The use of maximum likelihood methods to estimate the risk of tuberculous infection and disease in a Canadian First Nations population

Michael Clark¹ and Emilia Vynnycky²

Accepted 3 July 2003

Background Tuberculosis (TB) notification rates among First Nations people in British Columbia, Canada, are higher than those among non-First Nations people, although rates are declining more rapidly in the First Nations population. The epidemiology of tuberculous infection and disease during the period 1926–2000 in this population was investigated.

Methods The trend in the annual risk of infection (ARI) since 1926 was estimated using tuberculous meningitis mortality statistics and skin testing data. Risks of progression from infection to disease were estimated by fitting model predictions of disease incidence to TB notifications, using maximum likelihood methods. Infectious TB notifications were matched with ARI estimates to obtain the number of transmissions per infectious case over time.

Results We estimate that the ARI decreased from more than 10% during the prechemotherapy era to less than 0.1% by 2000. The risks of primary, reactivation, and exogenous re-infection disease among adults aged 25–44 years were 22%, 0.1%, and 6%, respectively. The number of transmissions per infectious case decreased from 16 to 2 from the early 1970s to the late 1990s.

Conclusions This study shows that the risk of infection among British Columbia First Nations people is decreasing, while the relative contribution of reactivation to disease incidence is increasing. Once infected, First Nations people may have a higher risk of developing disease than other populations.

Keywords Tuberculosis, epidemiology, model, Aboriginal, Canada

Tuberculosis (TB) is still a major public health problem in Canadian First Nations communities. During the 1990s, notification rates in the First Nations, on-reserve population were 8–10 times higher than the overall Canadian rates.¹ A high prevalence of latent tuberculous infection, and socioeconomic factors such as overcrowded housing and substance abuse, are known contributors to this disproportionate burden.²,³ During the period 1997–1999, the age-adjusted TB notification rate among First Nations people living on reserves in British Columbia was 32 per 100 000.³ In urban areas such as Vancouver, the risk of TB may be elevated among First Nations people due to human immunodeficiency virus (HIV)/AIDS, injection drug use, and homelessness.⁴

The Indian Act of 1867 defined a reserve as, ‘…any tract or tracts set apart by treaty or otherwise for the use and benefit of or granted to a particular band of Indians…’.⁵ In 2000, it was estimated that 110 529 First Nations people lived in British Columbia, and 51% of these people lived on reserves.⁶ The objective of this study was to examine the epidemiology of tuberculous infection and disease in the First Nations population of British Columbia over time, using available data. We attempt to estimate the following: the annual risk of infection (ARI) between 1926 and 2000; the age-specific prevalence of infection between 1972 and 2000; the risk of progression to disease from primary infection, latent infection (endogenous disease), and exogenous re-infection; and the number of new infections per infectious disease case between 1972 and 2000.
Methods

Study setting and general description

The analyses are based on data from the First Nations population of British Columbia. We first describe the methods used to estimate the long-term trend in the ARI with Mycobacterium tuberculosis, and then describe how these estimates were applied to estimate the risks of primary, endogenous reactivation, and exogenous re-infection disease.

Estimating the ARI during the prechemotherapy era

Given the absence of reliable tuberculin data for the prechemotherapy era, we used tuberculous meningitis mortality data among First Nations children aged 0–4 years to estimate ARI values during this period. We assumed that 1% of those infected develop meningitis and subsequently die, as estimated elsewhere.\(^7\)–\(^10\) These data are available from Statistics Canada since 1926. Use of streptomycin for TB treatment in Canada began in 1947,\(^11\) and thus mortality rates from meningitis after this year can no longer be assumed to reflect morbidity. Mass vaccination of newborns with bacille Calmette-Guérin (BCG) was not commenced in any Canadian province until after 1948.\(^12\) Therefore, effects of BCG on TB meningitis during the period before 1947 were assumed to be negligible.

Estimating the ARI after 1947

School screening data from 1991 to 2000 were supplied by the British Columbia Centre for Disease Control (BC CDC). This screening programme is supported by Health Canada as part of the National Tuberculosis Elimination Strategy implemented in First Nations, on-reserve communities. The data included the number of grade six children who were screened with a tuberculin skin test, and the number of these with a positive skin test, by year (Table 1). Since grade six children are generally aged from exactly 11 years to the end of 12 years (12 years and 11 months), we assume the prevalence of positive tests in these cohorts approximates the prevalence at age 12 years. Children were excluded from the data if they met any of the following criteria: the BC CDC had a record in their BCG registry for the child prior to the date of skin testing; the child had a BCG scar; or the parent or guardian reported the child had received BCG prior to the test. The criterion for a positive test was 10 mm induration to 5 TU (0.1 ml of tuberculin) of purified protein derivative (PPD). These data were used to estimate ARI values between 1979 (the year of birth for the earliest cohort represented in the data set) and 2000 (the last year of screening used in the analysis). We first assessed the nature of the trend in the ARI by calculating the average ARI experienced in each year of life of a cohort tested at time \(t\) using the formula:\(^13\)

\[
ARI_t = 1 - (1 - \hat{p}_{12,t})^{1/12}
\]  

(1)

where \(\hat{p}_{12,t}\) is the proportion of children aged 12 years found to be positive at time \(t\).

The risk of infection in 2000 and its trend since 1979 were then estimated using maximum likelihood methods, by minimizing the following standard binomial log-likelihood \(\chi^2\) deviance function:\(^14\)

\[
L = 2 \sum_t S_t \ln (1 - \hat{p}_{12,t}) + (N_t - S_t) \ln (\hat{p}_{12,t})
\]

\[
- S_t \ln (1 - p_{12,t}) - (N_t - S_t) \ln (p_{12,t})
\]

(2)

The fitting was carried out in Microsoft EXCEL using the Solver routine. In the above expression, \(S_t\) is the number of 12 year olds observed to be tuberculin negative (susceptible) at time \(t\), \(N_t\) is the total number who were tested at time \(t\), \(\hat{p}_{12,t}\) is the proportion found to be positive at time \(t\), and \(p_{12,t}\) is the expected proportion infected by age 12 years at time \(t\). The expected proportion infected \(p_{12,t}\) is given by the expression:\(^10\)

\[
p_{12,t} = 1 - \exp \left( - \int_{t-12}^{t} \lambda(u) du \right)
\]

(3)

where \(\lambda(u)\) is the annual infection rate at time \(u\). This rate is related to the annual risk of infection at time \(u\) by the expression: \(ARI = 1 - \exp(-\lambda(u))\). Throughout the study the

\begin{table}[h]
\centering
\caption{Summary of the proportion of First Nations individuals who were tuberculin positive at age 12 years (1991–2000) in British Columbia, and estimates of the average annual risk of infection (ARI) experienced during the lifetime of each cohort.}
\begin{tabular}{|l|l|l|l|}
\hline
Year & No tested & Proportion positive (%)\(^2\) & Average ARI (%) \\
\hline
1991 & 110 & 4.5 (1.5, 10.3) & 0.39 \\
1992 & 250 & 4.4 (2.2, 7.7) & 0.37 \\
1993 & 219 & 0.91 (0.1, 3.3) & 0.08 \\
1994 & 231 & 0.87 (0.1, 3.3) & 0.07 \\
1995 & 296 & 2.0 (0.8, 4.4) & 0.17 \\
1996 & 524 & 1.1 (0.4, 2.5) & 0.1 \\
1997 & 315 & 1.3 (0.3, 3.2) & 0.11 \\
1998 & 439 & 2.3 (1.1, 4.2) & 0.19 \\
1999 & 250 & 0.0 (0.0, 0.0) & 0.0 \\
2000 & 424 & 1.4 (0.5, 3.1) & 0.12 \\
\hline
\end{tabular}
\end{table}

\(^2\) Exact binomial 95% CI in brackets.
METHODS TO ESTIMATE THE RISK OF TUBERCULOUS INFECTION AND DISEASE

Estimating the risk of progression to TB disease

Reported TB case tallies for the British Columbia First Nations population from 1972 to 2000 for the 0–14, 15–24, 25–34, and 35–44 year age groups were provided by the Population and Public Health Branch of Health Canada. Owing to the high variance of reported TB cases from year to year among 0–24 year olds, and the possible impact of neonatal BCG vaccination on disease incidence in this age group, data from the 0–14 and 15–24 age groups were excluded from the analyses. These data include First Nations TB case diagnoses on and off reserve, as case report forms did not include this distinction prior to 1990. Between 1990 and 2000, the TB notification rate among First Nations people living on and off reserve (39 per 100 000) was only slightly higher than the on-reserve notification rate of 36 per 100 000.

Disease cases included in the model were new TB cases (pulmonary and extrapulmonary), diagnosed with bacteriological proof and/or clinical symptoms consistent with TB. Relapsed cases (those cases with recurrent active disease after a known period of inactivity) were excluded from the model. Estimates of the risk of disease were derived using methods similar to those of Sutherland et al., applied to the Dutch population, namely by fitting predictions of the number of cases expected in different age groups over time, expressed in terms of the risks of disease, to these data. Given small differences between the notification rates among males and females, the data were not stratified by sex. The method assumes that individuals develop disease either soon after primary infection (‘primary’ disease) or many years thereafter through endogenous reactivation, or after exogenous re-infection. ‘Endogenous disease’ refers to reactivation of a longstanding latent infection among those infected individuals who did not develop primary disease. ‘Exogenous disease’ refers to the development of disease soon after an individual with longstanding latent infection is re-infected with a new strain of M. tuberculosis. The number of TB disease cases in each age group at time \( t \) \( D(a, t) \) is given by the sum of the number of individuals experiencing primary disease \( P(a, t) \), disease through endogenous reactivation \( E_p(a, t) \), and exogenous disease \( E_x(a, t) \), as follows:

\[
D(a, t) = P(a, t) + E_p(a, t) + E_x(a, t)
\]

Assuming that individuals experience risks \( d_p \), \( d_e \), and \( d_x \) of developing primary, endogenous, and exogenous disease, this expression can be rearranged as follows:

\[
D(a, t) = I(a, t)d_p + L(a, t)d_e + R(a, t)d_x
\]

where \( I(a, t) \) and \( R(a, t) \) represent the numbers of people in age group \( A \) who have been initially infected (primary infection) and recently re-infected, respectively, and \( L(a, t) \) is the number of individuals with a longstanding latent infection. For simplicity—and following the approach used by Dye et al. \(^{20} \)—initial infection and re-infection are defined as that which has occurred during the previous year. Given that the risks of developing disease are highest during the first 5 years after infection, estimates of disease risk following initial infection and re-infection in our model are interpretable as the cumulative 5-year risks of developing disease after these events as described by Sutherland et al. \(^{19} \)

The numbers of individuals in age group \( A \) at time \( t \) (where \( A \) spans the ages \( a_j \) to \( a_{j+1} \)) were calculated using estimates of the ARI at different times \( t \) (\( ARI(t) \)) and estimates of the prevalence of infection among individuals of age \( a \) at time \( t \), as follows:

\[
I(A, t) = \sum_{a = a_j}^{a_{j+1}} ARI(t)[1 - p_{a-1,j-1}]n(a, t)
\]

\[
R(A, t) = \sum_{a = a_j}^{a_{j+1}} ARI(t)p_{a-1,j-1}n(a, t)
\]

\[
L(A, t) = \sum_{a = a_j}^{a_{j+1}} n(a, t)p_{a,t} - R(A, t) - I(A, t)
\]

where \( n(a, t) \) is the number of First Nations people of age \( a \) at time \( t \). Expressions 5–9 are consistent with those used elsewhere in analogous analyses.

The risks of developing primary, endogenous, and exogenous disease were estimated using maximum likelihood methods, by minimizing the following Poisson log-likelihood chi-square deviance function:

\[
D = -2 \sum_{A,t} O_{A,t} \ln(D(A, t)) - O_{A,t} \ln(O_{A,t}) + O_{A,t} - D(A, t)
\]

where \( O_{A,t} \) is the number of observed (reported) cases in age group \( A \) at time \( t \). The 95% CI for disease risk estimates were calculated using the profile likelihood method (see above). Data

\[
P_{12,t} = 1 - \exp \left\{ \frac{\lambda(t) - \lambda(t - 12)}{t} \right\}
\]

Confidence limits for the infection rate in 2000, and its trend from 1979, were calculated using the profile likelihood method. This method assumes that log-likelihoods calculated in the model approximate to a \( \chi^2 \) distribution. After finding a maximum likelihood fit to the data, a single parameter estimate was varied (increasing or decreasing it for upper and lower limits, respectively) until the deviance increased by 3.84, keeping the other parameter constant.\(^{16,17} \)

For simplicity, the ARI between 1947 and 1979 was assumed to decline at a constant rate.\(^{7,18} \) To assess the validity of this assumption, and the estimated trend between 1979 and 2000, we compared estimates of the expected prevalence of infection in 1992 and 1999 against the proportion of individuals aged 0–14, 15–24, and 25–44 years found to be tuberculin positive in cross-sectional surveys during the periods 1991–1993 and 1998–2000, respectively.
for individuals aged ≥45 years were excluded due to uncertainties in estimating the prevalence of infection in older age groups.

**Estimating the number of new infections per infectious case**

Three-year, moving rates of reported infectious TB were calculated from 1972–1974 to 1998–2000. Those cases considered infectious were respiratory cases in which *M. tuberculosis* was isolated from a sputum sample using smear microscopy and/or culture. This definition includes smear-negative respiratory cases with a positive sputum culture, due to a recent report showing that 17% of transmissions in San Francisco were attributed to such cases. The number of new infections per infectious case over time was calculated using the ratio between the ARI and the incidence of infectious TB cases.

**Population data sources**

Population estimates from the 1931, 1941, and 1951 censuses were used to calculate mortality rates from tuberculous meningitis during the pre-chemotherapy era. Population data between 1972 and 2000, by one-year age group, were obtained from the Department of Indian and Northern Affairs Canada.

**Results**

**Estimates of the ARI**

During the prechemotherapy era, the annual risk of infection, as estimated using tuberculous meningitis mortality data, was consistently high, ranging between 6% in 1926 and 21% in 1944. Skin testing results by year and average ARI experienced during the lifetime of individuals tested in school screening are shown in Table 1. Approximately 4.4% (95% CI: 2.5, 7.1) of 12 year olds were found to be tuberculin positive during the period 1991–1992; this proportion declined to 0.9% (95% CI: 0.33, 1.9) by 1999–2000.

According to best-fitting estimates (Figure 1), the ARI was 0.031% (95% CI: 0.023, 0.040) in the year 2000. The estimated annual rate of decline from 1979 to 2000 was 13.8% (95% CI: 11.5, 15.9). The estimated ARI since 1926 is presented in Figure 2. The assumed constant rate of decline between 1947 and 1979 was 9.1%.

Predictions of the prevalence of infection among those aged 0–24 years based on ARI estimates agree closely with the prevalence observed in cross-sectional tuberculin surveys during the 1991–1993 and 1998–2000 periods (Figure 3). Predictions are higher than observed values in the 25–44 age group.

Figure 1 Comparison between the predicted prevalence of infection and the observed prevalence of tuberculin sensitivity among children aged 12 years

Figure 2 Estimated annual risk of infection (%) during the period 1926–2000 among the First Nations people in British Columbia. The annual risk of infection during the period 1974–1979 has been calculated assuming that it declined at a constant rate (9.7% p.a.) from the level estimated in 1947 to that estimated in 1979

Figure 3 Comparison between the predicted prevalence of infection and the observed prevalence of tuberculin sensitivity by age group (1991–2000)
Estimates of the risks of developing disease

Figure 4 compares best-fitting model predictions of disease incidence against the observed notification rates among 25–44 year olds. The risk of primary disease (following initial infection) among adults aged 25–44 years was 21.5% (95% CI: 14.9, 28.6), and the estimated annual risk of reactivation was 0.09% (95% CI: 0.08, 0.10). The model estimates an elevated risk of disease following exogenous re-infection among adults aged 25–44 years, of 5.8% (95% CI: 2.8, 9.2).

The estimated proportion of total cases aged 25–44 years attributed to endogenous reactivation of latent infection increased from 48% in 1972 to 76% in 2000 (Figure 5). The proportion of cases due to exogenous re-infection disease decreased from 35% to less than 2% during the same period. The contribution of primary disease was 17% in 1972, and 22% in 2000. During the 29-year period, the model predicts that 73% of cases in the 35–44 age group were due to reactivation, compared with 50% in the 25–34 age group. Conversely, a much higher proportion of total cases in the 25–34 age group were due to primary disease (37%), when compared with the 35–44 age group (13%).

The 3-year reported rate of pulmonary, infectious TB decreased from 80.2 per 100 000 during the period 1976–1978 to 16.1 per 100 000 by 1998–2000 (Figure 6). Estimates of the number of new infections (transmissions) per infectious TB case decreased from 16 in 1973 to 2 during the late 1990s.

Discussion

We estimate the ARI among First Nations people in British Columbia has decreased from over 10% during the prechemotherapy era to less than 0.1% in 2000 (Figure 2). To our knowledge the results presented here represent the first attempts to estimate a long-term trend in ARI, and the risks of progression to disease, in this population. As with any study involving disease modelling, our results depend on assumptions. It was assumed that the ARI is not age-dependent, and that the risk of reactivation disease does not change with time since infection. It is possible that many people in First Nations communities progress to disease when their immune system is impaired, and the elevated risk they experience is not related to time since infection. The proportion of HIV cases attributed to First Nations people in Canada has been increasing for a decade, and the urban First Nations community in Vancouver is known to be at increased risk for HIV. The possible impact of increasing use of therapies to prevent progression to disease was also excluded from the model.

Estimates of the ARI during the prechemotherapy era have been calculated using a ratio between mortality rates from tuberculous meningitis among 0–4 year olds and the ARI.
which was derived from European studies. It is possible that the risk of TB disease among First Nations children was higher than in European populations, which would inflate our ARI estimates. In this context, it is worth noting that the average ARI estimates among two cohorts of First Nations children skin tested during the prechemotherapy era were 12% and 20%, which is consistent with the range of ARI values found in our study (6–21%). One would expect very high ARI values during the 1930s and 1940s, as death rates among British Columbia First Nations ranged over 581.5 and 1026.9 per 100 000. In the modern era, it is important to recognize the possible impact of BCG use on our interpretation of screening data. Some children with unrecognized BCG vaccination may have been included in the data set. The effects of this should be minimal, given the observation that BCG given in the first year of life rarely affects skin testing results beyond the age of 10 years. More importantly, the use of BCG may be more common in communities where TB incidence is higher. Another limitation in the school screening data was that the exact age of grade sixes screened was not reported. It is possible a small proportion of children in a given year are aged as low as 10 years, or as high as 14.

Despite the limitations described above, two findings indicate that the results of this study are plausible. First, the estimated annual rate of decline in ARI between 1979 and 2000 (13.8% p.a.) is identical to that observed in the Dutch population between 1940 and 1970. Predictions of infection prevalence among people aged 0–24 years in 1992 and 1999 are very similar to the proportion of individuals found to be tuberculin positive in these years (Figure 3), indicating that ARI estimates for the last three decades are in a plausible range. Our finding that predictions of infection prevalence for individuals aged over 25 years did not agree with survey results may be related to waning tuberculin sensitivity among adults who were infected during periods when the risk of infection was higher. In contrast with the analyses of Sutherland and colleagues, which included only pulmonary TB cases, our analyses considered all new pulmonary and extrapulmonary cases. The above study also included a wider age group (15–69 years) of male Europeans. However, it is striking that their estimates for risk of primary and exogenous disease fall within the 95% CI of our estimates for people aged 25–44 years. Both of the studies estimate an approximate risk of disease following primary infection of 22%.

It is often assumed that the risk of TB following infection is 5% within 2 years, with a 10% lifetime risk of developing disease, which is inconsistent with the findings of this and other studies. An increased risk of disease in young adults has implications for decisions on treatment of infection, particularly in light of the high incidence of infectious TB in this age group, and the possibility of secondary benefits due to reduced transmission. Treatment of infection in adults is controversial, due to poor adherence and the increased risk of adverse reactions such as isoniazid (INH) hepatitis.

It is interesting that our estimate of the risk of developing disease through reactivation is about threefold greater than that estimated in other populations (0.1% p.a. as compared with 0.03% p.a. for The Netherlands and the UK). This may be attributable to several factors including differences in immunosuppression, occurring as a result of poor nutrition and other socioeconomic influences. It is possible that First Nations people experience an elevated risk due to late exposure to European strains of bacilli, and resulting lower levels of innate immunity. A recent study showed that several First Nations TB cases had a gene deletion that may have predisposed them to developing active disease. The model predicts an increasing contribution of endogenous reactivation to total disease burden over time. The high prevalence of latent infection in many First Nations communities, coupled with an increased risk of disease, may result in cases of reactivation disease for many years to come.

These analyses show that the relative contribution of re-infection to disease incidence has decreased markedly over time in the 25–44 age group. Re-infection may still play a significant role in older age groups, in which a higher prevalence of latent infection can be expected. In a recent review of DNA fingerprinting data from Vancouver, no cases of re-infection by a different strain of M. tuberculosis were found among individuals with recurrent disease. Different results were found in a South African study, in which a high proportion relapsed cases had been re-infected.

The occurrence of outbreaks and paediatric disease indicates that M. tuberculosis transmission continues in some First Nations communities. Rapid progression to disease was reported during a 1992 outbreak among immunocompetent members of a British Columbia First Nations community, in which 13 of 21 TB cases had a conversion of their skin test result during the investigation. Despite these observations, we have estimated that the number of transmissions per infectious case is decreasing over time. This may be due in part to earlier case finding and treatment: Wang et al. reported that TB notification rates among British Columbia First Nations people declined more rapidly than those among non-First Nations people between 1992 and 1996, possibly due to greater use of directly observed therapy. In many industrialized countries decreasing TB incidence has led to the decentralization of responsibility and funding for TB programmes. This type of reform is often associated with the occurrence of microepidemics, in which a single source case infects a large number of young, previously uninfected contacts. Given the possibility of increased susceptibility to TB disease among First Nations people, and the occurrence of large outbreaks in First Nations communities, TB programme erosion must be prevented despite the decreasing risk of infection shown in this study.

Acknowledgements

We thank the First Nations and Inuit Health Branch (FNHIHB) of Health Canada and the British Medical Research Council for their support, the British Columbia Centre for Disease Control (BC CDC) for the provision of tuberculin skin testing data, and the Population and Public Health Branch (PPHB) of Health Canada for supplying TB disease data. The ongoing efforts of the BC CDC, FNHIHB Pacific Region, and primary health care workers in BC First Nations communities to eliminate TB should also be acknowledged.
KEY MESSAGES

- The burden of tuberculosis (TB) is greater among First Nations people than among other Canadians.
- It is likely that the annual risk of tuberculous infection in the British Columbia First Nations population is currently less than 0.1%.
- The risk of TB disease following primary infection may be higher than conventionally accepted values.
- The risk of TB disease, particularly reactivation disease, may be elevated among First Nations people when compared with other populations.

References


