# RISK FACTORS FOR TUBERCULOSIS IN A COHORT OF RECENT IMMIGRANTS TO ONTARIO

By

MARIA ISABELLA CREATORE

A thesis submitted to the Department of Community Health and Epidemiology in conformity with the requirements for the degree of Master of Science

Queen's University
Kingston, Ontario, Canada
January, 2000

## ABSTRACT

The World Health Organization estimates that one-third of the world's population is infected with TB. Eighty percent of TB in Ontario is among the foreign-born. Before effective targeted testing programs can be developed to control TB in Ontario local epidemiological profiles to identify high risk groups among foreign-born populations are needed. The objective of this study was to describe the effect of time since arrival on TB risk and to analyze determinants of risk for TB over time after immigrating to Ontario. Descriptive and bivariate analyses were conducted to describe the population and its risk for TB by age, sex and world region of origin. Actuarial life-table analysis and log-rank tests were used to evaluate risk of TB at different times and to determine contributors to TB risk. Factors of interest were age at arrival, sex, time since arrival, world region of origin and year of arrival. Cox proportional hazard and complementary log-log models were used to account for multiple factors simultaneously, and to establish survival probabilities at different time points and to generate risk ratios. Age-specific TB risk showed a bimodal pattern peaking at the 16 to 30 age group and over the age of 65. The highest risk was seen in the 16 to 30 age group from Sub-Saharan Africa. Males generally showed a higher level of risk in the oldest age group. A significant decrease in TB risk occurred soon after arrival to Ontario. The risk for most world regions remained stable through years 2 to 6 with some evidence of an increase in risk in years 7 and 8. The group contributing to this increased risk were young males from high endemic countries. Targeted TB testing activities among migrant populations should be directed to high risk groups only, which include people between 16-30 years, over the age of 65 and from countries with endemic TB. These people may be at significantly increased risk even 8 years after arrival. Future studies should investigate possible increases in risk over time possibly due to reunification of families or visiting the country of origin.

## ACKNOWLEDGEMENTS

It is with sincere gratitude that I acknowledge the contributions of my supervisors:

Dr. Wendy Wobeser for her mentoring, guidance and enthusiasm.
Dr. Miu Lam for his constant patience and equally valuable guidance.

This study was made possible through collaborations with both Dr. Neil Heywood from Citizenship and Immigration Canada and Dr. Monika Naus from the Ontario Ministry of Health.

Citizenship and Immigration Canada and the Queen's Department of Medicine provided financial support during the project.

The support and friendship from my classmates in the epidemiology program, and from Heather Kenney and Kim Merkley was always appreciated.

The deepest thanks to my husband and parents for their support, encouragement and confidence which made this project, and all things, possible.

## TABLE OF CONTENTS

## LIST OF TABLES

## LIST OF FIGURES

## LIST OF APPENDICES

## Glossary of Abbreviations and Technical Terms

**AIDS** - Acquired Immunodeficiency Syndrome
**ACET** - Advisory Council for the Elimination of Tuberculosis (U.S.A.)
**CIC** - Citizenship and Immigration Canada
**DOTS** - Directly Observed Therapy, Short-course. A specific type of tuberculosis-control program requiring, as two elements, a monitoring system for standardized chemotherapy administration to TB cases and program supervision.
**ECOT** - Expert Committee on Tuberculosis (Canada)
**'Elimination'** - Defined by the WHO as having been reached when the incidence of sputum smear-positive TB is 1 per million.
**EME** – Established Market Economies
**FSEE** – Former Socialist Economies of Europe
**HIV** - Human Immunodeficiency Virus
**IME** - Immigration Medical Exam
**IUATLD** – International Union Against Tuberculosis and Lung Disease
**LatCar** - Latin America and the Caribbean
**LIDS** – Landed Immigrant Data System
**'Low Incidence Countries'** - Term used to describe countries with incidence of all forms of active TB below 10 per 100,000 population (WHO definition).
**MDR-TB** – Multi-drug resistant TB, resistant to at least isoniazid and rifampin.
**MEC** – Middle Eastern Crescent
**MS** – Medical surveillance
**OthAsia** - Other Asia (All Asia except for China and Taiwan)
**RDIS** – Reportable Disease Information System
**SSA** – Sub-Saharan Africa
**TSRU** – Tuberculosis Surveillance Research Unit
**WHO** – World Health Organization

## 1.0 INTRODUCTION

TB causes more adult deaths than any other infectious disease worldwide [1,2]. In 1997, a global project by the World Health Organization to determine incidence, prevalence and mortality due to TB in 212 countries found that 32% of the world's population, or 1.8 billion people, were estimated to be infected with TB [3,4,5]. There are an average of 3 million deaths per year due to this disease [6-8,2]. Ninety-five percent of these cases and deaths occur in the developing world, 75% in Africa and Asia [9]. As tuberculosis is concentrated in the young and the middle-aged, the social and economic cost of this disease is tremendous [10,11]. There is a great need to study the natural history and behavior of this ancient disease in a modern context of increasing global migrations, HIV/AIDS, and multi-drug resistance.

Canada is composed of a very heterogeneous population with respect to risk for TB. Not only do we have pockets of high-prevalence populations, but we receive large numbers of immigrants from areas that have high rates of endemic TB [9]. Immigration is a major contributor to the Canadian TB incidence rate [4,9,12,13]. In Ontario, 80% of TB is among the foreign-born [14,9]. The decline of TB incidence has halted since 1985 and disease rates in the foreign-born in Canada have been increasing [15-17,13,2]. Large migrating populations must be studied with respect to burden of TB and identifiable risk factors in order to determine both the level of risk to the population itself, as well as to the population into which it is migrating. Only if these groups are studied closely can appropriate disease intervention programs be developed to target high-risk populations.

TB has the ability of once again becoming a significant threat to countries where the prevalence of the disease has, until now, been declining steadily. It is in countries like Canada that the disease can be most threatening due to its low incidence. Low incidence and prevalence is often associated with less experience and infrastructure in the medical profession for identifying and treating TB [9]. Failure to identify active cases and use preventive measures on high risk populations can lead to increased transmission from undiagnosed active cases. The potential for large outbreaks of TB is real in this setting [10,11].

This study is focussed on identifying high risk immigrant populations in Ontario, as well as describing the pattern of risk over time in these groups. Effective access to foreign-born persons for

treatment of latent tuberculosis infection is a major issue in current tuberculosis control [18]. Increasing preventive services in the community requires an in-depth understanding of who the high risk groups are and how and when to most effectively target them for programs. This information using recent data is greatly lacking in Ontario. This study provides profiles of populations at high risk in Ontario and the pattern of TB incidence among these populations.

## 2.0 LITERATURE REVIEW

### 2.1 Medical Overview

#### 2.1.1 Natural History of TB

Development of tuberculosis is a two-stage process. The initial step involves becoming infected with the bacillus *Mycobacterium tuberculosis*, which is primarily via the airborne route and in most cases attacks the lungs [2]. The second stage involves having sufficient multiplication of the infecting organism to overcome the host resistance and finally, to cause disease [19]. Following infection, a person can remain free from active disease for a few weeks, years, or the rest of their life [19,20]. Untreated, active TB has a fatality rate of between 40 and 60% [21,2,10].

The Tuberculosis Surveillance Research Unit (TSRU), which was set up by the International Union against Tuberculosis and Lung Disease (IUATLD) and the World Health Organization, suggests a method to subdivide people with TB infection that has some clinical meaning. That is, people with a) progressive primary tuberculosis from recent primary infection, b) endogenous reactivated tuberculosis due to distant primary infection (no reinfection) and c) exogenous reinfection tuberculosis from distant primary infection as well as recent reinfection [22]. Differentiating these clinical scenarios requires detailed and meticulous epidemiologic and clinical characterization [23]. For the sake of this study, "active TB" includes all three of these forms of tuberculosis.

These three TSRU risk groups may differ with respect to aetiology and risk factors for the disease. With respect to disease contribution in society, reactivated and remote infection disease are "continuous,

stable, and reasonably predictable phenomena. The segment arising from recent infection is labile, and may be subject to wide and rapid fluctuations in the future" [24]. In the context of immigration from endemic areas, this quote must be considered more carefully. Large numbers of people who have been exposed to TB are entering Canada which means that reactivation disease and remote infection may play a particularly important role in generating new active cases in Ontario [16]. A recent study conducted in San Francisco uses molecular evidence to support this contention that reactivation of infection that was acquired before immigrating is the most common form of TB in immigrants [25].

## RISK OF INFECTION

There is an observed pattern that male and female prevalence of infection is approximately equal until early adolescence when male prevalence begins to exceed infection in females [22,26-28,7,11]. This phenomenon may be biologically based, however, is likely due to the tendency in many countries for young men to have much more contact with the "outside world" and greater opportunity for infection than young women. An American study looking at prevalence of tuberculin positivity in recent immigrants found an increase in both men and women in those of child-bearing age. Nearly 50% of those between 25-34 were skin-test positive (n=2035) [29].

The association of risk of infection and origin from a high TB prevalence country is well established [15-17,13,.25,2]. Studies in the UK and Ontario show increased risk of TB infection in people immigrating from or visiting to high prevalence areas [30,31]. Crowding, poverty, homelessness and living in correctional facilities are also all associated with risk of infection [7,32,20,7,33-35]. Furthermore, there is an increased risk of exposure to, and infection with TB over a person's lifetime [22,20,7,33]. Styblo et al. found that most infections happen early in life in a high-prevalence country [28]. It has been shown that people arriving in the U.S. before the age of 5 had significantly lower risks of infection than people arriving over 5 years of age from a high prevalence country, making age at immigration a risk factor [36,16]. Thus, as a person ages, the probability of having been infected with TB increases proportionally to the annual risk of infection in that country.

## RISK OF PROGRESSION TO ACTIVE DISEASE

Early studies claimed that in 90% of people the infection becomes dormant, in 5% there is symptomatic disease immediately, and in the remaining 5% reactivation occurs at some point in life [17,28,2,37,34]. It is thought that the highest risk of developing active disease is generally shortly after exposure [7,38,16,22,24] with 80% of active cases occurring within the first two years after infection [6]. In Scandinavia and Britain, one study found that the proportion of people developing disease within the first year after infection varied from 11% to 48% [39]. An international review of studies compiled by the TSRU that the risk of developing disease within 5 years after infection is the highest [39]. There is some evidence that in young people, risk of developing active disease decreases with time after infection [35,38,40]. In one study, of those who developed disease, the proportion was between 58 - 61% within the first year, and 22-30% within the second year, and no cases occurring more than 8 years after infection [22].

Various factors are thought to influence whether a person develops active disease post-infection. As with risk of infection, age may play a role in susceptibility to active disease [19]. Studies have shown that children under 4 years of age have the highest rates of TB after being infected [19]. This is one reason that high rates of disease in children is indicative of high rates of active disease in the population, since TB in children signals recent transmission [8]. There is a second peak in late adolescence and the early twenties when there is again a very high rate of developing disease after infection [11,19,20]. In Ontario, the rate of disease for men rises steadily with increasing age up to a rate of 19 per 100, 000. For women, on the other hand, the curve is bimodal with two peaks, one at 20-30 years of age and one at over 70 years, with an overall average of 11 per 100 000 [9]. One study showed that foreign-born cases are mainly between the ages of 10 and 39 years, however, Canadian-born cases are predominantly over 40 or, in the case of the Aboriginal population, under 10 years of age [9]. This discrepancy is not clearly explained or documented in the literature. In general, tuberculosis incidence is concentrated in the ages 15 through 64 [10,29,22,41], although this may not be true of people born in developed countries where the burden of TB is primarily among the elderly [9].

Human Immunodeficiency Virus (HIV) has altered the epidemiology of TB. The greatest modern risk for developing active disease is HIV/TB coinfection [8,35,42]. TB incidence in people with Acquired Immunodeficiency Syndrome (AIDS) can be 500 times the incidence in the general population [8]. In people

with HIV, cell mediated immunity is compromised thus host resistance is too weak to overcome the disease. Unfortunately, people with HIV also have abnormal presentation of the disease, with more frequent occurrence of atypical pulmonary and extrapulmonary disease [8]. This only serves to hinder any timely diagnosis to increase the probability of survival.

Some studies found that risk of disease for women and people living in urban areas was greater than for men and rural residents [7,20,38,19]. Studies show that progression from infection to disease is faster among women of reproductive age than among men at the same age [43,26,11]. In Ontario, women between the ages of 10 and 34 years had up to 30% higher rates of disease progression than men. Men had rates that were twice as high, on the other hand, over 40 years of age [24,26,44]. Globally, there is an approximately 2:1 male-to-female case notification ratio [45,26]. This higher rate of disease in men probably reflects mainly the greater rate of infection in men. However, this could also be due to under-reporting of disease in women [26,45].

Being underweight is a risk factor for developing active disease [42,20,7] Poverty and malnutrition are also risks for disease activation [46,33]. The extent to which poverty accounts for ethnic discrepancies in TB rates has not been quantified to any great extent [32]. One recent study conducted in the U.S. found that by adjusting for SES, one could account for about 50% of the increased risk for active TB that is associated with U.S.-born African-Americans, Hispanics, and Native Americans [8]. This study also found that there was a significant interaction between SES and crowding, with the lowest SES group experiencing the greatest crowding (as measured by persons per room per household) [8]. Interestingly enough, they found that SES had less of an effect on the risk of foreign-born African-Americans and Asians.

## REACTIVATION

After the initial infection and inflammatory response to the bacilli, often a "primary complex" is left behind containing lesions in the lung as well as in the corresponding lymph nodes [6]. These lesions usually heal, leaving behind scars and dormant bacilli that can reactivate and cause active disease at any point in a person's life [6,2]. Studies conducted in the Netherlands revealed that the estimated annual risk of endogenous reactivation of disease was 0.0163% [22], whereas studies in Czechoslovakia among persons with identifiable lung lesions on chest Xray estimated disease at 0.5% [102].

There is conflicting evidence regarding risk factors for reactivation of latent TB. The one point of general agreement, is that people with latent disease are a high risk group for active disease at some point in the future [47,24,29]. Malnutrition,[9] inadequate chemotherapy for previous episodes of active disease, both in terms of compliance and effectiveness of the drugs[7,20,48], and age have been shown to be risk factors for reactivation over a person's lifetime [49]. HIV is a risk factor as well as end stage renal disease, silicosis, diabetes and some other comorbidities [49]. It is common in low incidence countries for TB rates to be highest in the elderly segment of the population, primarily due to reactivation of latent disease [42]. Some studies have shown that males have a higher reactivation rate than females, but studies contradicting this are also in existence [7].

### 2.1.2 Prevention and Treatment

There are various strategies for the control of tuberculosis. The first priority is to treat active cases and to prevent transmission from patients with active and contagious disease to others in order to decrease morbidity and infection rates [18]. The second priority involves contact tracing to identify people who may have already been infected and who are at risk for active disease [18]. The third priority is to try and prevent people who are already infected from developing disease, and therefore to prevent new, active cases [50]. To achieve the first priority, curative care with drugs must be implemented. To address the last, targeted testing and chemoprophylaxis of high risk groups are often used. Finally, in order to decrease risk of disease in the first place, BCG vaccination may be used, although there are numerous caveats with BCG which will be discussed below. This study attempts to increase the knowledge base about the local epidemiology in foreign-born groups in Ontario to facilitate targeted testing programs of high risk groups.

In a country such as Canada, where the risk of infection is very low, identifying active cases when they occur is a very high priority. A high risk group has been defined by the WHO as having an incidence of 100 per 100,000 or more [42]. This level was chosen primarily since it is at this rate of disease that active case-finding can be cost-effective [42]. Within high risk populations who experience very high rates of infection, only treating active cases does not address the underlying high burden of infection. Each strategy for addressing the priorities is associated with strengths and limitations, and it is generally felt that

no one strategy is sufficient [1].

## BCG VACCINE

In 1921 the bacillus of Calmette and Guerin (BCG) was developed [6]. Although the BCG vaccine is one of the most widely used globally, it's effectiveness is still controversial [10]. Clinical trials have found the vaccine any where from 0% to 80% effective in preventing TB [6,10,42,4]. There are many theories about the huge variance of effectiveness of this vaccine from differences in strains of BCG, loss of efficacy overt time, success of administration of vaccine, and nutritional factors[6,10]. Although significant, BCG cannot be relied on as the only method of TB control in a community [6,10]. Also to be considered, is that in a community with low infection rates of a disease, such as in the non-aboriginal Canadian-born, the protection that a vaccine confers on this population is negligible [20]. Effective vaccination programs against TB are limited to TB endemic populations.

Vaccination with BCG decreases the effectiveness of any policy to skin test immigrants due to its continued widespread use globally. This is due to the fact that BCG may cause a positive skin test which cannot be differentiated from a true TB infection. BCG immunization among known HIV-infected persons, and other immunocompromised persons is contraindicated as such persons are at high risk of developing disease from the BCG [10,4].

## DIAGNOSIS AND DRUG THERAPY

There are four principal diagnostic techniques utilized to detect TB infection and clinical disease. First, the tuberculin skin test (Mantoux test) identifies people who have developed a cell-mediated immune response to tuberculosis antigens, thus it is a good indicator of the prevalence of infection in a population [6]. This method, however, has low specificity if testing a BCG-vaccinated population, as well as people possibility infected with other non-tuberculous mycobacteria [40,10,6]. Second, sputum microscopy identifies those with tuberculosis bacilli in their sputum, and who are therefore, particularly infectious. Third, sputum culture identifies those with bacilli in their sputum in insufficient numbers to be detected by microscopy. Fourth, a chest x-ray can be performed in order to identify any changes or scarring in the lung

consistent with TB [10]. This method can identify both active disease and inactive disease for treatment or chemoprophylaxis, respectively.

Preventive therapy (treatment of latent infection) generally involves 6-9 months of isoniazid, although shorter drug regimens have been approved using other drugs [16]. Studies have shown isoniazid to have an efficacy of at least 50 to 90 percent [29,51] in preventing future progression to active disease. Studies in developed countries have also found that for a 6 month regimen, cost-effectiveness rates per case averted were $17,000 or more [10]. This cost-effectiveness is highly dependent on adherence with the preventive therapy.

Treatment for active TB involves using multiple drugs. Treatment can have serious side-effects as some people react poorly to the drug combinations. All the drugs have the potential for causing rash, nausea and fever, hepatitis symptoms and renal failure in some people [52]. The risk of toxic effects in the liver increases significantly after 35 years of age [19,48,20]. As well, the cost of drug therapy can be quite high. Isoniazid therapy alone for 6 months costs approximately $300 U.S. per patient (this includes physician costs) [29]. Therefore, a certain level of specificity to avoid unnecessary treatment is also required.

### 2.1.3 Active vs. Passive Case Detection

There are two strategies for case-detection, active and passive. Active case-finding involves attempting to screen the general population, or target populations, to find active cases. Passive case-detection involves diagnosis of those who present to the health care system with risks or symptoms suspicious of TB [10,38,6]. There has been general agreement globally to use the passive case-detection method for the identification of active TB. This policy is advocated by both the WHO and the IUATLD [6].

Containing active disease is an obvious and necessary first step towards eliminating TB. It is estimated that on average, a contagious, untreated TB case infects 10 to 14 people per year [15,52,10]. The longer a person goes untreated, the longer they remain infectious and the more secondary cases may arise. In an extreme case in Minneapolis, a man who was highly infectious infected 41 people from his local bar which he frequented before he was diagnosed [34]. Another recent report in North Dakota revealed that a nine-year-old child with infectious TB caused 56 people to result in positive skin tests [53]. In a typical

situation, it is expected that approximately 20 - 30% of close contacts with the infectious person have positive tuberculin skin tests [34]. Thus, the need for prompt identification of active cases and contact tracing is obvious.

Screening for latent infection is another method of TB control, that requires a greater understanding of the epidemiology of TB. Since TB can be inactive, or latent, it can be hard to diagnose in the general public until it reactivates and the person becomes symptomatic. Identifying and treating latent disease in people who have a history of active TB is important, for these people are at high risk for reactivation [47,20]. The risk for these people is significantly decreased if they have a history of effective TB chemotherapy [38]. Also, the incubation period for those infected with TB is highly variable and thus, not predictable with any accuracy [20,54]. It has been suggested that foreign-born people have a higher proportion of reactivations from remote infections [24,50,18]. In one study, immigrants with scars on their lungs had 13 times the incidence rate of active TB as immigrants with normal chest X-rays [47]. Screening for this group is reliant on the tuberculin skin test and the chest x-ray, unless they have a documented history of active TB.

Screening for tuberculin converters (people who have recently developed a reaction to a Mantoux skin test indicating recent infection) identifies people who are newly at risk for developing TB. This is primarily accomplished through contact tracing. Treating tuberculin converters is a rewarding tactic since it is highly effective in preventing disease in the individual, thus preventing both the individual and their contacts which would have been at risk for infection [10,11,18]. Recently infected people, as outlined in the previous section, are at particularly high risk for progression to active disease. However, mass screening for persons who have been recently infected with TB is costly both in time and money. Screening of high risk populations for latent infection or active disease may be considered when programs have high success rates with treatment of active cases [29, 51, 16].

The method of active screening in the population goes against the philosophy of passive detection. Regardless of how high the risk is in a population, active case detection costs more per case detected due to the low yield of screening [6]. One TB mathematical modeler, however, claims that treatment of active cases alone will never achieve eradication of TB. She claims that treatment of active disease as well as latent infection in high risk groups is the only way of eradicating TB [1]. This is based on the assumption

that we cannot sustain high enough and successful enough levels of treatment for eradication. Theories such as this advocate a role in TB control of active screening of high risk populations. Thus, the issue of eradication becomes a complex interplay of case-detection, treatment rates, and cure rates.

## 2.2 Global Trends

### 2.2.1 Recent Issues

There was considerable evidence that the problem of TB had been globally decreasing over time, especially in industrialized countries. This was likely due both to improved sanitation and the introduction of antituberculous chemotherapy in the 1950's [10]. TB infection rates in developing countries is now similar to infection rates in industrialized countries in the early part of the 20th century [26]. Since the mid-80's, however, the decline has stopped and in North America and Western Europe increases in the numbers of cases have been observed [9,55,8]. One striking pattern is that between 1985 and 1992, the segment of the U.S. population that had the largest percent increase in TB was the 25-44 year olds [8,56]. There are 4 very significant global trends that are complicating the control of TB [9]. These are:

1) An increase in TB/AIDS coinfection (which is also increasing the mortality rate from TB)
2) An increase in multi-drug resistant strains of TB (MDR-TB)
3) An increased migration of populations
4) Increasing urbanization.

### 2.2.2 HIV and MDR-TB

In the developing world in particular, HIV infection is resulting in a significant increase in tuberculosis prevalence and severity [10,57,42,40,4,58,56]. Transmission is facilitated by rapid progression from the incubation to the active stage [9]. It has been reported that in people with AIDS, the incidence of TB is almost 500 times that in the general population [8,59]. The World Health Organization estimated that in 1994 there were 1.3 million individuals with active TB in South-East Asia, which is also an area with high rates

of HIV infection [40]. People with AIDS can die as soon as 4 weeks after infection with the tubercle bacillus [9]. TB is now a leading cause of death among HIV positive people worldwide [26]. Multi-drug resistance and HIV often go together which increases both the infection rate and rate of death due to TB in these populations [7-9,29,26].

Multiple-drug resistance, defined as resistant to 2 or more drugs including the 2 most 'powerful' antituberculous drugs isoniazid and rifampin, is a result of poor TB control programs [60,61,101]. MDR-TB results in the necessity to use "second line" agents which are less effective, more toxic and more expensive than first line agents such as isoniazid [2,16]. The mortality rate for MDR-TB is high. Most developing countries cannot afford to treat cases of MDR-TB. Globally, it would appear that the burden is increasing and although a Directly Observed Therapy Short-course (DOTS) program may help to control it, the problem of what to do about MDR is not one with an easy solution [60,61,8].

MDR-TB is particularly associated with HIV, ineffective drug therapy, and non-adherence with treatment [4,9,8]. It has also been found to be more common in racial and ethnic minorities [16,8]. In Ontario, isolates resistant to 1 or more drugs reported to Tuberculosis Registry of Ontario has risen from 10% in 1980 to 14% in 1995 [9]. A global interest in the implementation of "Directly Observed Therapy, Short-course" (DOTS) is having an effect on the problem of non-compliance in certain well-documented areas [4,8,16].

### 2.2.3 Global Migration and its effect on TB in Developed Countries

Immigration has "become the most important factor influencing the stagnation in the decline in the incidence of TB see in many developed countries" [62]. Migration affects rates of TB in a socially complicated but salient way. Increasingly, the pattern of immigration to Canada and the U.S. is from countries with high incidences and burdens of TB [15,16,13,8,9,31,47,57,63]. As well, immigration has been steadily increasing to Canada since 1987 and is now averaging over 200,000 people per annum [9]. The result is an increasing number of new Canadian citizens being at risk for developing active disease [36,48]. Although the Canadian rate of active disease has been steadily decreasing, the proportion in foreign born people has been increasing every year [13,15,16,17]. The same pattern can be found in the U.S. [36,16,8,63,62,29,64,56,18], Italy [58], the

UK [65,43], Germany, Denmark, the Netherlands, Norway and Sweden [57].

The foreign-born population is made up of many small segments from all over the globe that experience widely different risks for TB [13,16]. Although parts of Asia and Africa can have high morbidities of TB, some populations from the U.S. or areas in western Europe have lower morbidity rates than the Canadian-born population [13]. It is thought that in the majority of developing countries the annual risk of infection with TB is between 1.0 and 2.0% [10,6]. These numbers are still only estimates. If they are to be biased, however, it is much more likely that they underestimate the truth, rather than overestimate. Even within high risk sub-populations, it is suggested that there are particular high risk groups [9]. One study showed that although a group of Southeast Asian refugees had an overall elevated incidence of TB, new cases were primarily from the group found to be tuberculin positive at immigration [44]. In fact, these people had a risk more than 13 times that of the tuberculin negative refugees tested at time of immigration [44]. In addition, the 5 year cumulative incidence of the tuberculin negative refugees was similar to that of the general U.S. population [44].

It has also been observed that rates are often lower in the people that emigrate compared to the national average rate in their country of birth. These people are often a special group who are younger, healthier and economically advantaged [13,14]. This phenomenon, however, is less likely to be seen in refugees. Thus, demographic and socioeconomic factors must be considered before establishing high risk populations.

### 2.2.4 Characteristics of TB Incidence in the Foreign-born

Incidence of tuberculosis among the foreign-born has been identified in Canada and globally as an important factor in the control and elimination of TB. The Canadian Expert Committee on Tuberculosis (ECOT) identifies all those arriving or returning from countries where tuberculosis is endemic as at possible risk, before and after arrival [114]. In 1997, the CDC Working Group on Tuberculosis among the Foreign-born published a report of recommendations for prevention and control of tuberculosis. The following was one of the recommendations:

"The epidemiology of TB among foreign-born populations differs considerably from area to area. To tailor TB-control efforts to local needs, TB-control programs should develop epidemiologic profiles to identify groups of foreign-born persons in their jurisdictions who are at high risk for TB." [16]

In a recent American study, it was found that among foreign-born persons, the most important risk factors for TB were world regions of origin and less than 5 years residence in the U.S. [62]. The probability of developing active TB among immigrants primarily reflects the infection acquired before immigration [48,20,12,7,42,13]. The tendency for infection to occur prior to immigration has been shown by molecular epidemiologic studies of TB clusters and the country of origin of people in these clusters [36]. Clustering is thought to be a good indication of recent infection by *M. tuberculosis* [36]. It has been found that foreign-born cases are less likely to be infected with similar strains of TB as nonforeign-born cases [36]. This would imply that infection was likely remote and the person has reactivated [16,18].

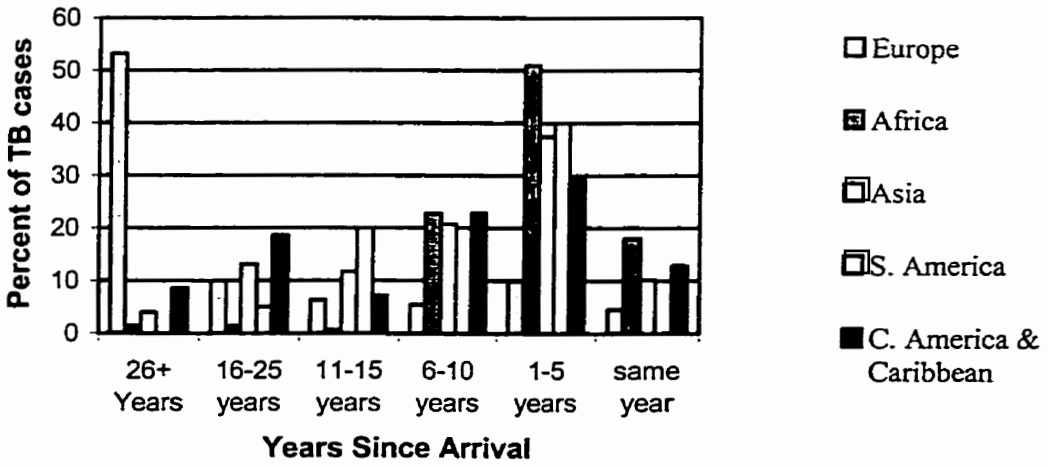
There are differing reports as to when the risk is highest, and for how long the increased risk is present. In one study it was shown that many immigrants from high prevalence areas had TB rates higher than 20 per 100 000 even more than 20 years after immigration [36,16,40]. The British Thoracic and Tuberculosis Association found that immigrants from Pakistan, India and Africa had an incidence of TB 5 years after immigration that was 78 to 95% the incidence experienced upon arrival [66]. Perhaps diagnosis is not being made on incident cases due to the fact that they had latent disease that reactivated, or perhaps the illness was not detected immediately.

On the other hand it was observed, in the Southeast Asia refugee study, by the fourth and fifth year after immigration, annual incidence was only 10% of that at time of immigration [44]. In Vietnamese refugees in Denmark (that has an annual incidence the same as Canada) TB rates decreased by a factor of 14 in the years following immigration, from 1.14% to 0.07%, and no TB cases were seen after 11 years [40].

It is generally believed that rates of TB are much higher among recently arrived persons [36,44,11,48]. A person's risk of developing TB following migration is usually highest in the first years and then starts to

decline but still reflects the risk in their country of origin [13,15,47,66,7,11,42]. The sharpest decrease in risk is seen in immigrants from countries experiencing high burdens of TB infection and disease and this decrease is sharpest in the first two to five years after arrival [16,63,12]. Canadian TB data for 1996 [67] was used to generate Figure1 in which the different pattern of diagnosis depending on endemicity of TB are visible. For instance, over 50% of all people from Africa diagnosed with TB in 1996 arrived within 5 years of diagnosis [67]. On the other hand, over 50% of European cases diagnosed in 1996 had arrived over 26 years ago [67]. Different trends are observed in different studies suggesting that pattern of disease incidence among the foreign-born may depend on characteristics of the immigrating population.

**Figure 1. Cases of TB reported in Canada in 1996 in the Foreign-born. Percent of total cases from each region that were diagnosed in each time period and time between arrival and diagnosis are presented.**




*Data Courtesy of Health Canada [67]

## 2.3 The Canadian Situation

### 2.3.1 The Status of TB Research in Canada

Canada has one of the lowest TB rates in the world that has been essentially stable since 1987 at between 6 and 8 cases per 100,000 [68,9] . Unfortunately, however, the rates of disease vary among subgroups in the Canadian population. For instance, although Canadian-born people have a disease rate of only 1.6 per 100,000, Aboriginal people have a rate of 24 per 100,000 and immigrants of Asian origin have a disease rate of 50 per 100,000 [9,15,67]. The epidemiology of TB has been significantly changing over the past 10-15 years with much higher proportions of TB in the foreign-born and an increase seen among Aboriginals including the Inuit [51,68,15,9,69].

A descriptive study was conducted recently in Ontario to examine trends over the past 25 years. This study found that 80% of TB cases in Ontario live in urban centres, primarily Metro Toronto and Ottawa-Carleton areas[9]. The greatest risk factor for TB in Ontario was identified as travel to, or residence in a country with high endemic rates [9,31]. In Ontario, 80% of cases are in the foreign-born [14,9]. In Montreal, studies looking at factors affecting tuberculin reactivity and drug resistance in the foreign-born found that annual incidence in the foreign-born was 10 times the rate for Canadian-born people and that incidences reflected the incidence in their region of birth [12]. Factors affecting tuberculin reactivity were rates in the country of origin, age at immigration, BCG vaccination, and residence in poorer neighbourhoods in Montreal [69]. Recent Asian immigrants to British Columbia were studied to determine their risk of TB and they were found to have an annual incidence rate 8 times the British Columbia annual rate [47]. In Alberta a case-control study was conducted looking at foreign-born patients between 1990 and 1994 [48]. They found that immigrants accounted for 70.6 percent of TB diagnosed in Alberta and were mainly of Asian origin. They also found that 50 percent of patients presented within 7 years of arrival. According to the Ministry of Health, among foreign-born cases diagnosed between 1989 and 1995, the mean time between arrival in Canada and diagnosis was 9 years, and over half were diagnosed in the first 5 years [9] . In Manitoba they found most cases occur within 4 years of arrival in Canada [15].

Between 1968 and 1995 in Ontario, the proportion of cases of TB in the foreign-born has increased from 42% to 81%. Thus, although our overall number of cases per year has been in decline over the past 4

decades, certain subgroups of the population in Canada are at very high risk and have high incidences of active disease [13]. This disease must be prevented or at least controlled if we are to avoid outbreak situations in dense urban areas.

### 2.3.2 Immigration and Social Implications

According to Citizenship and Immigration Canada (CIC), in 1993, of the 255,819 immigrants that entered Canada, 52.5% were destined for Ontario, and 28.4% for Toronto [9]. It has recently been suggested that Canada increase it's annual immigration to 300,000 people annually [75]. There is a definite trend for people entering the country to settle down in big cities [7]. There are many reasons for this, more job opportunities, greater diversity of cultures, races and religions in larger urban centres, thus promoting a greater feeling of comfort and support, as well as the simple reason of being more familiar with the names and locations of big cities. On top of that, one, or two family members often precede the arrival of an extended family some time later. The result of this is that our urban centres are becoming more populated and more ethnically diverse. Along with ethnic diversity and the cultural and economic wealth it bestows upon us, come new diseases. In Quebec, Montreal is accounting for a higher and higher proportion of annual TB cases in the province, from 36.7% in 1985, to presently over 60% [12]. With over 77% of these cases being in the foreign-born [12], and 92% in Toronto being in the foreign-born [31], it suggests that the urban centres of Canada are already experiencing an increase in this infectious disease.

Something that must not be lost sight of, however, is the inextricable link between TB and poverty, and immigration and poverty, especially for refugees and non-skilled labourers. The highest rates of TB are in large cities, due to crowding, and poverty [7,31]. A study completed in England and Wales investigated the association between TB in London boroughs, poverty and immigration. It was found that immigration was correlated with social deprivation and overcrowding which all influenced tuberculosis rates in urban areas of England [33]. In another European study, both low education and not speaking the native language were associated with higher rates of TB [58].

A new immigrant experiencing symptoms of tuberculosis may avoid seeking medical care due to varying attitudes towards the health care system ranging from simple ignorance to fear [16,15,18,29]. This

attitude can affect both case finding success, as well as compliance with treatment once someone with active TB is identified [15]. The health care system is not always accessible to people who do not speak the language, are unaware of the procedures and practices in this country [16]. There is some evidence that immigrants' and refugees' use of health care is less than that of other Canadians [70]. New or non-English speaking immigrant women to Canada have been found to be less likely to receive screening for cervical and breast cancer, or to participate in prenatal care [70-72]. Some immigrants come from countries where having TB is associated with incarceration, loss of employment and social isolation. In many countries TB is often related to HIV infection which comes with its own prejudices [16,73]. At best, TB is usually associated with stigma and fear [73].

### 2.3.3 Canadian Immigration Screening

Pre-immigration screening for TB is performed for all immigrants, all convention refugees (refugees applying for refugee status from outside Canada), students from "designated" countries seeking a visa for greater than 6 months, and visitors on long-term visas from "designated" countries [73]. All applicants for immigration are screened before entering Canada by chest X-ray (if they are over 11 years of age), and medical history to determine those who have active or inactive pulmonary TB [47,15]. The notable exception for TB screening are "refugee claimants" which consist of people arriving in Canada who have not applied for landing status, and who then apply for asylum as refugees. An increasing proportion of immigrants are applying from within Canada. These people do not receive any kind of medical exam prior to entry. Among those who do apply from without, no TB skin test is performed to determine those who are infected with TB. If the radiograph is suspicious for TB, active disease must be excluded prior to permission being granted for immigration to Canada [47,15].

Anyone who has active TB is inadmissible to Canada (unless they are seeking refugee status or entry on humanitarian grounds) but can be reassessed after completing treatment [15]. After arriving in Canada, a person is referred for medical surveillance if they have an abnormal chest X-ray, or have a history of treated TB [15,47,48]. Medical surveillance (MS) involves reporting to provincial health authorities within 30 days of arrival in the intended city or town for a medical evaluation focussing on TB and the option of treatment. Once

a person is legally landed, however, there are no consequences of not following up with their surveillance [15,14].

Although persons referred for medical surveillance (MS) are at an increased risk for TB, they contribute a relatively small proportion of future cases of TB annually [7,4,6,15]. One recent study showed that persons referred for MS only accounted for 11-13% of all TB cases resulting from this screened population in Ontario [14]. Thus, the number of persons referred for MS does not reflect the number of actual active TB cases accrued from immigrating populations. A method to access other high risk groups is necessary. In order to do this a detailed understanding of the foreign-born population and its associated risk for TB are needed. Epidemiological profiling of this high risk group must be done to better design targeted testing programs.

Part of the current problem lies in the fact that the immigration screening program is designed to identify people with active disease, or with suspicious chest X-rays suggestive of once active disease, not all people infected. Therefore, you may have a large group of people remotely infected with TB who have inactive disease that reactivates at some point after arrival to Canada. Latent TB is particularly dangerous with respect to reactivation if the person has never received any therapy, which in immigrants accounts for approximately 50% of inactive cases [47]. The other scenario is that you have people who are recently infected who are not yet showing symptoms of active disease, who develop active disease after their arrival. Another controversial problem with the current system is that there are no repercussions to the new citizen or landed immigrant to report for further testing, surveillance or treatment [15]. It is left up to the discretion of the individual whether they adhere to suggested conduct. As well, it would appear that in provinces such as Ontario where this evaluation is often undertaken by family physicians, there are very low rates of utilization of preventive therapy. Due to reasons outlined above about social stigmas and fear attached to TB, it is understandable that a significant proportion of these people are never seen again, until they develop active disease, or some other health outcome [15].

## 2.4 Summary

The World Health Organization has estimated that one-third of the world's population are infected with tuberculosis. The distribution of infection, however, is not equal across the globe. There tends to be a different TB epidemiology in developed countries, where the burden of disease is primarily found in the elderly, than in people from developing countries, where the burden is well-distributed across the entire population. Most infections happen early in life in a high-prevalence country and this is associated with high rates of disease in younger segments of the population. High rates of HIV among young, reproductive age adults in some developing countries is contributing to high rates of TB among this population. One clinical aspect of TB that complicates control strategies, is that it is a disease that can become latent and reactivate at any time during a persons' lifetime. High burdens of infection in the young and unpredictable reactivation patterns are two factors that make it a difficult disease to eradicate, globally.

There are 4 very significant global issues that are complicating the control of TB. These are:

1) An increase in multi-drug resistant strains of TB
2) An increase in TB/AIDS coinfection
3) Increased urbanization
4) Increased migration of populations

The decline of TB incidence in Canada has halted since 1987 and disease rates in the foreign-born have been increasing. In both the U.S. and Canada, it is becoming recognized that the most important factor influencing the "stagnation in the decline in TB incidence" is immigration from high burden countries [62]. Canada receives large numbers of immigrants from areas that have high rates of endemic TB. It has recently been suggested that Canada increase it's annual immigration to 300,000 people (from approximately 200,000 currently) [75]. Immigration of large numbers of people with high rates of infection means that the health care system must focus greater attention on preventing active cases by treating latent infection. The only other alternative, which is unacceptable economically, ethically and socially, is to stop immigration.

Ontario receives the majority of immigrants. In Ontario 80% of TB cases are in the foreign-born, in Montreal, the same population accounts for 77% of TB. Ontario also does not have a centralized TB

control program which means that treatment is monitored primarily by the treating physician and not monitored at a provincial level [60]. Because of this, there is little power to implement a province-wide standard of treatment and targeted testing programs. It leaves open the possibility that this reemergence of TB, especially in urban areas, will have an effect by increasing the overall Canadian rate. To re-establish the earlier pattern of decline in TB rates across North America, a greater unified public health effort to identify and treat active cases, and to use preventive measures on high risk populations will have to be undertaken

There are 2 strategies for case-detection, active and passive. The current screening program used by immigration authorities is not designed to identify people at high risk for TB due to infection. It is meant to identify people who have active disease or who are at high risk due to old, latent disease (identified by chest X-ray as scarring in the lung tissue). The outcome of not preventing disease in someone who is at high risk impacts both the individual and the society. The longer a person has untreated active disease, the longer they remain infectious and the more secondary cases may arise. Seeking treatment for TB when someone comes from an entirely different political climate, different cultural belief regarding disease or does not speak the language may not occur in a timely fashion. Acceptance and completion of treatment of latent TB infection among these high risk groups is often low [18]. In order to even begin tackling the problem of accessing high risk populations of various ethnic backgrounds, a clear understanding of the epidemiology of TB is needed.

It has been suggested that TB cannot be eradicated through the treatment of active cases alone, and that chemoprophylaxis of high risk groups is needed [1,18]. Treating tuberculin converters (people who are recently infected) is a rewarding tactic since it is highly effective in preventing disease in the individuals, and their contacts which would have been at risk for infection. In order for treatment of latent TB infection to increase in importance in TB control programs [18], more needs to be known about the epidemiology of TB in immigrants to Ontario. Information about risk factors associated with disease in this population, the most effective timing of an intervention or screening program after immigration, as well as how to access high-risk immigrating populations needs to be improved [18].

Predicting the future burden of TB would facilitate the designation of sufficient funds and

research impetus towards the containment and prevention of any serious public health threat due to TB in this province. Understanding the risk of TB among immigrant groups and the behaviour of risk over time in these populations is the first step towards predictive TB modeling and designing appropriate interventions.

## 3.0 METHODS

### 3.1 Study Objectives

There were two objectives of primary interest in this study.

1. To describe the TB occurrence among immigrants to Ontario from 1990-1997 according to their demographic characteristics such as age, sex and world region of origin.
2. To describe the effect of time since arrival on risk of TB and to analyze determinants of risk of developing TB over time after immigrating to Ontario.

### 3.2 Overview of Study Design and Methods

This study was a population-based retrospective cohort study. The population under study were immigrants to Ontario from 1990-1997 and study outcome of interest was development of active TB in the 1990 to 1998 study period. The study variables were: World Region of Origin (WRO), age at arrival, sex and year of landing. Primary data sources included immigration data from Citizenship and Immigration Canada and disease surveillance information from the Ontario Reportable Disease Information System.

Objective 1 involved both investigating the data using descriptive statistics, as well as studying the risk and pattern of developing active TB among people immigrating between 1990-1997 using incidence rates and risk ratios. The numerator included the number of TB cases in foreign-born people occurring between 1990 and 1998 who arrived during the study period 1990-1997. The denominator included all people legally immigrating to Ontario between the years 1990 and 1997.

Descriptive statistics and bivariate analyses were performed looking at variations in age distribution, sex, immigration patterns from the different world regions and socioeconomic status. Incidence rates were calculated for different world region of origin strata. Age and sex-specific rates were calculated for the various world regions. Age and sex standardized incidence rates were presented by World region of origin. Risk ratios were generated to compare the risk of various world regions to the Canadian-born, non-Aboriginal population.

Objective 2 focussed on the time to developing active TB since arrival in Canada. This involved determining if there were differences between the world regions, sexes, age groups and calendar years of

arrival with respect to risk over time since immigration. Actuarial life-table analysis and log-rank tests were used to evaluate the chance of developing active TB at different times and to identify the significance (at 0.05 level) of each of the study factors in relation to their risk of developing active TB. The factors of interest were age at arrival, sex, time since arrival, WRO and calendar year of arrival.

Cox proportional hazard and complementary log-log models were further applied to account for multiple factors such as age, sex, WRO, and calendar year of arrival simultaneously (including those found significant in the log-rank test) to establish survival probabilities at different time points.

### 3.3 Data Sources

Citizenship and Immigration Canada (CIC) have a main administrative database, containing information on all persons entering Canada. The Landed Immigrant Data System (LIDS) is a batch file compiled from this main database, however, containing only information obtained from the landing documents of persons migrating to Canada. "Facts and Figures" is a non-nominal analytical database assembled from the LIDS database that is available for the purpose of research. The information it contains for all immigrants, derived from the landing document, is: year of landing, year of birth, sex, educational qualifications, country of birth, country of last permanent residence, country of citizenship, occupation, and immigration class.

The Ministry of Health, Ontario, maintains the second database, the Reportable Disease Information System (RDIS). Included is the Tuberculosis Registry for Ontario. For research purposes, information from this database is available by applying to the Ministry of Health for access under the Freedom of Information Act. Confidentiality requirements must be met before access is given. This database contains a record of all cases of reported tuberculosis in Ontario. Anyone diagnosed with TB is included in this database by requirement of the Health Protection and Promotion Act regardless of whether they are Canadian-born, landed immigrant, refugee, or student. Visitors are not included in the registry unless they have an official Ontario residence (Communication with Kingston Health Unit). Within this database we have access to the age, sex, country of birth and year of arrival of each person. Site of the disease and drug resistance is also available through this database. The case definition for inclusion in this database is described below.

For the purpose of age and sex standardizing our data, and in order to produce risk ratios, basic demographic data on the Canadian population was also necessary. The Canadian-born, non-Aboriginal data was accessed through the "public use" microdata files (PUMF) from the 1996 Census of Canada compiled by Statistics Canada. This data can be accessed through the Queens University **Social Science Data Centre (SSDC)**, part of the Documents Unit. These microdata files are an anonymized sample of raw census records.

3.3.1 Inconsistencies Between the Databases

The study cohort, and hence the denominator data, was comprised of people who are legally landed with an intended destination of Ontario. The numerator data, includes all foreign-born persons who were diagnosed with active TB in Ontario, including non-landed persons such as students or refugees. Obtaining reliable denominator information for such persons was not deemed possible. Thus, we would expect to over-estimate TB rates by a certain amount. The results of a previous study which determined the proportion of foreign-born persons with TB who were landed, were used to make crude adjustments to our rates [14]. This was done in order to reduce the potential for overestimation and/or bias due to the numerator including a non-random population that was not in the denominator.

### 3.4 Definition of TB Cases and Study Variables

It is required that all active cases of TB be reported to the Public Health Branch in the Ontario Ministry of Health. These cases in turn are reported in the Tuberculosis Registry in RDIS. Cases are defined either by the presence of *Mycobacterium tuberculosis* on any culture from sputum, body fluids or tissues, or, in the absence of bacteriological proof, by the presence of radiological or pathological symptoms or signs. These symptoms or signs are preferably accompanied by a positive tuberculin skin test and/or acid-fast bacilli in sputum or other body fluid smears, and/or response to antituberculous treatment.

The Facts and Figures database contains only immigrants who are legally landed, which excludes visitors, students or refugees. For the purpose of the survival analysis, "start time" for observation of that subject was considered as the date when "landed" status is given. In RDIS, the field "date of arrival" was

used as a proxy for when the person became a landed immigrant in Canada. It is the only date available for this purpose in the RDIS database. Date of diagnosis was considered "finish time" as when the outcome of interest occurred.

The three study variables of interest were World Region of Origin, age at landing, and sex. The countries of the world were broken down into 'World Regions of Origin' (WRO) using a schema developed by the World Bank [98]. These are (along with the short-forms used in this study): Sub-Saharan Africa (SSA), India, China, Other Asia and Islands (OthAsia or Asia), Latin America and the Caribbean (LatCar), Middle East crescent (MEC), formerly socialist economies of Europe (FSEE) and established market economies (EME)[42]. People were assigned to a world region based on their recorded "country of birth" in Facts and Figures and RDIS.

In addition, age was divided into 5 age groupings. These are: 0-15, 16-30, 31-45, 46-65, 66+. Sex, was divided into male and female.

Once the denominator (Facts and Figures) and numerator (RDIS) data were categorized by our study variables, the two databases were merged. Then the data was linked by category to include both numerator and denominator information for each strata of the variables of interest.

### 3.5 Objective 1

#### 3.5.1 Exclusion Criteria and Initial Cleaning

The RDIS data was received for the years 1990-1997 which included anyone diagnosed with active TB within the time period January 1990 to March 1998. Cases of TB in the "foreign-born" were selected from the database. This was done by deleting any cases with the "Origin" field as "Born Canada", "Reg Indian" (registered Indian), "Unspecified" , "Unknown" or "Unregistered". Of those cases occurring in the "foreign-born", only TB cases claiming "date of arrival" as being between 1990 and 1997 were kept. The cases were then evaluated for sex, date of birth, date of arrival and country of birth, entries containing blanks for any of those fields were deleted. As well, any entries containing a date of birth of "9999" or a date earlier than 1870 were deleted. All TB cases (pulmonary and extrapulmonary) were included in this study.

"Facts and Figures" was assessed for the years 1990 to 1997. Only entries with a "06" code for the "province of destination" field (the Ontario code) were included. Anyone with a blank "province of destination" or "sex" field were deleted. Anyone with a blank "Country of birth", "Country of Last Permanent Residence" and "Citizenship" entry (thus no country affiliation at all) or with "Country of Birth" as Canada were also deleted. The overall criteria for inclusion in the denominator were: being born outside Canada, arriving between 1990 and 1997, a destination of Ontario and with complete demographic information for date of birth, sex and either country of birth, country of last permanent residence, or citizenship.

The immigration data (from Facts and Figures) was in the form of individuals and this was organized into the strata of interest. This task was done using ACCESS [76] to filter and organize the data. The data in this grouped form was then transferred into SAS [77] for cleaning and analysis.

3.5.2 Bivariate and Descriptive Analyses

The characteristics of cases that were included and excluded in the study were compared. Reason for exclusion was also presented for both the excluded foreign-born and Canadian-born.

The overall proportions of TB cases in the study captured from 1990 to 1997 among the immigrant population to Ontario were compared according to sex, age groups and world regions using Chi-square tests. Using this method, disease rate varying according to world region of origin, age and sex was investigated. The differences between age distributions among Ontario immigrants by world were evaluated.

The number of years of schooling was used as a proxy for variation in socioeconomic status (SES) across the world regions for all landed immigrants to Ontario. The justification for using schooling as a proxy for SES is discussed in previous papers [58,80]. This analysis included only those of age 15 years or older. Only a basic ecologic analysis of SES was possible due to the inability to link data at the individual level for any socioeconomic information. Thus, the mean number of years of education for each world region were correlated with its age and sex-standardized rate of TB and the Pearson correlation coefficient was calculated.

3.5.3 Calculation and Analysis of Incidence Rates

The incidence rates for the first year after landing were calculated for each year of immigration from 1990 to 1997. These rates were compared (using Chi-square tests) to assess homogeneity of TB incidence across different yearly cohorts. If the years were significantly different, the data would be kept separate. If the underlying incidence rate was homogenous across the different years of arrival, then the data from all years of arrival could be pooled.

Since immigrants in different years had different follow-up periods, incidence rates were calculated according to the person years. Age and sex specific incidence rates per 100,000 person-years were calculated for each world region with. ninety-five percent confidence intervals. Then, a chi-square test for trend in the age and sex specific incidence rate data was performed. The individual world region rates were then age and sex standardized to the Canadian-born, non-aboriginal population using 1996 Census data. Direct standardization was used to generate risk ratios by dividing the standardized rates by the rates of the Canadian-born, non-aboriginal population.

The Canadian-born, non Aboriginal rate of active TB in Canada in 1996 (1.6 per 100,000) was used as the "unexposed" comparison group for generating risk ratios. This was the rate of *reported* new active and relapsed tuberculosis cases in Canada from the Health Canada 1996 "Tuberculosis in Canada" report published in 1998 [67]. This number was confirmed using the number of cases in this study and Census data from Statistics Canada. These results were considered sufficiently similar and the highest estimate was used in order to obtain the most conservative results. The risk ratios of the age and sex standardized rates for the World Regions were then presented along with 95% confidence intervals.

A further step was performed in order to correct the estimates for possible overestimation in the rates due to some cases not being landed immigrants (and therefore not being included in the denominator). Results were used from a previous study on a two year sample from this population [14]. This study was case control in design, and involved linking the CIC and RDIS databases. The results of that linkage were used in this study to determine how many people in each world region strata could be expected to be found in the numerator, but *not* to be included in the denominator. This proportion of cases that were believed to not be officially "landed" at the time of diagnosis were not included in the adjusted rate.

### 3.6 Objective 2

3.6.1 Overview

Using the merged databases (Facts and Figures and RDIS), time from arrival to active TB or from arrival to the end of 1997 was calculated for each immigrant coming to Ontario. The survival distribution was estimated by the actuarial life table method. The survival distribution describes the probability of developing active TB beyond any specific time interval since arrival in Canada [78].

The hazard rate at different times since arrival in Canada were also estimated by the actuarial life table method. To compare the survival distributions (hazard rates) between different sexes, age groups and world regions, log-rank tests were applied [79].

Proportional hazard regression models and complementary log-log models were further applied to incorporate all the variables concurrently in a model as predictors of survival. The advantage of modeling is to take into account possible confounding factors such as age and sex while focussing on comparing different WROs [78]. The appropriateness of the Cox proportional hazard model was investigated through checking the proportional hazard assumption. The significance of each of the variables was investigated both as predictors of survival time and as confounders or effect modifiers.

3.6.2 Data Organization

The RDIS data was organized by age, sex, world region of origin and time from arrival to time of diagnosis. The date of diagnosis minus the date of arrival was the Survival time (in years to the nearest tenth of a year). The survival times were then grouped into one year intervals starting at "0-1 years of survival" to "7-8 years of survival". Anyone who had a date of diagnosis before the date of arrival, or had "date of diagnosis" missing, had their survival time re-calculated using "episode date" as their "date of diagnosis". "Episode date" refers to a period of time during which the person experienced symptoms of active disease and came into contact with a physician due to the illness. If their survival time was still computed as a negative, or if there was no "episode date", they were not included in the survival analysis. These people were included in a descriptive analysis of characteristics of people who are not landed at the time of diagnosis.

The Facts and Figures (population at risk) data was organized using December 31, 1997 minus date of arrival as the time in years "at risk" for TB. Since the time in years that a person was in the country was the maximum amount of time that they could have been "observed" as being in the study, this time "at risk" was used as the contributed survival time for that strata. For example, everyone who arrived in 1997 made up the number of people "censored" after 1 year, minus the number of people who got TB (regardless of the date they landed) after being in Ontario for 1 year or less. Therefore, survival time for the denominator data, since individual information on disease status was not included in this data, was based on observation time i.e. time since landing.

SAS [77] was used to merge these two data bases containing the disease data and the immigration data.

3.6.3 Confounders and other Covariates

Confounding by age and sex was addressed by using stratification. There is some suggestion in the literature that the effects of gender may vary depending on age[28]. Thus, this possible interaction was investigated.

Socioeconomic status is a possible covariate of interest, however, no information on this variable was available for cases. Number of years of schooling among the denominator population and prevalence of disease in the study population was plotted across the different strata as part of Objective 1. Only this descriptive report on the visible trends of socioeconomic status with respect to WRO and year of landing among the study population was undertaken. The limitations of not having information on socioeconomic status will be discussed further in the discussion of this paper.

For this objective, the three time-independent variables of interest, age, sex and world region of origin, were all included in the proportional hazards model building procedure to investigate the contribution to survival of all three as well as to look at confounding and interaction among and between them. The time-dependent covariates, year of arrival and duration of time in Canada, were investigated with the complementary log-log model.

### 3.6.4 Life Table and Survival Estimates

The SAS procedure "lifetest" [77] was used to generate survival distribution functions for the factors of interest using actuarial life table methods. Survival estimates were applied to determine the survival distributions according to different world regions, sex and age groups. This type of event time data is called the group survival data [81]. The survival distribution describes the chance of not developing active TB within any specific time interval since arrival in Canada. These duration times were grouped into one year intervals. Those who did not develop TB (the majority) were considered as censored event times.

Equivalent to the survival distribution, the hazard rate can also be used to describe the TB development pattern among immigrants. It indicates the TB risk (instantaneous incidence rate) at any specific point in time since arrival in Canada. The survival distributions (and hazard rates) can be graphically displayed by plotting the corresponding estimates against time (year). The hazard rate plots were presented due to the ability of presenting more visual information with hazard rates in the case of a rare disease than survival rates. Hazard estimate plots also allow us to ignore the confusing information of censored information, since the majority of the population did not experience disease [81].

Log-rank tests were applied to see if the risks of developing TB over time is different for different sexes, age groups or world regions. These tests were used to evaluate the homogeneity of the survival functions across strata and to identify the significance (at 0.05 level) of each of the study factors in relation to their causing a differing risk of developing active TB.

*Censored* subjects included all those who did not get disease by the end of the study period. The actuarial method was used instead of the Product-Limit method since the data was organized by discrete time intervals of disease instead of exact times in the form of continuous data. As well, we assume that the censored times were distributed uniformly over the observed time intervals.

### 3.6.5 Cox Proportional Hazard Model and Complementary log-log Model for Group Survival Data

To incorporate all the time-independent variables, including age, sex and world region of origin, simultaneously and relate them to the TB risk at various times, the Cox's proportional hazards model for group survival data [81] was applied. Using this model, hazard ratios can be generated which can be interpreted the same as an odds ratio [78]. An alternative equivalent approach using the Complementary log-

log model was applied to estimate the regression parameters and risk ratio of each factor and covariate including the time-dependent variable, calendar year. The advantages of using the Complementary log-log model are: (i) the parameter estimates agree with the Cox proportional hazards model; (ii) it also provides an evaluation of the patterns of TB relative risk, including by year of arrival, for any immigrant over the time since his/her arrival in Canada [82].

Parameter estimates were generated using the method of maximum likelihood [78]. Dummy variables were defined for all the levels of our covariates. A model was fitted one covariate at a time and the null hypothesis that the variable did not improve the model was tested by comparing the likelihood ratio test [-2log(L)] values. The model was simplified by excluding those variables that were not significant at the 0.05 level of significance (by the likelihood ratio test). Possible interactions were assessed and included to improve the model fitting. All the analyses were performed using the different procedures of the Statistical Analysis System (SAS) [77].

3.6.6 Appropriateness of the Models

Before using a Cox Proportional hazard model for the data, the proportional hazard assumption had to be tested. Namely, that covariates are time-independent, and if all the covariates were fixed at time 0, the hazard rates of 2 subjects with distinct values for the covariates would be proportional [100]. This assumption was tested using the method of graphing the log(-log) of the survival curves for the different covariates and expecting them to be parallel if the assumption was met.

When there are many ties in data, however, due to an inability to accurately assign failure time, Cox regression models may not provide an accurate approximation by using partial likelihood methods [82]. A solution to the problem of many ties is to use full likelihood methods (as opposed to partial likelihood) such as the complementary log-log model. In this model the parameter estimates are identical to the coefficients in the underlying proportional hazards model [82]. As well the complementary log-log coefficients have a relative risk interpretation, just like the Cox model coefficients [82].

## 4.0 RESULTS

### 4.1 Database Cleaning and Summarizing Included and Excluded Cases

The RDIS data for the years 1990-1997 contained a total of 6400 people. Cases that did not have "foreign-born" indicated in the "Origin" field, were deleted. Of those cases occurring in the "foreign-born", 2074 TB cases claimed "date of arrival" as being between 1990 and 1997. Any cases with blank entries for the fields sex, date of birth, date of arrival, or country of birth, were deleted. The remaining 2039 TB cases (all types) were included in this study. The following tables show the total number of TB cases received from RDIS, and the characteristics of those included and excluded from the study (Table 1).

**Table 1. Characteristics of Included and Excluded Cases.**

| *Characteristics of Cases* | *Included in Study (% total) n=2039* | *Excluded foreign-born (% total) n=2999* | *Excluded Canadian-born (% total) n=1139* |
|---|---|---|---|
| **Age (at diagnosis)** | | | |
| 0-15 | 210 (10.3) | 38 (1.2) | 114 (10.0) |
| 16-30 | 949 (46.5) | 665 (22.2) | 144 (12.6) |
| 31-45 | 426 (20.9) | 836 (27.8) | 245 (21.5) |
| 46-65 | 290 (14.2) | 663 (22.1) | 317 (27.8) |
| 66+ | 164 (8.0) | 798 (26.6) | 318 (27.9) |
| Missing | - | 3 (0.1) | 2 (0.2) |
| **Gender** | | | |
| Male | 1013 (49.7) | 1658 (55.3) | 699 (61.4) |
| Female | 1026 (50.3) | 1343(44.7) | 439 (38.6) |
| **Site of Disease** | | | |
| Pulmonary | 1176 (57.7) | 1636 (54.5) | 890 (78.1) |
| Extra-pulmonary | 840 (41.2) | 1295 (43.2) | 208 (18.3) |
| Not Available | 23 (1.1) | 71 (2.3) | 41 (3.6) |
| **Microbiologically confirmed** | 1466 (71.9) | 2233 (74.4) | 609 (53.5) |
| **Rate of Resistance*** | 250 (12) | 338 (11.3) | 28 (2.5) |
| **WRO** | | | |
| China | 190 (9.3) | 303 (10.1) | - |
| EME | 27 (1.4) | 391 (13.0) | - |
| FSEE | 64 (3.2) | 280 (9.4) | - |
| India | 213 (10.4) | 278 (9.3) | - |
| LatAmer | 102 (5.0) | 211 (7.0) | - |
| MEC | 124 (6.0) | 93 (3.1) | - |
| OthAsia | 758 (37.1) | 1125 (37.4) | - |
| SSA | 561 (27.5) | 274 (9.1) | - |
| Blank | - | 47 (1.6) | - |
| **Method of Detection** | | | |
| Symptoms | 1602 (78.6) | 2492 (83.1) | 831 (72.9) |
| Immigration Screening | 166 (8.1) | 28 (0.9) | - |
| Other | 248 (12.2) | 385 (12.8) | 301 (26.4) |
| Missing | 23 (1.1) | 97 (3.2) | 7 (0.7) |

* resistance to at least one first-line antituberculous agent.

The proportion of foreign-born cases excluded from the study that were diagnosed with extrapulmonary TB was slightly higher than in the study population. The rate of resistance was similar, however, for both included and excluded foreign-born cases, the rate of resistance was higher than for Canadian-born cases. More excluded foreign-born were detected due to symptoms, versus the immigration screening program. Excluded Canadian-born cases were also primarily detected due to symptoms.

More males were excluded from the study (61%) than females, and they were primarily excluded due to being born in Canada (not Aboriginal), unknown, or blank entry. Only 120 (11%) of cases were Aboriginal. The excluded group also tended to be older with between 57-68% 46 years of age or older. Only a small proportion of cases excluded from the study were under the age of 15. The Aboriginal population were evenly distributed across sex and age groups. The largest number of people were excluded due to being born in Canada (68%). The table below describes the Canadian-born excluded cases by age and sex as well as by reason for exclusion (Table 2).

**Table 2. Characteristics of Canadian-born or unknown origin cases excluded from study cohort.**

| *Reason for exclusion* | *Sex* | | *Age at Diagnosis* | | | | | *Total* |
|---|---|---|---|---|---|---|---|---|
| | *#Male (%)* | *#Female (%)* | *0-15 (%)* | *16-30 (%)* | *31-45 (%)* | *46-65 (%)* | *66+ (%)* | |
| "Born Canada" | 439 (61) | 282(39) | 79(11) | 80(11) | 152(21) | 194(27) | 218(30) | 723 |
| "Reg Indian" or "Inuit" | 64(53) | 56(47) | 16(13) | 26(22) | 24(20) | 30(25) | 24(20) | 120 |
| Unspecified, unknown, or unregistered | 114(67) | 57(33) | 2(1) | 20(12) | 34(20) | 51(30) | 64(37) | 171 |
| Blank | 35 (65) | 18 (35) | 1(2) | 7(14) | 8(16) | 14(28) | 20(40) | 50 |
| Date of birth before 1870 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Total | 652 | 413 | 98 | 133 | 218 | 289 | 326 | 1064 |

The age and sex distribution of the excluded foreign-born cases is described below (Table 3). There was a relatively even distribution of missing data across the sexes, people arriving before 1990 were slightly more likely to be female (56%). Those excluded due to missing data, tended to be over 65 or between the age of 16 and 30. Date of arrival before 1990 was distributed evenly across the ages except for the 0 to 15 age group which only accounted for 1% of pre-1990 arrivals.

**Table 3. Sex and age of foreign-born cases excluded from the study cohort.**

| *Reason for exclusion* | *Sex (row%)* | | *Age (row%)* | | | | | *Total* |
|---|---|---|---|---|---|---|---|---|
| | *Male* | *Female* | *0-15* | *16-30* | *31-45* | *46-65* | *66+* | |
| No date of arrival | 230(52) | 209(48) | 9(2) | 109(25) | 102(23) | 84(19) | 135(31) | 439 |
| No COB | 8(57) | 6(43) | 0(0) | 5(36) | 3(21) | 1(7) | 5(36) | 14 |
| No date of Diagnosis | 8(62) | 5(38) | 0(0) | 3(23) | 2(15) | 3(23) | 5(38) | 13 |
| Date of arrival before 1990 | 1414 (56) | 1125 (44) | 29(1) | 550(22) | 730(29) | 576(23) | 654(25) | 2539 |
| Total | 1660 | 1345 | 38 | 667 | 837 | 664 | 799 | 3005 |

The distribution of excluded cases by World Region of birth is presented below in table 4. Other Asia and Islands had the largest number of people excluded due to missing data or arrival prior to 1990. One hundred-and-fourteen cases had no date of arrival. The next largest number of cases with no data of arrival were people from EME. It is likely that many of these people arrived very remotely from Europe and date of arrival was therefore not considered important with respect to disease aetiology.

**Table 4. World region of birth of foreign-born cases excluded from the study cohort.**

| *Reason for exclusion* | *World Region of Origin (row%)* | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | *China* | *EME* | *FSEE* | *India* | *LatAmer* | *MEC* | *OthAsia* | *SSA* | *Blank* | *Total* |
| No date of arrival | 31(7) | 75(17) | 54(12) | 40(9) | 26(6) | 13(3) | 114(26) | 52(12) | 34(8) | 439 |
| No date of Diagnosis | 1(8) | 0(0) | 2(15) | 3(23) | 1(8) | 0(0) | 4(31) | 2(15) | 0(0) | 13 |
| Date of arrival before 1990 | 271(11) | 315(12) | 224(9) | 237(9) | 184(7) | 80(3) | 1007(40) | 220(9) | 1(0) | 2539 |
| Total | 303 | 390 | 280 | 280 | 211 | 93 | 1125 | 274 | 35 | 2991 |

The majority of cases (63%) that were excluded due to having arrived in Canada before the 1990 to 1997 study period, arrived between 1980 and 1989 (See Table 5 below). Of the remainder, 19% arrived between 1970 and 1979.

**Table 5. Arrival dates of persons not included in the study cohort due to date of arrival being before 1990.**

| *Date of arrival Interval* | *# of people (%)* |
|---|---|
| < 1950 | 92 (4) |
| 1950 - 1959 | 191 (7) |
| 1960 - 1969 | 174 (7) |
| 1970 - 1979 | 484 (19) |
| 1980 - 1989 | 1598 (63) |
| Total | 2539 (100) |

Cases that were included in the study are described below (Tables 6-8). Overall, most cases (46.5%) were between the ages of 16 and 30 and from either Other Asia and Islands, or Sub-Saharan Africa (Tables 6 & 7). The incidence of TB was generally distributed evenly between the sexes although there were differences within some of the world regions, namely Other Asia and China (Table 7 & 8).

**Table 6. Distribution of Cases Included in the Study from RDIS between 1990 and 1997 by age group and World Region of Origin (% of total).**

| *Age Grp.* | *China* | *EME* | *FSEE* | *India* | *Lat Amer* | *MEC* | *OthAsia* | *SSA* | *Total* |
|---|---|---|---|---|---|---|---|---|---|
| 0-15 | 4 | 4 | 3 | 8 | 18 | 11 | 60 | 102 | 210 (10.3) |
| 16-30 | 51 | 12 | 25 | 108 | 39 | 41 | 331 | 342 | 949 (46.5) |
| 31-45 | 39 | 3 | 14 | 34 | 28 | 29 | 202 | 77 | 426 (20.9) |
| 46-65 | 58 | 1 | 11 | 43 | 9 | 30 | 106 | 32 | 290 (14.2) |
| 66+ | 38 | 7 | 11 | 20 | 8 | 13 | 59 | 8 | 164 (8) |
| Total | 190 (9) | 27 (1.4) | 64 (3.2) | 213 (10.4) | 102 (5) | 124 (6) | 758 (37) | 561 (28) | 2039 (100) |

**Table 7. Distribution of Cases from RDIS Included in the Study between 1990 and 1997 by sex and age group.**

| ***Sex*** | *0-15* | *16-30* | *31-45* | *46-65* | *66+* | ***Total*** |
|---|---|---|---|---|---|---|
| Male | 100 | 454 | 223 | 145 | 91 | 1013 (49.7) |
| Female | 110 | 495 | 203 | 145 | 73 | 1026 (50.3) |
| Total | 210(10.3) | 949 (46.5) | 426 (20.9) | 290 (14.2) | 164 (8) | 2039 (100) |

**Table 8. Distribution of Cases from RDIS Included in the Study from 1990 to 1997 by sex and World Region of Origin.**

| | *China* | *EME* | *FSEE* | *India* | *LatAmer* | *MEC* | *OthAsia* | *SSA* | *Total* |
|---|---|---|---|---|---|---|---|---|---|
| Male | 113 | 16 | 34 | 111 | 48 | 62 | 346 | 283 | 1013 (49.7) |
| Female | 77 | 11 | 30 | 102 | 54 | 62 | 412 | 278 | 1026 (50.3) |
| Total | 190 (9.3) | 27 (1.4) | 64 (3.2) | 213 (10.4) | 102 (5.0) | 122 (6.0) | 756 (37.1) | 561 (27.5) | 2039 (100) |

From the Facts and Figures data (denominator data), yearly landed immigrants to Ontario for 1990 to 1997 were extracted. This resulted in between 113,476 and 138,791 people included in the study per year who fulfilled all the previously mentioned criteria (See Table 9 & 10). The differences between age distributions among Ontario landed immigrants by world region are reported below (Table 9). In general it was found that China and India had the oldest immigrant age distribution with China having 1/3 of the population immigrating being over 46 years of age (India had 27% over this age). Chinese immigrants under the age of 31 only accounted for another 1/3 of the population. All the other world regions had between 51 and 63% of their immigrating population under the age of 31. Sixty-three percent of immigrants from Sub-Saharan Africa, in particular, were under the age of 31. All the world regions, however, had the majority of immigrants between the ages of 16 and 45.

**Table 9. Age distribution of landed immigrants between 1990 and 1997 by World Region of Origin (%). N=975,435**

| *Age* | *China* | *EME* | *FSEE* | *India* | *Latin Car.* | *MEC* | *Other Asia* | *SSA* |
|---|---|---|---|---|---|---|---|---|
| 0-15 | 12 | 26 | 24 | 12 | 23 | 28 | 20 | 23 |
| 16-30 | 22 | 34 | 31 | 39 | 37 | 33 | 32 | 40 |
| 31-45 | 33 | 25 | 33 | 22 | 26 | 27 | 34 | 28 |
| 46-65 | 23 | 9 | 9 | 22 | 11 | 10 | 11 | 7 |
| 66+ | 10 | 6 | 3 | 5 | 3 | 2 | 3 | 2 |

The averaged annual crude rate for the different years of arrival was 47 cases per 100,000 people (95% C.I.s 43-50 per 100,000). The chi-square test of the homogeneity of the rates was conducted on the year 1 rate for all years of arrival (See Table 10). The result of this test was not significant ($X^2 = 3.1$, p=0.87). Thus it was considered reasonable to combine the data since the year 1 rates for 1990 to 1997 were comparable.

**Table 10. Distribution of TB cases and study population by year of arrival, along with incidence rates in the first year after arrival.**

| | *# of TB cases* | *Study Pop.** | *# TB cases in 1st year* | *1st year incidence rate (per 100,000)* | *95% Confidence Intervals* |
|---|---|---|---|---|---|
| 1990 | 460 | 113476 | 54 | 47.6 | 35.8-62.1 |
| 1991 | 372 | 119612 | 59 | 49.3 | 37.6-63.6 |
| 1992 | 324 | 138791 | 60 | 43.2 | 33.0-55.6 |
| 1993 | 296 | 134582 | 66 | 49.0 | 37.9-62.4 |
| 1994 | 253 | 117433 | 69 | 58.8 | 45.7-74.4 |
| 1995 | 163 | 115827 | 48 | 41.4 | 30.6-54.9 |
| 1996 | 117 | 119674 | 48 | 40.1 | 29.6-53.2 |
| 1997 | 54 | 118053 | 49 | 41.5 | 30.7-54.9 |
| Total | 2039 | 977448 | 453 | 46.4 | 42.2-50.8 |

* Study population is the total number of landed immigrants for that year from Facts and Figures.

### 4.2 Bivariate Analyses

#### 4.2.1 Disease status, sex , age and world region of origin

To determine whether the proportion of disease in cases varied by sex, a Chi-square test was performed and resulted in a non-significant association (p>0.05) (Table 11). The null hypothesis that disease status does not vary by sex could not be discarded. To evaluate whether sex contributed significantly to risk of disease at any level of the other variables, more chi-square statistics were performed (Tables 12 &13). The only significant (p<0.01) results were for the world region China and Taiwan (p<0.005) and the over 65 age group (p=0.001) in which cases more males experienced disease.

**Table 11. Unadjusted Chi-square for disease status and sex.**

| *Variable (Sex)* | *Number of TB Cases* | *Population* | *P-value* |
|---|---|---|---|
| Male | 1013 | 475,589 | 0.389 |
| Female | 1026 | 501,859 | |

**Table 12. Chi-square Test for Disease Status and Sex adjusting for age.**

| *Age* | *Sex* | *Number of TB cases* | *Population* | *P-value* |
|---|---|---|---|---|
| 0-15 | M | 100 | 107,299 | 0.647 |
| | F | 110 | 100,364 | |
| 16-30 | M | 454 | 150,207 | 0.590 |
| | F | 495 | 170,008 | |
| 31-45 | M | 223 | 147,129 | 0.931 |
| | F | 203 | 143,106 | |
| 46-65 | M | 145 | 53,682 | 0.100 |
| | F | 145 | 65,827 | |
| 66+ | M | 91 | 17,272 | 0.001 |
| | F | 73 | 22,554 | |

**Table 13. Chi-square test for disease status by sex adjusting for world region of origin**

| *WRO* | *Sex* | *Number of TB cases* | *Population* | *P-Value* |
|---|---|---|---|---|
| China | M | 113 | 47,655 | 0.002 |
| | F | 77 | 51,524 | |
| EME | M | 17 | 38,313 | 0.089 |
| | F | 10 | 43,868 | |
| FSEE | M | 35 | 62,692 | 0.318 |
| | F | 29 | 66,692 | |
| India | M | 111 | 40,214 | 0.683 |
| | F | 102 | 39,078 | |
| Latin America | M | 48 | 63,697 | 0.987 |
| | F | 54 | 71,425 | |
| MEC | M | 61 | 70,047 | 0.234 |
| | F | 63 | 58,462 | |
| Asia | M | 343 | 123,373 | 0.622 |
| | F | 415 | 143,527 | |
| SSA | M | 283 | 29,598 | 0.449 |
| | F | 278 | 27,283 | |

Age showed a significant association with disease status according to a chi-square test (p=0.001). (Table 14). A chi-square test was used to evaluate whether disease status varied by world region of origin and also yielded a significant result (p= 0.001) (Table15).

**Table 14. Unadjusted Chi-square for disease status by age.**

| *Variable (AGE)* | *Number of TB cases* | *Population* | *P-Value* |
|---|---|---|---|
| 0-15 | 210 | 207,663 | 0.001 |
| 16-30 | 949 | 320,215 | |
| 31-45 | 426 | 290,235 | |
| 46-65 | 290 | 119,509 | |
| 66+ | 164 | 39,826 | |

**Table 15. Unadjusted Chi-square test for disease status and world region of origin.**

| *Variable (WRO)* | *Number of TB cases* | *Population* | *P-Value* |
|---|---|---|---|
| China | 190 | 99,179 | 0.001 |
| EME | 27 | 82,181 | |
| FSEE | 64 | 129,384 | |
| India | 213 | 79,292 | |
| Latin America | 102 | 135,122 | |
| MEC | 124 | 128,509 | |
| Asia | 758 | 266,900 | |
| SSA | 561 | 56,881 | |

### 4.3 Calculation of Rates

#### 4.3.1 Age and Sex-specific Rates

The age and sex-specific rates per 100,000 person-years were calculated for each world region (See Figures 2a &b). Both sexes from Sub-Saharan Africa and women from India showed a peak rate of TB in the 16-31 years of age group. Men from the Former Socialist Economies of Europe had a maximum rate of TB in the 46-65 age group. Except for SSA and women from India, both sexes and all world regions showed the highest rate of disease in the over 66 years of age group.

The world regions tended to show a bimodal effect with TB rate peaks in the 16-30 and over 65 age groups. This phenomenon was seen at different levels of magnitude for all world regions except for the Former Socialist Economies of Europe (FSEE) and females from the Established Market Economies (EME) which showed relatively constant patterns of TB risk. The effect is apparent in both sexes not including the exceptions mentioned above. Within the bimodal pattern, the only variability is among the Sub-Saharan African female population which experiences their second peak at 46-65 years of age.

SSA had the highest rates consistently for both males and females following a very similar pattern and magnitude until the oldest age group, when male and female rates diverged. The next highest TB rates were in India, Other Asia and China, although the rates in women from China were lower that those of males. Although in the younger age groups the pattern in these 3 countries was similar between males and females (excepting women between 16 and 30 from China), among the oldest age groups there was a much higher increase in rates among males. The difference between men and women from Other Asia was over 100 cases per 100,000 in the over 66 age group, the males had double the female rate. Men from China also had twice the rate as their female counterparts at that age. Men from India had a rate that was approximately one-third higher than the rate of females from India in the oldest age group. Only Latin America and FSEE showed a pattern of females having a higher rate in the oldest age group. The difference was small for Latin America & Caribbean, but larger for the FSEE where females had almost double the rate as males in this age group.

**Figure 2a Age-specific rates of all forms of TB among women immigrating between 1990 and 1997.**

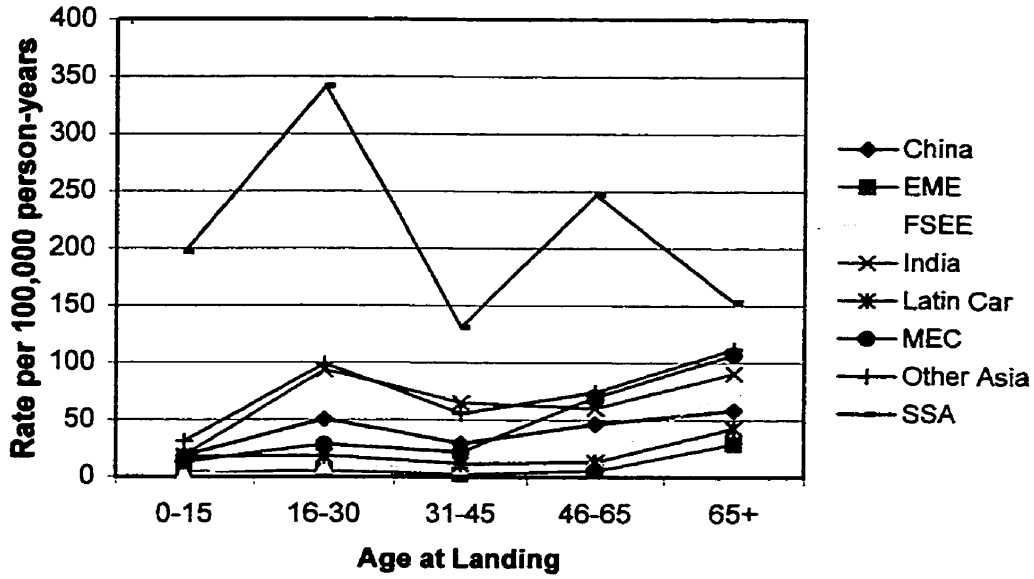
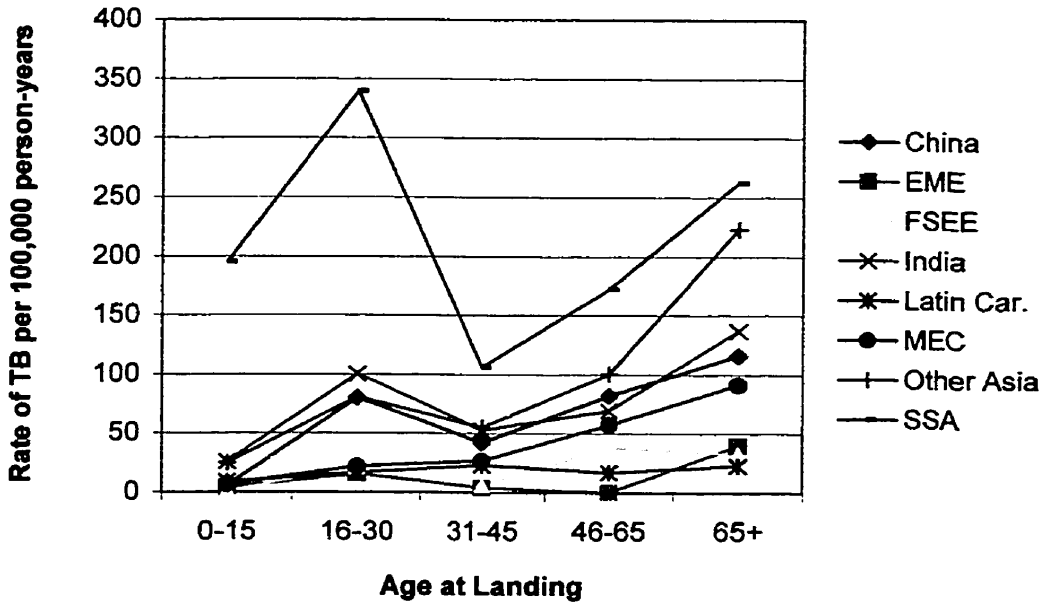



**Figure 2b Age-specific TB rates of all forms of TB among men immigrating between 1990 and 1997.**



4.3.2 Standardized Rates and Risk Ratios

Once the category-specific rates were calculated, the TB rates for the immigrants from the world regions were directly standardized to the Canadian-born, 1996 non-Aboriginal population. The age and sex-standardized rates for world region along with the risk ratios of each world region compared to the Canadian-born, non-Aboriginal population are reported below (Table 16). Standardized risk ratios were generated by dividing the standardized rate by the Canadian-born, non-Aboriginal rate. The ninety-five percent confidence intervals are also presented for the standardized risk ratio.

**Table 16. Age and sex Standardized Rates and Risk Ratios for TB by World Region of Origin**

| *World Region of Origin* | *Total # of TB cases* | *Total Population (total person-years)* | *Crude rate (per 100,000 person-years) & 95% Confidence Intervals* | *Age & Sex Standardized Rate (per 100,000 person-years)* | *Standardized Risk Ratio * (95% Confidence Intervals)* |
|---|---|---|---|---|---|
| China | 190 | 99179 (354190) | 52.8 (45.5, 60.9) | 45.7 | 28.7 (20.8, 39.3) |
| Established Market Economies | 29 | 82181 (382421) | 7.3 (4.9, 10.6) | 7.7 | 4.9 (2.2, 10.0) |
| Former Socialist Europe | 66 | 129384 (543926) | 11.6 (8.9, 14.8) | 16.4 | 10.3 (5.7, 16.6) |
| India | 213 | 79292 (286006) | 73.1 (63.5, 83.7) | 63.7 | 40.0 (30.3, 52.6) |
| Latin America & the Caribbean | 102 | 135122 (600500) | 16.8 (13.7, 20.4) | 17.5 | 11.0 (6.6, 18.1) |
| Middle Eastern Crescent | 122 | 128509 (445335) | 27.2 (22.6, 32.5) | 35.5 | 22.3 (15.5, 31.9) |
| Other Asia & Islands | 756 | 266900 (1089020) | 68.6 (63.8, 73.7) | 71.5 | 44.9 (34.1, 57.6) |
| Sub-Saharan Africa | 561 | 56881 (238316) | 235.4 (216.3, 255.7) | 210.1 | 132.0 (110.8, 156.8) |
| Total immigrant population | 2039 | 977448 (3939715) | 51.2 (49.0, 53.5) | 50.8 | 31.9 (28.4, 35.5) |

*Comparing world regions with the Canadian-born, non-Aboriginal rate of TB in 1996 of 1.59 per 100,000.

4.3.3 Correction of Rates for Overestimation due to Non-landed Immigrants in the Numerator

The data was then corrected for the potential discrepancy between people who appeared in the numerator of the rate as cases, and the people who made up the denominator as the population at risk. Using the former study [14], the proportion of cases from each of the world regions were determined that would likely not be found in the denominator. Every world region was found to have a certain proportion of people who were not "landed citizens" upon diagnosis with TB, or who were simply not linked due to missing information and who were therefore, not included in the denominator. The adjusted rates are presented below (Table 17). Sub-Saharan Africa had by far the largest decrease in rate after adjustment for non-landed cases, followed by MEC and China, although SSA still maintained the highest rate among the world regions. EME was least affected by the adjustment.

**Table 17. Crude Rate per 100,000 person-years, before and after Correction for non-landed Cases.**

| *WRO* | *Crude Rate* | *95% C.I.s* | *Correction Factor (%) ** | *Corrected Crude Rate* |
|---|---|---|---|---|
| China | 52.8 | 45.5-60.6 | 20.8 | 41.8 |
| EME | 7.3 | 4.9-10.6 | 1.4 | 7.2 |
| FSEE | 11.6 | 8.9-14.8 | 18.1 | 9.5 |
| India | 73.1 | 63.5-83.7 | 14.9 | 62.2 |
| Latin America | 16.8 | 13.7-20.4 | 19.6 | 13.5 |
| MEC | 27.2 | 22.6-32.5 | 20.9 | 21.5 |
| Other Asia | 68.6 | 63.7-73.7 | 17.0 | 56.9 |
| SSA | 235.4 | 216.3-255.7 | 38.9 | 143.9 |

* The percent of cases that were not landed immigrants in our numerator. These correction factors were generated from a previous study looking at this phenomenon in the 1994-95 immigrant population to Ontario.

4.3.4 Socioeconomic Status

Number of years of schooling was used as a proxy for a descriptive evaluation of variation in socioeconomic status across the world regions. All landed immigrants to Ontario of age 15 years or older were included in this analysis. The region with the highest mean years of schooling was the Former Socialist Europe group with a mean of 13.09 (95% C.I. : 13.07-13.11). The group with the lowest number of years of schooling was Latin America and the Caribbean with a mean of 10.87 years (95% C.I.: 10.85-10.90). Census Canada reports the average amount of schooling for Canadians 15 years of age or older, is

between 9 and 13 years (1996 Census). Thus, the difference between years of education in the Canadian population and the immigration population mean was small.

Due to the restrictions of no socioeconomic information available at the individual level, only ecologic descriptions could be used. The Pearson correlation coefficient for the mean number of years of education and the age and sex-standardized rate of TB for each world region was not significantly different from zero (r=0.89, p=0.057). Therefore, no significant correlation was found between SES and TB rate in this population using this ecologic data.

**Table 18. Mean Number of Years of Schooling among Landed Immigrants to Ontario between 1990 and 1997 by WRO with 95% Confidence Intervals and rate of TB in corresponding study population.**

| *World Region of Origin* | *Total Population N* | *Mean (95% Confidence Interval)* | *Age & Sex Standardized Rate (per 100,000 person-years)* |
|---|---|---|---|
| Other Asia | 247725 | 11.9 (11.87-11.90) | 71.5 |
| SSA | 55278 | 12.1 (12.12-12.19) | 210.1 |
| Latin America | 121810 | 10.9 (10.85-10.90) | 17.5 |
| MEC | 102881 | 12.7 (12.67-12.73) | 35.5 |
| FSEE | 126277 | 13.1 (13.07-13.11) | 16.4 |
| EME | 78271 | 11.6 (11.55-11.62) | 7.7 |
| India | 82881 | 11.5 (11.44-11.52) | 63.7 |
| China & Taiwan | 102671 | 11.7 (11.67-11.73) | 45.7 |
| Total | 917794 | 11.9 (11.93-11.95) | 50.8 |

### 4.4 Survival Analysis of Objective 2

#### 4.4.1 Life Table and Survival Distribution

Any subject that had a negative "survival" value due to developing active disease before they were given legal landed status, was deleted from the database. A perfunctory look at the number and distribution of cases that were deleted was conducted to ensure that there was no selection bias due to deleting people from one strata more than from another (See table 19). Although there was *some* difference in the

proportions of people who were deleted due to temporal inconsistency in their landing and diagnosis dates, the numbers were so low as to not suggest any significant bias. There was essentially no difference between males and females, with 12 and 13 people deleted respectively. Overall, only 25, or 1.2% of the total cases in this study were deleted due to not conforming with inclusion criteria.

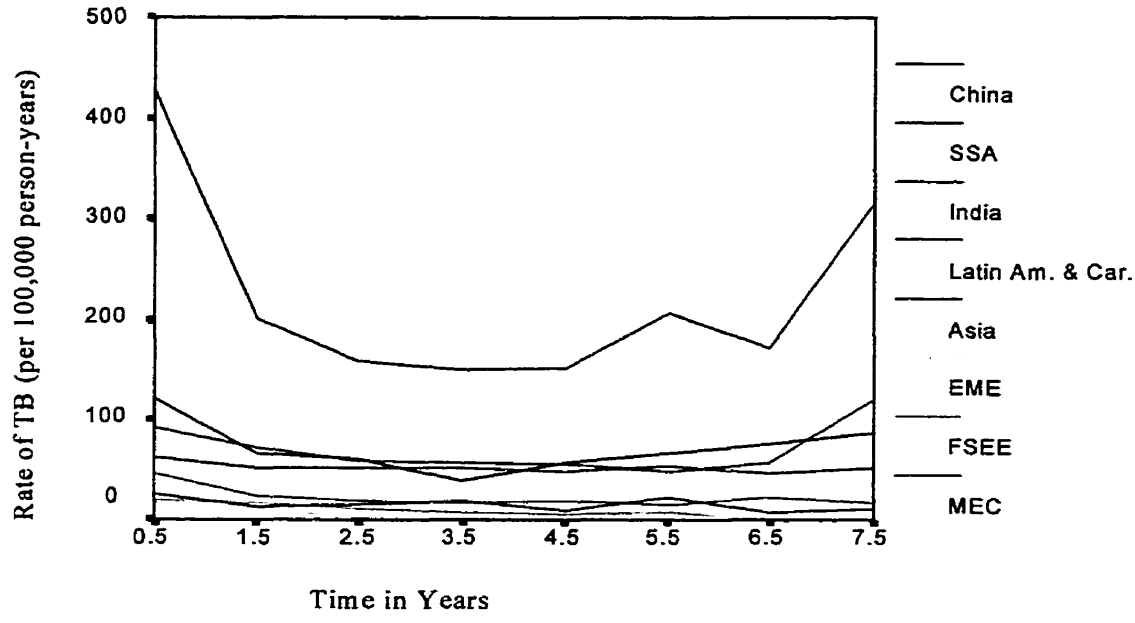
**Table 19. Summary of Characteristics of TB cases with date of diagnosis and episode date *before* date of arrival. Female/male. (%) of Total Cases for that strata that were deleted.**

| *WRO* / *Age Gr.* | *China* | *EME* | *FSEE* | *India* | *LatCar* | *MEC* | *Asia (Other)* | *SSA* | *Totals (Row %)* |
|---|---|---|---|---|---|---|---|---|---|
| *0-15* | - | - | - | - | - | - | *1 / 0* | - | *1 (0.5)* |
| *16-30* | - | - | *0 / 1* | *1 / 0* | *0/1* | - | *2 / 1* | *2 / 1* | *9 (0.95)* |
| *31-45* | *0 / 1* | - | *1 / 0* | *0 / 1* | - | *1 / 1* | *2 / 1* | *0 / 1* | *9 (2.1)* |
| *46-65* | - | - | - | - | - | - | *1 / 0* | - | *1 (0.3)* |
| *66+* | - | *0 / 1* | - | *1 / 0* | - | *1/0* | *0 / 2* | - | *5 (3.0)* |
| Totals (Column %) | 1 (0.5) | 1 (3.4) | 2 (3.0) | 3 (1.4) | 1 (1.0) | 3 (2.4) | 10 (1.3) | 4 (0.7) | 25 (1.2) |

4.4.2 Hazard Estimates and Graphs

The Hazard rates by world region, sex and age are plotted below (see Figures 3 and 4a-o). Survival estimates and log rank statistics are presented in Appendix A. The risk of TB is highest for SSA followed by India, Asia and to a lesser extent, China. It is evident that the initial high risk of disease decreases over the first few years after arrival, however, in some cases tends to increase around year seven. This trend is observed primarily with the population from Sub-Saharan Africa and "Other Asia".

**Figure 3. Hazard Estimates over Time for all the WRO. Hazards are measured as a rate per 100,000.**

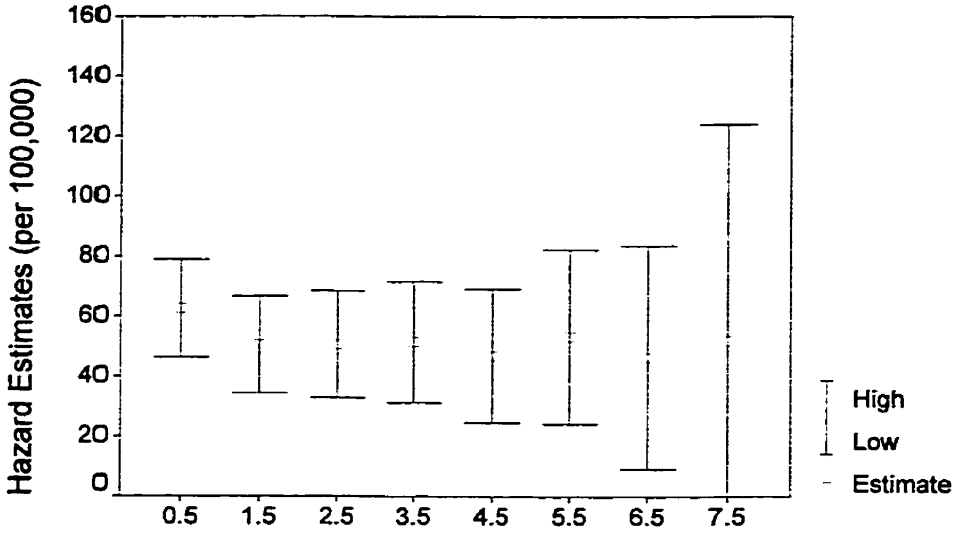



This increase after 6 or 7 years could be a real phenomenon or there could be some low power effect due to a decreased number of people available for observation over the longest survival periods. In order to evaluate the contribution of power to this observed trend the individual world region, sex and age hazard estimates along with 95% confidence intervals are plotted below (See Figures 4a-o). When looking at the hazard estimates of the final year of observation (year 7-8), along with 95% confidence intervals, this trend almost disappears for most of the strata. The suggestion of the trend remains strongest for Other Asia and SSA, the 16 to 30 age group and male sex which all show what could be a "real" increase.

The primary trend that was predicted is the initial decrease in risk soon after arrival. This pattern is seen most strongly in the SSA and India population. Other Asia, China, EME, FSEE, Latin America & Caribbean show a more constant hazard rate across time, although they all show some decrease in the first 2 to 3 years with the exception of China. When the population was broken down by age groups, the pattern of a decrease soon after landing was still evident although seen most strongly in the oldest (66+) and youngest (0-15) age groups, with 16-30 showing the next most evident drop in rate at the beginning. The ages 31-65 showed less of a dramatic decrease of risk of disease over time.

Figure 4a. Incidence Rates for China and Taiwan

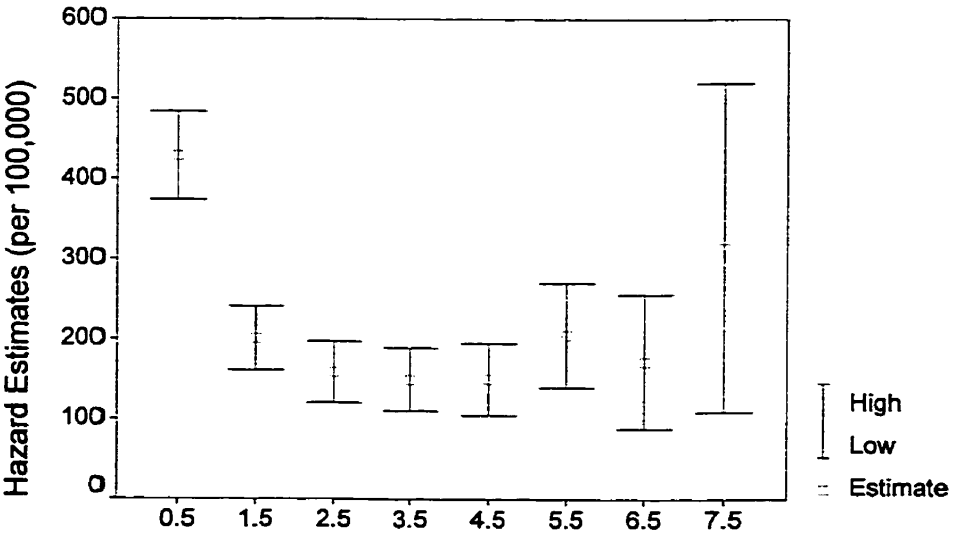
with 95% Confidence Intervals




Time in Years

Figure 4b. Incidence Rates for Sub-Saharan Africa

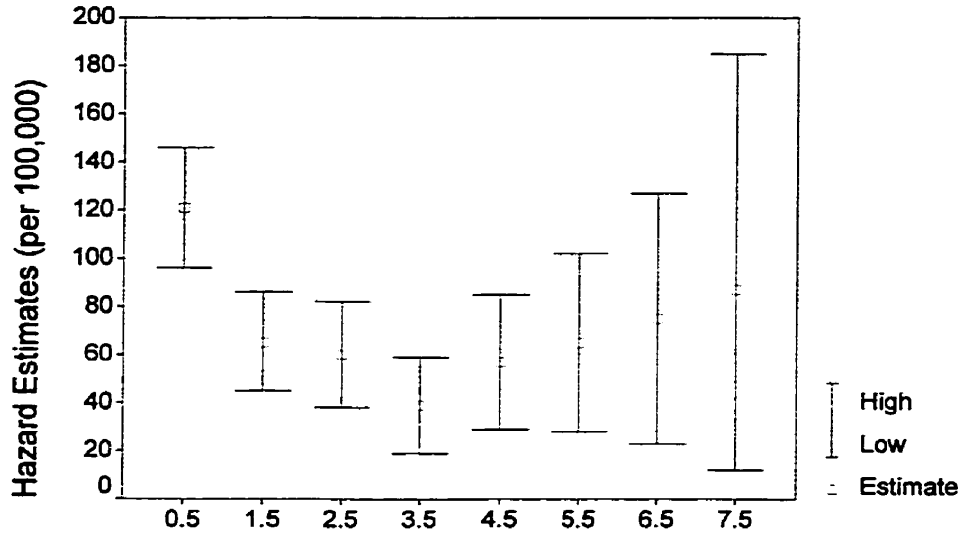
with 95% Confidence Intervals




Time in Years

Figure 4c. Incidence Rates for India

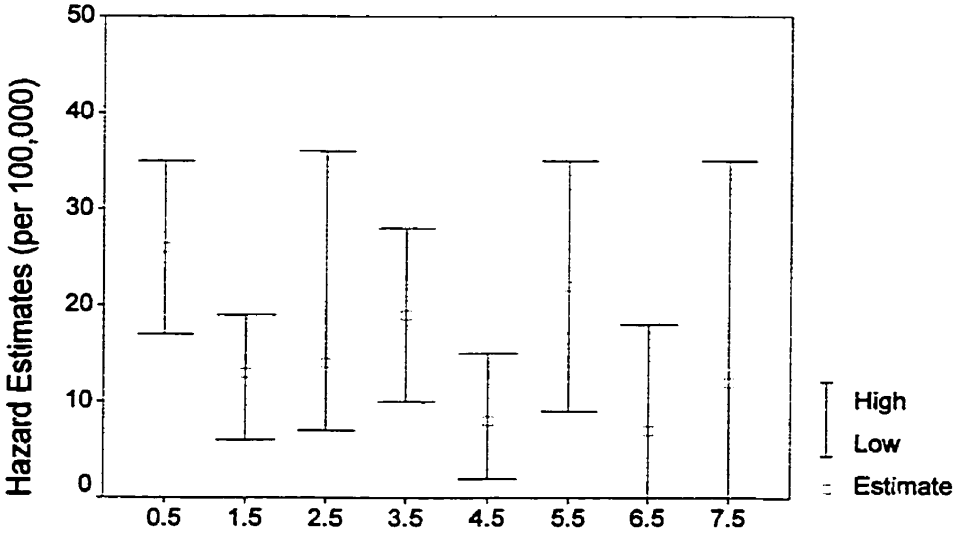
with 95% Confidence Intervals




Time in Years

Figure 4d. Incidence Rates for Latin America & Caribbe

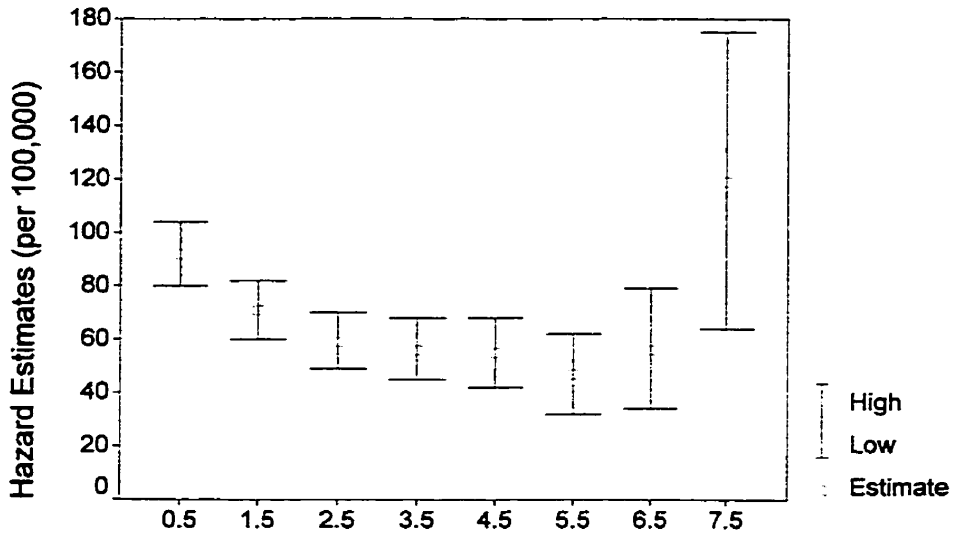
with 95% Confidence Intervals




Time in Years

Figure 4e. Incidence Rates for Other Asia and Islands

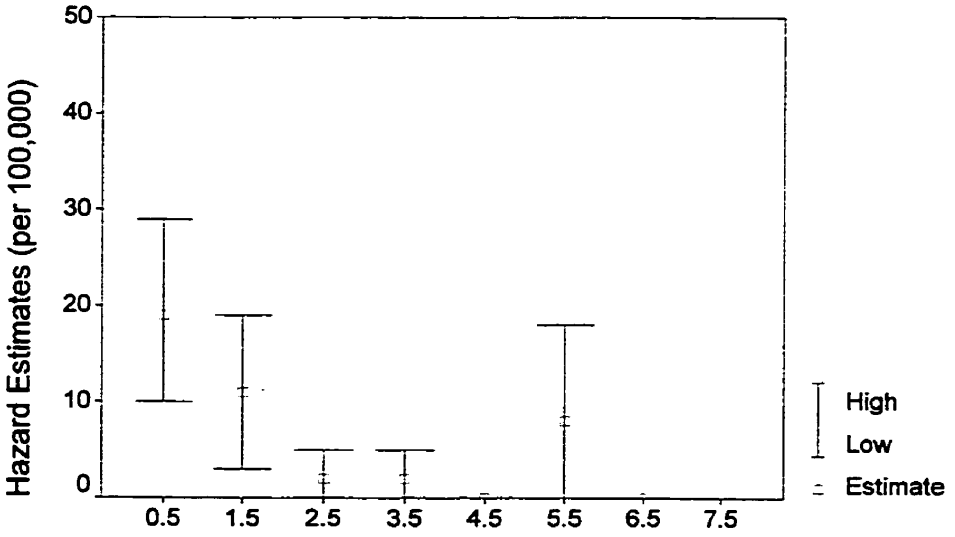
with 95% Confidence Intervals




Time in Years

Figure 4f. Incidence Rates for EME

with 95% Confidence Intervals




Time in Years

Figure 4g. Incidence Rates for FSEE

with 95% Confidence Intervals

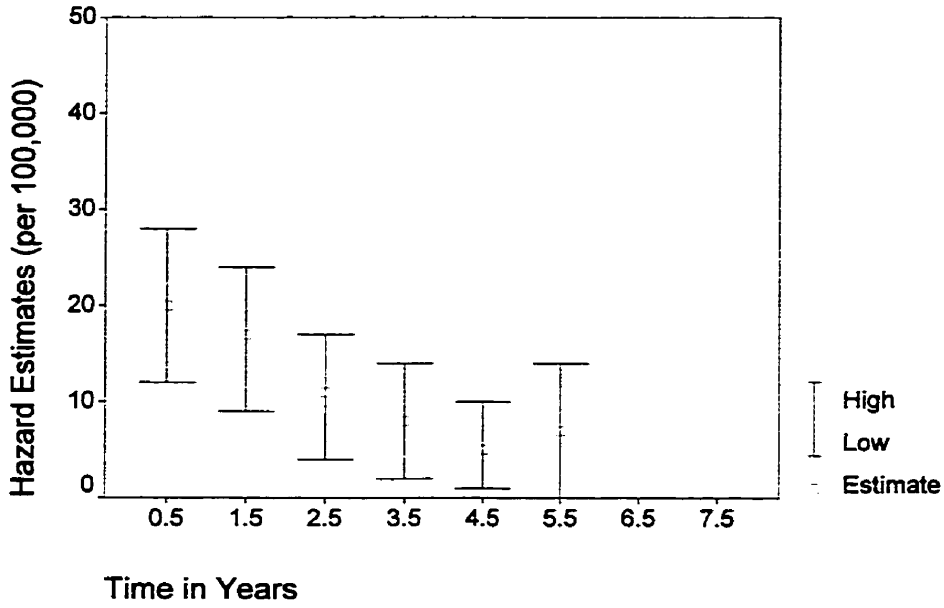



Figure 4h. Incidence Rates for MEC

with 95% Confidence Intervals

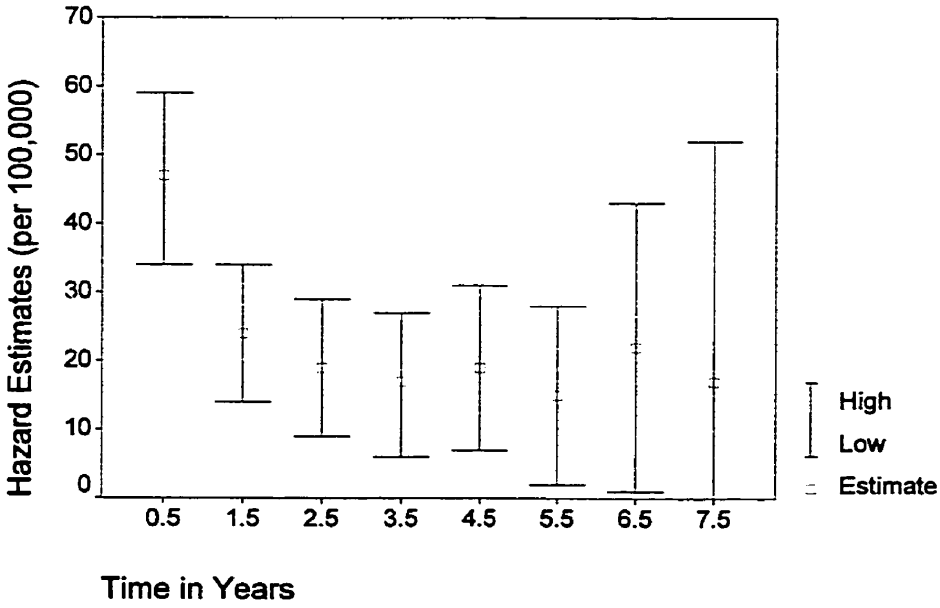
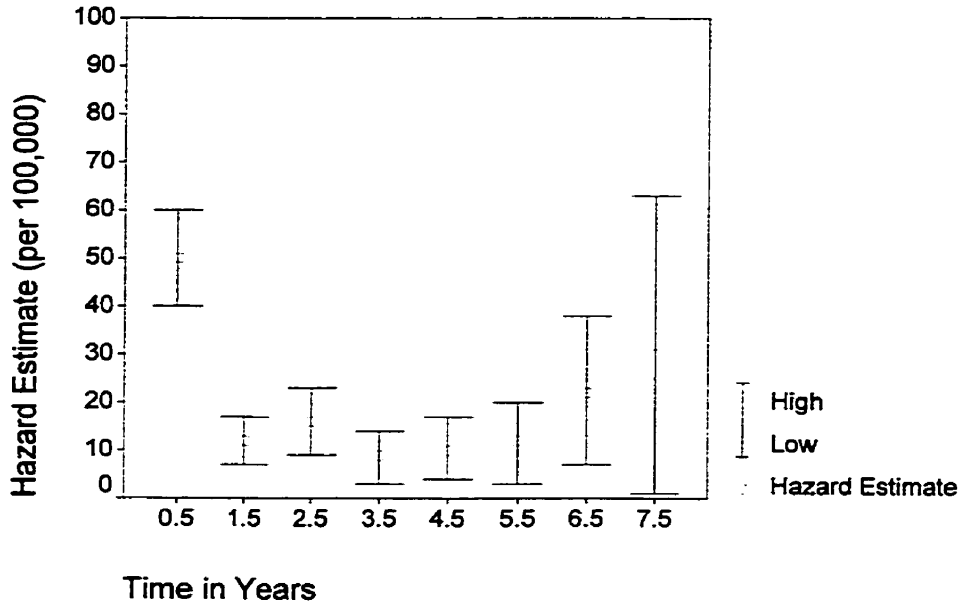


Figure 4i. Incidence Rates for 0-15 age group

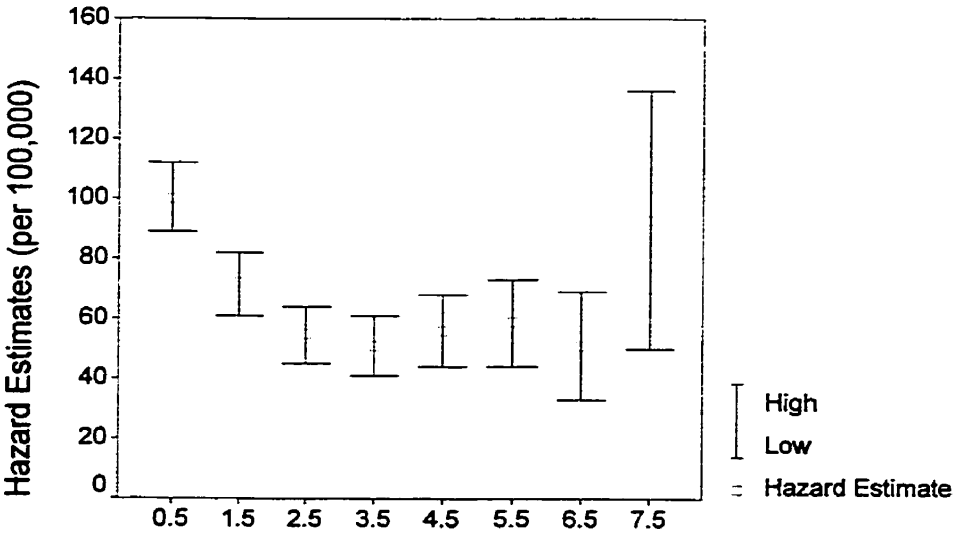
with 95% Confidence Intervals




Time in Years

Figure 4j. Incidence Rates for 16-30 age group

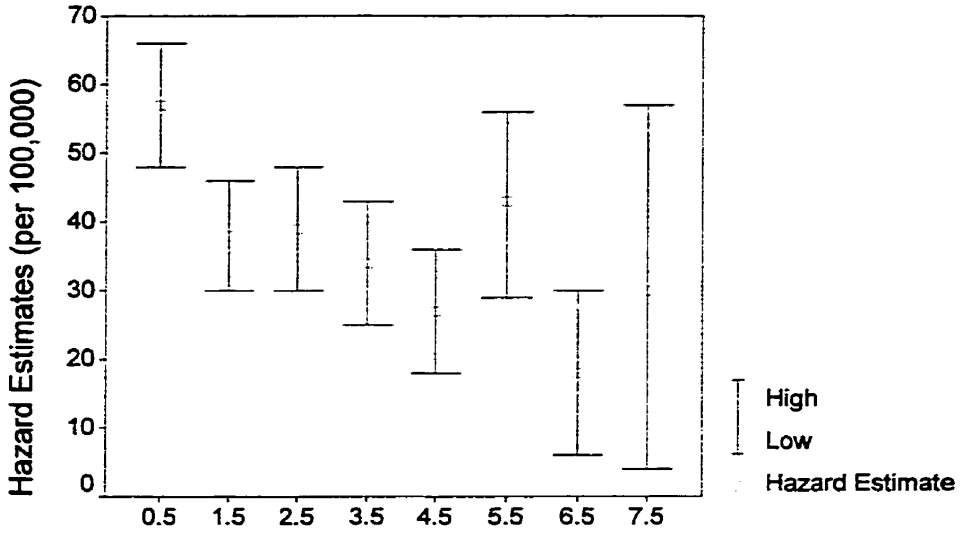
with 95% Confidence Intervals




Time in Years

Figure 4k. Incidence Rates for 31-45 age group

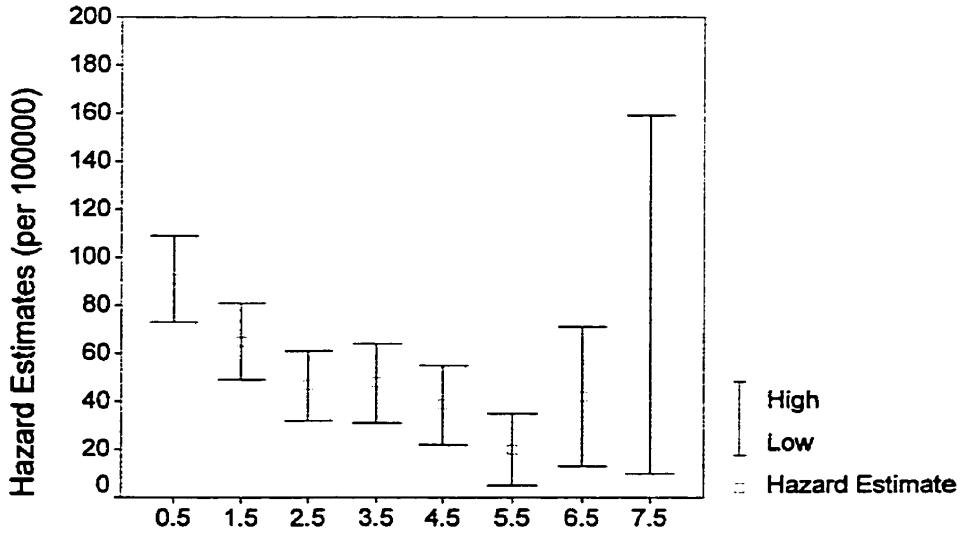
with 95% Confidence Intervals




Time in Years

Figure 4l. Incidence Rates for 46-65 age group

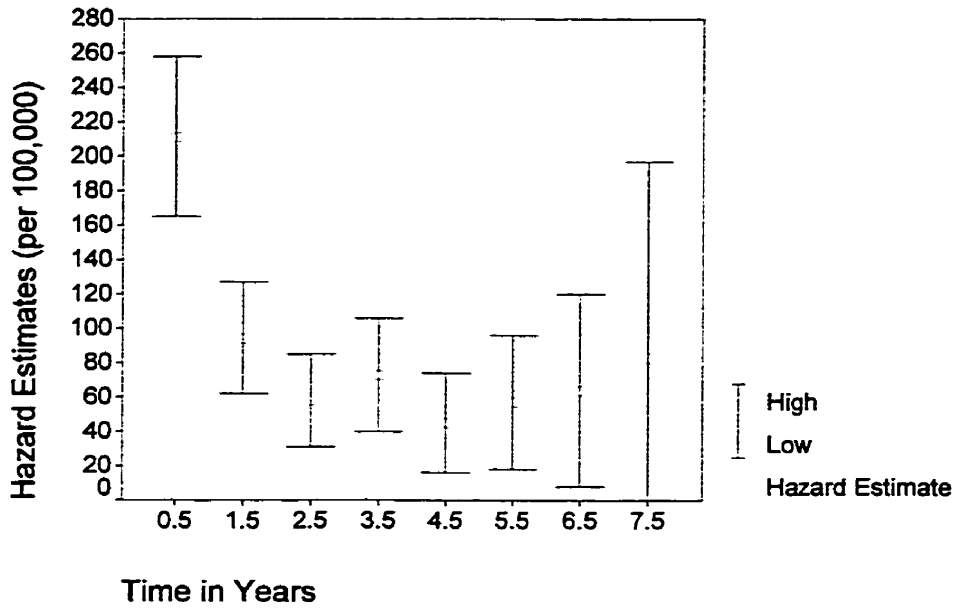
with 95% Confidence Intervals




Time in Years

Figure 4m. Incidence Rates for 66+ age group

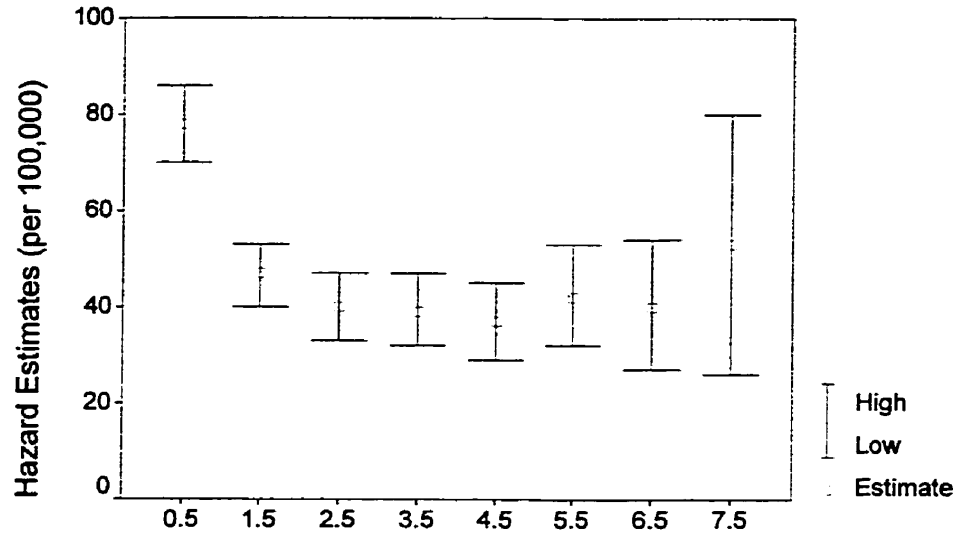
with 95% Confidence Intervals




Time in Years

Figure 4n. Incidence Rates for Females

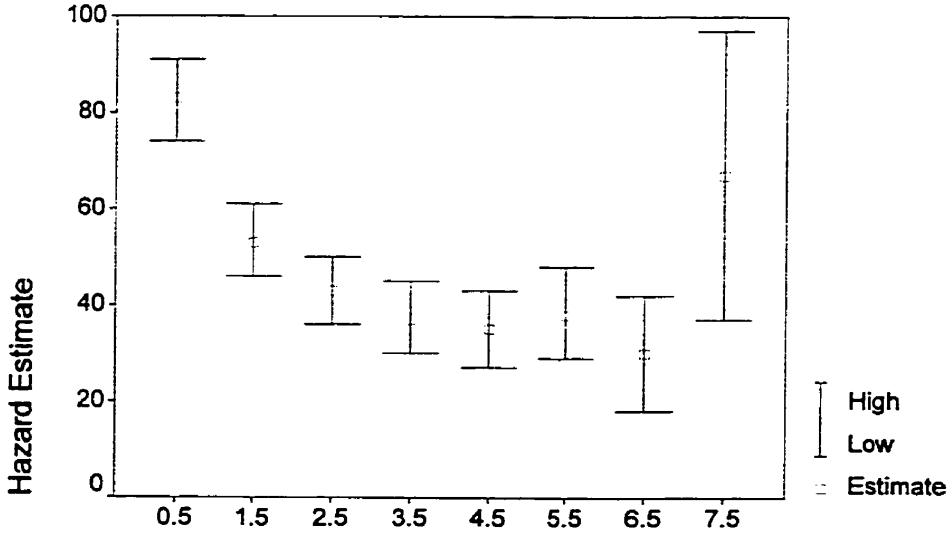
with 95% Confidence Intervals




Time in Years

Figure 4o. Incidence Rates for Males

with 95% Confidence Intervals




Time in Years

4.4.3 Log-rank Tests

The log-rank test to evaluate the homogeneity of the survival functions across strata indicate that there is a significant effect ($p<0.0001$) of world region and age on survival function over time since arrival (See Table 20). Sex did not produce a significant effect ($p=0.406$) on the survival or hazard functions and the curves were not significantly different. The greatest risk of disease was in the oldest age group (66+ years) with the next highest risk being in the 15-30 years of age group.

**Table 20. Log-rank test results for effect of study variables on the Hazard distributions.**

| *Variable* | *Log-rank Chi-Square Value* | *Df* | *P-Value* |
|---|---|---|---|
| Sex | 0.69 | 1 | 0.406 |
| Age | 348.3 | 4 | 0.0001 |
| WRO | 2136.8 | 7 | 0.0001 |

4.4.4 Cox Proportional Hazard Model and Complementary log-log model

Two models were constructed with the data. First, a Cox proportional hazard model looking at sex, age at landing and world region of origin was performed. Second, a complementary log-log model was performed including calendar year of arrival and duration of stay in Canada to take into account time-dependent variables and to more efficiently incorporate large numbers of ties in the data. Baseline values for the Cox model were designated as Established market economies (EME), 0-15 age group, and male sex. EME and 0-15 age group were chosen on the basis of the results from objective 1. For the complementary log-log model, baseline was automatically generated as 1997 arrival year, SSA and 66+ age group.

The proportional hazards model (no time-dependent variables), without interaction terms revealed that WRO and age were significant ($p \leq 0.05$) in predicting TB after arrival. All of the world regions and all the age groups were associated with a significant risk for TB relative to baseline (EME and 0-15 age group). Sex did not significantly improve the model at this, or even a less conservative level of significance ($p \leq 0.1$). Age was a confounding variable on the effect of world region of origin on survival time. The –2Log(L) score and the Wald statistic for the model with just WRO and the model with both

WRO and age were significantly different, using a $X^2$ test ($p \leq 0.0001$). There was no other confounding relationship in this model.

When a complementary log-log model was run with year of arrival and years spent in Canada, both variables significantly contributed to the fit of the model. The highest risk for TB was associated with 1990 and the lowest risk with 1997. The highest risk was associated with living in Canada for 1 year or less, and the lowest risk was associated with more than 7 years of residence (see Table 21). The risk ratios for age, wro and sex were not greatly changed by including calendar year of landing into the model. The complementary log-log and Cox models had nearly identical results for the model when the variable duration of stay was not included. The results of an initial model including all the variables but no interaction terms is presented below (Table 21).

**Table 21. Model of TB risk over time since arrival to Ontario without any interaction terms.**

| *Variable* | | *Parameter Estimate* | *Standard Error* | *Risk Ratio* | *P-value* |
|---|---|---|---|---|---|
| Sex | M | 0.043 | 0.045 | 1.04 | 0.334 |
| | F | 0 (baseline) | - | - | - |
| Age | 0-15 | 0 (baseline) | - | - | - |
| | 16-30 | 1.04 | 0.082 | 2.83 | 0.0001 |
| | 31-45 | 0.54 | 0.089 | 1.71 | 0.0001 |
| | 46-65 | 0.96 | 0.097 | 2.62 | 0.0001 |
| | 66+ | 1.63 | 0.109 | 5.09 | 0.0001 |
| WRO | EME | 0 (baseline) | - | - | - |
| | China | 1.85 | 0.203 | 6.36 | 0.0001 |
| | FSEE | 0.53 | 0.227 | 1.69 | 0.0200 |
| | India | 2.19 | 0.202 | 8.93 | 0.0001 |
| | LatCar | 0.86 | 0.214 | 2.37 | 0.0001 |
| | MEC | 1.32 | 0.210 | 3.73 | 0.0001 |
| | Asia | 2.26 | 0.193 | 9.58 | 0.0001 |
| | SSA | 3.51 | 0.194 | 33.5 | 0.0001 |
| Arrival Year | 1990 | 0 (baseline) | - | - | - |
| | 1991 | -0.26 | 0.071 | 0.77 | 0.0002 |
| | 1992 | -0.61 | 0.075 | 0.54 | 0.0001 |
| | 1993 | -0.49 | 0.078 | 0.61 | 0.0001 |
| | 1994 | -0.36 | 0.082 | 0.70 | 0.0001 |
| | 1995 | -0.59 | 0.096 | 0.56 | 0.0001 |
| | 1996 | -0.64 | 0.110 | 0.53 | 0.0001 |
| | 1997 | -0.88 | 0.150 | 0.41 | 0.0001 |
| Years in Canada | $\leq 1$ | 1.62 | 0.213 | 5.08 | 0.0001 |
| | 1.1-2 | 1.09 | 0.215 | 2.96 | 0.0001 |
| | 2.1-3 | 0.85 | 0.217 | 2.35 | 0.0001 |

| | | | | | |
|---|---|---|---|---|---|
| | 3.1-4 | 0.72 | 0.220 | 2.06 | 0.0010 |
| | 4.1-5 | 0.62 | 0.223 | 1.87 | 0.0050 |
| | 5.1-6 | 0.64 | 0.227 | 1.89 | 0.0050 |
| | 6.1-7 | 0.30 | 0.243 | 1.36 | 0.2085 |
| | 7+ | 0 (baseline) | - | - | - |

When all the interaction terms were introduced into the proportional hazards (time-independent) model, FSEE was no longer significant at a p=0.05 level. Since an interaction term including FSEE was significant, FSEE was forced into the model. SSA was a significant effect modifier for all the different age groups. This model including interaction terms is presented below in Table 22. Although sex did not significantly improve the model, nor was it a confounder, its role as an effect modifier was investigated. Forcing sex into the model resulted in 2 more significant ($p \leq 0.05$) interaction terms – sex and over 66 age group, and sex and China ($p = 0.045$ and $p = 0.009$ respectively).

If a WRO-Age interaction term was significant, the Risk Ratio (RR) was calculated including the interaction value, in the form of :

$$RR\ (X1, X2 \mid x_1=0, x_2=0) = e^{b1X1 - b2X2 - b3X1*X2}$$

If the combination was not included in the model as a significant interaction term, then a multiplicative model was assumed of the form :

$$RR\ (X1, X2 \mid x_1=0, x_2=0) = e^{b1X1 + b2X2}.$$

**Table 22. Risk Ratios† of Age and WRO Covariates with 95% Confidence Intervals and Interaction Terms where significant ($p \leq 0.05$)**

| *AGE* | *EME* | *SSA* | *MEC* | *FSEE* | *OthAsia* | *LatCar* | *India* | *China* |
|---|---|---|---|---|---|---|---|---|
| 0-15 | 1.0 baseline | 57.2 (35.7-91.4) | 2.8 (1.8-4.3) | 1.3 (0.8-2.1) | 8.9 (6.1-13.0) | 3.1 (1.9-4.9) | 6.8 (4.5-10.4) | 6.8 (4.4-10.3) |
| 16-30 | 3.5 (2.8-4.4) | 124.4* (80.8-191.4) | 9.9 (7.7-12.6) | 4.6 (3.5-6.0) | 31.2 (24.9-39.0) | 6.3* (4.8-8.3) | 34.1* (27.0-43.1) | 23.8 (18.6-30.4) |
| 31-45 | 2.4 (1.9-3.0) | 49.5* (31.0-78.8) | 6.8 (5.3-8.7) | 3.2 (2.4-4.2) | 21.5 (17.2-26.9) | 7.5 (5.8-9.7) | 16.5 (12.9-21.1) | 16.4 (12.9-20.9) |
| 46-65 | 3.5 (2.7-4.5) | 66.4* (50.0-88.4) | 21.8* (16.4-28.9) | 4.6 (3.5-6.2) | 31.4 (25.0-39.6) | 4.8* (3.2-7.2) | 24.1 (18.9-30.8) | 24.0 (18.8-30.6) |
| 66+ | 7.8 (5.7-10.5) | 124.9* (86.4-180.7) | 42.1* (29.5-60.1) | 27.3* (18.6-39.9) | 69.2 (54.1-88.6) | 24.1 (18.2-32.0) | 53.1 (40.9-69.0) | 52.8 (41.0-68.0) |

† baseline values were 0-15 age group from EME
* Risk Ratios calculated including interaction terms that were significant to the model ($p \leq 0.05$)

Using the complementary log-log model, interaction between year of arrival and the other variables was investigated, as well as interaction between duration of stay in Canada and age at arrival. The other possible interaction terms could not be investigated due to small numbers and instability of the estimates when these interaction terms were introduced into the model. When all the terms were included, once again, sex was not significant. Sex was removed from the model leaving the final model as reported below (Table 23).

**Table 23. ANOVA table for log-rank statistics of variables in final model**

| *Variable* | *DF* | *Chi-Square* | *Chi-Square P-value* |
|---|---|---|---|
| Years in Canada† | 7 | 289.3 | 0.0001 |
| AGE | 4 | 85.5 | 0.0001 |
| YEAR | 7 | 26.5 | 0.0004 |
| WRO | 7 | 425.2 | 0.0001 |
| Years in Canada*AGE | 28 | 65.6 | 0.0001 |
| AGE*YEAR | 28 | 46.8 | 0.0144 |
| AGE*WRO | 28 | 99.1 | 0.0001 |
| YEAR*WRO | 49 | 92.1 | 0.0002 |

† refers to number of years spent in Canada (since immigration).

The risk ratios for the interaction terms of the final model including calendar year of arrival and duration of stay in Canada are presented in Appendix B (Tables I-III). Calendar trend seems strongest for China and SSA. Including time-dependent variables in the model decreased the power to identify significant associations, so emphasis should be put on patterns of risk in this data, not actual risk ratio values. The model investigating age and world region interactions using the Cox proportional hazard model was presented earlier and the results of this interaction in a complementary log-log model is not presented here (See Table 22).

*Final Model including Interaction Terms*

Disease was significantly predicted by WRO, age, calendar year of arrival and years since arrival. Confounding and interaction with WRO occurred at some levels of the variable age. Sex was not a significant predictor of disease, nor a confounder. Interaction terms involving age and WRO were predictive of the "hazard" of developing TB among immigrants over time in the Cox proportional hazard model. The complementary log-log model, revealed that: i) The risk associated with world regions and age were modified by year of landing. ii) The risk associated with age was affected by number of years since arrival with the greatest effect being in the oldest age group. iii) The initial drop in risk associated with arriving in Canada was seen most dramatically in the oldest age groups. The interactions of the variable "duration of time in Canada" could only be investigated with age at landing due to small numbers, and this interaction was significant.

To help visualize any age-related trends occurring within the risk for TB among immigrants from the different world regions, the data was also presented as a line graph below (see figure 5). It is clear that the over 66 years of age group is by far the highest risk group for all world regions except for SSA. People from SSA exhibited an equally high, age-dependent risk in the 16 to 30 age group. People from India also showed a high peak of risk in the 16 to 30 age group, although lower than that seen in the eldest group. When the age data for risk of TB was plotted by duration of stay in Canada, the over 65 age group is still the highest upon arrival followed by the 46-65 group (See Figure 6). By the second year after arrival, the over 65 age group and the 46 to 65 are similar. The over 65 group also shows the most dramatic decrease in risk in the first 2 years. The 16 to 30 age group experiences a lower risk than both the elder age groups

when duration of stay is included in the model, which differs from the hazard estimates generated by the life table analysis.

The log-rank statistics identified year as a significant contributor to the model so a visible effect would be expected. A time-dependent effect related to year of landing and world region is apparent, particularly in people from SSA and China (Figure7).

**Figure 5. Risk Ratios for TB since arrival Among Ontario Immigrants by World Region of Birth and Age at Landing. Model without time-dependent variables.**

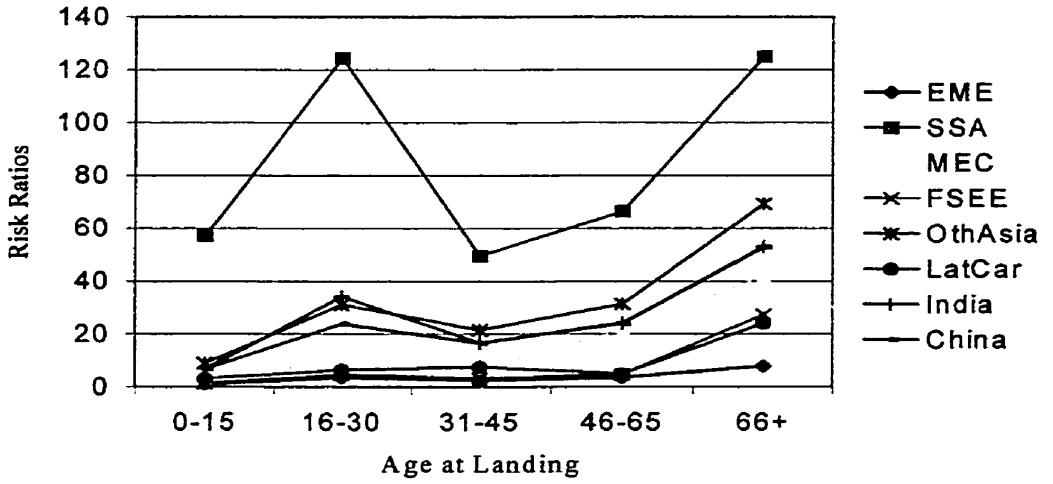



**Figure 6 Risk Ratios of Risk of TB by duration of stay in Canada and age at Landing. Over 65 age group and 8 years of stay in Canada are baseline.**

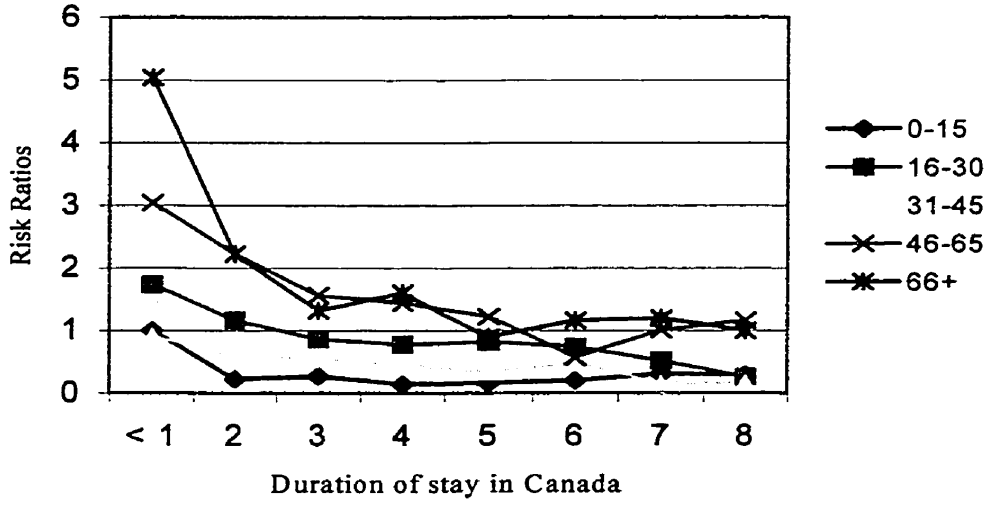
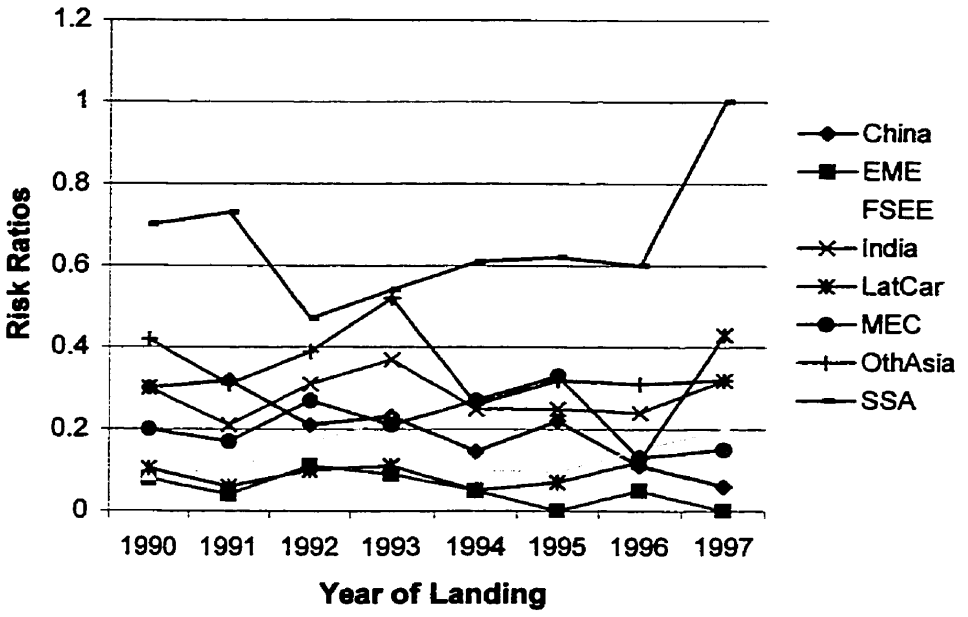


**Figure 7. Risk Ratios of World Region of Origin by Year of Landing. SSA and 1997 Landing date were used as baseline.**




***Testing the Model Assumptions***

The previously mentioned log-log plot of the survival probabilities for the different covariates revealed that the proportional hazard assumption of time-independent covariates was met (see Appendix C). The graphs show that the different values for the covariates are relatively parallel, they do not cross at any point. Only world region of origin shows some possible non-parallel patterns over time with established market economies and former socialist economies. At the longest survival time group there is also some convergence between the male and female function. Both these examples are probably an instability of the estimates due to a small numbers phenomenon. Established market economies, former socialist economies and the group "followed up" for the longest period are all small populations. With this in mind, the conclusion that the proportional hazard assumption was met, was accepted. For the time-dependent variables, the Complementary log-log model is appropriate. The number of strata that resulted from including time-dependent variables in the model, however, decreased the ability to interpret the findings as conclusive.

## 5.0 DISCUSSION

### 5.1 Study Rationale

Tuberculosis causes more adult deaths than any other infectious disease worldwide [1,2]. The World Health Organization estimated in 1993 that one-third of the world's population is infected with TB [3,4]. Developed countries such as Canada are in no way insulated from the effects of TB. Immigration is a major contributor to the Canadian TB incidence rate [9,12,13,4]. The decline of TB incidence has halted and was beginning to increase until a plateau of about 7 per 100,000 was reached in 1987 [13,2,9]. The proportion of foreign-born cases has been steadily increasing [15,13,17,2]. Failure to identify active cases results in transmission from undiagnosed active cases. In order to move towards elimination of TB and continue the decline in rates of TB seen in previous decades, a public health effort is necessary not only to identify and treat active cases but also to reduce the future incidence of disease in high risk populations by identifying and treating latent infection.

There are several aspects of tuberculosis that make it a disease that is hard to control and even harder to eradicate. One aspect is that the disease can be inactive, or latent, and it can be hard to diagnose in the general public until it reactivates and the person becomes symptomatic and possibly infectious. As well, the incubation period for TB is highly variable and thus, not predictable with any accuracy [20]. These and other factors make testing for, and treatment of, latent infection with TB an attractive strategy. If the disease can be identified in a pre-active, latent stage, and if a high-risk group can be identified screening may be cost-effective [10]. Since active TB is a fatal and contagious disease when left untreated, and its treatment is associated with a possible serious health risk, a screening tool with both high sensitivity and specificity is necessary [10].

Eighty percent of TB in Ontario is among the foreign-born [14,9]. It is often reported that foreign-born people have a higher proportion of reactivation disease than their native-born counterparts, as a result of greater numbers of remote infections acquired before immigration [24,47,25,18]. In one study, immigrants with inactive pulmonary disease on chest Xray had 13 times the incidence rate of active TB as immigrants with normal chest Xrays [47]. Many cases of reactivated disease could be prevented through targeted testing and treatment of latent infection [18]. However, targeting high risk communities and ensuring that there is

high uptake of preventive, risk reducing interventions, are key to developing a cost-effective strategy [18]. Identifying and treating high risk groups is a complicated and controversial topic and generally targeted testing is discouraged unless the proper infrastructure is available. Passive case detection, involving treatment of people who appear with symptoms, will not alter the natural history or reduce the risk posed by people who are remotely infected but without active disease, and who are at a high risk to develop active disease at any time in their life [84,16].

In order to control this disease and move toward elimination of TB in a country such as Canada, an effective intervention strategy would need to be implemented within high-risk groups. First, however, knowing more about the epidemiology in this population is necessary including the pattern of disease over time. In 1998 the CDC published certain recommendations with respect to control of TB among the foreign-born [16]. One of their key recommendations was to develop epidemiologic profiles to identify groups of foreign-born persons in jurisdictions who are at high risk for TB in order to tailor TB-control efforts to local needs [16]. This study attempts to look at the particular aspects of TB epidemiology among recent immigrants to Ontario.

Previously, only one study (Wobeser, 1998) focussed on risk of TB in immigrants to Ontario and this study was case-control in design [14]. Before that, the last in depth study focussing on immigrants to Ontario (although it was descriptive in design) was published in 1979 by Enarson et al. [13]. The current study is the first to look at risk for immigrants over time since arrival in Ontario in the past 25 years. This study is also the first cohort design study looking at such a large population in Ontario to investigate TB. Ontario-specific information on TB in immigrants would be useful for developing profiles for community-based programs. The Canadian Expert Committee on Tuberculosis (ECOT) identifies those arriving from a country where tuberculosis is endemic as possible candidates for screening using the tuberculin skin test [114]. The Advisory Council for the Elimination of Tuberculosis (ACET) 1995 recommendations emphasized the need for community-based targeted testing for latent TB infection among the same groups [18]. This current study could provide the lacking Ontario-specific information for this purpose.

### 5.2 Key Findings

The risk of TB among all recently arrived immigrants to Ontario was thirty-two times higher than the Canadian-born, non-aboriginal risk. Risk varied by World Region of Origin and age at landing, and although there were some sex-dependent trends overall, the variable sex was not predictive of TB. Persons from Sub-Saharan Africa and people over the age of 65 were the highest risk groups. Sub-Saharan Africa had a risk ratio of over 100 times the Canadian-born, non-aboriginal population. The risk of TB by age showed a bimodal pattern with an initial high risk in the 16 to 30 age group followed by another peak of risk over the age of 65. A significant decrease in risk of TB occurs over the first 1-2 years after arrival to Ontario. The risk for most world regions, however, still remains elevated even seven years after immigration. There is some suggestion of an increase in risk after 7 or 8 years in certain groups. The relationship between risk of TB and world region of birth was affected by age. There was an effect on risk by calendar year of arrival which may affect the results of the Cox proportional hazard model for predicting risk over time.

### 5.3 Characteristics of TB Cases Not Included in the Study

As predicted for a country with a current low risk of TB, the Canadian-born population showed a greater proportion of cases among the elderly, although it is surprising that the proportion was not greater. Only 31% of the Canadian-born cases were 66 and over. At first glance the explanation would seem that the relatively large number of younger people with TB probably came from the "registered Indian" or "Inuit" population which has a much higher burden of TB than the non-Aboriginal population. This is not the case. The greatest number of cases among the younger age groups are from among the "Born in Canada" population, excluding Aboriginals. Of the 133 cases in the 16-30 age group, only 26 are Aboriginals, the rest are other Canadian-born, unspecified or unknown people, or those with the origin field blank. This draws our attention to two other potentially high risk groups, the homeless population and persons infected with HIV. Likely, a large contributor to TB in younger, Canadian-born people is homelessness or HIV. The other notable characteristic is that 61% of TB cases among the Canadian-born were male, compared to 50% in the foreign-born study population.

The excluded foreign-born population also included more males. The highest proportion of cases were in the 66+ age group, or between 31 and 45. By far, the largest proportion were excluded due to arrival before 1990 (84.5%). If we only look at the population excluded due to missing information, the majority were excluded for not having a date of arrival recorded. It is possible that a significant number of people who are included in this group are refugees who have not officially "landed" or people who have been in the country for a very long time. It may be that some of the population of refugees and "non-landed" persons that are of concern for our numerator estimates, may in fact not be in our numerator due to not having a "date of arrival".

The primary differences between the foreign-born cases included in this study and the Canadian-born cases is that Canadian-born cases were older (largest proportion over 66) and more likely to be male (61%). Foreign-born cases that were *not* included in the study due to incomplete information tended to be somewhat older than the study cases, more likely to be male, and to come from "Other Asia". Of those not included in the study due to date of arrival being before 1990, 63% arrived between 1980 and 1989. In total, 3005 foreign-born cases were excluded.

### 5.4 Possible Causes of Instability In the Risk Estimates

The denominator data used in this study includes information on landed immigrants and convention refugees only, not visitors, students, refugee claimants or people on work visas. In contrast, the numerator includes all refugees (both convention and refugee claimants), and students, and any foreign-born person with a place of residence in Ontario. It is possible that this latter group may be an important risk group. A previous study found that between 1970 and 1974 31% of non-Canadian born people who developed active TB were not legally landed immigrants [26]. Recently, in Ontario, it was found that 24% of recently arrived TB cases were not legally landed [36]. Students, however, likely do not contribute a large number of cases since a 1997 study by Health Canada evaluating the screening program of foreign-born students did not find a single active TB case among those screened [85]. Socioeconomic status and age may be an important element in the low rate of TB among foreign-born students.

Since the non-landed immigrants described above *will* appear in the provincial TB registry but *will not* appear in the immigration data, an adjustment for this discrepancy was attempted in order to not

overestimate the incidence of active TB in the legally landed population This adjustment, however, was only possible on the rates generated for objective one and was not applied to the survival data. Thus, the results of objective 2 are likely inflated for some of the world regions that might have had a higher representation among the cases of refugees and/or students. For instance, a large number of refugees from Sub-Saharan Africa arrived over the study period to Canada. As well, China tends to provide large numbers of students to North America [75]. These trends may bias the results by preferentially overestimating the incidence among certain immigrant groups. Although information regarding legal immigration status of the cases was not available, a previous study by Enarson et al. in 1979 [13] stressed that the majority of cases among non-Canadian-born persons were in legally landed immigrants who were screened at entry and found to be disease free at the time.

Another possible discrepancy in the denominator data is that some of the people who say they are coming to live in Ontario at the time of immigration may not stay in Ontario. Likewise, the denominator data will not include people who said they were going elsewhere and who then took up residence in Ontario. This is unlikely to cause any kind of bias unless the people who are not in our denominator are in some way different from those who are with respect to risk factors for TB. In the same way, people who say they are staying in Ontario and then leave appear to be a relatively unusual phenomenon [86], nor is there any evidence that they are different in some risk factor for TB.

### 5.5 Summary of Results

The risk of TB among recently arrived immigrants to Ontario was significantly higher than the Canadian-born, non-aboriginal risk. Risk varied by World Region of Origin and age at landing, but was not significantly affected by gender. Among the world regions, only China showed a significant difference in risk according to sex with men being at a consistently higher risk than women at all ages. Men over the age of 65 had a significantly higher risk than women for all regions except for FSEE, MEC and Latin America (& the Caribbean) where women had a higher risk. The risk of TB by age showed a bimodal pattern among all the world regions except EME and FSEE, with an initial peak in the 16-30 age group followed by a second peak in the over 65 age group. Only women from SSA showed a different bimodal

pattern with the second peak in the 46-65 age group. Sub-Saharan Africa, most parts of Asia and India were responsible for a large proportion of active TB cases seen in Ontario between 1990 and 1998.

The immigrating population had an overall relative risk of 32 times the risk of the Canadian-born, non-aboriginal population. Sub-Saharan Africa, by far, had the highest relative risk of 132, followed by Other Asia and Islands, India and China with relative risks of 45, 40 and 29, respectively. The lowest risk was associated with the Established Market Economies which had a relative risk of 5 compared to the Canadian-born, non-Aboriginal population. When the rates for the world regions were adjusted for over-estimation due to non-landed immigrants in the numerator, the rates per 100,000 population were deflated for SSA, Other Asia, India and China. The rate for SSA showed the greatest decreased from 235 to 144 per 100,000. For Other Asia and Islands the rate decreased from 69 to 57 per 100,000. Established Market Economies had an adjusted rate of 7 cases per 100,000, which is roughly equivalent to the overall Canadian rate.

Survival analysis revealed that a significant drop in TB risk occurs over the first few years after arrival to Canada. This pattern was shown most strongly in the highest risk groups, that is, for SSA, India, and the over 65 age group. The 0-15 age group showed a strong initial decrease even though they do not experience a high risk overall. This likely is related to the fact that disease in children is generally due to recent infection and leaving an endemic country and coming to Canada is associated with a substantial decrease in annual risk of infection. There is some suggestion of an increase in risk after 6 or 7 years in males, people from Asia, SSA, and less strongly, in 16 to 30 year olds. Log-rank tests determined that the risk for TB over time since immigration differed significantly by world region and age.

The Cox Proportional Hazard and the complementary log-log models revealed that TB was significantly predicted over time since arrival by world region, age and year of arrival. Sex was not a significant predictor, except over the age of 65 when men had a higher risk, nor did sex confound any relationship between world region and disease, or age and disease. The association between world region and disease over time varied according to age at landing for many world regions, the notable exception being Asia and China which both exhibited a relatively constant risk across all ages. As well, the association between age group and disease varied depending on world region of birth. Sub-Saharan Africa had the greatest variability of risk of TB since arrival by age group, with every age group over 15 resulting

in a significant interaction term. Even after including all the variables in a model, including interaction terms, most of the world regions with high rates of TB showed evidence of a bimodal effect relating disease to age. There was an increase in risk of TB in the 16-30 age group, followed by a decline in the 31-45 age group and then risk increasing again up to the oldest group.

The risks associated with the different world regions compared to the risk associated with the Established Market Economies varied considerably, but all had a risk ratio greater than 1. Sub-Saharan Africa displayed the greatest risk ratios, peaking in the 16 to 30 age group at 125 times the risk of the Established Market Economies in that same age group. People over the age of 65 all had elevated risk ratios ranging from 20 to 103 times the risk of the EME, from Latin America and Sub-Saharan Africa, respectively. Before including interaction in the model, Sub-Saharan Africa showed the highest peak in TB rates in the 16-30 age group. After allowing for the effects of age-world region interaction, and calculating risk ratios at the various ages, the 16-30 and over 65 age groups had almost identical risk ratios.

When calendar year of landing and number of years spent in Canada were included in a model, both were significant for predicting disease in the study group. Risk for TB was not equivalent for all years of arrival, with 1990 associated with the highest risk and 1997 associated with the lowest. Similarly shortest duration of stay in Canada was associated with the highest risk and longest duration with the lowest risk. This is similar to the results of the Cox proportional hazard model that revealed a drop in risk after the first year. It calls into question, however, the validity of the increase in risk seen in year 7-8 using the lifetable hazard estimates. Risk associated with year of arrival varied by WRO, with SSA having the highest risk for all years followed by Other Asia (except for in 1997 when Latin America was higher).

## 5.6 Interpretation of Results

### 5.6.1 World Region of Origin and TB Risk

The overall risk ratio for immigrants in this study was 32 times the Canadian-born, non-Aboriginal TB rate. This number is comparable to a previous study looking at immigrants in Alberta from 1990 to 1994 which found a risk ratio for TB that was 21 times the Alberta, Canadian-born, non-Aboriginal rate [48]. The slightly higher rate in this study may reflect a different immigrant population in Ontario, or it may reflect our inability to adjust for non-landed persons in our numerator. For all the world regions to show a

higher risk than the Established Market Economies was expected. Rates in parts of Europe and Scandinavia are similar to the rates in Canada. Denmark, Iceland, Italy, Malta, the Netherlands, Norway, Sweden, Switzerland and the UK, all have rates of 10 cases per 100,000 or less [87]. Sub-Saharan Africa, India and parts of Asia, on the other hand, have much higher rates than any industrialized countries [49]. In 1989, the overall U.S. TB rate was 9.5 per 100,000, although among recently arrived, foreign-born persons the estimated TB rate was 124 per 100,000 [88]. The incidence of tuberculosis in Sub-Saharan Africa is considered to be one of the highest in the world [89,6], and many be increasing as a result of HIV[49].

As found in the study by Enarson et al. twenty years ago [13], rates in people coming to Ontario from these regions appear to be lower than in their native countries, except for MEC and SSA (See Table 24). A 1998 study conducted in Montreal found similar results and was provided for comparison (Table 24). The caveat in using WHO TB estimates for the world regions is to understand the limitations of this data. One of the difficulties in gaining a clear picture of the global tuberculosis situation is the lack of good quality data from developing countries in many cases. Incidence rates are almost impossible to obtain, so average annual risk of infection, incidence of smear-positive pulmonary tuberculosis and case-fatality rates are used [10]. Both the World Health Organization and the World Bank provide global estimates of TB burden using results from these methods and statistical imputations.

The risk among SSA immigrants is substantially higher in this study than the reported rate in that region according to the WHO. It is comparable, however, to both the Montreal study which found a rate of 200 per 100,000 (see Table 24) as well as a study conducted in Sweden which found a rate of 229 per 100,000 among African immigrants in 1990 [57]. TB estimates from the SSA region may be particularly unreliable so these results may or may not be due to a maintained (or increased) high risk of TB. The most likely explanation is that this is a result of under-notification of TB in the African, and Middle East region and thus an underestimation of the "true" rate of disease [87]. The WHO estimates that throughout Africa, the percent case detection of TB cases is about 53% and 27% in the Eastern Mediterranean (which corresponds closely to our MEC) [87]. If the WHO world region estimate is adjusted by this number, the estimated (corrected) rate for SSA becomes 234 per 100,000, which is almost exactly the rate found in this study. If the corrected rate is substituted for the notification rate, the rate of TB among SSA immigrants in Ontario is the same as in their region of birth. If the WHO rate for MEC is corrected for under-notification,

the true MEC rate becomes 107. This would mean that compared to the true rate in the Middle Eastern Crescent, rates among immigrants from that region experience a significantly lower risk.

For the rest of the study population, a decrease in risk compared to their country of origin was seen. Some of this decrease in risk must be attributed to pre-immigration screening for active disease [13,14]. However, it is logical that by removing the high risk of infection, rates of disease should decrease due to reducing the risk of developing active primary disease from recent infection. By moving to a country with very low annual risks of infection (ARI), the risk of becoming newly infected with TB decreases. In fact, we do see in the survival analysis that the very high rate associated with coming from SSA decreases significantly in the first two years.

**Table 24. Rate of TB Incidence (notified cases) in World Regions and among immigrants from World Regions in the current study (per 100,000)**

| *World Region* | *Rate of TB (in WRO)** | *Rate of TB (in current study)* | *Rate of TB in 1998 Montreal Study†* |
|---|---|---|---|
| China | 85 | 53 | - |
| EME | 23 | 7 | 7 |
| FSEE | 63 | 12 | 24 |
| India | 138 | 73 | 111 |
| Latin Car. | 46 | 17 | 34 - 88 |
| MEC | 29*** | 27 | 16.5 |
| MEC(corrected) | 107 | - | - |
| Other Asia | 83** | 69 | 57 - 116 |
| SSA | 124 | 235 | 200 |
| SSA (corrected) | 234**** | - | - |

*Rates in World Regions calculated from the WHO tuberculosis surveillance and monitoring project report, "Global Tuberculosis Control", 1998.
**Includes all of West Pacific and SE Asia except for China, Taiwan, Japan, New Zealand and Australia.
***Corrected for the WHO estimate of 27% case detection.
****Corrected for the WHO estimate of 53% case detection.
† From 1998 study by Rivest et al. in Montreal looking at TB cases between 1992 and 1995 [12]

A "healthy immigrant" concept that people who immigrate may be healthier than their compatriots is not a new one. It has been observed that rates of disease are often lower in the people that emigrate compared to the national average rate in their country of birth [13,14]. These people are often a special group who are younger, healthier and economically advantaged [13,14]. If this is true, a lower rate

among people coming to Ontario than the rate experienced in the home country would be expected. The initial high risk followed by a decrease in the first couple of years could be explained by the importation of prevalent disease. This may have been a result of TB not being detected at the immigration medical exam (IME) or the person not having an IME. As was mentioned in the literature review, a recent study by Wobeser (1998) found that 13% of all TB cases arising from a landed immigrant population were people originally identified by the medical exam as "high risk" [14]. Another possible factor for increasing risk in the first years since migration may be the stress of immigrating and adapting to a new country [24,14]. It may be that after the initial shock of arrival and associated high risk of disease passes, these people tend to experience a lower risk than those in their home countries.

Another theory is that case-detection is lower among immigrants and we are simply not seeing all the active disease in this population. The health care system is not always accessible to people who do not speak the language, and who are unaware of the procedures and practices in this country, especially regarding the treatment of people with an infectious disease [16,18]. There is some evidence that immigrants' and refugees' use of health care is less than that of other Canadians [70]. A recent study, however, found no difference in Canada for rates of contact with GPs or specialists among immigrants [90]. In fact, this study found that contact with GPs was higher than in the Canadian-born population [90]. There could be a complex combination of factors clouding this issue.

The initial decrease after arrival seen in the current study has been documented [47,14,12]. It is generally believed that a person's risk of developing TB following migration is usually highest in the first years and then starts to decline but still reflects the risk in their country of origin [36,13,15,47,66,7,11,42]. It has also been observed that the sharpest decrease in risk is seen in immigrants from countries experiencing high burdens of TB infection [16,63,12]. This study found that the greatest decrease in initial risk was among the SSA and over 65 population who also showed the highest rates of disease overall. There are two potential explanations for a sharper decline in these particularly high risk groups. One is that the importation of prevalent active disease is greater and thus a more dramatic drop in risk is seen as active cases are identified and treated. Another possibility is that these people experience a higher annual risk of infection in their regions of origin and the effective removal of this high risk of infection results in a sudden decrease in active cases. The incidence rate remains elevated, however, which suggests that either removing the risk

of recent infection is not enough and that reactivation disease is a high risk, or that recent infection is not being removed either due to associations here in Ontario or revisiting endemic countries.

One study of immigrants to Alberta found that 50% of TB cases developed more than 7 years after arrival [48], which corresponds to the findings of this study that even after 7 or 8 years, the risk is still not negligible. A cross-sectional study in the U.S. found that people from countries with a high prevalence of TB maintained incidence rates higher than 20 per 100,000 (about 3 times the national average) for more than 20 years after arrival [36]. In the current study, other than the initial drop in risk seen in the first 1-2 years, only FSEE and EME show a visible pattern of decreasing risk over time. The TB risk for all the other world regions in this study seem to plateau over time and for some, seem to have small increases around the 7th year. However, when time since arrival was included in a model, there was a decreasing level of risk of TB with time since arrival. People who were in Canada 1 year or less experienced 5 times the risk of people who were in Canada for 8 years (the baseline). Unfortunately, this pattern could not be investigated in the model at the level of world region, or sex due to small numbers of cases.

The recent study by Wobeser [14] on foreign-born populations in Ontario, found different levels of risk for active TB for different years of arrival. A variation in risk by calendar year has been described before in the context of varying annual risks of infection in populations [28,99], but little has been written on annual variations in incidence of active disease. To investigate whether this occurred in our population, TB risk by calendar year of arrival was plotted and it was found that there was a difference in risk by calendar year of arrival. People arriving from SSA had a higher risk associated with arrival in 1990, 1991 and 1997. People from China arriving in 1990 or 1991 also had a higher risk for TB. An increased risk associated with 1990 and 1991 arrival could contribute to the apparent surge in risk seen after 7 or 8 years in Canada by some groups. The observation could also be an effect due to small numbers, and future studies would have to follow a cohort for a longer period to investigate this phenomenon.

### 5.6.2 TB Risk and Age

Another issue to consider when interpreting pattern of risk, is the concept of age. Age has been defined in this study as the age at arrival. Age has a bearing on infection rates since the assumption is that the more years someone lives in a TB endemic country the higher is the risk of being infected [91,28]. In other

words, there is a "cumulative" risk of infection over a lifetime. A study conducted in Montreal found that tuberculin reactivity (likely TB infection) among immigrants was associated with age at immigration among those from countries with high and intermediate TB rates [69]. There were higher rates of significant reactions among those who were older at immigration [69].

Another "age effect" is that of biological aging. It must be kept in mind that over the period of the study the people are getting older. The combination, in our survival analysis, of measuring age at landing and time since landing, inherently includes effect of biological aging without directly measuring it. In the 16 to 30 age group, for instance, there could be more than one level of risk which is masked by grouping together 14 years of age. In that time period a person may have left home and started to work, or gone to university, or had children or a number of different life scenarios. There tend to be biologically based differences in rates of progression associated with different increasing age intervals [10,29,22,41,19,11].

The observed increase of risk after 7-8 years of arrival was observed most strongly in males in the 16-30 age group. This increase in risk at year 7-8 disappears when duration of stay in included in the model except for the 46-65 age group which shows a small increase. When age was entered into the model along with duration of stay in Canada, the eldest age group had the most dramatic drop in risk over time. By year 2 after arrival, the over 65 age group and the 46-65 age group were very similar. The numbers were quite small for the time-dependent analysis so further study should be undertaken to look at the possible effect of an increase of risk over time. Treating age as a continuous variable and observation over a longer period may provide more information on this phenomenon.

One theory to account for the lack of decrease in risk (and suggestion of an increase) in males, is the "Creatore second wave of risk" hypothesis. It is a common immigration practice for one member of a family (usually the male) to immigrate first, to try to find work and settle down before bringing the other family members [92]. One person arriving first also facilitates immigration for the rest of the (often extended) family due to the Canadian immigrant sponsorship program. Therefore, it is possible that males in particular experience a resurgence of risk of infection as multiple family members join him over time. These family members may have been recently infected, or have reactivation disease. Most studies claim that particularly among immigrants, endogenous reactivation of disease is the most common method of disease development [23,16,25]. If the "second wave of risk" hypothesis is correct, this study would support the

idea that reinfection may be an important cause of disease. This would also have implications with respect to targeting newly arrived family members for testing and treatment for TB as well as following-up people who have been in Ontario for several years but who may have contact with newly arrived family members.

A second cause of a "second wave of risk" would be revisiting the country of origin and becoming infected, or reinfected, at that point [30]. The risk of exogenous reinfection is thought to be significant in areas with high incidence of disease [23]. It is conceivable that for many recent immigrants from countries with high burdens of TB, the financial ability to travel back to their country of origin would only be possible after 7 or 8 years of work in Canada. Again, this could cause the interesting increase in risk after a period of time in Ontario.

The plateau in risks fits in well with the assumption that the majority of disease in the foreign-born is due to reactivation from remote infection. If the disease observed over time is due to reactivations of infection that were acquired before immigration, then it would make sense that the risk would remain elevated for the foreign-born and relatively stable. After the initial decrease due to the decline in risk of new infection or reinfection, and treatment of prevalent, active disease, it would logically follow that the risk would stabilize. The initially high rate has been described before as the initial "skimming off" of particularly unstable cases, leaving behind the more stable annual relapse rate due to latent disease [47].

According to some of the literature, there is a surge in morbidity in the reproductive years of both sexes [11,93,19,55]. Such a phenomenon is particularly characteristic of populations with a high annual risk of infection [49,7]. This study found a high rate in the 16-30 age group, for both sexes, particularly in people from the higher incidence countries. SSA, India, Other and Islands, and China all showed a peak rate in the 16 to 30 age group. The second peak was found in the over 65 age group. The particularly high risk in the 16 to 30 age group from SSA may be contributed to by high rates of HIV in this population. No information on HIV status, however, was obtained for this study. The bimodal pattern of risk seen in this study has been documented [14,24,12,69], particularly in women [24,14,9]. This sex-related pattern will be discussed below.

The high risk associated with the oldest age group is also a well documented phenomenon [9,22,55,42,57,11,26]. As a person ages, their chance of being infected with TB, particularly in countries where the disease is endemic, accumulates [28,91]. There is also a cohort effect, which was first described by Frost in

1939, causing high rates in older age groups due to experience of higher rates of annual infection in earlier life [98]. Therefore, someone who is in their seventies experiences a "residual" increased risk due to higher annual rates of infection from earlier periods. There is also some evidence that the risk of progression to active disease is greater in old age [42]. A pattern of high risk in young people is frequently reported in countries where the risk of disease is high, whereas in countries where TB was once common, but is now rare, the older segment of the population tends to be mostly affected [26,11]. One eminent TB researcher stated that as risk of infection in developed countries declines, " TB shifts from being a disease of young women to a disease of old men" [11] . An important point to keep in mind, is that Canada receives large numbers of immigrants from countries with both types of TB epidemiologies.

A bimodal effect could be interpreted as signifying an overlap of two epidemiological patterns of disease from countries experiencing different "stages of the TB epidemic". Sally Blower et al. discuss in their paper the effects of the development of TB by 2 pathogenic mechanisms. The first is direct progression, resulting in a "young epidemic", the second is endogenous reactivation resulting in a "mature epidemic" [84]. They explain how as the proportion of individuals who experience disease by the 2 mechanisms changes, so does the age-distribution of cases seen in a community since "latently infected individuals develop disease slowly through endogenous reactivation and therefore (in general) have to be older than individuals who develop disease quickly through primary progression." [84] In theory, if you overlap a "young" and a "mature" epidemic pattern, you would result in the observed bimodal pattern of disease. The significance of this theory becomes clear when it is extrapolated to explain the bimodal observation. That is, we are observing not only the well-documented residual effect of high rates of infection of the past, but a reemergence of high rates of infection. This is resulting in the "young" epidemic which contributes the second, more worrisome peak in the 16 to 30 age group.

### 5.6.3 TB Risk and Sex

Some studies show that progression from infection to disease is faster among women of reproductive age than among men at the same age [43,49,26,11]. In Ontario, women between the ages of 10 and 34 years had up to 30% higher rates of disease progression than men [26]. In the current study, there was no significant association with sex in any predictive model, although some sex-related trends were found. The

lack of an effect associated with sex except within the oldest age group, has also been documented in the literature, although conflicting opinions exist. There are various explanations for the effect of sex on TB risk. One is a tendency for under-reporting of disease in women, to explain men having higher rates post-reproductive age [45,26,6,49]. Another is that there is a reduction in immunological resistance during the reproductive years in women, to explain the higher peak in women during those years [26,94]. There are still no conclusive findings in human models regarding an increased biological susceptibility to active disease during pregnancy, which is the only physiological sex difference that has yet to be explored [94,49]. Researchers also tend to quote the sociological phenomenon of men having more "out of the house" contacts found in many cultures, and thus a higher risk of infection [26]. This sociological phenomenon is generally accepted to explain why *infection* rates are similar for men and women up to the early teens, after which men tend to have higher rates of infection [26-28,22,11,7]. A higher burden of infection in the male population could result in a higher number of active cases. Generally it has been found that males show higher rates of disease after the reproductive years and women show higher rates during those years [7,24,9,26,22,43]. This pattern has been found even after controlling for notification rates [45].

The fact that this study did not find a sex difference in TB risk below the age of 65 may signify a true lack of a difference, or some effect of passive surveillance tactics. Perhaps rates of TB among non-reproductive age males would be increased if an active approach to case-finding was used and diagnosis was not dependent on voluntary visits to the doctor. Men may access the health care system less frequently than women of the same age during the reproductive years [26]. Perhaps young women also do not have equal access to health care. This theory has often been considered due to greater time and financial constraints on women, as well as lower social status and educational attainment [26]. These issues may be particularly salient in women coming from countries where there are greater societal gender inequalities.

There is also the possibility that there is not actually a higher risk in females during the reproductive period, but that females are more likely to be in contact with the health care system during these years and therefore disease can be more readily detected. In other words, some form of detection bias may exist. Perhaps in developing countries child-bearing is one of the only reasons for a person to have frequent contact with a health care provider, resulting in higher notification-rates among women in this age group. On the other hand, studies finding a higher rate among females during the reproductive years have

also been conducted in the high income settings of Ontario, Denmark, Norway and the U.S., all of which found a rate 25-30% higher in women [26]. If we assume that higher income countries promote more equal access to health care between the sexes, then the fact that the same pattern is found in high income settings suggests that there is more at work than notification rates. Perhaps the risk factor for higher rates in women is *recent* infection during the child-bearing years which carries with it a high risk of progression to active disease. Immigrating to a low TB burden country should significantly decrease the risk of infection and thus rates decline. Perhaps TB rates under "natural conditions" among reproductive-age women may be less due to reactivation disease, and more likely due to primary progression of recent infection.

### 5.6.4 The Interpretation of Interactions Between Study Variables

A significant interaction term between two variables in the model means that the risk of TB associated with one variable varies according to different levels of the second variable. We found that there was some interaction between certain world regions and age, specifically Sub-Saharan Africa, Middle East, Latin America, India and the Former Socialist Europe. There were two cases of interaction involving sex: Sex and the over 65 age group, and sex and China.

Interaction with sex is easier to explain. The risk of TB was most significantly different between males and females in the oldest age group when risk for males exceeded the risk for females. Similarly, males from China had a higher risk of TB than females at every age group over 15 years of age. Both phenomena are likely related to the tendency for greater rates of infection in males due to higher rates of "out of the house" contacts in adult life as described above [26]. The phenomenon of higher rates of disease in elderly males is probably explained by the cohort effect described first by Frost [93]. There are numerous socio-behavioural factors that caused higher annual risks of infection to be experienced by males in earlier decades. This could result in a concentration of the cohort effect in males from certain generations that later in life may be manifested in higher rates of reactivation disease in elderly males.

The interaction terms involving world region of origin and age are more complicated to consider. It was expected that different ages would experience different rates of disease due to the complex relationship between age and risk of infection, age at emigration and cumulative risk of infection (which is related to annual risk of infection for that country) and finally, age and progression to active disease. What

is more difficult to interpret is why some countries exhibited more levels of interaction than others. For instance, for the 16 to 30 age group, interaction was seen with Sub-Saharan Africa, Latin America and India. For these 3 world regions, the 16-30 group has a distinct relative risk for TB that followed a different pattern than for the other ages. SSA and India experience high infection rates of TB. It would be expected that there would be a higher risk associated with the 16 to 30 age group due to primary progression as a result of recent infection [49]. The question is why is this effect not seen for Asia and China which also have high rates? China exhibits a relatively constant level of risk until the over 65 age group when the risk is elevated. Asia also shows a high level of risk for all ages, with a peak in the over 65 group. Perhaps effect of HIV in the 16 to 30 age group, particularly with SSA is showing up as a strong interaction term between age and world region.

There was interaction between the over 65 age group and Sub-Saharan Africa, the Middle East, and Former Socialist Europe. The over 65 age group was generally associated with a higher risk of TB than the other age groups, except with Sub-Saharan Africa which had an almost identically high risk in the 16-30 group. Interaction between risk for TB in immigrants from these regions and the eldest age group may be due to a residual effect of high rates of TB in the past which are still affecting the oldest age groups [49]. Sub-Saharan Africa, the Middle East and Latin America also experienced interaction at the 46 to 65 age group. This is likely explained by the large increase in risk from the 31-45 group to the 46-65 age group shown by both SSA and MEC. The other world regions show a slower increase between the latter two age groups and then a jump in risk in the over 65 group. Latin America is the only world region-age interaction resulting in a decrease in risk from the 31-45 age group through the 46-65 age group. All the other world regions show an increase in risk in each successively older age group.

How are these differences explained when interpreting risk for TB in immigrants since arrival in Ontario from these world regions? It would seem that the interaction between world region and age for SSA, MEC, Other Asia, India and China, all show the pattern of residual high risk in the older segment of the population from earlier periods, as well as continued high rates of TB strongly affecting the 16 to 30 age group. As already mentioned, the two high risk periods in adult life to develop active disease are during the reproductive prime, and old age. On the other hand, to have a true cohort effect due to residual effects of high rates of infection, there have to be distinct changes in annual risk of infection. Particularly

for high incidence regions such as SSA, it is unlikely that there have been large changes in infection rates over the past decades. The Established Market Economies, Former Socialist Europe, and Latin America to a certain extent, show the pattern of a country with relatively low rates at the moment, but which are still seeing the vestigial pattern of old disease in the elderly from times of higher risk for TB. The pattern exhibited by these regions is a relatively constant and low pattern of risk until the over 65 age group when there is a significant increase.

Interaction involving the time-dependent variables could be investigated in a rather limited way due to small sample size and a lack of power. It does seem, however, that risk of TB from the different world regions may vary by calendar year. Different levels of risk were seen for the world regions depending on year of arrival. Sub-Saharan African revealed a strong pattern of high risk associated with immigrating in 1990, 1991 and 1997. Other Asia experienced an increase in risk in people arriving in 1993. Duration of stay in Canada and associated risk for TB over time since immigration tends to follow a different pattern depending on age at arrival. For instance, the older age groups show a greater initial decrease in risk after arrival (refer back to Figure 6). This is likely due to prevalent cases in newly arrived immigrants which are detected and treated soon after arrival.

### 5.6.5 The Possible Role of Socioeconomic Status

Crowding, poverty, homelessness and living in correctional facilities are all associated with risk of TB infection [7,32-35,20]. Poverty and malnutrition are also risks for disease activation [46,33]. The extent to which poverty accounts for ethnic discrepancies in TB rates has not been quantified to any great extent [32], particularly in Canada. In Montreal case investigations have suggested that many TB patients suffer from poverty, homelessness and alcohol or drug use, however, there is no data quantifying this observation [12]. One recent study conducted in the U.S. found that by adjusting for SES, one could account for about 50% of the increased risk for active TB that is associated with U.S.-born African-Americans, Hispanics, and Native Americans [32]. This study also found that there was a significant correlation between SES and crowding, with the lowest SES group experiencing the greatest crowding (as measured by persons per room per household) [32].

A 1992 study conducted in Montreal found that after adjusting for other factors, tuberculin reactivity (possible infection) was higher among immigrants living in poorer neighbourhoods in Montreal [69]. A 1998 case-control study in Ontario found that immigrants with greater than 12 years of education were at decreased risk for TB compared to those with equal or less than 12 years of education [14]. This latter study measured SES at the time of immigration, which may not accurately reflect the risk experienced post-landing although there should be some correlation. These two studies are the only ones to our knowledge that addressed the issue of SES and its effect on TB in Canada. It is difficult to say to what extent the results in the current study may have been explained by an association with crowding or poverty if there had been some available measure of SES, other than correlating years of schooling with disease. The correlation in this current study suffers from the "ecological fallacy" limitation common to ecologic analyses. That is, we do not know whether the mean number of years of schooling for their country of birth was representative of the people who actually developed TB. It is possible that within world regions, the TB cases in this study had lower or higher levels of schooling than the mean from that region. Further study among this population in Canada including a reliable measure of SES at immigration and over time would be useful.

The previous study by Wobeser[14] conducted in Ontario found that refugees and family class immigrants had a significantly higher risk of TB than other classes of immigrants. This present study did not use information about immigrant class in the analysis. If refugees are contributing a significant number of cases to the numerator, and we have not included them in the denominator, the risk ratio will likely be inflated. There are 2 types of refugees, convention refugees and refugee claimants. Convention refugees generally apply for asylum from abroad, and therefore are subject to a medical evaluation before entry. When convention refugees arrive in Canada, they arrive as a landed immigrant. Refugee claimants are those who apply once they are in Canada for asylum and who may or may not have undergone a thorough medical exam. The numerator will only be inflated by the number of refugee claimants developing TB, since convention refugees are included in the "legally landed" population. There are approximately 15,000 refugee claimants and between ten and seventeen thousand convention refugees to Canada annually [95] (See table 25).

The effect that including refugee claimants in the numerator and not the denominator may have on the results of objective 2 depends on the distribution of cases by world region and age that are refugee claimants. It is known that a significant portion of convention refugees to Canada come from Africa and Asia and the Pacific (see Table 26), but whether the same pattern is seen among refugee claimants, and what the case distribution among these refugees is, is not known beyond the results of the aforementioned study [14]. Thus, the high risk associated with SSA, Other Asia and India may in part be explained by a higher proportion of refugee claimants in this group. The study by Wobeser, however, found that SSA and Other Asia were a high risk group even after excluding persons who were not landed at the time of diagnosis [14].

**Table 25. Proportion of immigrants to Ontario in 1995-97 that were convention refugees.**

| | *Total number of Immigrants** | *Total number that were refugees** | *Proportion refugees (%)* |
|---|---|---|---|
| 1995 | 115,901 | 16,405 | 14.2 |
| 1996 | 119,713 | 13,933 | 11.6 |
| 1997 | 118,085 | 11,670 | 9.9 |

*Numbers courtesy of "Facts and Figures 1997" by Citizenship and Immigration Canada [95].

**Table 26. Distribution of convention refugees to Canada in 1997 by the 2 world regions with the highest risk.**

| | *# of Refugees* | *% of total # of refugees from all World Regions* |
|---|---|---|
| Africa & Middle East | 7,935 | 32.8 |
| Asia & Pacific | 7,167 | 29.6 |
| Total | 15,102 | 62.3 |

*Numbers courtesy of "Facts and Figures 1997" by Citizenship and Immigration Canada [95].

5.6.6 Case Detection of Active Disease and Targeted Testing of Latent Infection

The two strategies for case-detection, active and passive surveillance are both used for diagnosis of TB among immigrants to Canada. Active case-finding occurs when the entire immigrating population is screened for active disease before receiving legally landed status. There is also some active case-finding among persons referred to public health officials as a result of immigration medical evaluation. Passive

case-detection occurs after immigration by diagnosing only those who present themselves to the health care system with risks or symptoms suspicious of TB [10,38,6]. There is wide-spread agreement to use the passive case-detection method for the identification of active TB among the general population. This policy is advocated by both the WHO and the IUATLD [6,10].

The argument against instituting an active targeted testing program for latent infection is based on 3 points. First, appropriate care should be physically, socially and economically available, which is why developing countries generally do not actively screen for latent infection. Second, people who are identified as infected, but who are not suffering from any symptoms of active disease, may be less likely to comply with a drug regimen. This, in turn may make them more susceptible to acquired drug resistant strains of TB if they do develop active disease [6]. Targeted testing of people to identify latent infection without establishing the infrastructure to offer effective treatment is discouraged by the Advisory Council for the Elimination of Tuberculosis (ACET) [18]. The final issue is cost. The issue of cost-effectiveness and access is central to most debates over testing for latent infection.

The argument for targeted testing and active case detection methods among high risk groups is simply that more ill or infected people may be diagnosed and subsequently offered treatment for active disease and latent infection, respectively. Sometimes by the time someone comes to see a physician, they may have been ill for some time. This delay in diagnosis affects both their own health, and their close contacts. A study was conducted recently in Nepal looking at notification rates among women. This study revealed that by using passive case-finding methods women made up 28% of the cases, whereas by using active methods, the percentage of women rose to 46% [96].

In order to reduce the risk posed by TB, methods to identify infected people and offer them treatment (targeted testing) are important. Studies such as the current one are necessary to break down high risk groups further into even higher risk groups, and to describe the pattern of disease over time in foreign-born populations. There are still reservations about active case-finding among immigrant and refugee populations based on low cost-effectiveness and politics [6,18]. Identifying specific high risk groups and behaviour of the disease over time in these populations would provide valuable information for developing community-based TB control programs. There is also a need to improve access to high risk populations, and to increase compliance with drug therapy among these groups and among physicians. Both these

priorities are particular challenges among the foreign-born population that may have linguistic, cultural and social barriers to overcome in these areas [18]. It has been suggested that by increasing community-based programs dependent on local agencies and community-integrated programs for immigrants and non-English speaking persons, a more efficient targeted testing program could be implemented [18].

Containing active disease is the first priority of TB control and is an obvious and necessary first step towards eliminating TB. Identifying people who are infected and offering them treatment is also important since recently infected people are at particularly high risk for progression to active disease. As outlined by Blower et al., treatment of active cases alone will never achieve eradication of TB. Treatment of active disease, as well as of latent infection in high risk groups is the only way of eradicating TB [1]. This is based on the assumption that we cannot sustain high enough and successful enough levels of treatment for eradication. Thus, the issue of eradication becomes a complex interplay of case-detection, treatment rates, and cure rates.

### 5.6.7 Global Trends

As stated earlier, there are 4 very significant global trends that are complicating the control of TB [9]. These are:

1) An increase in TB/HIV coinfection (which is also increasing the mortality rate from TB)
2) An increase in multi-drug resistant strains of TB
3) An increased migration of populations
4) Increasing urbanization.

Unlike many other infectious diseases TB is intrinsically difficult to predict due to a delay between infection and disease which is extremely variable [54,49]. What is clear, however, is that countries with a low risk of TB experiencing immigration from countries with higher risks of TB infection will continue to experience cases of TB.

There is no way that a study such as this one can contribute formulae, or mathematical models to account for the changing patterns of migration, new diseases, the behaviour of drug-resistant mutations of disease or new vaccines. These phenomenon, however, may greatly contribute to the epidemiology of TB over the next few decades. This study describes the pattern of TB risk among immigrants over the past 8

years in Ontario. The pool of persons at risk is dynamic and varies from region to region. Certain barriers to tuberculosis control may be overcome by better understanding of the local epidemiologic profiles to tailor needs and to identify groups who are at high risk for TB. The goal of the current study is to describe the epidemiology of TB in Ontario among recently arrived, foreign-born persons.

Using the World Bank designations of "World Regions of Origin" has its limitations. First of all, by dividing the world up into broad regions, localized variations in risk due to sociologic phenomenon such as wars and political instability, or unpredictable phenomenon such as natural disasters, cannot be teased out from the regional average. For example, in the current data, a large number of Somali refugees, experiencing a high risk of TB, likely increased the estimates for Sub-Saharan Africa [14]. In the same way that more recently, large numbers of Tibetan refugees will likely inflate the risk for TB among immigrants from China. The instability in the former Soviet Union and breakaway republics is causing high rates of TB to reappear. The combination of HIV/TB is causing higher TB mortalities globally as well as facilitating the spread of drug-resistant strains [8,9,29,49]. The trend of increased and denser urbanization is associated with high risk of transmission due to higher rates of contact and greater homelessness [49]. All these variables interact in a way to make definitive predictive models among immigrants to Ontario difficult.

There are two primary types of information that are gained from this study. First, an understanding of the current status of TB among immigrants in this province, including risk factors and incidence rates for this disease. Secondly, this study highlights trends occurring over time that are dependent on region of origin and age in a way that will continue to impact the Ontario TB situation over the coming decade. The possibility that rates may increase due to a "second wave of risk" phenomenon has implications for follow-up of high risk groups, as well as widening the window of time in which screening for disease may be recommended. Further research specifically focussing on this theory is needed. Global trends are not static, however they do follow patterns and have certain similarities with historical events that can make planning prevention programs a more informed process.

### 5.7 Strengths and Limitations

The primary strength of this study is that this information is lacking for Ontario. This is the first study to look at Ontario and develop not only incidence rates for immigrating groups, but to describe the risk over time and consider what variables affect the probability of developing active TB in the first 7-8 years after arrival. As well, certain assumptions are likely to be stable at least over the next decade, such as sex and age-pattern risks. Understanding the epidemiology of TB in high risk immigrant populations is the first step towards considering issues that affect the health of new Canadians. This is important in order to aggressively screen and prevent serious health outcomes in these populations. The current study is unique in quantifying the change in risk after migration from a high to a low burden country. This kind of research can contribute a great deal by directing attention and funds to the appropriate areas.

Other cohort studies which have been conducted among immigrants at risk for TB have either involved significantly fewer people, or used census data for the denominator. Most studies in Canada and the U.S. have been conducted using disease data from TB registries and immigration data from Census information [13,24,47,69,32,36,55]. One problem with Census data is that it is a sample of the true population. It is not accurate to classify people by their self-reported immigration status and ethnicity. On the other hand, census data is more likely than Immigration data to include people who migrate into Ontario from other provinces. This study is the first immigrant TB cohort study in Ontario to use denominator data from Citizenship and Immigration Canada landing documents.

One of the primary weaknesses of this study is that we do not have the ability to directly link the data from the two databases since RDIS is non-nominal. Thus, we cannot determine which individuals within certain risk groups became cases and what was their individual sociodemographic, marital status, or immigration class. Individuals from the same country could have vastly different risks of infection depending on their personal risk factors [42]. In objective one we try to consider this problem by looking at trends in education among the world regions. Information on sociodemographic status and immigration class would surely contribute further useful information and likely more detailed risk assessments.

The second major issue is also related to the inability to link the databases. The numerator and denominator data used in this study do not correspond perfectly to include the same population. As mentioned above, the numerator included a potentially important risk group that may not be included in the

denominator. Previous studies have shown that a quarter to one-third of non-Canadian born people who develop active TB were not legally landed immigrants, [13,14] and therefore would not be included in the denominator of this study. This study cannot separately estimate the risk of non-landed or landed persons. These trends may bias the results to overestimate the incidence among certain immigrant groups. It could not be investigated whether there was uneven distribution among the non-landed cases across world regions, as found in the study by Wobeser [14], nor whether there was a relation between when a person developed disease and their landing status.

On the other hand, considering the large number of immigrants to Ontario every year (between 115,000 and 120,000) [95], the effect may not be so large. The results are likely inflated, but not to an unrealistic number. The interpretation of the results in objective 2 should be to consider the risk as representing the risk for all new arrivals, not just landed immigrants. It may be more accurate to consider not that cases should be excluded, but that the denominator is an underestimate of the true number of people arriving to Canada.

Another difficulty lies in the inability to control the accuracy of our exposure assumptions. A certain amount of measurement error is expected in our grouping of people into exposure levels by world region of origin. We group people according to their country of birth, and have no information on how many years they may have spent elsewhere, or the last country of permanent residence. This is information that may be particularly relevant for people who left a country for political or economic reasons and who spent time as refugees, homeless persons, or persons institutionalized in another country before arriving in Canada. The previous study by Wobeser (1998) on the immigrant population between 1994 and 1995 found that the above phenomenon was quite uncommon [14].

Once a person is legally landed we know nothing of their behaviors in Canada. We will not know whether the person returns frequently or for long periods of time to their country of origin. Or, if they remain in Canada associating with a group with a high rate of disease, and thus higher transmission rates than in the rest of Canada. Therefore, the interpretation that risk for TB among people from SSA and MEC remains the same as in the country of origin may reflect that these groups are more likely to associate with other people from the same region and therefore maintain a high risk of infection here in Ontario. We can only attribute high rates of disease to the variables we have available to us in the study.

### 5.8 Conclusions and Recommendations

Infectious diseases arriving from other countries through the migration of populations is no new concept. Europeans were responsible for the deaths of millions of Native Canadians in the 16th century simply by colonizing the New World and exposing those who already lived there to foreign infectious diseases [97]. One of the classic causes of the reemergence of old diseases or the emergence of new ones is changes in patterns of human behaviour, particularly migration or travel [97]. This will have important implications on future efforts to effectively target populations for risk reducing interventions.

This study suggests that although people arriving generally show a decrease in risk over the first couple of years, the high risk associated with being foreign-born does not dissipate, at least not within the first 8 years. This is particularly the case with high burden countries such as Sub-Saharan Africa, Asia and Pacific Islands and India. Furthermore, for some world regions, in particular Africa and the Middle East, the risk does not stray significantly from the risk experienced in the world region of birth. This goes against the popular beliefs found in the literature that the risk remains somewhat elevated, however does decrease from the rate in the country of origin [36,13]. This fact has important implications for targeted testing of people who immigrated 5 to 8 years previously from endemic countries, namely that they should still be considered a high risk group.

Information from this and other similar studies should be used to tailor prevention programs to the Ontario-specific immigration experience. We recommend community-based, active screening of selective high risk groups as identified by this study. It should be kept in mind, however, that the primary priority is treatment of all passively detected active cases. However, many factors can decrease the yield of such an undertaking [18] and need to be considered when planning any TB control program. Developing targeted testing programs is important, though should only be done once there is an effective program in place treating all active cases. Implementing more community-based screening programs and tailoring the program to local high-risk population profiles would increase the cost-effectiveness of actively screening immigrants and refugees. As well, further studies specifically looking at these populations in Ontario are needed. Specifically, looking at a cohort for a longer period of time and direct linkage of the 2 databases to provide sociodemographic and other personal information is needed. Exploring the epidemiology of TB in the long-term, along with family reunification patterns and revisiting of the country of birth would be

interesting to determine if these factors may have an impact. Failure to identify active cases and use preventive measures on high risk populations will lead to increased transmission from undiagnosed active cases. The potential for large outbreaks of TB is real in this setting [6,11].

We have shown that: a) the epidemiology of TB is hypervariable amongst different migrating populations; b) that the highest risk period is during the first couple of years after arrival; c) that risk may not always decline permanently following immigration. Although the initial high risk is familiar to TB epidemiologists, this is the first study to document the variability in risk by world region, and age and assign risk ratios. It is important to recognize that this study represents a "snapshot" of the current situation and that future trends in the epidemiology of TB will be determined by multiple factors. The most important of which, are the number and characteristics of the incoming population. Ongoing studies will be needed to capture and track the dynamic populations at risk. The hypervariability in population epidemiology underscores the importance of designing population specific interventions.

## References


1. Blower SM, Small PM, Hopewell PC. Control strategies for tuberculosis epidemics: New models for old problems. Science 1996; 273: 497-500.

2. Bloom BR, Murray CJL. Tuberculosis: Commentary on a reemergent killer. Science 1992; 257: 1055-63.

3. McKenna MT, McCray E, Onorato I. The epidemiology of tuberculosis among foreign-born persons in the United States, 1986 t 1993. N Engl J Med 1995;332(16):1071-6.

4. Castro KG. Global Tuberculosis Challenges. Emerg Infect Dis 1998; 4(3): 1-3.

5. Dye C, Scheele S, Dolin P, Pathania V, Raviglione MC. Global burden of tuberculosis – Estimated incidence, prevalence, and mortality by country. JAMA 1999; 282: 677-86.

6. Murray C, Styblo K, Rouillon A. Jamison DT, Mosle6y WH, Measham AR, Bobadilla JL, editors. Disease Control Priorities in Developing Countries. Oxford University Press; 1993; 11, Tuberculosis. P.233-59.

7. Comstock GW, Cauthen GM. Reichman LB, Hershfield ES, editors. Tuberculosis: a comprehensive international approach. New York: Marcel Dekker, Inc; 1993; 2, Epidemiology of tuberculosis. P. 23-48.

8. Huebner RE, Castro KG. The changing face of tuberculosis. Annu Rev Med 1995; 46: 47-55.

9. Kerbel D. Epidemiology of tuberculosis in Ontario 1995. PHERO 1997; 8(4): 81-93.

10. Murray CJL, Styblo K, Rouillon A. Tuberculosis in developing countries: burden, intervention and cost. Bull. Int. Un. Tuberc. 1990; 65(1): 6-24.

11. Murray CJL. Social, economic and operational research on tuberculosis: Recent studies and some priority questions. Bull. Int. Un. Tuberc. 1991; 66: 149-56.

12. Rivest P, Tannenbaum T, Bedard L. Epidemiology of tuberculosis in Montreal. CMAJ 1998; 158(5): 605-9.

13. Enarson D, Ashley MJ, Grzybowski S. Tuberculosis in immigrants to Canada. Am Rev Respir Dis 1979; 119: 11-8.

14. Wobeser WL. Risk factors for active tuberculosis among foreign-born persons arriving in Ontario 1998; Graduate Department of Community Health, University of Toronto.

15. Orr PH, Manfreda J, Hershfield ES. Tuberculosis surveillance in immigrants to Manitoba. CMAJ 1990; 142(5): 453-8.

16. Centers for Disease Control and Prevention. Recommendations for prevention and control of tuberculosis among foreign-born persons: Report of the working group on tuberculosis among foreign-born persons. MMWR 1998; 47(RR-16): 1-29.

17. Fanning EA. Globalization of tuberculosis. CMAJ 1998; 158(5): 611-2.

18. Nolan CM. Community-wide Implementation of Targeted Testing for and Treatment of Latent Tuberculosis Infection. Clin Infect Dis 1999; 29: 880-7.

19. Comstock GW, Livesay VT, Woolpert SF. The prognosis of a positive tuberculin reaction in childhood and adolescence. Am J Epidemiol 1974; 99(2): 131-7.

20. Comstock GW. Modern epidemiology of tuberculosis. Am J Epidemiol 1975; 101(5): 367-382.

21. Iseman MD, Madsen LA. Clin Chest Med 1989; 10: 341.

22. Sutherland I. Recent studies in the epidemiology of tuberculosis, based on the risk of being infected with Tubercle bacilli. Advances in tuberculosis research 1976; 19: 1-63.

23. Van Rie A, Warren R, Richardson M, Victor TC, Gie RP, Enarson DA, Beyers N, van Helden PD. Exogenous reinfection as a cause of recurrent tuberculosis after curative treatment. N Engl J Med 199; 341(16): 1174-9.

24. Grzybowski S. and Allen, EA. The challenge of tuberculosis in decline – a study based on the epidemiology of tuberculosis in Ontario, Canada. Toronto, Ont. Dept. of Health. 1963; p.707.

25. Chin DP, Deriemer K, Small PM, deLeon AP, Steinhart R, Schecter GF, Daley CL, Moss AR, Paz EA, Jasmer RM, et al. Differences in contributing factors to tuberculosis incidence in U.S.-born and Foreign-born persons. Am J Repir Crit Care Med 1998; 158(6): 1797-803.

26. Holmes CB, Hausler H, Nunn P. A review of sex differences in the epidemiology of tuberculosis. Int J Tuberc Lung Dis 1998; 2(2): 96-104.

27. Nyboe J. Interpretation of tuberculosis infection age curves. Bull WHO 1957; 17: 319-39.

28. Styblo K, Meuer J, Sutherland I. The transmission of tubercle bacilli: Its trend in a human population. Bull Int Union Tuberc 1969; 42: 1-104.

29. Blum RN, Polish LB, Tapy JM, Catlin BJ, Cohn DL. Results of screening for tuberculosis in foreign-born persons applying for adjustment of immigration status. Chest 1993; 103(6): 1670-4.

30. McCarthy OR. Asian immigrant tuberculosis – the effect of visiting Asia. Br. J. Dis. Chest 1984; 78(3): 248-53.

31. The Lung Association. TB in Toronto: Information for physicians. www.oma.org/phealth/tbinfo.htm. Ontario Medical Association. 1999; p.1 Tb info.

32. Cantwell MF, McKenna MT, McCray E, Onorato IM. Tuberculosis and race/ethnicity in the United States. Am J Respir Crit Care Med 1998; 157: 1016-20.

33. Tocque K, Doherty MJ, Bellis MA, Spence DPS, Williams CSD, Davies PDO. Tuberculosis notifications in England: the relative effects of deprivation and immigration. Int J Tuberc Lung Dis 1998; 2(3): 213-8.

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

35. Ferguson JS, Hornick DB, Dayton CS. Patients with an abnormal chest radiograph and latent tuberculosis. Am Fam Phys 1996; 54(8): 2495-500.

36. Zuber PLF, McKenna MT, Binkin NJ, Onorato IM, Castro KG. Long-term risk of tuberculosis among foreign-born persons in the United States. JAMA 1997; 278(4): 304-7.

37. American Thoracic Society. Diagnostic standards and classifications of tuberculosis. Am Rev Respir Dis 1990; 142: 725-35.

38. Horwitz O, Wilbek E, Erickson PA. Epidemiological basis of tuberculosis eradication. Bull WHO 1969; 41: 95-113.

39. Styblo K. Flenley DC, Retly JL, editors. Advances in Respiratory Medicine. Edinburgh: Churchill Livingstone; 1986; Tuberculosis control and surveillance. P. 77-108.

40. Wilcke JTR Poulsen S, Askgaard DS, Enevoldsen HK, Ronne T, Kok-Jensen A. Tuberculosis in a cohort of Vietnamese refugees after arrival in Denmark 1979-1982. Int J Tuberc. Lung Dis 1998; 2(3): 219-24.

41. Powell DE, Meador MP, Farer LS. Foreign-born persons with tuberculosis in the United States. Am J Public Health 1981; 71(11): 1223-7.

42. Clancy L, Rieder HL, Enarson DA, Spinaci S. Tuberculosis elimination in the countries of Europe and other industrialized countries. Eur Respir J 1991; 4: 1288-95.

43. Bakkshi SS, Hawker J, Ali S. The epidemiology of tuberculosis by ethnic group in Birmingham and its implications for future trends in tuberculosis in the U.K. Ethnicity and Health 1997; 2 (3): 147-53.

44. Nolan CM, Elarth AM. Tuberculosis in a cohort of Southeast Asian refugees. Am Rev Respir Dis 1988; 137: 805-9.

45. Gender and case detection. 1999; Madrid: World Conference on Lung Health.

46. Elender F, Bentham G, Langford I. Tuberculosis mortality in England and Wales during 1982-1992: its association with poverty, ethnicity and AIDS. Soc Sci Med 1998; 46(6): 673-81.

47. Wang JS, Allen EA, Enarson DA, Grzybowski S. Tuberculosis in recent Asian immigrants to British Columbia, Canada: 1982-1985. Tubercle 1991;72: 277-83.

48. Cowie RL,Sharpe JW. Tuberculosis among immigrants : interval from arrival in Canada to diagnosis – a 5-year study in southern Alberta. CMAJ 1998; 158(5): 599-602.

49. Rieder HL. Epidemiologic basis of tuberculosis control. Conference proceedings. Paris: International Union Against Tuberculosis and Lung Disease. 1999.

50. Menzies R, Tannenbaum T, FitzGerald JM. Tuberculosis: 10. Prevention. CMAJ 1999; 161(6): 717-24.

51. Yuan L, Richardson E, Kendall RW. Evaluation of a tuberculosis screening program for high-risk students in Toronto schools. CMAJ 1995; 153(7): 925-31.

52. Anonymous. Essentials of tuberculosis control for the practicing physician. CMAJ 1994; 150(10): 1561-71.

53. Nine-Year-old Immigrant Spurs Tuberculosis Outbreak in North Dakota. N Engl J Med 1999; 341: 1491-5.

54. Fine PEM, Small PM> Exogenous reinfection in tuberculosis. N Engl J Med 1999; 341(16): 1226-7.

55. Tocque K, Bellis MA, Tam CM, Shiu LC, Qutub S, Remmington T, Davies PDO. Long-term trends in tuberculosis. Am J Respir Crit Care Med 1998; 158: 484-8.

56. Cantwell MF, Snider DE, Cauthen GM, Onorato IM. Epidemiology of tuberculosis in the United States, 1985 through 1992. JAMA 1994; 272(7): 535-9.

57. Raviglione MC, Sudre P, Rieder HL, Spinaci S. Kochi A. Secular trends of tuberculosis in Western Europe. Bull Wld Hlth Org 1993; 71(3/4): 297-306.

58. Rapiti E, Fano V, Forastiere F, Agabiti N, Geraci S, Scano M, Alichino F, Rinnenburger D. Determinants of tuberculosis in an immigrant population in Rome: a case-control study. Int J Tuberc Lung Dis 1998; 2(4): 479-83.

59. Murray CJL, Lopez AD. Alternative projections of mortality and disability by cause 1990-2020: Global burden of disease study. Lancet 1997; 349: 1498-504.

60. Wobeser WL, Yuan L, Naus M. Outcome of pulmonary tuberculosis treatment in the tertiary care setting – Toronto 1992/93. CMAJ 1999; 160(6): 789-94.

61. Porter JDH, McAdam KPWJ, editors. Tuberculosis – Back to the Future. 3rd ed. New York: John Wiley & Sons; 1994; 11, Multidrug resistance. P. 225-37.

62. McKenna MT, McCray E, Onorato I. Immigration and the resurgence of tuberculosis in the United States. Tuber Lung Dis 1995; 76(2): 124.

63. Binkin NJ, Zuber PLF, Wells CD, Tipple MA, Castro KG. Overseas screening for tuberculosis in immigrants and refugees to the United States: Current status. Clin Infect Dis 1996; 23: 1226-32.

64. Liu Z, Shilkret KL, Tranoti J, Freund CG, Finelli L. Distinct trends in tuberculosis morbidity among foreign-born and US-born persons in New Jersey, 1986 through 1995. Am J Public Health 1998; 88(7): 1064-7.

65. Ormerod P. Tuberculosis and Immigration. Br J Hosp Med 1996; 56(5): 209-12.

66. Somner AR. Tuberculosis among immigrants related to length of residence in England and Wales. BMJ 1975; (3): 698-9.

67. Laboratory Centre for Disease Control – Health Canada. Tuberculosis in Canada 1996. Ottawa. Minister of Public Works and the Government Services Canada. 1998; 2. P.1 Annual report on tuberculosis morbidity and mortality.

68. Wilkins K. Tuberculosis, 1994. Health Reports 1996; Published by Health Statistics Division at Statistics Canada, 8(1): 33-9.

69. Menzies R, Vissandjee B, Amyot D. Factors associated with tuberculin reactivity among the foreign-born in Montreal. Am Rev Respir Dis 1992; 146: 752-6.

70. Ford-Jones EL. Immigrant and refugee mothers. Paediatr Child Health 1999; 4(3): 193-4.

71. Goel V. Factors associated with cervical cancer screening: results from the Ontario Health Survey. Can J Public Health 1994; 85: 125-7.

72. Edwards N. Factors influencing prenatal class attendance among immigrants in Ottawa-Carlton. Can J Public Health 1994; 85: 245-8.

73. Medical surveillance of immigrant populations. Tuberculosis: A call to action, progress through partnerships; Toronto, Ont. 1998.

74. Lipsitch M, Levin BR. Population dynamics of tuberculosis treatment: mathematical models of the roles of non-compliance and bacterial heterogeneity in the evolution of drug resistance. Int J Tuberc Lung Dis 1998; 2(3): 187-99.

75. INS. Naturalization, Deportations. Migration News 1999; 6(12).

76. ACCESS 97 [computer program]. Microsoft Corporation. 1996.

77. SAS Software [computer program]. SAS Institute Inc. Cary, NC: 1998.

78. Kleinbaum DG. Survival Analysis – A self-learning text. New York: Springer-Verlag; 1996.

79. Rosner B. Kugushev A, editors. Fundamentals of Biostatistics. 4th ed. Boston: Wadsworth Publishing Company; 1995; 13, Hypothesis Testing: Person-Time Data. p. 585-631.

80. Krieger N, Williams DR, Moss NE. Measuring Social Class in US Public Health Research: Concepts, Methodologies, and Guidelines. Annu Rev Public Health 1997; 18: 341-78.

81. Klein and Moeschberger, editors. Survival Analysis – Techniques for Censored and Truncated Data. New York: Springer-Verlag; 1997.

82. Allison PD. Survival Analysis Using the SAS system: A Practical Guide. Cary, N.C. SAS Institute, Inc. 1995; 7, Analysis of tied or discrete data with the LOGISTIC, PROBIT, and GENMOD procedures. P. 211-31.

83. Immigration Medical Services. Medical Officer's Handbook. Ottawa. Medical Services Branch, Dept. of National Health and Welfare. 1981; 4. P.410.

84. Blower SM, McLean AR, Porco TC, Small PM, Hopewell PC, Sanchez MA, Moss AR. The intrinsic transmission dynamics of tuberculosis epidemics. Nature Medicine 1995; 1(8): 815-21.

85. Heywood N, Personal communication. Citizenship and Immigration Canada, January 5, 1999.

86. Statistics Canada. The longitudinal Immigration Data Base (IMDB). 1998; Ottawa: Citizenship and Immigration Canada.

87. WHO tuberculosis surveillance and monitoring project. Global Tuberculosis Control. World Health Organization. 1998;2.

88. CDC. Tuberculosis among foreign-born persons entering the United States. Recommendations of the Advisory Committee for Elimination of Tuberculosis. MMWR 1990; 39(RR-18): 1-21.

89. Greene VW, Dolberg OT, Alkan ML, Schlaeffer FC. Tuberculosis cases in the Negev 1978-1987; ethnicity, sex, and age. Public Health Reviews 1992; 20(1-2): 53-60.

90. Wen S, Goel V, Williams, JI. Utilization of health care services by immigrants and other ethnic/cultural groups in Ontario. Toronto. Institute for Clinical and Evaluative Sciences. 1999; 30 ICES Working paper.

91. Styblo K. The relationship between the risk of tuberculous infection and the risk of developing infectious tuberculosis. Bull. Int. Un. Tuberc. 1985; 60(3-4): 117-9.

92. Heywood N, Personal Communication. Citizenship and Immigration Canada, December 10, 1998.

93. Frost WH. The age selection of mortality from tuberculosis in successive decades. Am J Hygiene 1939; 30: 91-6.

94. Diwan VK, Thorson A. Sex, gender, and tuberculosis. Lancet 1999; 353: 1000-1.

95. Citizenship and Immigration Canada. Facts and Figures 1997. Ottawa. Minister of Public Works and Government Services Canada. 1998.

96. Cassels A, Heineman E, LeClerq S. Tuberculosis case-finding in Eastern Nepal. Tubercle 1982; 63: 173-185.

97. Krause RM. Dynamics of Emergence. J Infect Dis 1994; 170: 265-71.

98. World Bank. World development report 1993 - investing in health. Oxford University Press. 1993.

99. Vynnycky E, Fine PEM. The annual risk of infection with Mycobacterium tuberculosis in England and Wales since 1901. Int J Tuberc Lung Dis 1997; 1(5): 389-96.

100. Cox DR. Regression models and life-tables. Journal of the Royal Statistical Society 1972; 2: 187-201.

101. Farmer, P., Becerra, M., Kim, J.Y., et al. The Global Impact of Drug-Resistant Tuberculosis. Boston: Program in Infectious Disease and Social Change, Harvard Medical School , 1999.

102. Styblo, K., Dankova, D., Drapela, J., Galliova, J., Jezek, Z., Krivanek, J., Kubik, A., Langerova, M., and J. Radkovsky. Epidemiological and clinical study of tuberculosis in the district of Kolin, Czechoslovakia. Bull Wld Hlth Org 1967; 37: 819-874.

## Appendix A - WRO

Life Table Survival Estimates
**WRO = China**

| Interval [Lower, | Upper) | Number Failed | Number Censored | Effective Sample Size | Conditional Probability of Failure |
|---|---|---|---|---|---|
| 0 | 1 | 57 | 16153 | 91102.5 | 0.000626 |
| 1 | 2 | 38 | 15760 | 75089.0 | 0.000506 |
| 2 | 3 | 31 | 12089 | 61126.5 | 0.000507 |
| 3 | 4 | 25 | 13057 | 48522.5 | 0.000515 |
| 4 | 5 | 17 | 11353 | 36292.5 | 0.000468 |
| 5 | 6 | 13 | 12444 | 24377.0 | 0.000533 |
| 6 | 7 | 6 | 10433 | 12925.5 | 0.000464 |
| 7 | 8 | 2 | 7701 | 3852.5 | 0.000519 |

| Interval [Lower, | Upper) | Conditional Probability Standard Error | Survival | Failure | Survival Standard Error |
|---|---|---|---|---|---|
| 0 | 1 | 0.000083 | 1.0000 | 0 | 0 |
| 1 | 2 | 0.000082 | 0.9994 | 0.000626 | 0.000083 |
| 2 | 3 | 0.000091 | 0.9989 | 0.00113 | 0.000117 |
| 3 | 4 | 0.000103 | 0.9984 | 0.00164 | 0.000148 |
| 4 | 5 | 0.000114 | 0.9978 | 0.00215 | 0.00018 |
| 5 | 6 | 0.000148 | 0.9974 | 0.00262 | 0.000213 |
| 6 | 7 | 0.000189 | 0.9968 | 0.00315 | 0.000259 |
| 7 | 8 | 0.000367 | 0.9964 | 0.00361 | 0.00032 |

Evaluated at the Midpoint of the Interval

| Interval [Lower, | Upper) | PDF | PDF Standard Error | Hazard | Hazard Standard Error |
|---|---|---|---|---|---|
| 0 | 1 | 0.000626 | 0.000083 | 0.000626 | 0.000083 |
| 1 | 2 | 0.000506 | 0.000082 | 0.000506 | 0.000082 |
| 2 | 3 | 0.000507 | 0.000091 | 0.000507 | 0.000091 |
| 3 | 4 | 0.000514 | 0.000103 | 0.000515 | 0.000103 |
| 4 | 5 | 0.000467 | 0.000113 | 0.000469 | 0.000114 |
| 5 | 6 | 0.000532 | 0.000147 | 0.000533 | 0.000148 |
| 6 | 7 | 0.000463 | 0.000189 | 0.000464 | 0.00019 |
| 7 | 8 | 0.000517 | 0.000366 | 0.000519 | 0.000367 |

## Appendix A - WRO

Life Table Survival Estimate
**WRO = EME**

| Interval [Lower, | Upper) | Number Failed | Number Censored | Effective Sample Size | Conditional Probability of Failure |
|---|---|---|---|---|---|
| 0 | 1 | 15 | 7041 | 78660.5 | 0.000191 |
| 1 | 2 | 8 | 7894 | 71178.0 | 0.000112 |
| 2 | 3 | 1 | 7472 | 63487.0 | 0.000016 |
| 3 | 4 | 1 | 7985 | 55757.5 | 0.000018 |
| 4 | 5 | 0 | 10471 | 46528.5 | 0 |
| 5 | 6 | 3 | 11093 | 35746.5 | 0.000084 |
| 6 | 7 | 0 | 14091 | 23151.5 | 0 |
| 7 | 8 | 0 | 16106 | 8053.0 | 0 |

| Interval [Lower, | Upper) | Conditional Probability Standard Error | Survival | Failure | Survival Standard Error |
|---|---|---|---|---|---|
| 0 | 1 | 0.000049 | 1.0000 | 0 | 0 |
| 1 | 2 | 0.00004 | 0.9998 | 0.000191 | 0.000049 |
| 2 | 3 | 0.000016 | 0.9997 | 0.000303 | 0.000063 |
| 3 | 4 | 0.000018 | 0.9997 | 0.000319 | 0.000065 |
| 4 | 5 | 0 | 0.9997 | 0.000337 | 0.000068 |
| 5 | 6 | 0.000048 | 0.9997 | 0.000337 | 0.000068 |
| 6 | 7 | 0 | 0.9996 | 0.000421 | 0.000083 |
| 7 | 8 | 0 | 0.9996 | 0.000421 | 0.000083 |

Evaluated at the Midpoint of the Interval

| Interval [Lower, | Upper) | PDF | PDF Standard Error | Hazard | Hazard Standard Error |
|---|---|---|---|---|---|
| 0 | 1 | 0.000191 | 0.000049 | 0.000191 | 0.000049 |
| 1 | 2 | 0.000112 | 0.00004 | 0.000112 | 0.00004 |
| 2 | 3 | 0.000016 | 0.000016 | 0.000016 | 0.000016 |
| 3 | 4 | 0.000018 | 0.000018 | 0.000018 | 0.000018 |
| 4 | 5 | 0 | . | 0 | . |
| 5 | 6 | 0.000084 | 0.000048 | 0.000084 | 0.000048 |
| 6 | 7 | 0 | . | 0 | . |
| 7 | 8 | 0 | . | 0 | . |

## Appendix A - WRO

Life Table Survival Estimates
**WRO = FSEE**

| Interval [Lower, | Upper) | Number Failed | Number Censored | Effective Sample Size | Conditional Probability of Failure |
|---|---|---|---|---|---|
| 0 | 1 | 24 | 14489 | 122139.5 | 0.000196 |
| 1 | 2 | 18 | 14271 | 107735.5 | 0.000167 |
| 2 | 3 | 10 | 14894 | 93135.0 | 0.000107 |
| 3 | 4 | 6 | 14876 | 78240.0 | 0.000077 |
| 4 | 5 | 3 | 17710 | 61941.0 | 0.000048 |
| 5 | 6 | 3 | 16909 | 44628.5 | 0.000067 |
| 6 | 7 | 0 | 17928 | 27207.0 | 0 |
| 7 | 8 | 0 | 18243 | 9121.5 | 0 |

| Interval [Lower, | Upper) | Conditional Probability Standard Error | Survival | Failure | Survival Standard Error |
|---|---|---|---|---|---|
| 0 | 1 | 0.00004 | 1.0000 | 0 | 0 |
| 1 | 2 | 0.000039 | 0.9998 | 0.000196 | 0.00004 |
| 2 | 3 | 0.000034 | 0.9996 | 0.000364 | 0.000056 |
| 3 | 4 | 0.000031 | 0.9995 | 0.000471 | 0.000066 |
| 4 | 5 | 0.000028 | 0.9995 | 0.000548 | 0.000073 |
| 5 | 6 | 0.000039 | 0.9994 | 0.000596 | 0.000078 |
| 6 | 7 | 0 | 0.9993 | 0.000663 | 0.000087 |
| 7 | 8 | 0 | 0.9993 | 0.000663 | 0.000087 |

Evaluated at the Midpoint of the Interval

| Interval [Lower, | Upper) | PDF | PDF Standard Error | Hazard | Hazard Standard Error |
|---|---|---|---|---|---|
| 0 | 1 | 0.000196 | 0.00004 | 0.000197 | 0.00004 |
| 1 | 2 | 0.000167 | 0.000039 | 0.000167 | 0.000039 |
| 2 | 3 | 0.000107 | 0.000034 | 0.000107 | 0.000034 |
| 3 | 4 | 0.000077 | 0.000031 | 0.000077 | 0.000031 |
| 4 | 5 | 0.000048 | 0.000028 | 0.000048 | 0.000028 |
| 5 | 6 | 0.000067 | 0.000039 | 0.000067 | 0.000039 |
| 6 | 7 | 0 | . | 0 | . |
| 7 | 8 | 0 | . | 0 | . |

## Appendix A - WRO

Life Table Survival Estimates
**WRO = India**

| Interval [Lower, | Upper) | Number Failed | Number Censored | Effective Sample Size | Conditional Probability of Failure |
|---|---|---|---|---|---|
| 0 | 1 | 88 | 12964 | 72810.0 | 0.00121 |
| 1 | 2 | 39 | 13263 | 59608.5 | 0.000654 |
| 2 | 3 | 29 | 9808 | 48034.0 | 0.000604 |
| 3 | 4 | 15 | 9588 | 38307.0 | 0.000392 |
| 4 | 5 | 16 | 11207 | 27894.5 | 0.000574 |
| 5 | 6 | 12 | 7863 | 18343.5 | 0.000654 |
| 6 | 7 | 8 | 7473 | 10663.5 | 0.00075 |
| 7 | 8 | 3 | 6916 | 3461.0 | 0.000867 |

| Interval [Lower, | Upper) | Conditional Probability Standard Error | Survival | Failure | Survival Standard Error |
|---|---|---|---|---|---|
| 0 | 1 | 0.000129 | 1.0000 | 0 | 0 |
| 1 | 2 | 0.000105 | 0.9988 | 0.00121 | 0.000129 |
| 2 | 3 | 0.000112 | 0.9981 | 0.00186 | 0.000166 |
| 3 | 4 | 0.000101 | 0.9975 | 0.00246 | 0.0002 |
| 4 | 5 | 0.000143 | 0.9971 | 0.00286 | 0.000224 |
| 5 | 6 | 0.000189 | 0.9966 | 0.00343 | 0.000266 |
| 6 | 7 | 0.000265 | 0.9959 | 0.00408 | 0.000325 |
| 7 | 8 | 0.0005 | 0.9952 | 0.00483 | 0.000419 |

Evaluated at the Midpoint of the Interval

| Interval [Lower, | Upper) | PDF | PDF Standard Error | Hazard | Hazard Standard Error |
|---|---|---|---|---|---|
| 0 | 1 | 0.00121 | 0.000129 | 0.001209 | 0.000129 |
| 1 | 2 | 0.000653 | 0.000105 | 0.000654 | 0.000105 |
| 2 | 3 | 0.000603 | 0.000112 | 0.000604 | 0.000112 |
| 3 | 4 | 0.000391 | 0.000101 | 0.000392 | 0.000101 |
| 4 | 5 | 0.000572 | 0.000143 | 0.000574 | 0.000143 |
| 5 | 6 | 0.000652 | 0.000188 | 0.000654 | 0.000189 |
| 6 | 7 | 0.000747 | 0.000264 | 0.000751 | 0.000265 |
| 7 | 8 | 0.000863 | 0.000498 | 0.000867 | 0.000501 |

## Appendix A - WRO

Life Table Survival Estimates
**WRO = LatCar**

| Interval [Lower, | Upper) | Number Failed | Number Censored | Effective Sample Size | Conditional Probability of Failure |
|---|---|---|---|---|---|
| 0 | 1 | 34 | 11008 | 129618.0 | 0.000262 |
| 1 | 2 | 15 | 11913 | 118123.5 | 0.000127 |
| 2 | 3 | 15 | 14218 | 105043.0 | 0.000143 |
| 3 | 4 | 17 | 14487 | 90675.5 | 0.000187 |
| 4 | 5 | 6 | 22218 | 72306.0 | 0.000083 |
| 5 | 6 | 11 | 23868 | 49257.0 | 0.000223 |
| 6 | 7 | 2 | 20154 | 27235.0 | 0.000073 |
| 7 | 8 | 1 | 17155 | 8578.5 | 0.000117 |

| Interval [Lower, | Upper) | Conditional Probability Standard Error | Survival | Failure | Survival Standard Error |
|---|---|---|---|---|---|
| 0 | 1 | 0.000045 | 1.0000 | 0 | 0 |
| 1 | 2 | 0.000033 | 0.9997 | 0.000262 | 0.000045 |
| 2 | 3 | 0.000037 | 0.9996 | 0.000389 | 0.000056 |
| 3 | 4 | 0.000045 | 0.9995 | 0.000532 | 0.000067 |
| 4 | 5 | 0.000034 | 0.9993 | 0.000719 | 0.000081 |
| 5 | 6 | 0.000067 | 0.9992 | 0.000802 | 0.000088 |
| 6 | 7 | 0.000052 | 0.9990 | 0.00103 | 0.00011 |
| 7 | 8 | 0.000117 | 0.9989 | 0.00110 | 0.000122 |

Evaluated at the Midpoint of the Interval

| Interval [Lower, | Upper) | PDF | PDF Standard Error | Hazard | Hazard Standard Error |
|---|---|---|---|---|---|
| 0 | 1 | 0.000262 | 0.000045 | 0.000262 | 0.000045 |
| 1 | 2 | 0.000127 | 0.000033 | 0.000127 | 0.000033 |
| 2 | 3 | 0.000143 | 0.000037 | 0.000143 | 0.000037 |
| 3 | 4 | 0.000187 | 0.000045 | 0.000187 | 0.000045 |
| 4 | 5 | 0.000083 | 0.000034 | 0.000083 | 0.000034 |
| 5 | 6 | 0.000223 | 0.000067 | 0.000223 | 0.000067 |
| 6 | 7 | 0.000073 | 0.000052 | 0.000073 | 0.000052 |
| 7 | 8 | 0.000116 | 0.000116 | 0.000117 | 0.000117 |

## Appendix A - WRQ

Life Table Survival Estimates
**WRQ = MEC**

| Interval [Lower, | Upper) | Number Failed | Number Censored | Effective Sample Size | Conditional Probability of Failure |
|---|---|---|---|---|---|
| 0 | 1 | 54 | 26098 | 115460.0 | 0.000468 |
| 1 | 2 | 22 | 21176 | 91769.0 | 0.00024 |
| 2 | 3 | 14 | 15133 | 73592.5 | 0.00019 |
| 3 | 4 | 10 | 11822 | 60101.0 | 0.000166 |
| 4 | 5 | 9 | 13405 | 47477.5 | 0.00019 |
| 5 | 6 | 5 | 15289 | 33121.5 | 0.000151 |
| 6 | 7 | 4 | 13990 | 18477.0 | 0.000216 |
| 7 | 8 | 1 | 11477 | 5739.5 | 0.000174 |

| Interval [Lower, | Upper) | Conditional Probability Standard Error | Survival | Failure | Survival Standard Error |
|---|---|---|---|---|---|
| 0 | 1 | 0.000064 | 1.0000 | 0 | 0 |
| 1 | 2 | 0.000051 | 0.9995 | 0.000468 | 0.000064 |
| 2 | 3 | 0.000051 | 0.9993 | 0.000707 | 0.000082 |
| 3 | 4 | 0.000053 | 0.9991 | 0.000897 | 0.000096 |
| 4 | 5 | 0.000063 | 0.9989 | 0.00106 | 0.00011 |
| 5 | 6 | 0.000068 | 0.9987 | 0.00125 | 0.000126 |
| 6 | 7 | 0.000108 | 0.9986 | 0.00140 | 0.000143 |
| 7 | 8 | 0.000174 | 0.9984 | 0.00162 | 0.000179 |

Evaluated at the Midpoint of the Interval

| Interval [Lower, | Upper) | PDF | PDF Standard Error | Hazard | Hazard Standard Error |
|---|---|---|---|---|---|
| 0 | 1 | 0.000468 | 0.000064 | 0.000468 | 0.000064 |
| 1 | 2 | 0.00024 | 0.000051 | 0.00024 | 0.000051 |
| 2 | 3 | 0.00019 | 0.000051 | 0.00019 | 0.000051 |
| 3 | 4 | 0.000166 | 0.000053 | 0.000166 | 0.000053 |
| 4 | 5 | 0.000189 | 0.000063 | 0.00019 | 0.000063 |
| 5 | 6 | 0.000151 | 0.000067 | 0.000151 | 0.000068 |
| 6 | 7 | 0.000216 | 0.000108 | 0.000217 | 0.000108 |
| 7 | 8 | 0.000174 | 0.000174 | 0.000174 | 0.000174 |

## Appendix A - WRO

Life Table Survival Estimates
**WRO = OthAsia**

| Interval [Lower, | Upper) | Number Failed | Number Censored | Effective Sample Size | Conditional Probability of Failure |
|---|---|---|---|---|---|
| 0 | 1 | 234 | 24183 | 254808.5 | 0.000918 |
| 1 | 2 | 162 | 29217 | 227874.5 | 0.000711 |
| 2 | 3 | 116 | 35701 | 195253.5 | 0.000594 |
| 3 | 4 | 89 | 39541 | 157516.5 | 0.000565 |
| 4 | 5 | 65 | 39038 | 118138.0 | 0.00055 |
| 5 | 6 | 37 | 40024 | 78542.0 | 0.000471 |
| 6 | 7 | 25 | 28322 | 44332.0 | 0.000564 |
| 7 | 8 | 18 | 30128 | 15082.0 | 0.00119 |

| Interval [Lower, | Upper) | Conditional Probability Standard Error | Survival | Failure | Survival Standard Error |
|---|---|---|---|---|---|
| 0 | 1 | 0.00006 | 1.0000 | 0 | 0 |
| 1 | 2 | 0.000056 | 0.9991 | 0.000918 | 0.00006 |
| 2 | 3 | 0.000055 | 0.9984 | 0.00163 | 0.000082 |
| 3 | 4 | 0.00006 | 0.9978 | 0.00222 | 0.000099 |
| 4 | 5 | 0.000068 | 0.9972 | 0.00279 | 0.000115 |
| 5 | 6 | 0.000077 | 0.9967 | 0.00333 | 0.000134 |
| 6 | 7 | 0.000113 | 0.9962 | 0.00380 | 0.000154 |
| 7 | 8 | 0.000281 | 0.9956 | 0.00437 | 0.000191 |

Evaluated at the Midpoint of the Interval

| Interval [Lower, | Upper) | PDF | PDF Standard Error | Hazard | Hazard Standard Error |
|---|---|---|---|---|---|
| 0 | 1 | 0.000918 | 0.00006 | 0.000919 | 0.00006 |
| 1 | 2 | 0.00071 | 0.000056 | 0.000711 | 0.000056 |
| 2 | 3 | 0.000593 | 0.000055 | 0.000594 | 0.000055 |
| 3 | 4 | 0.000564 | 0.00006 | 0.000565 | 0.00006 |
| 4 | 5 | 0.000549 | 0.000068 | 0.00055 | 0.000068 |
| 5 | 6 | 0.00047 | 0.000077 | 0.000471 | 0.000077 |
| 6 | 7 | 0.000562 | 0.000112 | 0.000564 | 0.000113 |
| 7 | 8 | 0.00119 | 0.00028 | 0.001194 | 0.000281 |

## Appendix A - WRO

Life Table Survival Estimates
**WRO = SSA**

| Interval [Lower, | Upper) | Number Failed | Number Censored | Effective Sample Size | Conditional Probability of Failure |
|---|---|---|---|---|---|
| 0 | 1 | 232 | 5379 | 54191.5 | 0.00428 |
| 1 | 2 | 97 | 5781 | 48379.5 | 0.00200 |
| 2 | 3 | 67 | 6229 | 42277.5 | 0.00158 |
| 3 | 4 | 54 | 5860 | 36166.0 | 0.00149 |
| 4 | 5 | 43 | 9021 | 28671.5 | 0.00150 |
| 5 | 6 | 38 | 11179 | 18528.5 | 0.00205 |
| 6 | 7 | 16 | 7160 | 9321.0 | 0.00172 |
| 7 | 8 | 9 | 5716 | 2867.0 | 0.00314 |

| Interval [Lower, | Upper) | Conditional Probability Standard Error | Survival | Failure | Survival Standard Error |
|---|---|---|---|---|---|
| 0 | 1 | 0.00028 | 1.0000 | 0 | 0 |
| 1 | 2 | 0.000203 | 0.9957 | 0.00428 | 0.00028 |
| 2 | 3 | 0.000193 | 0.9937 | 0.00628 | 0.000345 |
| 3 | 4 | 0.000203 | 0.9921 | 0.00785 | 0.000395 |
| 4 | 5 | 0.000229 | 0.9907 | 0.00933 | 0.000443 |
| 5 | 6 | 0.000332 | 0.9892 | 0.0108 | 0.000497 |
| 6 | 7 | 0.000429 | 0.9872 | 0.0128 | 0.000595 |
| 7 | 8 | 0.00104 | 0.9855 | 0.0145 | 0.000729 |

Evaluated at the Midpoint of the Interval

| Interval [Lower, | Upper) | PDF | PDF Standard Error | Hazard | Hazard Standard Error |
|---|---|---|---|---|---|
| 0 | 1 | 0.00428 | 0.00028 | 0.00429 | 0.000282 |
| 1 | 2 | 0.00200 | 0.000203 | 0.002007 | 0.000204 |
| 2 | 3 | 0.00157 | 0.000192 | 0.001586 | 0.000194 |
| 3 | 4 | 0.00148 | 0.000201 | 0.001494 | 0.000203 |
| 4 | 5 | 0.00149 | 0.000226 | 0.001501 | 0.000229 |
| 5 | 6 | 0.00203 | 0.000329 | 0.002053 | 0.000333 |
| 6 | 7 | 0.00169 | 0.000423 | 0.001718 | 0.00043 |
| 7 | 8 | 0.00309 | 0.00103 | 0.003144 | 0.001048 |

## Appendix A - WRO

Summary of the Number of Censored and Uncensored Values

| WRO | Total | Failed | Censored | %Censored |
|---|---|---|---|---|
| China | 99179 | 189 | 98990 | 99.8094 |
| EME | 82181 | 28 | 82153 | 99.9659 |
| FSEE | 129384 | 64 | 129320 | 99.9505 |
| India | 79292 | 210 | 79082 | 99.7352 |
| LatCar | 135122 | 101 | 135021 | 99.9253 |
| MEC | 128509 | 119 | 128390 | 99.9074 |
| OthAsia | 266900 | 746 | 266154 | 99.7205 |
| SSA | 56881 | 556 | 56325 | 99.0225 |
| Total | 977448 | 2013 | 975435 | 99.7941 |

Testing Homogeneity of Survival Curves over Strata
Time Variable TIME

Rank Statistics

| WRO | Log-Rank | Wilcoxon |
|---|---|---|
| China | 0.48 | -7058915 |
| EME | -157.95 | -1.143E8 |
| FSEE | -209.87 | -1.562E8 |
| India | 60.47 | 46674199 |
| LatCar | -196.76 | -1.503E8 |
| MEC | -120.05 | -9.276E7 |
| OthAsia | 188.82 | 1.3282E8 |
| SSA | 434.86 | 3.4108E8 |

Test of Equality over Strata

| Test | Chi-Square | DF | Pr > Chi-Square |
|---|---|---|---|
| Log-Rank | 2136.8298 | 7 | 0.0001 |
| Wilcoxon | 2005.0900 | 7 | 0.0001 |
| -2Log(LR) | 1615.9654 | 7 | 0.0001 |

## Appendix A - AGE

Life Table Survival Estimates
**AGEGR = 0-15**

| Interval [Lower, | Upper) | Number Failed | Number Censored | Effective Sample Size | Conditional Probability of Failure |
|---|---|---|---|---|---|
| 0 | 1 | 97 | 28699 | 193313.5 | 0.000502 |
| 1 | 2 | 20 | 28302 | 164716.0 | 0.000121 |
| 2 | 3 | 22 | 26051 | 137519.5 | 0.00016 |
| 3 | 4 | 10 | 24457 | 112243.5 | 0.000089 |
| 4 | 5 | 9 | 27043 | 86483.5 | 0.000104 |
| 5 | 6 | 7 | 25984 | 59961.0 | 0.000117 |
| 6 | 7 | 8 | 21952 | 35986.0 | 0.000222 |
| 7 | 8 | 4 | 24998 | 12503.0 | 0.00032 |

| Interval [Lower, | Upper) | Conditional Probability Standard Error | Survival | Failure | Survival Standard Error |
|---|---|---|---|---|---|
| 0 | 1 | 0.000051 | 1.0000 | 0 | 0 |
| 1 | 2 | 0.000027 | 0.9995 | 0.000502 | 0.000051 |
| 2 | 3 | 0.000034 | 0.9994 | 0.000623 | 0.000058 |
| 3 | 4 | 0.000028 | 0.9992 | 0.000783 | 0.000067 |
| 4 | 5 | 0.000035 | 0.9991 | 0.000872 | 0.000073 |
| 5 | 6 | 0.000044 | 0.9990 | 0.000976 | 0.000081 |
| 6 | 7 | 0.000079 | 0.9989 | 0.00109 | 0.000092 |
| 7 | 8 | 0.00016 | 0.9987 | 0.00131 | 0.000121 |

Evaluated at the Midpoint of the Interval

| Interval [Lower, | Upper) | PDF | PDF Standard Error | Hazard | Hazard Standard Error |
|---|---|---|---|---|---|
| 0 | 1 | 0.0000502 | 0.000051 | 0.000502 | 0.000051 |
| 1 | 2 | 0.000121 | 0.000027 | 0.000121 | 0.000027 |
| 2 | 3 | 0.00016 | 0.000034 | 0.00016 | 0.000034 |
| 3 | 4 | 0.000089 | 0.000028 | 0.000089 | 0.000028 |
| 4 | 5 | 0.000104 | 0.000035 | 0.000104 | 0.000035 |
| 5 | 6 | 0.000117 | 0.000044 | 0.000117 | 0.000044 |
| 6 | 7 | 0.000222 | 0.000079 | 0.000222 | 0.000079 |
| 7 | 8 | 0.00032 | 0.00016 | 0.00032 | 0.00016 |

## Appendix A - AGE

Life Table Survival Estimates
**AGEGR = 16-30**

| Interval [Lower, | Upper) | Number Failed | Number Censored | Effective Sample Size | Conditional Probability of Failure |
|---|---|---|---|---|---|
| 0 | 1 | 304 | 34007 | 303211.5 | 0.00100 |
| 1 | 2 | 192 | 35012 | 268398.0 | 0.000715 |
| 2 | 3 | 127 | 34899 | 233250.5 | 0.000544 |
| 3 | 4 | 101 | 37096 | 197126.0 | 0.000512 |
| 4 | 5 | 87 | 46590 | 155182.0 | 0.000561 |
| 5 | 6 | 63 | 49325 | 107137.5 | 0.000588 |
| 6 | 7 | 31 | 43597 | 60613.5 | 0.000511 |
| 7 | 8 | 18 | 38766 | 19401.0 | 0.000928 |

| Interval [Lower, | Upper) | Conditional Probability Standard Error | Survival | Failure | Survival Standard Error |
|---|---|---|---|---|---|
| 0 | 1 | 0.000057 | 1.0000 | 0 | 0 |
| 1 | 2 | 0.000052 | 0.9990 | 0.00100 | 0.000057 |
| 2 | 3 | 0.000048 | 0.9983 | 0.00172 | 0.000077 |
| 3 | 4 | 0.000051 | 0.9977 | 0.00226 | 0.000091 |
| 4 | 5 | 0.00006 | 0.9972 | 0.00277 | 0.000104 |
| 5 | 6 | 0.000074 | 0.9967 | 0.00333 | 0.00012 |
| 6 | 7 | 0.000092 | 0.9961 | 0.00392 | 0.000141 |
| 7 | 8 | 0.000219 | 0.9956 | 0.00443 | 0.000168 |

Evaluated at the Midpoint of the Interval

| Interval [Lower, | Upper) | PDF PDF | PDF Standard Error | Hazard Hazard | Hazard Standard Error |
|---|---|---|---|---|---|
| 0 | 1 | 0.00100 | 0.000057 | 0.001003 | 0.000058 |
| 1 | 2 | 0.000715 | 0.000052 | 0.000716 | 0.000052 |
| 2 | 3 | 0.000544 | 0.000048 | 0.000545 | 0.000048 |
| 3 | 4 | 0.000511 | 0.000051 | 0.000512 | 0.000051 |
| 4 | 5 | 0.000559 | 0.00006 | 0.000561 | 0.00006 |
| 5 | 6 | 0.000586 | 0.000074 | 0.000588 | 0.000074 |
| 6 | 7 | 0.000509 | 0.000091 | 0.000512 | 0.000092 |
| 7 | 8 | 0.000924 | 0.000218 | 0.000928 | 0.000219 |

## Appendix A - AGE

Life Table Survival Estimates
**AGEGR = 31-45**

| Interval [Lower, | Upper) | Number Failed | Number Censored | Effective Sample Size | Conditional Probability of Failure |
|---|---|---|---|---|---|
| 0 | 1 | 153 | 39224 | 270623.0 | 0.000565 |
| 1 | 2 | 89 | 38638 | 231539.0 | 0.000384 |
| 2 | 3 | 76 | 35130 | 194566.0 | 0.000391 |
| 3 | 4 | 54 | 32783 | 160533.5 | 0.000336 |
| 4 | 5 | 34 | 37309 | 125433.5 | 0.000271 |
| 5 | 6 | 37 | 40138 | 86676.0 | 0.000427 |
| 6 | 7 | 9 | 33522 | 49809.0 | 0.000181 |
| 7 | 8 | 5 | 33034 | 16522.0 | 0.000303 |

| Interval [Lower, | Upper) | Conditional Probability Standard Error | Survival | Failure | Survival Standard Error |
|---|---|---|---|---|---|
| 0 | 1 | 0.000046 | 1.0000 | 0 | 0 |
| 2 | 3 | 0.000045 | 0.9991 | 0.00095 | 0.000061 |
| 3 | 4 | 0.000046 | 0.9987 | 0.00134 | 0.000076 |
| 4 | 5 | 0.000046 | 0.9983 | 0.00168 | 0.000088 |
| 5 | 6 | 0.00007 | 0.9981 | 0.00195 | 0.0001 |
| 6 | 7 | 0.00006 | 0.9976 | 0.00237 | 0.000122 |
| 7 | 8 | 0.000135 | 0.9974 | 0.00255 | 0.000136 |

Evaluated at the Midpoint of the Interval

| Interval [Lower, | Upper) | PDF | PDF Standard Error | Hazard | Hazard Standard Error |
|---|---|---|---|---|---|
| 0 | 1 | 0.000565 | 0.000046 | 0.000566 | 0.000046 |
| 1 | 2 | 0.000384 | 0.000041 | 0.000384 | 0.000041 |
| 2 | 3 | 0.00039 | 0.000045 | 0.000391 | 0.000045 |
| 3 | 4 | 0.000336 | 0.000046 | 0.000336 | 0.000046 |
| 4 | 5 | 0.000271 | 0.000046 | 0.000271 | 0.000046 |
| 5 | 6 | 0.000426 | 0.00007 | 0.000427 | 0.00007 |
| 6 | 7 | 0.00018 | 0.00006 | 0.000181 | 0.00006 |
| 7 | 8 | 0.000302 | 0.000135 | 0.000303 | 0.000135 |

## Appendix A - AGE

Life Table Survival Estimates
**AGEGR = 46-65**

| Interval [Lower, | Upper) | Number Failed | Number Censored | Effective Sample Size | Conditional Probability of Failure |
|---|---|---|---|---|---|
| 0 | 1 | 103 | 12435 | 113291.5 | 0.000909 |
| 1 | 2 | 65 | 13645 | 100148.5 | 0.000649 |
| 2 | 3 | 40 | 15084 | 85719.0 | 0.000467 |
| 3 | 4 | 33 | 17385 | 69444.5 | 0.000475 |
| 4 | 5 | 20 | 17583 | 51927.5 | 0.000385 |
| 5 | 6 | 7 | 16775 | 34728.5 | 0.000202 |
| 6 | 7 | 8 | 14512 | 19078.0 | 0.000419 |
| 7 | 8 | 5 | 11809 | 5909.5 | 0.000846 |

| Interval [Lower, | Upper) | Conditional Probability Standard Error | Survival | Failure | Survival Standard Error |
|---|---|---|---|---|---|
| 0 | 1 | 0.00009 | 1.0000 | 0 | 0 |
| 1 | 2 | 0.00008 | 0.9991 | 0.000909 | 0.00009 |
| 2 | 3 | 0.000074 | 0.9984 | 0.00156 | 0.00012 |
| 3 | 4 | 0.000083 | 0.9980 | 0.00202 | 0.000141 |
| 4 | 5 | 0.000086 | 0.9975 | 0.00250 | 0.000163 |
| 5 | 6 | 0.000076 | 0.9971 | 0.00288 | 0.000184 |
| 6 | 7 | 0.000148 | 0.9969 | 0.00308 | 0.000199 |
| 7 | 8 | 0.000378 | 0.9965 | 0.00350 | 0.000248 |

Evaluated at the Midpoint of the Interval

| Interval [Lower, | Upper) | PDF | PDF Standard Error | Hazard | Hazard Standard Error |
|---|---|---|---|---|---|
| 0 | 1 | 0.000909 | 0.00009 | 0.00091 | 0.00009 |
| 1 | 2 | 0.000648 | 0.00008 | 0.000649 | 0.000081 |
| 2 | 3 | 0.000466 | 0.000074 | 0.000467 | 0.000074 |
| 3 | 4 | 0.000474 | 0.000083 | 0.000475 | 0.000083 |
| 4 | 5 | 0.000384 | 0.000086 | 0.000385 | 0.000086 |
| 5 | 6 | 0.000201 | 0.000076 | 0.000202 | 0.000076 |
| 6 | 7 | 0.000418 | 0.000148 | 0.000419 | 0.000148 |
| 7 | 8 | 0.000843 | 0.000377 | 0.000846 | 0.000379 |

## Appendix A - AGE

Life Table Survival Estimates
**AGEGR = 66+**

| Interval [Lower, | Upper) | Number Failed | Number Censored | Effective Sample Size | Conditional Probability of Failure |
|---|---|---|---|---|---|
| 0 | 1 | 81 | 2950 | 38351.0 | 0.00211 |
| 1 | 2 | 33 | 3678 | 34956.0 | 0.000944 |
| 2 | 3 | 18 | 4380 | 30894.0 | 0.000583 |
| 3 | 4 | 19 | 5495 | 25938.5 | 0.000733 |
| 4 | 5 | 9 | 5898 | 20223.0 | 0.000445 |
| 5 | 6 | 8 | 6447 | 14041.5 | 0.00057 |
| 6 | 7 | 5 | 5968 | 7826.0 | 0.000639 |
| 7 | 8 | 2 | 4835 | 2419.5 | 0.000827 |

| Interval [Lower, | Upper) | Conditional Probability Standard Error | Survival | Failure | Survival Standard Error |
|---|---|---|---|---|---|
| 0 | 1 | 0.000234 | 1.0000 | 0 | 0 |
| 1 | 2 | 0.000164 | 0.9979 | 0.00211 | 0.000234 |
| 2 | 3 | 0.000137 | 0.9969 | 0.00305 | 0.000286 |
| 3 | 4 | 0.000168 | 0.9964 | 0.00363 | 0.000317 |
| 4 | 5 | 0.000148 | 0.9956 | 0.00436 | 0.000358 |
| 5 | 6 | 0.000201 | 0.9952 | 0.00481 | 0.000387 |
| 6 | 7 | 0.000286 | 0.9946 | 0.00537 | 0.000436 |
| 7 | 8 | 0.000584 | 0.9940 | 0.00601 | 0.00052 |

Evaluated at the Midpoint of the Interval

| Interval [Lower, | Upper) | PDF | PDF Standard Error | Hazard | Hazard Standard Error |
|---|---|---|---|---|---|
| 0 | 1 | 0.00211 | 0.000234 | 0.002114 | 0.000235 |
| 1 | 2 | 0.000942 | 0.000164 | 0.000944 | 0.000164 |
| 2 | 3 | 0.000581 | 0.000137 | 0.000583 | 0.000137 |
| 3 | 4 | 0.00073 | 0.000167 | 0.000733 | 0.000168 |
| 4 | 5 | 0.000443 | 0.000148 | 0.000445 | 0.000148 |
| 5 | 6 | 0.000567 | 0.0002 | 0.00057 | 0.000201 |
| 6 | 7 | 0.000635 | 0.000284 | 0.000639 | 0.000286 |
| 7 | 8 | 0.000822 | 0.000581 | 0.000827 | 0.000585 |

## Appendix A - AGE

The LIFETEST Procedure

Summary of the Number of Censored and Uncensored Values

| AGEGR | Total | Failed | Censored | %Censored |
|---|---|---|---|---|
| 0-15 | 207663 | 177 | 207486 | 99.9148 |
| 16-30 | 320215 | 923 | 319292 | 99.7118 |
| 31-45 | 290235 | 457 | 289778 | 99.8425 |
| 46-65 | 119509 | 281 | 119228 | 99.7649 |
| 66+ | 39826 | 175 | 39651 | 99.5606 |
| Total | 977448 | 2013 | 975435 | 99.7941 |

The LIFETEST Procedure

Testing Homogeneity of Survival Curves over Strata
Time Variable TIME

Rank Statistics

| AGEGR | Log-Rank | Wilcoxon |
|---|---|---|
| 0-15 | -238.83 | -1.788E8 |
| 16-30 | 245.10 | 1.7692E8 |
| 31-45 | -128.87 | -1.042E8 |
| 46-65 | 34.61 | 30729432 |
| 66+ | 88.00 | 75352887 |

Test of Equality over Strata

| Test | Chi-Square | DF | Pr > Chi-Square |
|---|---|---|---|
| Log-Rank | 348.3815 | 4 | 0.0001 |
| Wilcoxon | 328.0297 | 4 | 0.0001 |
| -2Log(LR) | 339.5813 | 4 | 0.0001 |

## Appendix A - SEX

Life Table Survival Estimates
**SEX = F**

| Interval [Lower, | Upper) | Number Failed | Number Censored | Effective Sample Size | Conditional Probability of Failure |
|---|---|---|---|---|---|
| 0 | 1 | 369 | 59027 | 472345.5 | 0.000781 |
| 1 | 2 | 192 | 60686 | 412120.0 | 0.000466 |
| 2 | 3 | 141 | 60372 | 351399.0 | 0.000401 |
| 3 | 4 | 114 | 62484 | 289830.0 | 0.000393 |
| 4 | 5 | 83 | 72373 | 222287.5 | 0.000373 |
| 5 | 6 | 64 | 69260 | 151388.0 | 0.000423 |
| 6 | 7 | 35 | 60207 | 86590.5 | 0.000404 |
| 7 | 8 | 15 | 56437 | 28233.5 | 0.000531 |

| Interval [Lower, | Upper) | Conditional Probability Standard Error | Survival | Failure | Survival Standard Error |
|---|---|---|---|---|---|
| 0 | 1 | 0.000041 | 1.0000 | 0 | 0 |
| 1 | 2 | 0.000034 | 0.9992 | 0.000781 | 0.000041 |
| 2 | 3 | 0.000034 | 0.9988 | 0.00125 | 0.000053 |
| 3 | 4 | 0.000037 | 0.9984 | 0.00165 | 0.000063 |
| 5 | 6 | 0.000053 | 0.9976 | 0.00241 | 0.000083 |
| 6 | 7 | 0.000068 | 0.9972 | 0.00283 | 0.000099 |
| 7 | 8 | 0.000137 | 0.9968 | 0.00324 | 0.00012 |

Evaluated at the Midpoint of the Interval

| Interval [Lower, | Upper) | PDF | PDF Standard Error | Hazard | Hazard Standard Error |
|---|---|---|---|---|---|
| 0 | 1 | 0.000781 | 0.000041 | 0.000782 | 0.000041 |
| 1 | 2 | 0.000466 | 0.000034 | 0.000466 | 0.000034 |
| 2 | 3 | 0.000401 | 0.000034 | 0.000401 | 0.000034 |
| 3 | 4 | 0.000393 | 0.000037 | 0.000393 | 0.000037 |
| 4 | 5 | 0.000373 | 0.000041 | 0.000373 | 0.000041 |
| 5 | 6 | 0.000422 | 0.000053 | 0.000423 | 0.000053 |
| 6 | 7 | 0.000403 | 0.000068 | 0.000404 | 0.000068 |
| 7 | 8 | 0.00053 | 0.000137 | 0.000531 | 0.000137 |

## Appendix A - SEX

Life Table Survival Estimates

**SEX = M**

| Interval [Lower, | Upper) | Number Failed | Number Censored | Effective Sample Size | Conditional Probability of Failure |
|---|---|---|---|---|---|
| 0 | 1 | 369 | 58288 | 446445.0 | 0.000827 |
| 1 | 2 | 207 | 58589 | 387637.5 | 0.000534 |
| 2 | 3 | 142 | 55172 | 330550.0 | 0.00043 |
| 3 | 4 | 103 | 54732 | 275456.0 | 0.000374 |
| 4 | 5 | 76 | 62050 | 216962.0 | 0.00035 |
| 5 | 6 | 58 | 69409 | 151156.5 | 0.000384 |
| 6 | 7 | 26 | 59344 | 86722.0 | 0.0003 |
| 7 | 8 | 19 | 57005 | 28521.5 | 0.000666 |

| Interval [Lower, | Upper) | Conditional Probability Standard Error | Survival | Failure | Survival Standard Error |
|---|---|---|---|---|---|
| 0 | 1 | 0.000043 | 1.0000 | 0 | 0 |
| 1 | 2 | 0.000037 | 0.9992 | 0.000827 | 0.000043 |
| 2 | 3 | 0.000036 | 0.9986 | 0.00136 | 0.000057 |
| 3 | 4 | 0.000037 | 0.9982 | 0.00179 | 0.000067 |
| 4 | 5 | 0.00004 | 0.9978 | 0.00216 | 0.000077 |
| 5 | 6 | 0.00005 | 0.9975 | 0.00251 | 0.000086 |
| 6 | 7 | 0.000059 | 0.9971 | 0.00289 | 0.0001 |
| 7 | 8 | 0.000153 | 0.9968 | 0.00319 | 0.000116 |

Evaluated at the Midpoint of the Interval

| Interval [Lower, | Upper) | PDF | PDF Standard Error | Hazard | Hazard Standard Error |
|---|---|---|---|---|---|
| 0 | 1 | 0.000827 | 0.000043 | 0.000827 | 0.000043 |
| 1 | 2 | 0.000534 | 0.000037 | 0.000534 | 0.000037 |
| 2 | 3 | 0.000429 | 0.000036 | 0.00043 | 0.000036 |
| 3 | 4 | 0.000373 | 0.000037 | 0.000374 | 0.000037 |
| 4 | 5 | 0.00035 | 0.00004 | 0.00035 | 0.00004 |
| 5 | 6 | 0.000383 | 0.00005 | 0.000384 | 0.00005 |
| 6 | 7 | 0.000299 | 0.000059 | 0.0003 | 0.000059 |
| 7 | 8 | 0.000664 | 0.000152 | 0.000666 | 0.000153 |

## Appendix A - SEX

Summary of the Number of Censored and Uncensored Values

| SEX | Total | Failed | Censored | %Censored |
|---|---|---|---|---|
| F | 501859 | 1013 | 500846 | 99.7982 |
| M | 475589 | 1000 | 474589 | 99.7897 |
| Total | 977448 | 2013 | 975435 | 99.7941 |

The LIFETEST Procedure

Testing Homogeneity of Survival Curves over Strata
Time Variable TIME

Rank Statistics

| SEX | Log-Rank | Wilcoxon |
|---|---|---|
| F | -18.629 | -2.06E7 |
| M | 18.629 | 20601654 |

Test of Equality over Strata

| Test | Chi-Square | DF | Pr > Chi-Square |
|---|---|---|---|
| Log-Rank | 0.6904 | 1 | 0.4060 |
| Wilcoxon | 1.3072 | 1 | 0.2529 |
| -2Log(LR) | 0.5543 | 1 | 0.4566 |

## Appendix B

**Table I. Risk Ratios for WRO - Year of Landing combinations. Model uses SSA and 1997 as the baseline. * terms significant at $p \leq 0.05$.**

| *Year* | *WRO* | *Parameter Estimate* | *Std Error* | *Chi-Square P-value* | *RR* |
|---|---|---|---|---|---|
| 1990 | China | -1.1964 | 0.6185 | 0.0531 | 0.30 |
| 1991 | China | -1.1329 | 0.6078 | 0.0623 | 0.32 |
| 1992 | China | -1.5493 | 0.6070 | 0.0107 | 0.21* |
| 1993 | China | -1.4781 | 0.6226 | 0.0176 | 0.23* |
| 1994 | China | -1.9283 | 0.6395 | 0.0026 | 0.145 |
| 1995 | China | -1.5013 | 0.6483 | 0.0206 | 0.22 |
| 1996 | China | -2.1945 | 0.7342 | 0.0028 | 0.11 |
| 1997 | China | -2.7658 | 0.8210 | 0.0008 | 0.06 |
| 1990 | EME | -2.4981 | 0.7537 | 0.0009 | 0.08 |
| 1991 | EME | -3.2779 | 0.8520 | 0.0001 | 0.04* |
| 1992 | EME | -2.2233 | 0.7761 | 0.0042 | 0.11 |
| 1993 | EME | -2.4346 | 0.8247 | 0.0032 | 0.09 |
| 1994 | EME | -3.0402 | 0.9975 | 0.0023 | 0.05* |
| 1995 | EME | -21.4001 | 9055.142 | 0.9981 | 0.00 |
| 1996 | EME | -2.9957 | 1.2294 | 0.0148 | 0.05* |
| 1997 | EME | -21.3024 | 15939.45 | 0.9989 | 0.00 |
| 1990 | FSEE | -2.4246 | 0.7362 | 0.0010 | 0.09* |
| 1991 | FSEE | -2.4119 | 0.7345 | 0.0010 | 0.10* |
| 1992 | FSEE | -1.7933 | 0.7161 | 0.0123 | 0.17* |
| 1993 | FSEE | -1.4936 | 0.6945 | 0.0315 | 0.22* |
| 1994 | FSEE | -2.3298 | 0.7694 | 0.0025 | 0.10* |
| 1995 | FSEE | -2.3515 | 0.8299 | 0.0046 | 0.095* |
| 1996 | FSEE | -1.9246 | 0.8127 | 0.0179 | 0.15* |
| 1997 | FSEE | -1.6090 | 0.7456 | 0.0309 | 0.20* |
| 1990 | India | -1.2176 | 0.6444 | 0.0588 | 0.30 |
| 1991 | India | -1.5766 | 0.6595 | 0.0168 | 0.21* |
| 1992 | India | -1.1664 | 0.6470 | 0.0714 | 0.31 |
| 1993 | India | -0.9841 | 0.6418 | 0.1252 | 0.37 |
| 1994 | India | -1.3875 | 0.6656 | 0.0371 | 0.25* |
| 1995 | India | -1.3670 | 0.6990 | 0.0505 | 0.25 |
| 1996 | India | -1.4183 | 0.7242 | 0.0502 | 0.24 |
| 1997 | India | -1.1398 | 0.5499 | 0.0382 | 0.32* |
| 1990 | LatCar | -2.2528 | 0.7058 | 0.0014 | 0.105* |
| 1991 | LatCar | -2.7536 | 0.7248 | 0.0001 | 0.06* |
| 1992 | LatCar | -2.2639 | 0.7114 | 0.0015 | 0.10* |
| 1993 | LatCar | -2.2334 | 0.7210 | 0.0020 | 0.11* |

| | | | | | |
|---|---|---|---|---|---|
| 1994 | LatCar | -3.0796 | 0.8151 | 0.0002 | 0.05* |
| 1995 | LatCar | -2.6842 | 0.8408 | 0.0014 | 0.07* |
| 1996 | LatCar | -2.1081 | 0.8440 | 0.0125 | 0.12* |
| 1997 | LatCar | -0.8362 | 0.6246 | 0.1806 | 0.43 |
| 1990 | MEC | -1.5913 | 0.6802 | 0.0193 | 0.20* |
| 1991 | MEC | -1.7498 | 0.6860 | 0.0108 | 0.17* |
| 1992 | MEC | -1.3234 | 0.6743 | 0.0497 | 0.27* |
| 1993 | MEC | -1.5599 | 0.6945 | 0.0247 | 0.21* |
| 1994 | MEC | -1.3042 | 0.6914 | 0.0593 | 0.27 |
| 1995 | MEC | -1.1075 | 0.7152 | 0.1215 | 0.33 |
| 1996 | MEC | -2.0159 | 0.7935 | 0.0111 | 0.13* |
| 1997 | MEC | -1.8784 | 0.6868 | 0.0062 | 0.15* |
| 1990 | OthAsia | -0.8748 | 0.6093 | 0.1510 | 0.42 |
| 1991 | OthAsia | -1.1826 | 0.6101 | 0.0526 | 0.31 |
| 1992 | OthAsia | -0.9500 | 0.6016 | 0.1143 | 0.39 |
| 1993 | OthAsia | -0.6521 | 0.5976 | 0.2752 | 0.52 |
| 1994 | OthAsia | -1.3627 | 0.6215 | 0.0283 | 0.26* |
| 1995 | OthAsia | -1.1481 | 0.6394 | 0.0726 | 0.32 |
| 1996 | OthAsia | -1.1718 | 0.6747 | 0.0824 | 0.31 |
| 1997 | OthAsia | -1.1338 | 0.4909 | 0.0209 | 0.32* |
| 1990 | SSA | -0.3576 | 0.5523 | 0.5173 | 0.70 |
| 1991 | SSA | -0.3161 | 0.5517 | 0.5667 | 0.73 |
| 1992 | SSA | -0.7587 | 0.5503 | 0.1680 | 0.47 |
| 1993 | SSA | -0.6124 | 0.5494 | 0.2649 | 0.54 |
| 1994 | SSA | -0.5007 | 0.5707 | 0.3802 | 0.61 |
| 1995 | SSA | -0.4786 | 0.5927 | 0.4194 | 0.62 |
| 1996 | SSA | -0.5032 | 0.6328 | 0.4265 | 0.60 |
| 1997 | SSA | 0.0000 | 0.0000 | . | 1.00 |

## Appendix B

**Table II. Risk Ratios for Years in Canada - Age at landing combinations. Baseline for the model is 8years in Canada and over 65 age group. * terms significant at $p \leq 0.05$.**

| *Duration of stay in Canada (years)* | *Age at Landing* | *Parameter Estimate* | *Std Error* | *Chi-Square* | *Chi-Square P-value* | *Risk Ratio* |
|---|---|---|---|---|---|---|
| Intercept | - | -6.3043 | 0.9197 | 46.9871 | 0.0001 | - |
| ≤ 1 | 0-15 | -0.0036 | 1.0020 | 0.0000 | 0.9971 | 1.00 |
| 2 | 0-15 | -1.5282 | 1.0320 | 2.1928 | 0.1387 | 0.22 |
| 3 | 0-15 | -1.3088 | 1.0303 | 1.6136 | 0.2040 | 0.27 |
| 4 | 0-15 | -1.9738 | 1.0570 | 3.4872 | 0.0618 | 0.14 |
| 5 | 0-15 | -1.8493 | 1.0633 | 3.0252 | 0.0820 | 0.16 |
| 6 | 0-15 | -1.6024 | 1.0810 | 2.1971 | 0.1383 | 0.20 |
| 7 | 0-15 | -1.1656 | 1.0751 | 1.1754 | 0.2783 | 0.31 |
| 8 | 0-15 | -1.2469 | 1.1381 | 1.2005 | 0.2732 | 0.29 |
| ≤1 | 16-30 | 0.5489 | 0.9234 | 0.3534 | 0.5522 | 1.73 |
| 2 | 16-30 | 0.1457 | 0.9281 | 0.0247 | 0.8752 | 1.16 |
| 3 | 16-30 | -0.1560 | 0.9297 | 0.0282 | 0.8667 | 0.86 |
| 4 | 16-30 | -0.2452 | 0.9309 | 0.0694 | 0.7922 | 0.78 |
| 5 | 16-30 | -0.2037 | 0.9320 | 0.0478 | 0.8270 | 0.82 |
| 6 | 16-30 | -0.2977 | 0.9346 | 0.1015 | 0.7501 | 0.74 |
| 7 | 16-30 | -0.6525 | 0.9437 | 0.4781 | 0.4893 | 0.52 |
| 8 | 16-30 | -1.3529 | 0.9928 | 1.8569 | 0.1730 | 0.26 |
| ≤ 1 | 31-45 | -0.0934 | 0.9358 | 0.0100 | 0.9205 | 0.91 |
| 2 | 31-45 | -0.5526 | 0.9456 | 0.3415 | 0.5589 | 0.58 |
| 3 | 31-45 | -0.6092 | 0.9468 | 0.4140 | 0.5199 | 0.54 |
| 4 | 31-45 | -0.8026 | 0.9499 | 0.7138 | 0.3982 | 0.45 |
| 5 | 31-45 | -1.0735 | 0.9560 | 1.2608 | 0.2615 | 0.34 |
| 6 | 31-45 | -0.7389 | 0.9553 | 0.5982 | 0.4393 | 0.48 |
| 7 | 31-45 | -1.8267 | 0.9991 | 3.3428 | 0.0675 | 0.16 |
| 8 | 31-45 | -1.8555 | 1.0445 | 3.1557 | 0.0757 | 0.16 |
| ≤ 1 | 46-65 | 1.1125 | 0.9395 | 1.4024 | 0.2363 | 3.04 |
| 2 | 46-65 | 0.7994 | 0.9535 | 0.7028 | 0.4018 | 2.22 |
| 3 | 46-65 | 0.4463 | 0.9590 | 0.2166 | 0.6416 | 1.56 |
| 4 | 46-65 | 0.3632 | 0.9621 | 0.1425 | 0.7058 | 1.44 |
| 5 | 46-65 | 0.1973 | 0.9734 | 0.0411 | 0.8394 | 1.22 |
| 6 | 46-65 | -0.5384 | 1.0211 | 0.2780 | 0.5980 | 0.58 |
| 7 | 46-65 | 0.0096 | 1.0144 | 0.0001 | 0.9925 | 1.01 |
| 8 | 46-65 | 0.1468 | 1.0552 | 0.0193 | 0.8894 | 1.16 |
| ≤ 1 | 66+ | 1.6169 | 0.7402 | 4.7710 | 0.0289 | 5.04* |
| 2 | 66+ | 0.7979 | 0.7516 | 1.1270 | 0.2884 | 2.22 |
| 3 | 66+ | 0.2799 | 0.7677 | 0.1329 | 0.7154 | 1.32 |
| 4 | 66+ | 0.4720 | 0.7650 | 0.3808 | 0.5372 | 1.60 |
| 5 | 66+ | -0.1036 | 0.8013 | 0.0167 | 0.8972 | 0.90 |
| 6 | 66+ | 0.1447 | 0.8075 | 0.0321 | 0.8577 | 1.16 |
| 7 | 66+ | 0.1862 | 0.8480 | 0.0482 | 0.8262 | 1.20 |
| 8 | 66+ | (baseline) | 0.0000 | . | . | 1.00 |

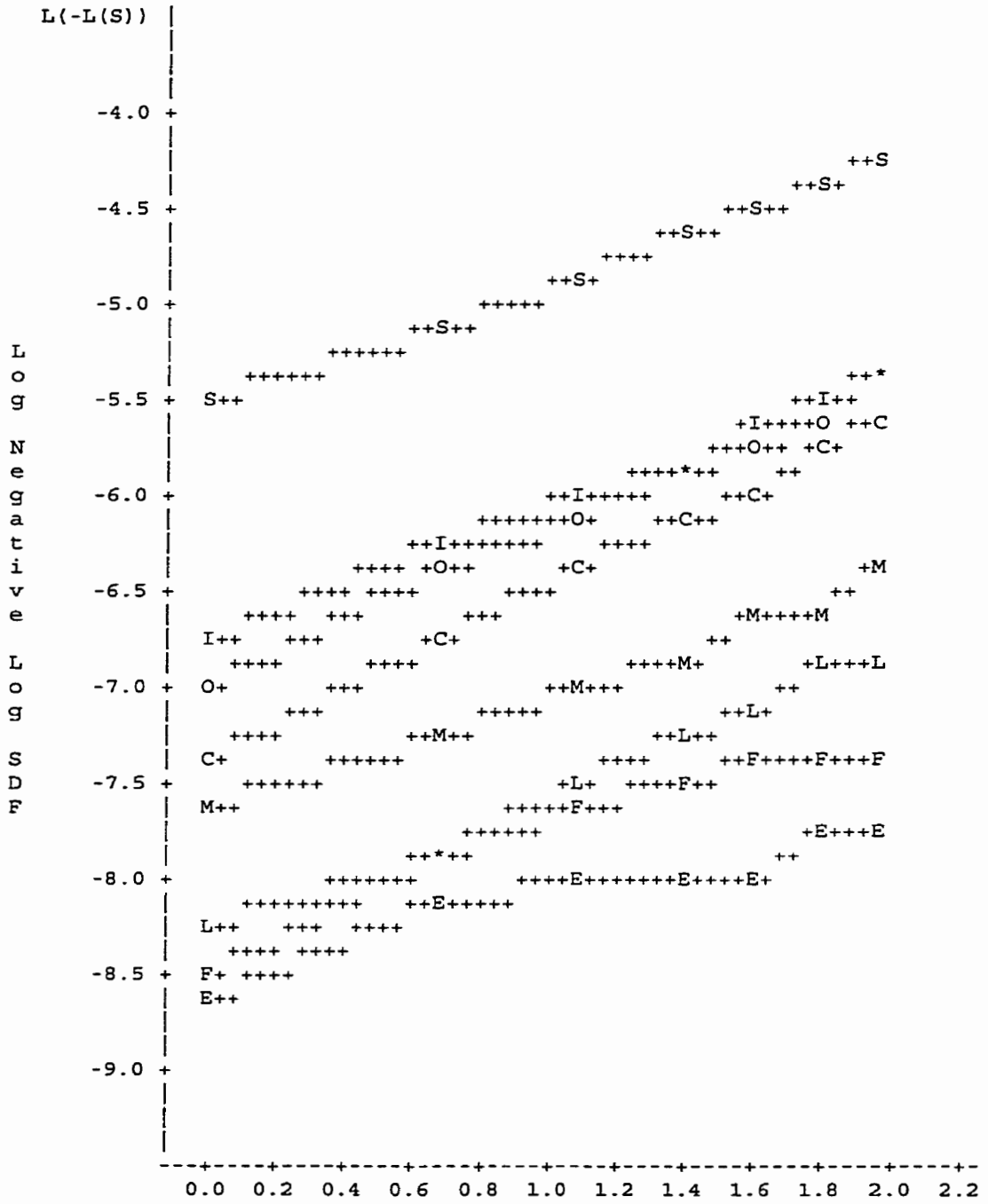
## Appendix B

**Table III. Risk Ratios for Year of Landing - Age at Landing combinations. Year of landing 1997 and over 65 age group are used as baseline. * terms significant ($p \leq 0.05$).**

| *Year of Landing* | *Age at Landing* | *Parameter Estimate* | *Std Error* | *ChiSquare* | *Chi-Square P-value* | *Risk Ratio* |
|---|---|---|---|---|---|---|
| Intercept | - | -6.3043 | 0.9197 | 46.9871 | 0.0001 | - |
| 1990 | 0-15 | -0.2339 | 1.0539 | 0.0493 | 0.8243 | 0.79 |
| 1991 | 0-15 | -0.2333 | 1.0824 | 0.0464 | 0.8294 | 0.79 |
| 1992 | 0-15 | -1.2229 | 1.0953 | 1.2465 | 0.2642 | 0.29 |
| 1993 | 0-15 | -0.3555 | 1.0831 | 0.1078 | 0.7427 | 0.70 |
| 1994 | 0-15 | -0.2613 | 1.0887 | 0.0576 | 0.8103 | 0.77 |
| 1995 | 0-15 | -0.8582 | 1.1056 | 0.6025 | 0.4376 | 0.42 |
| 1996 | 0-15 | -0.6473 | 1.1114 | 0.3393 | 0.5603 | 0.52 |
| 1997 | 0-15 | -1.2469 | 1.1381 | 1.2005 | 0.2732 | 0.29 |
| 1990 | 16-30 | -0.4792 | 0.9886 | 0.2349 | 0.6279 | 0.62 |
| 1991 | 16-30 | -0.5295 | 0.9931 | 0.2842 | 0.5939 | 0.59 |
| 1992 | 16-30 | -1.3480 | 0.9965 | 1.8301 | 0.1761 | 0.26 |
| 1993 | 16-30 | -1.4587 | 1.0013 | 2.1222 | 0.1452 | 0.23 |
| 1994 | 16-30 | -0.7104 | 0.9999 | 0.5047 | 0.4774 | 0.49 |
| 1995 | 16-30 | -0.9204 | 1.0053 | 0.8382 | 0.3599 | 0.40 |
| 1996 | 16-30 | -0.9487 | 1.0115 | 0.8797 | 0.3483 | 0.39 |
| 1997 | 16-30 | -1.3529 | 0.9928 | 1.8569 | 0.1730 | 0.26 |
| 1990 | 31-45 | -1.1277 | 1.0321 | 1.1937 | 0.2746 | 0.32 |
| 1991 | 31-45 | -1.2296 | 1.0405 | 1.3964 | 0.2373 | 0.29 |
| 1992 | 31-45 | -2.1849 | 1.0446 | 4.3751 | 0.0365 | 0.11* |
| 1993 | 31-45 | -2.1175 | 1.0494 | 4.0716 | 0.0436 | 0.12* |
| 1994 | 31-45 | -1.4463 | 1.0500 | 1.8974 | 0.1684 | 0.24 |
| 1995 | 31-45 | -1.7664 | 1.0589 | 2.7828 | 0.0953 | 0.17 |
| 1996 | 31-45 | -1.9769 | 1.0723 | 3.3986 | 0.0653 | 0.14 |
| 1997 | 31-45 | -1.8555 | 1.0445 | 3.1557 | 0.0757 | 0.16 |
| 1990 | 46-65 | -0.2439 | 1.0469 | 0.0543 | 0.8158 | 0.78 |
| 1991 | 46-65 | -0.4565 | 1.0690 | 0.1824 | 0.6693 | 0.63 |
| 1992 | 46-65 | -1.0720 | 1.0740 | 0.9963 | 0.3182 | 0.34 |
| 1993 | 46-65 | -1.0840 | 1.0768 | 1.0135 | 0.3141 | 0.34 |
| 1994 | 46-65 | -0.1450 | 1.0714 | 0.0183 | 0.8924 | 0.86 |
| 1995 | 46-65 | -1.0548 | 1.0964 | 0.9256 | 0.3360 | 0.35 |
| 1996 | 46-65 | -0.4884 | 1.0974 | 0.1981 | 0.6563 | 0.61 |
| 1997 | 46-65 | 0.1468 | 1.0552 | 0.0193 | 0.8894 | 1.16 |
| 1990 | 66+ | -0.3576 | 0.5523 | 0.4193 | 0.5173 | 0.70 |
| 1991 | 66+ | -0.3161 | 0.5517 | 0.3282 | 0.5667 | 0.73 |
| 1992 | 66+ | -0.7587 | 0.5503 | 1.9010 | 0.1680 | 0.47 |
| 1993 | 66+ | -0.6124 | 0.5494 | 1.2427 | 0.2649 | 0.54 |
| 1994 | 66+ | -0.5007 | 0.5707 | 0.7699 | 0.3802 | 0.61 |
| 1995 | 66+ | -0.4786 | 0.5927 | 0.6520 | 0.4194 | 0.62 |
| 1996 | 66+ | -0.5032 | 0.6328 | 0.6324 | 0.4265 | 0.60 |

## Appendix C - World Region of Origin

**Log(-Log(Survival Function)) Estimates**

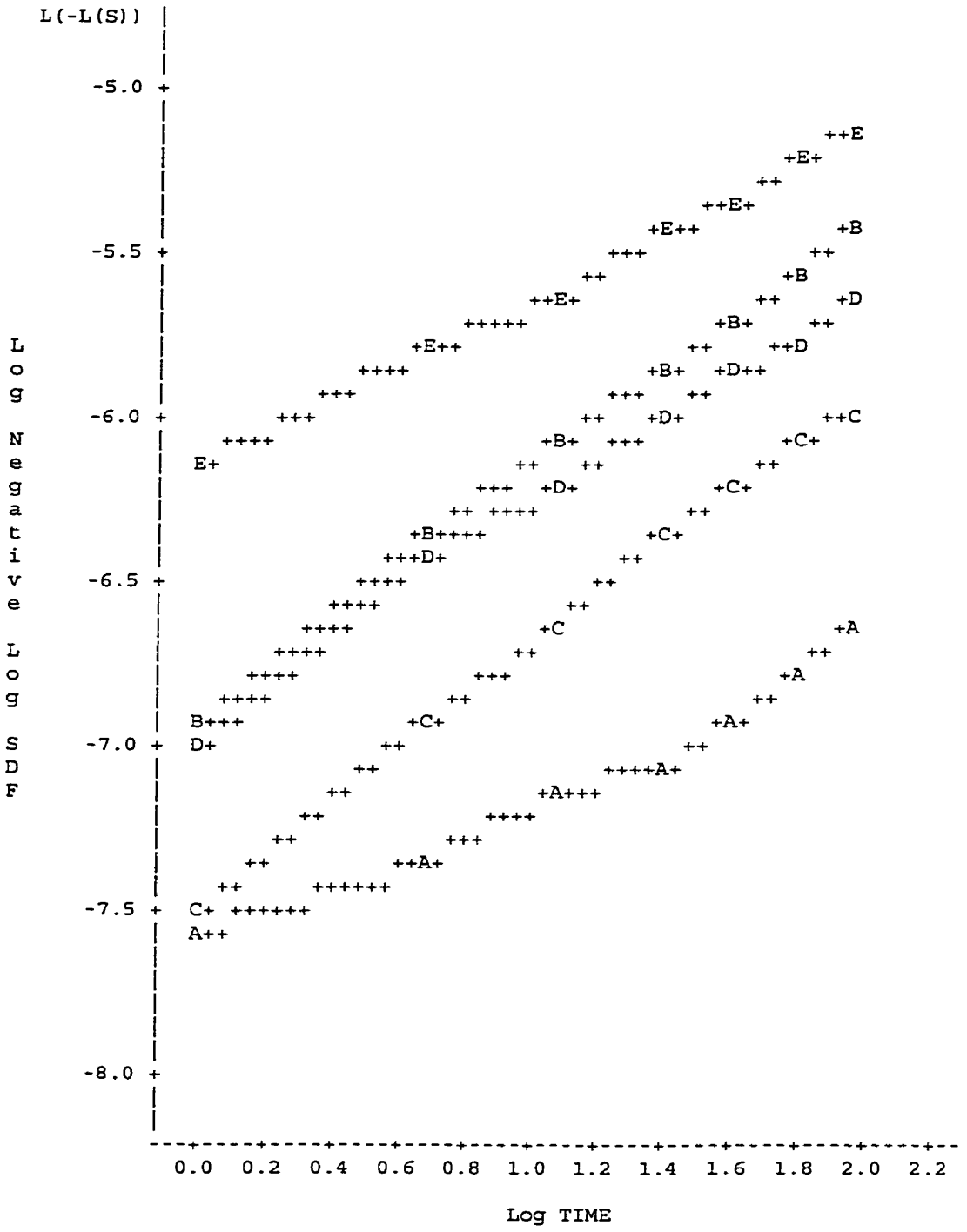



**Legend for Strata Symbols**

C:WRO=China E:WRO=EME F:WRO=FSEE I:WRO=India
L:WRO=LatCar M:WRO=MEC O:WRO=OthAsia S:WRO=SSA

## Appendix C - Age at Landing

**Log(-Log(Survival Function)) Estimates**

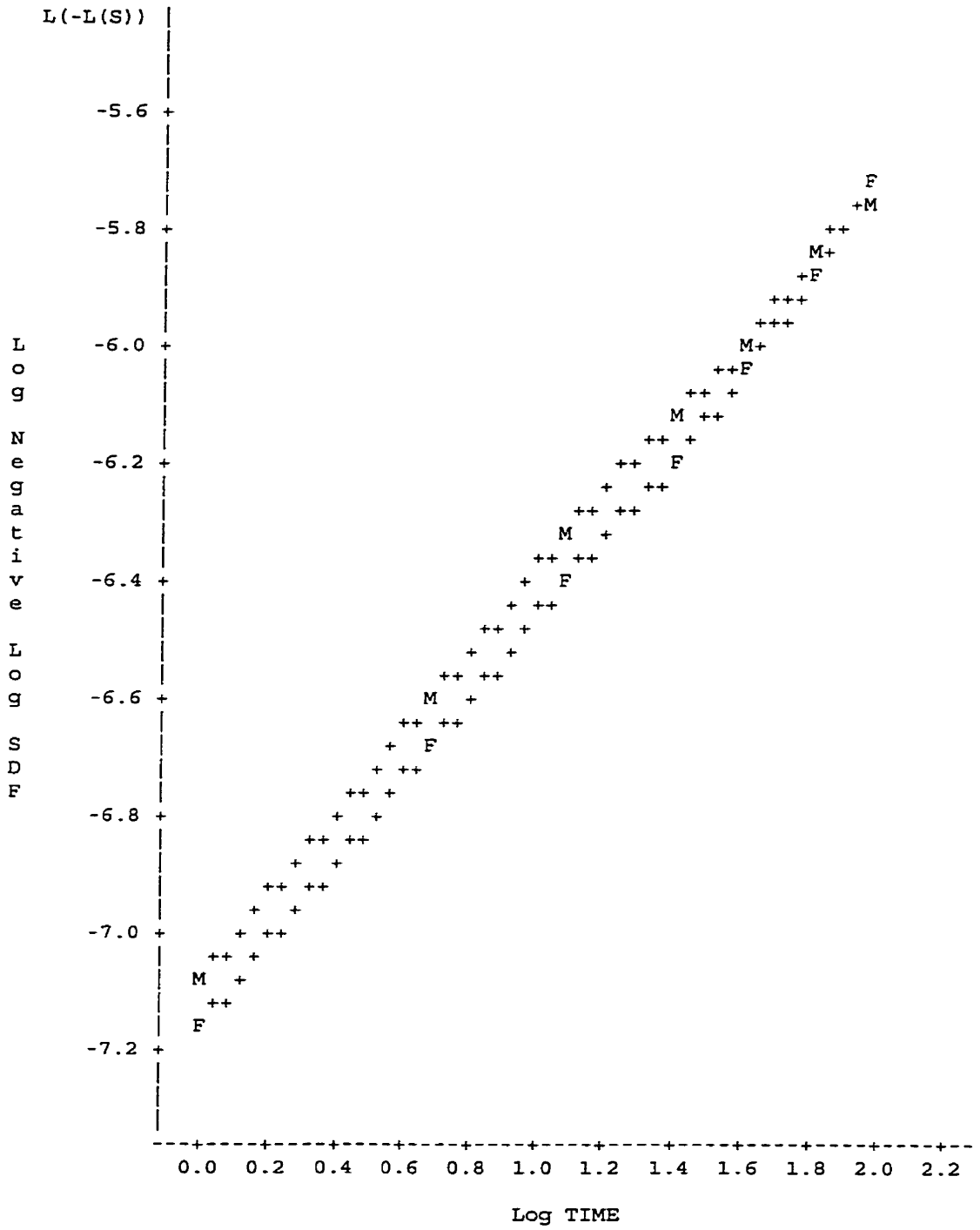



**Legend for Strata Symbols**

A: AGEGR=0-15 B: AGEGR=16-30 C: AGEGR=31-45 D: AGEGR=46-65
E: AGEGR=66+

## Appendix C - Sex

**Log(-Log(Survival Function)) Estimates**




**Legend for Strata Symbols**

F: SEX=F M: SEX=M