1000 person years) in Saskatchewan demonstrates the difference in average population disease rates. Ethnic origin had the strongest association with progression from TB infection and disease in the Cox regression analysis. This observation is compatible with Grigg's

theory of the natural history of a TB epidemic within a defined population or ethnic group (3).

The age of individuals with TB disease changes throughout the epidemic, as described by Grigg (3;46). Early in the course of an epidemic, children and young adults are the most vulnerable to a new exposure and have increased risk of developing TB disease. These individuals will be expected to progress from infection to disease more quickly if there is an immune naive response to TB infection. In a more immune competent population further along their epidemic curve, older individuals develop reactivation TB from a previous infection as their immunity wanes in advancing years. In the Saskatchewan Aboriginal population, the mean age of those with TB $\geq$ 1 month following a positive skin test was 20.8 years. The average age in the Caucasian population was 49.7 years. Status Indians less than 19 years of age were at increased risk of progressing to TB disease (Table 4.15 Adjusted Hazard ratios for Individual Variables). The epidemiology of TB in the Saskatchewan Aboriginal population is following the epidemic profile of an immunologically naïve population hypothesized by Grigg.

In Saskatchewan, progression from infection to disease was highest among non-Status Indians and Metis. This group were predominantly from two communities (47%). Although this group could have had increased transmission and rates of infection by living in a similar socio-economic setting in the same communities, this study examined the rate of progression to disease following infection. The increased rate of disease in these two communities suggests that many of the individuals in the non-Status Indian and Metis study population may have common ancestors and it is possible that they share a

genetically determined immune response to tuberculosis. The non-Status Indian and Metis population in this study were very small compared to the Status Indian and Canadian-born Caucasian populations. Although this may have affected the results, the non-Status Indian and Metis in this study, possibly a remotely extended family, may share a genetic susceptibility to TB disease resulting in the increased progression from TB infection to disease. Families at risk of developing TB have been identified (14). In the Bronte family, all of the siblings developed TB disease presumably from their father who had a life long chronic cough. The French royal Bourbon family (Louis XIII and Louis XIV), and ten generations of Ralph Waldo Emerson's family were all documented to have tuberculosis disease (14). Identical twins have been found to have the same susceptibility to disease following infection compared to each other (42). Within families there may be inherited susceptibility to tuberculosis disease.

TB epidemics are initiated and potentially accelerated by factors other than the immune response of the general population (3). Although tuberculosis may have been present among the different ethnic groups such as European Caucasians or Status Indians in Saskatchewan prior to the onset of the epidemic, something would have been required to initiate the TB epidemic in populations where scattered cases of TB were presumably already occurring. Urbanization, as occurred during the Industrial Revolution in Europe or the loss of traditional lifestyle and movement onto reserves by the First Nations population, may have been a key factor (3). According to Grigg, urbanization, leading to overcrowding, resulted in more individuals being exposed to each case of infectious TB and "more people participated in the beginning of the TB phenomenon"(3). "The speed with which the TB epidemic wave progresses depends on urbanization of the affected

population" (3). In an immune naive or susceptible population, more individuals are at risk of progressing to TB disease once infected. Mortality and morbidity peak quickly. As immunity within a population improves, infectious cases will have less of an impact as fewer individuals progress to disease if infected. The effect of urbanization described by Grigg refers specifically to the role of overcrowding as the socio-economic factor that accelerates the TB epidemic. Overcrowding would provide a neighbourhood effect as it would be a factor in disease transmission.

Social factors have been attributed to the epidemiology of TB; however, risk factors for acquisition of infection differ than risk factors for progression to disease. TB is described as a disease of poverty (2;28). Grigg maintained that social factors were included in the 'generic term' of urbanization (3). Overcrowding was the social factor of urbanization that resulted in an increased risk of exposure. The role of overcrowding in the epidemiology of TB has been demonstrated by previous studies that looked at individual social factors (10;27). Studies using composite indices or definitions of poverty demonstrated an association with TB because overcrowding was one of the composite measures (9;28). Cantwell *et al* identified that the quartile with the most overcrowding had the strongest association with the development of TB (13).

In the Saskatchewan study person per room, a measure of overcrowding, was associated with an increased risk of TB disease only among those who developed TB disease $\geq$ 1 month of a documented TB infection and only in the hexatile that had the enumeration areas with the highest number of people per room (>0.89) (OR 3.9 95% CI 2.0 – 7.4). Overcrowding in Status Indians in Canada has been associated with increased TB rates (48). This association was thought to be due to increased transmission of

infection in a crowded environment. Urbanization, overcrowding, promotes the natural history of a TB epidemic by enhancing disease transmission and acquisition of infection, but it does not theoretically have a role in the progression from infection to disease. No social factors have been consistently documented in the literature or in the Saskatchewan study to contribute to the progression from TB infection to disease.

The differences in TB rates between Saskatchewan Caucasians and Status Indians can be explained by Grigg's natural history curve of TB epidemiology (3). The Caucasian population in Saskatchewan is far to the right on the decline of the epidemic curve, the point where TB becomes 'a medical novelty' (3). At this point reactivation TB disease is seen among previously infected older adults. The Aboriginal population of Saskatchewan is much earlier in its epidemic as the disease rates are higher and children are more at risk of developing TB disease. As Status Indians, non-Status Indians and Metis in Saskatchewan are earlier in the epidemic curve, it is the individual factor of ethnic origin that explains the increased risk of TB disease following infection. Socio-economic factors have no significant contribution to the second stage, progression from infection to disease.

Although there is a natural history to the TB epidemic, TB Control programs aim to further decrease the incidence of TB disease and disease transmission by the early identification of active cases and provision of prophylactic therapy to those with TB infection and who have increased risk of developing TB disease. By identifying individuals who are at increased risk of progressing from infection to disease, high risk individuals can be screened, treatment of latent TB infection can be provided to decrease the individual's risk of disease and to prevent future disease transmission.

6.0 Conclusion

Examining both individual and area-level measures are important in understanding the epidemiology of TB. The socio-economic factor of crowding is best measured as an area-level measure as it results in a neighbourhood effect of potential disease transmission. No gradient for socio-economic factors were identified in the progression from TB infection to disease. No consistent association was identified between socio-economic factors, as identified by area-level measures, and the progression from infection to TB disease. Ethnic origin (Status Indian and non-Status Indian and Metis) and age (< 19 years of age) had the strongest association with progression from infection to TB disease.

Several theories have been postulated for the mechanism of genetic susceptibility and immune response to the development of TB. Although social factors have been associated with the epidemiology of TB, the natural history of TB in a population is more dominant than socio-economic factors in the progression from infection to disease. The differences in the rates of progression from infection to disease, between the ethnic populations of Saskatchewan, is due to the same natural history of the disease in each population, and not socio-economic factors. Defining TB as a disease of poverty could potentially limit the identification of groups at risk of progressing to TB disease. Identifying those at risk of developing TB and understanding the determinants of TB epidemiology are important for establishing successful TB control programs.

6.0 Reference List

(1) Tuberculosis : A Comprehensive International Approach. 1 ed. New York: Marcel Dekker, Inc, 1993.

(2) A Clinician's Guide to Tuberculosis. Philadelphia: Lippincott Williams and Wilkins, 2000.

(3) GRIGG ER. The arcana of tuberculosis with a brief epidemiologic history of the disease in the U.S.A. Am Rev Tuberc 1958; 78(2):151-172.

(4) Diamond J. Guns, Germs, and Steel. New York: W.W. Norton & Company, 1997.

(5) Ferguson R.G. Tuberculosis among the Indians of the Great Canadian Plains. London: Adlard & Son, LImited, 1928.

(6) World Health Organization. world wide web . 2004.
Ref Type: Electronic Citation

(7) Stead W.W. Pathogenesis of a first episode of chronic pulmonary tuberculosis in man: recrudesence of residuals of the primary infectin or exogenous reinfection? American Review of Respiratory Disease 1967; 95(5):729-744.

(8) Canadian Tuberculosis Standards. 5th ed. Ottawa: Health Canada. Canadian Lung Association/Canadian Thoracic Society and Centre of Infectious Disease Prevention and Control, 2000.

(9) Bhatti N, Law MR, Morris JK, Halliday R, Moore-Gillon J. Increasing incidence of tuberculosis in England and Wales: a study of the likely causes. BMJ 1995; 310(6985):967-969.

(10) Mangtani P, Jolley DJ, Watson JM, Rodrigues LC. Socioeconomic deprivation and notification rates for tuberculosis in London during 1982-91. BMJ 1995; 310(6985):963-966.

(11) Ward HA, Marciniuk DD, Hoeppner VH. Extent of pulmonary tuberculosis in patients diagnosed by active compared to passive case finding. Int J Tuberc Lung Dis 2004; 8(5):593-597.

(12) Saskatchewan Tuberculosis Control Program. Saskatchewan Tuberculosis Control Program Annual Reports. 1991. Saskatoon, Canada, Saskatchewan Health, Government of Saskatchewan.
Ref Type: Report

(13) Cantwell MF, McKenna MT, McCray E, Onorato IM. Tuberculosis and race/ethnicity in the United States: impact of socioeconomic status. Am J Respir Crit Care Med 1998; 157(4 Pt 1):1016-1020.

(14) Dubos R, Dubos Jean. The White Plague. New Brunswick, New Jersey: Rutgers University Press, 1958.

(15) Lynch J, Kaplan G. Socioeconomic position. In: Berkman LF, Kawachi I, editors. Social Epidemiology. New York: Oxford University Press, 2000: 13-35.

(16) Kaplan G. People and places: contrasting perspectives on the association between social class and health. International Journal of Health Services 1996; 26(3):507-519.

(17) Krieger N, Williams DR, Moss NE. Measuring social class in U.S. public health research: concepts, methodologies and guidelines. Annu Rev Public Health 1997; 18:341-378.

(18) Carstairs V. Socio-economic factors at area level and their relationship with health. In: Elliott P, Wakrefield JC, Best NG, Briggs DJ, editors. Spatial Epidemiology: Methods and Applications. London: Oxford University Press, 2000: 51-67.

(19) Macintyre S, Ellaway A. Ecological approaches: rediscovering the role of the physical and social environment. In: Berkman LF, Kawachi I, editors. Social Epidemiology. New York: Oxford University Press, 2000: 332-346.

(20) Locker D, Ford J. Evaluation of an area-based measure as an indicator of inequalities in oral health. Community Dent Oral Epidemiol 1994; 22(2):80-85.

(21) Haan M, Kaplan GA, Camacho T. Poverty and health. Prospective evidence from the Alameda County Study. Am J Epidemiol 1987; 125(6):989-998.

(22) Sin DD, Svenson LW, Man SF. Do area-based markers of poverty accurately measure personal poverty? Can J Public Health 2001; 92(3):184-187.

(23) Geronimus AT, Bound J. Use of census-based aggregate variables to proxy for socioeconomic group: evidence from national samples. Am J Epidemiol 1998; 148(5):475-486.

(24) Krieger N. Overcoming the absence of socioeconomic data in medical records: validation and application of a census-based methodology. Am J Public Health 1992; 92:703-710.

(25) Locker D, Ford J. Using area-based measures of socioeconomic status in dental health services research. J Public Health Dent 1996; 56(2):69-75.

(26) Townsend P, Phillimore P, Beattie A. Health and deprivation : inequality and the north. London: Croom Helm, 1988.

(27) Hawker JI, Bakhshi SS, Ali S, Farrington CP. Ecological analysis of ethnic differences in relation between tuberculosis and poverty. BMJ 1999; 319(7216):1031-1034.

(28) Spence DP, Hotchkiss J, Williams CS, Davies PD. Tuberculosis and poverty. BMJ 1993; 307(6907):759-761.

(29) Tocque K, Doherty MJ, Bellis MA, Spence DP, Williams CS, Davies PD. Tuberculosis notifications in England: the relative effects of deprivation and immigration. Int J Tuberc Lung Dis 1998; 2(3):213-218.

(30) Marmot M. Multilevel approaches to understanding social determinants. In: Berkman LF, Kawachi I, editors. Social Epidemiology. New York: Oxford University Press, 2000: 349-367.

(31) Stead WW, Senner JW, Reddick WT, Lofgren JP. Racial differences in susceptibility to infection by Mycobacterium tuberculosis. N Engl J Med 1990; 322(7):422-427.

(32) Parslow R, El Shimy NA, Cundall DB, McKinney PA. Tuberculosis, deprivation, and ethnicity in Leeds, UK, 1982-1997. Arch Dis Child 2001; 84(2):109-113.

(33) Kushigemachi M, Schneiderman LJ, Barrett-Connor E. Racial differences in susceptibility to tuberculosis: risk of disease after infection. J Chronic Dis 1984; 37(11):853-862.

(34) Canada Census. world wide web . 2001.
Ref Type: Electronic Citation

(35) Health Canada FNIaHBSR, Saskatchewan Health. Tuberculosis Control: A reference guide to the tuberculosis program in Saskatchewan. 2002. Health Canada FNIHB, Saskatchewan Region and Saskatchewan Health.
Ref Type: Report

(36) American Thoracic Society., Centers for Disease Control. Diagnostic standards and classification of tuberculosis in adults and children. Am J Respir Crit Care Med 2000; 161:1376-1395.

(37) Canada Post. world wide web . 2004.
Ref Type: Electronic Citation

(38) Norman GR, Streiner David L. Biostatistics: The Bare Essentials. 2 ed. Hamilton, ON, Canada: B.C. Decker Inc., 2000.

(39) Kleinbaum DG. Survival Analysis: A Self-Learning Text. New York, NY: Springer-Verlag New York Inc, 1996.

(40) Hosmer David W, Lemeshow S. Applied Logistic Regression. 2 ed. New York: John Wiley and Sons, 1989.

(41) Collett D. Modelling Survival Data in Medical Research. 2 ed. Boca Raton, Florida: Chapman & Hall/CRC CRC Press LLC, 2003.

(42) Davies PD, Grange JM. Factors affecting susceptibility and resistance to tuberculosis. Thorax 2001; 56 Suppl 2:ii23-ii29.

(43) Bellamy R, Ruwende C, Corrah T, McAdam KP, Whittle HC, Hill AV. Variations in the NRAMP1 gene and susceptibility to tuberculosis in West Africans. N Engl J Med 1998; 338(10):640-644.

(44) Sousa AO, Salem JI, Lee FK, Vercosa MC, Cruaud P, Bloom BR et al. An epidemic of tuberculosis with a high rate of tuberculin anergy among a population previously unexposed to tuberculosis, the Yanomami Indians of the Brazilian Amazon. Proc Natl Acad Sci U S A 1997; 94(24):13227-13232.

(45) Bretscher PA. A strategy to improve the efficacy of vaccination against tuberculosis and leprosy. Immunol Today 1992; 13(9):342-345.

(46) Hoeppner VH. Tuberculosis in Saskatchewan treaty Indians. Prairie Medical Journal 1995; 65(1):18-20.

(47) Health Canada, Sante Canada. Tuberculosis in First Nations Communities. 1999. Minister of Public Works and Government Services of Canada.
Ref Type: Report

(48) Clark M, Riben P, Nowgesic E. The association of housing density, isolation and tuberculosis in Canadian First Nations communities. Int J Epidemiol 2002; 31(5):940-945.

Appendix 8.1 Glossary

Tuberculosis (35)

Tuberculosis (**Tuberculosis Disease)** Tubercle bacilli have entered the body and are replicating (growing). Evidence of disease is present as positive cultures, symptoms, or chest radiograph changes.

Contacts Individuals who have breathed the same indoor air ten hours or more one month prior to the date of diagnosis of a person with primary tuberculosis or smear positive tuberculosis. The emphasis is on the time and proximity of exposure.

Cure Client is considered cured when 90% or greater of prescribed doses have been taken.

Infectious tuberculosis smear positive pulmonary tuberculosis can transmit infection to other people

Mantoux A skin test used to determine persons who have been infected with (Tuberculin Skin Test) Mycobacterium tuberculosis

Non-Infectious tuberculosis Smear negative pulmonary tuberculosis is much less likely to transmit infection to others. Non-pulmonary tuberculosis will not transmit infection to other people.

PPD – 5TU A tuberculin, purified protein derivative for intradermal tuberculin testing. The standardized dose is 5 Tuberculin Units (5TU). This is equal to 0.1 ml.

Primary tuberculosis Tuberculosis which develops during the first two years following initial infection with M. tuberculosis (active disease).

Reactivation (Secondary) Tuberculosis Tuberculosis which develops more than 2 years after the primary infection

Significant Skin Test Reaction Equal to or greater than 5 mm in contact to infectious TB or persons with HIV/AIDS. Equal to or greater than 10 mm in all others.

Tuberculosis Infection Tubercle bacilli have entered the body but are dormant (not growing). The skin test shows a significant reaction but there is no evidence of disease.

Biostatistics and Epidemiology

B a standardized regression coefficient (the amount of change that will on the average take place in one characteristic when the other characteristic changes by one unit)

Confidence Interval the range within which the true magnitude of effects lies with a certain degree of assurance

Confounding a variable that is associated with both the independent and dependent (outcome) variable being considered and so can cause an association between the independent and dependent variables which would not otherwise exist, or conversely, can negate an association

Epidemiology the study of distribution of a disease or condition in a population and of the factors that influence that distribution

Hazard Ratio the potential per unit time for the event to occur, given that the individual has survived to time t.

Incidence Rate the number of new cases, per unit population, occurring during a stated period of time

Interaction an effect modifier in which the effect of one variable on the outcome is dependent on another variable

Non-parametric Statistics Statistical methods for nominal or ordinal data that are identified by counts (eg, gender) or measured data that is not distrubted along a normal bell-shaped curve

Odds Ratio a measure of association or rate among people who have a disease of having been exposed in the past to the rate of exposure in the past among people who do not have the disease.

Parametric Statistics statistical methods for analyzing measured data that is distributed along a normal bell curve

Prevalence Rate the number of new and existing cases, per unit population, occurring during a stated period of time

Relative Risk The ratio or risk of disease in exposed compared to the risk of disease in nonexposed individuals

Standard Error (of the mean) describes the distribution of mean values around the population mean (standard deviation divided by the square root of the sample size)

Appendix 8.2 Saskatchewan TB Control Program