

The intrinsic transmission dynamics of tuberculosis epidemics

SALLY M. BLOWER¹, ANGELA R. MCLEAN², TRAVIS C. PORCO¹, PETER M. SMALL³,
PHILIP C. HOPEWELL⁴, MELISSA A. SANCHEZ⁵ & ANDREW R. MOSS¹

¹Department of Epidemiology & Biostatistics, University of California San Francisco,
San Francisco General Hospital, Box 1347, San Francisco, California 94143-1347, USA

²Zoology Department, Oxford University, Oxford UK

³Division of Infectious Diseases & Geographic Medicine, Stanford Medical School, Stanford, California 94305, USA

⁴Medicine Service and Department of Medicine, University of California San Francisco,
San Francisco General Hospital, San Francisco, California 94143-1347, USA

⁵School of Public Health, University of California Berkeley, Berkeley, California 94720, USA

Correspondence should be addressed to S.M.B.

In developed countries the major tuberculosis epidemics declined long before the disease became curable in the 1940s. We present a theoretical framework for assessing the intrinsic transmission dynamics of tuberculosis. We demonstrate that it takes one to several hundred years for a tuberculosis epidemic to rise, fall and reach a stable endemic level. Our results suggest that some of the decline of tuberculosis is simply due to the natural behaviour of an epidemic. Although other factors must also have contributed to the decline, these causal factors were constrained to operate within the slow response time dictated by the intrinsic dynamics.

Tuberculosis is an ancient disease that is caused by *Mycobacterium tuberculosis*. In developed countries major epidemics began several centuries ago.¹ These epidemics peaked and declined long before tuberculosis became curable in the late 1940s (Fig. 1)^{1–4}. Several hypotheses have been proposed to explain why tuberculosis epidemics declined in the absence of medical interventions. The most widely accepted hypothesis is that the epidemics declined because of a progressive improvement in the standard of living⁵. Other hypotheses suggest that the epidemics declined because of the segregation of infectious cases in sanatoria or workhouse infirmaries⁶; that there has been attenuation in the genetic virulence of *M. tuberculosis*; or that natural selection has increased host resistance^{3,7}. Here we address two specific questions concerning the historical epidemiology of tuberculosis: Why did major epidemics suddenly arise and why did these epidemics decline dramatically in developed countries prior to the availability of an effective therapy? We have addressed these questions by formulating and analysing two mathematical models — one simple and one more detailed — that reflect the intrinsic transmission dynamics of *M. tuberculosis*. Mathematical models are required to understand the transmission dynamics, because tuberculosis epidemics are complex non-linear systems.

Tuberculosis transmission models

Our simple transmission model consists of three ordinary differential equations that represent the current biological understanding of tuberculosis. The model captures the temporal dynamics of three groups: susceptible individuals (X), latently infected individuals (that is individuals who have been infected with *M. tuberculosis*, but have no clinical illness and hence are noninfectious) (L) and active infectious tuberculosis cases (T).

We assume that infected individuals can develop tuberculosis by either of two pathogenic mechanisms: direct progression (the disease develops soon after infection) or endogenous reactivation (the disease can develop many years after infection)⁸. Consequently, we model two types of tuberculosis: primary progressive tuberculosis (which we shall refer to as fast tuberculosis) and reactivation tuberculosis (which we shall refer to as slow tuberculosis). The majority (estimated to be approximately 90%) of infected individuals never develop tuberculosis⁸. The model's assumptions and equations are described in more detail in the Methods section.

The detailed model expands upon the simple model by the addition of three refinements that reflect further biological complexities of the natural history of tuberculosis: (1) only a certain fraction of cases are assumed to be infectious; (2) a case may be spontaneously cured (that is, without treatment) and move into the recovered non-infectious state; and (3) an individual in the recovered state may either relapse and develop tuberculosis again or may never relapse and die of other causes. A flow diagram of the detailed transmission model is given in Fig. 2. Equations specifying this model are given in the Methods section, and the model's parameters are described in Table 1.

The simple and the detailed models were developed to understand the historical epidemiology of tuberculosis. The models are therefore appropriate for assessing the transmission dynamics in populations of immunocompetent individuals. Neither model includes the phenomenon of reinfection, because clinical and experimental data indicate that reinfection by *M. tuberculosis* is uncommon in immunocompetent individuals⁹. Since reinfection can occur in heavily exposed and/or immunocompromised individuals^{10,11}, these models are not appropriate for assessing the transmission dynamics of tuberculosis in populations that are

heavily infected by the human immunodeficiency virus (HIV).

Basic reproductive number of tuberculosis

A useful summary parameter that quantifies the transmission potential of a pathogen is the basic reproductive number (R_0), the average number of secondary infectious cases produced when one infectious individual is introduced into a population where everyone is susceptible^{12,13}. The derivation of an analytical expression for R_0 permits the evaluation of the transmission threshold. When $R_0 > 1$, an epidemic can occur (the predicted endemic levels of tuberculosis can be calculated from the equilibrium results, see Methods), and when $R_0 < 1$, the infectious disease will be eliminated^{12,13}. The expression for R_0 for tuberculosis (calculated from the simple model) is given in equation 1:

$$R_0 = R_0^{\text{FAST}} + R_0^{\text{SLOW}} \quad (1)$$

where:

$$R_0^{\text{FAST}} = \left(\frac{\beta \Pi}{\mu} \right) \left(\frac{1}{\mu + \mu_T} \right) p$$

$$R_0^{\text{SLOW}} = \left(\frac{\beta \Pi}{\mu} \right) \left(\frac{1}{\mu + \mu_T} \right) \left(\frac{(1-p)v}{v + \mu} \right)$$

The R_0 for tuberculosis (calculated from the detailed model) is given in equation 2:

$$R_0 = R_0^{\text{FAST}} + R_0^{\text{SLOW}} + R_0^{\text{RELAPSE}} \quad (2)$$

where:

$$R_0^{\text{FAST}} = \left(\frac{\beta \Pi}{\mu} \right) \left(\frac{1}{\mu + \mu_T + c} \right) p f$$

$$R_0^{\text{SLOW}} = \left(\frac{\beta \Pi}{\mu} \right) \left(\frac{1}{\mu + \mu_T + c} \right) \left(\frac{q(1-p)v}{v + \mu} \right)$$

$$R_0^{\text{RELAPSE}} = \left(\frac{\beta \Pi}{\mu} \right) \left(\frac{1}{(\mu + \mu_T + c) ((\mu + \mu_T + c) - ((2\omega c) / (2\omega + \mu)))} \right) \cdot \left(\left(p + \frac{(1-p)v}{v + \mu} \right) \left(\frac{\omega c}{2\omega + \mu} \right) \right)$$

These results illustrate that a tuberculosis epidemic can be viewed as a series of linked subepidemics. Equations 1 and 2 show that the aggregate R_0 is simply the sum of the R_0 values in each subepidemic. The aggregate R_0 for the simple model can be decomposed into two R_0 values: R_0^{FAST} due to the subepidemic driven by direct progression and R_0^{SLOW} due to the subepidemic driven by endogenous reactivation. The aggregate R_0 for the detailed model includes a third component, representing the R_0 for the subepidemic driven by relapse tuberculosis. Equations 1 and 2 show that the value of R_0 in each of the subepidemics is determined by the product of three components: (1) The average number of susceptibles that one infectious case infects per unit time; (2) the average time that a case is infectious (which is the same for fast and slow tuberculosis, but is different for relapse tuberculosis); and (3) the probability that an infected individual will develop into an infectious case (which is different for fast, slow or relapse tuberculosis).

The value of R_0 when tuberculosis epidemics first arose could

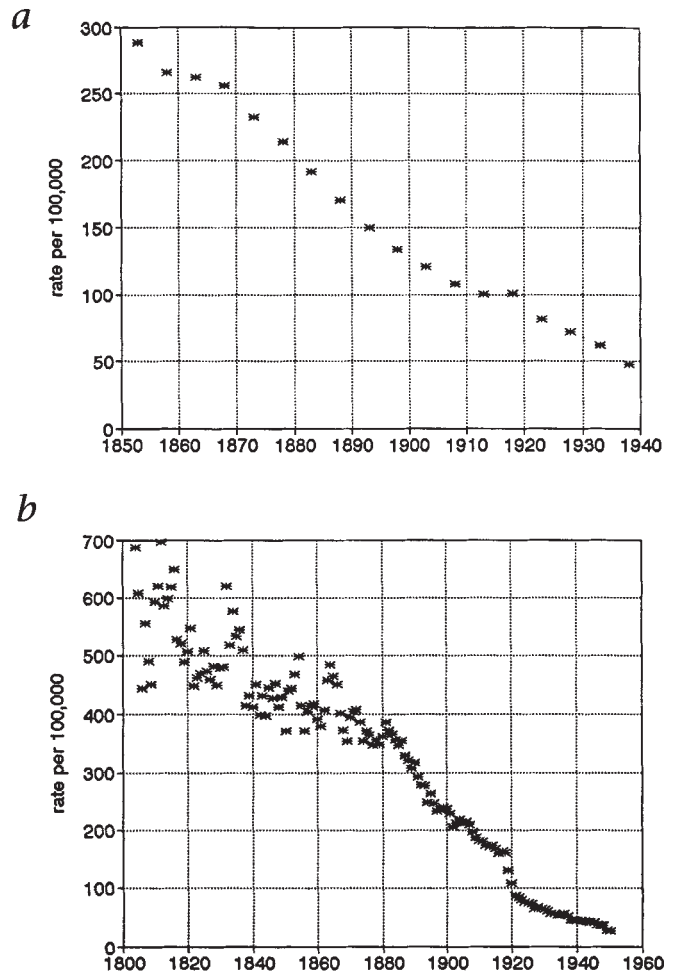


Fig. 1 Annual mortality rates due to pulmonary tuberculosis per 100,000 persons in England and Wales from 1851 to 1938 (a) and in New York City from 1804 to 1950 (b) demonstrating decreasing rates long before the introduction of effective medical interventions in the late 1940s. In the United States the highest tuberculosis mortality rates occurred around the beginning of the nineteenth century when the rates could have been as high as 1,500 deaths per 100,000 individuals³.

not be determined accurately, because of the considerable estimation uncertainty for many of the parameters. We assessed the imprecision in estimating R_0 by performing an uncertainty analysis using Latin Hypercube Sampling (see Methods) and determining a distribution function for R_0 . The distribution function for R_0 shown in Fig. 3a reveals the probable range of values for R_0 when the major tuberculosis epidemics were initiated. It may be seen that the estimates of R_0 ranged from 0.74 to 18.58 with a median value for R_0 of 4.47 (see Fig. 3a legend for further statistics).

The epidemic doubling time

We calculated the doubling time of a tuberculosis epidemic during the early exponential growth phase. The doubling time (t_d) is plotted as a function of the average number of secondary infections produced per infectious case per year (Fig. 3b). Although the doubling time is extremely long if the average number of secondary infections produced per infectious case per year is low, it decreases dramatically as the average number of secondary infec-

Table 1 Model parameters: Biological interpretation and values used in the Latin Hypercube Sampling for the uncertainty analysis

Symbol	Biological Interpretation	Units	Parameter Values			Notes
			Min	Peak	Max	
$(\beta\Pi)/\mu$	Average number of infections caused by one case	/year	3	7	13	See reference 4
$1/\mu$	Average life expectancy	years	25		75	
β	Transmission coefficient	/person/year				Derived from estimate of $(\beta\Pi)/\mu$
Π	Recruitment rate	people/year				Derived from estimate of $(\beta\Pi)/\mu$
p	Proportion of new infections that develop TB within a year		0	0.05	0.30	See references 4, 8
v	Progression rate to TB	/person/year	0.00256		0.00527	This range of values corresponds to a range of 5–10% progression in 20 years ¹⁵
f	Probability of developing infectious TB (if one develops fast TB)		0.50	0.70	0.85	See references 4, 8
q	Probability of developing infectious TB (if one develops slow TB)			0.50	0.85	1.0 See references 4, 8
μ_r	Mortality rate due to TB	/person/year	0.058	0.139	0.461	The peak value corresponds to a 50% death rate in 5 years ³¹
2ω	Rate of relapse to active TB	/person/year	0	0.01	0.03	See reference 4
c	Natural cure rate	/person/year	0.021	0.058	0.086	The peak value corresponds to a 25% cure rate in 5 years ³¹

For all of the parameters except v and $1/\mu$ we used triangular distribution functions for the LHS, for v and $1/\mu$ we used uniform distributions.

tions increases. Doubling time lengthens as the epidemic progresses and exponential growth ceases.

The threshold population size

We also derived an analytical expression for the threshold population size. The threshold population size (N_r) is the minimum number of susceptibles that have to be present before a tuberculosis epidemic can occur. We then calculated historical values of the threshold population size by assuming that the historical values of several of the parameters were similar to their current values (see Fig. 4a for the parameter values). The threshold population size is plotted as a function of the transmission coefficient (β) (Fig. 4a).

Tuberculosis epidemics have very slow dynamics

Stability analysis of the detailed model demonstrates that (under constant conditions) a tuberculosis epidemic will attain a stable endemic level. The values of the endemic levels of the incidence rate and the prevalence of infection and disease can be calculated from the derived equilibrium expressions. We carried out an uncertainty analysis on the detailed model to determine the likely values of the length of a tuberculosis epidemic, where epidemic length is defined as the time it takes for an epidemic to rise, fall and reach the endemic equilibrium level. The results of the uncertainty analysis are presented as a distribution function for the epidemic length (see Fig. 4b). Tuberculosis epidemics have very slow dynamics. The length of a tuberculosis epidemic ranges from a minimum of 31 years to a maximum of 7,524 years with a median epidemic length of 100 years. Only 3% of the simulations had an epidemic length of less than fifty years (see Fig. 4b for further statistics).

Tuberculosis epidemics age

Further features of the temporal dynamics of a tuberculosis epi-

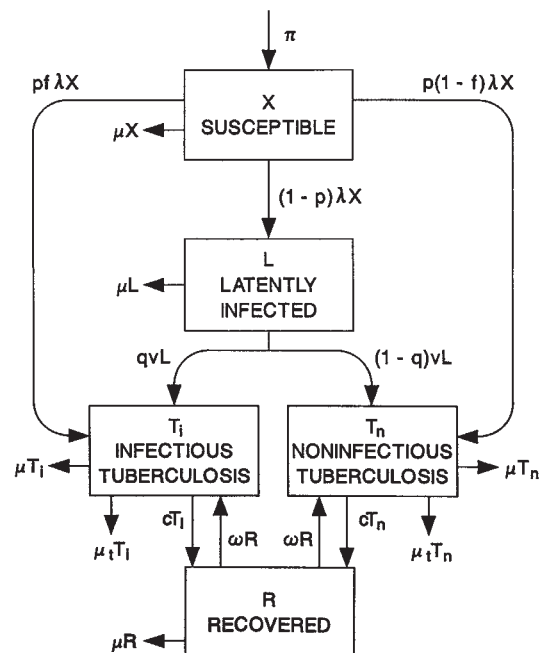


Fig. 2 The detailed tuberculosis transmission model. The model captures the temporal dynamics of five groups: susceptibles (X), latently infecteds (L), infectious cases (T_i), non-infectious cases (T_n) and recovered cases (R). The parameters are defined in Table 1 and the model is further described in Methods.

demic are illustrated by a single numerical simulation of the detailed model (Fig. 5, parameter values are listed in the legend). In this simulation, one of 1,000 performed for the uncertainty

analysis, the length of the epidemic is approximately 100 years, the declining phase of the epidemic takes approximately 30 years and the epidemic stabilizes at a high endemic level. Many of the 1,000 simulations in the uncertainty analysis declined over a longer time period and stabilized at lower endemic levels than the Fig. 5 simulation. Frequency distributions for the endemic levels of infection and disease for all 1,000 simulations will be presented elsewhere.

These simulated data illustrate that a tuberculosis epidemic can be viewed as a series of three linked time-lagged subepidemics because of the three types of infectious tuberculosis (the peak in the disease incidence for each wave can be several decades apart). Furthermore, the relative proportions of the types of tuberculosis (fast vs. slow vs. relapse) that occur during the course of an epidemic can change dramatically (fast cases will predominate in a young epidemic, whereas slow cases will predominate in a mature epidemic). Finally, the age-distribution of cases will shift towards older individuals as the relative proportion of cases that are due to endogenous reactivation increases as the epidemic matures. Age is implicitly included in our models, because 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. Consequently, the age-distribution of cases reflects the relative proportions of the type of tuberculosis. Thus the age-distribution of cases is a function of the age of the epidemic. Fig. 5 illustrates that a young and a mature epidemic have different characteristics.

Discussion

Major tuberculosis epidemics arose in Europe in the early 1600s, spread for almost two hundred years and then peaked at the end of the eighteenth century or at the beginning of the nineteenth century^{1,3}. The major epidemics in North America began after the European epidemics^{1,2} and spread westward from the eastern coastal cities³. In Europe and North America these epidemics have been in decline since at least the beginning of this century^{1,4} (Fig. 1). Our results shed some light on certain features of the historical epidemiology of tuberculosis.

Our threshold result provides a theoretical framework for understanding the historical transition to epidemic tuberculosis. Tuberculosis is an ancient disease¹⁴; however, the great epidemics of tuberculosis did not occur until the seventeenth and eighteenth centuries^{1,3}. Before these epidemics, tuberculosis could have been maintained in human populations by repeated cross-infections from other species. Population growth, urbanization and industrialization have all been proposed as factors that contributed to the rise of tuberculosis¹⁻³. Our threshold result illustrates how these three factors could have acted in concert to initiate major epidemics. Population growth would have resulted in the population size exceeding the threshold, and at this critical point R_0 (the basic reproductive number) would have become greater than one. Urbanization (by crowding people together) and industrialization (by increasing poverty, malnutrition and pollution) would have increased the transmission coefficient and thus reduced the threshold population size necessary for initiating an epidemic (see Fig. 4a). The simultaneous effects of both these processes would have resulted in the threshold population size being suddenly and dramatically exceeded. R_0 could therefore have quickly become significantly greater than one, generating major epidemics. Our uncertainty analysis results suggest that R_0 was fairly high, hence the tuberculosis epidemics were fairly severe.

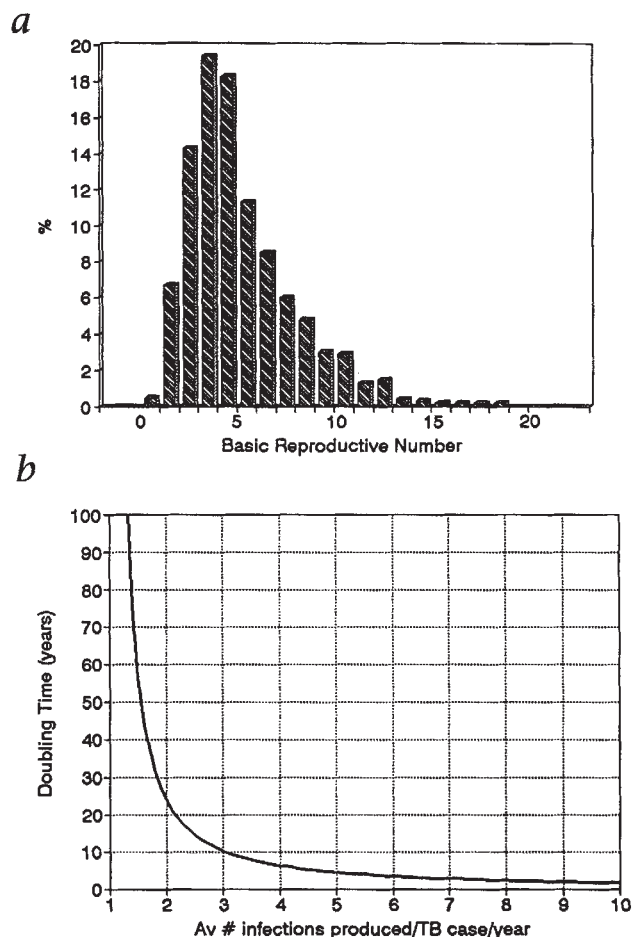


Fig. 3a, The results from the uncertainty analysis of R_0 . R_0 was calculated using equation 2. Statistics for this distribution for R_0 are as follows: minimum = 0.74, maximum = 18.58, median = 4.47, mean = 5.16, standard deviation = 2.82. b, The relationship between the initial doubling time of a tuberculosis epidemic and the average number of secondary infections caused by one infectious case per year ($\beta\pi/\mu$). To calculate this relationship we assumed that 5% of the newly infected develop tuberculosis within the first year after infection ($p = 0.05$), that 5% develop tuberculosis over the next 20 years ($v = 0.00256$), that 50% of untreated cases die within 5 years ($\mu_r = 0.139$) and that the average life expectancy ($1/\mu$) when tuberculosis epidemics first arose was 45 years. R_0 ranges from 1.01 (when $\beta\pi/\mu = 1.10$) to 9.19 (when $\beta\pi/\mu = 10$).

Epidemic time scales

Our uncertainty analysis also revealed that tuberculosis epidemics operate on extremely long time scales, in the order of one hundred to several hundred years. The declining phase of such slow epidemics will take many decades, and (if conditions remain constant) a stable endemic level of tuberculosis will be reached. Our results, therefore, suggest that part of the observed dramatic decline of tuberculosis in developed countries may simply reflect the natural behaviour of an epidemic, but that (because the epidemics did not stabilize at endemic equilibrium levels) other causal factors must have also influenced the decline. The very slow dynamics are the result of the gradual development (over several generations) of a large pool of latently infected non-infectious individuals. These latently infected individuals gradually

develop disease (and become infectious) over their lifetime, hence the disease incidence rate at any specific time reflects (in part) the infection incidence rate of several decades earlier. Therefore, any factor that acts to decrease the infection incidence rate can not produce an immediate and dramatic decline in the concurrent disease incidence rate, because the majority of new cases of tuberculosis (during the mature stage of the epidemic) are generated by the latently infected pool. Consequently, any of the factors (such as socio-economic improvement or isolation of infectious cases in sanatoria) that contributed to the decline of the major epidemics were constrained to operate within the slow response time dictated by the intrinsic transmission dynamics of tuberculosis.

Contemporary tuberculosis rates

In the United States, before the mid 1980s, tuberculosis incidence rates had been declining for decades and the United States was experiencing a mature tuberculosis epidemic¹⁵. Before 1985, the highest case rate was in individuals between 45 and 64 years of age (the oldest age group tabulated), and it was generally believed that disease was predominantly the result of endogenous reactivation^{15,16}. In 1985 the trend in declining incidence rates was reversed and the incidence rates increased between 1985 and 1992 (ref. 15). These incidence data suggest that a young epidemic (that arose because of a variety of causal factors including HIV epidemics, increased poverty, the dismantling of tuberculosis control programs and increased immigration from countries where tuberculosis is prevalent) had been superimposed upon a mature epidemic. The results from our models indicate that a young and a mature tuberculosis epidemic should behave very differently. We can make two qualitative predictions when a young epidemic is superimposed upon a mature epidemic: First, the proportion of cases due to primary progression should increase. Second, the age distribution of cases should shift to younger individuals. Surveillance data indicate that since 1985 the age distribution of cases has begun to shift to younger individuals¹⁷. Furthermore, recent molecular epidemiological studies of the transmission of tuberculosis in San Francisco and New York City have demonstrated that an increasing proportion of newly diagnosed cases result from recent transmission and direct progression^{18,19}. Hence our two qualitative predictions are substantiated by the surveillance data and the results from the molecular epidemiological studies.

We have used transmission models as tools^{20,21} to obtain a qualitative and a quantitative understanding of the intrinsic transmission dynamics of tuberculosis. The first tuberculosis transmission model was published in 1962 by Waaler, Geser and Andersen²². In the intervening thirty years many other tuberculosis models have been published (see for example refs 23,24). Our analyses differ from previous analyses in two important respects: Previous studies have used models to evaluate control strategies rather than to understand the intrinsic dynamics of an epidemic and they have focused upon computer simulations of scenario analyses for specific parameter values rather than upon deriving R_0 or other analytical results. The approach that we have adopted has enabled us to obtain quantitative answers to explain the historical epidemiology of tuberculosis and to identify the mechanisms that drive tuberculosis epidemics. We have used simple models and now more complex models may be analysed. However, our results are likely to remain robust, because simple and complex disease models often generate similar predictions²⁵⁻²⁷.

We have developed a theoretical framework for understanding the historical rise and fall of tuberculosis epidemics. However,

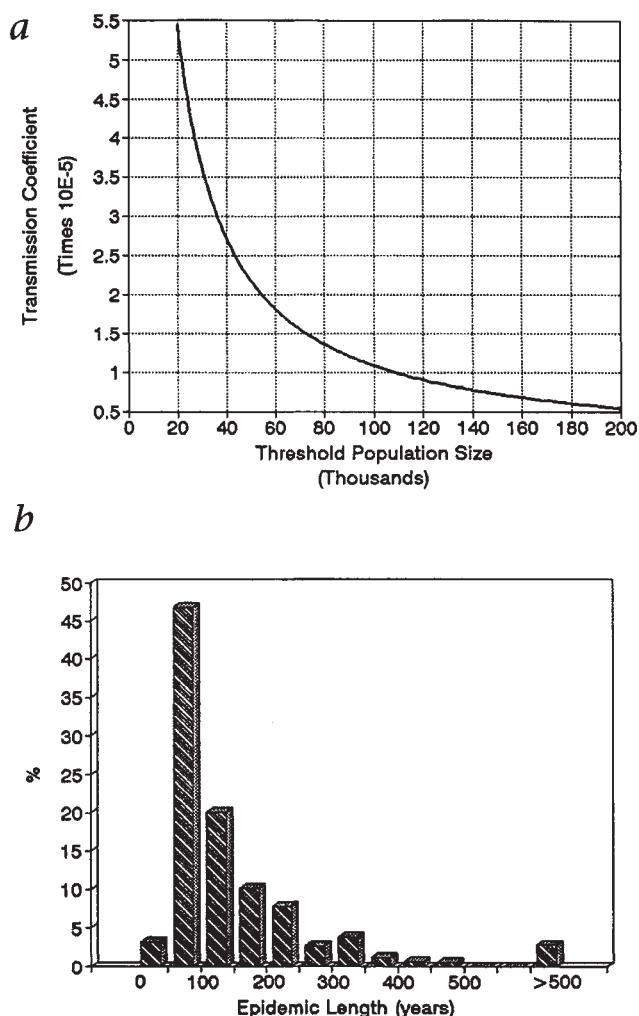


Fig. 4a, The relationship between the threshold population size and the transmission coefficient. The transmission coefficient is given in units of transmissibility per person per year. To calculate this relationship we assumed that 5% of infected individuals developed tuberculosis within a year ($p = 0.05$), 5% of infected individuals developed tuberculosis over a period of 20 years ($v = 0.00256$), 50% of untreated tuberculosis cases died in 5 years ($\mu_1 = 0.139$) and that the average life expectancy when tuberculosis epidemics first arose was 45 years. b, The results from the uncertainty analysis of the frequency distribution of the length of a tuberculosis epidemic are shown. The length of a tuberculosis epidemic was calculated from the time of the start of the epidemic to the point where the infection incidence rate approached (within one person year) the value of the analytically derived equilibrium infection incidence rate. Statistics for this distribution are as follows: minimum = 31 years, maximum = 7,524 years, median = 100 years, mean = 163 years, standard deviation = 320 years.

this framework may also provide insight into the dynamics of tuberculosis in developing countries where treatment of tuberculosis is not yet the norm. Our results also have significant implications for disease control. We have shown that tuberculosis epidemics have very slow intrinsic dynamics. These dynamics ensure that if control strategies focus only on preventing the occurrence of new infections it may take decades to eliminate tuberculosis. We have demonstrated that tuberculosis epidemics

may be viewed as a series of linked time-lagged subepidemics. These results imply that different control strategies may be necessary for controlling each subepidemic and that different control strategies may have to be employed as an epidemic ages. The theoretical framework that we have developed can now be used to understand the contemporary dynamics of tuberculosis and to design new disease control strategies.

Methods

A simple tuberculosis transmission model. The model captures the temporal dynamics of three groups: susceptible individuals (X), latently infected individuals (L) and active infectious tuberculosis cases (T). Recruitment to the susceptible population occurs at a constant rate Π . The incidence rate of infection is calculated as the product of the number of susceptibles that are present at time t ($X(t)$) and the *per capita* force of infection at time t ($\lambda(t)$); where $\lambda(t)$ is defined as the *per-susceptible* risk of becoming infected with *M. tuberculosis*. Since tuberculosis is a directly transmitted air-borne infection, $\lambda(t)$ is calculated as the product of the number of infectious cases that are present at time t ($T(t)$) and the transmission coefficient (β) of the pathogen (where β reflects the likelihood that an infectious case will successfully transmit the infection to a susceptible individual). The *per capita* average non-tuberculosis mortality rate is μ . Hence, the following equation specifies the instantaneous rate of change in the number of susceptibles: $dX/dt = \Pi - \lambda X - \mu X$.

Infected individuals remain non-infectious until they develop disease by one of two pathogenic mechanisms: direct progression (soon after infection with *M. tuberculosis*) or endogenous reactivation (several years after infection). The two pathogenic mechanisms are modelled by allowing a proportion (p) of the newly infected to develop tuberculosis directly and a proportion ($1 - p$) of the newly infected to enter the latent class. Latently infected individuals will either develop tuberculosis slowly at an average rate v or they will die of other causes at an average rate μ before developing tuberculosis. Only a small fraction of latently infected individuals will develop tuberculosis, because the progression rate to disease is extremely slow. The number of latently infected individuals, L , therefore, obeys the equation: $dL/dt = (1 - p)\lambda X - vL - \mu L$.

Two types of tuberculosis (where the types are defined on the basis of the pathogenesis of the infection) contribute to the incidence rate of tuberculosis disease: one type of tuberculosis develops through endogenous reactivation in the latently infected individuals (slow tuberculosis) and the other type of tuberculosis develops because of direct progression soon after infection (fast tuberculosis). Active infectious tuberculosis cases die, either because of tuberculosis at an average rate μ , or because of other causes at an average rate μ . Hence, the third and final equation of the model is: $dT/dt = vL + p\lambda X - (\mu + \mu_r)T$.

Expressions for the equilibrium number of susceptibles ($\hat{X}$), latently infecteds ($\hat{L}$) and active tuberculosis cases ($\hat{T}$) were calculated from the model; these expressions enable the determination of the expected endemic levels of tuberculosis: $\hat{X} = \Pi/(\mu R_0)$, $\hat{L} = [(1 - p)\Pi(1 - (1/R_0))]/(v + \mu)$ and $\hat{T} = (\mu/\beta)(R_0 - 1)$.

A detailed tuberculosis transmission model. The detailed model captures the temporal dynamics of five groups: susceptibles (X), latently infecteds (L), infectious cases (T_i), non-infectious cases (T_n) and recovered cases (R). A flow diagram of the model is shown in Fig. 2. The detailed model differs from the simple model in that it includes three refinements that reflect additional biological complexities of the natural history of tuberculosis: (1) Only a certain fraction of cases are assumed to be infectious (a fraction f of cases that develop tuberculosis because of primary progression and a frac-

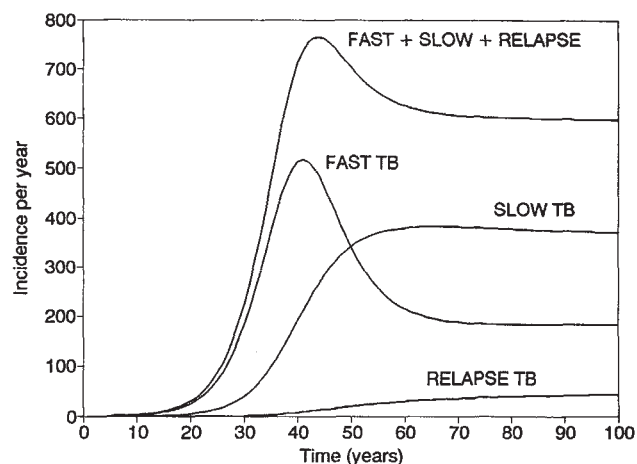


Fig. 5 A numerical simulation of a tuberculosis epidemic using the detailed transmission model is shown. The simulation illustrates the relative contribution of the three types (fast, slow and relapse) of tuberculosis to the disease incidence rate. It should be noted that the decline occurs in the absence of any changes in any of the parameters, and are simply due to the intrinsic transmission dynamics. The following parameter values were used: $\Pi = 4,400$, $\mu = 0.0222$, $p = 0.05$, $v = 0.00256$, $f = 0.70$, $q = 0.85$, $\omega = 0.005$, $\mu_r = 0.139$, $c = 0.058$, $\beta = 0.00005$. These parameter values result in an average number of infections produced per tuberculosis case per year ($\beta\Pi/\mu$) of 10. The epidemic was initiated by entering one infectious case of tuberculosis at time zero in a susceptible population of 200,000 (which is the disease free equilibrium level).

tion q of cases that develop tuberculosis because of endogenous reactivation); (2) a case may be spontaneously cured (that is, without treatment) at a *per capita* rate c and move into the recovered non-infectious state (R); and (3) an individual in the recovered state may either relapse (and develop, with equal probability, either infectious or non-infectious tuberculosis at the *per capita* rate ω , hence the rate of relapse to active tuberculosis is 2ω) or may never relapse and die of other causes at the average rate μ . The following five equations specify the model ($\lambda = \beta T_i$ and the remaining parameters are as previously defined): $dX/dt = \Pi - \lambda X - \mu X$, $dL/dt = (1 - p)\lambda X - (v + \mu)L$, $dT_i/dt = pf\lambda X + qvL + \omega R - (\mu + \mu_r + c)T_i$, $dT_n/dt = p(1 - f)\lambda X + (1 - q)vL + \omega R - (\mu + \mu_r + c)T_n$ and $dR/dt = c(T_i + T_n) - (2\omega + \mu)R$.

The epidemic doubling time. We calculated (equation 3, below) the initial growth rate (Λ) during the early exponential phase of a tuberculosis epidemic using methodology described elsewhere²⁸.

$$\Lambda = - \frac{(D_1(1 - R_0^f) + D_2) \pm \sqrt{(D_1(1 - R_0^f) + D_2)^2 + 4D_1D_2(R_0^f + R_0^s - 1)}}{2D_1D_2} \quad (3)$$

where:

$$D_1 = \left[\frac{1}{\mu + v} \right]$$

$$D_2 = \left[\frac{1}{\mu + \mu_r} \right]$$

$$R_0^f = R_0^{FAST}$$

$$R_0^s = R_0^{SLOW}$$

The initial growth rate was used to calculate the initial epidemic doubling time (t_d) as shown below:

$$t_d = \frac{\ln 2}{\Lambda} \quad (4)$$

The threshold population size. The analytical expressions for R_0 (calculated from the simple model) were rearranged to assess the threshold population size (N_T):

$$N_T = \left(\frac{\mu + \mu_r}{\beta} \right) \left(\frac{v + \mu}{\mu p + v} \right) \quad (5)$$

Uncertainty analysis: R_0 and epidemic length. Latin Hypercube Sampling (LHS) has been used previously to determine prediction imprecision in estimates derived from transmission models^{29,30}. We performed an uncertainty analysis by using LHS to obtain 1,000 samples from each of the parameter distribution functions (pdfs) given in Table 1; the LHS method is described in detail elsewhere^{29,30}. LHS ensured that each pdf was independently randomly sampled without replacement; before sampling, each pdf was stratified into 1,000 equiprobability areas. The value of R_0 was then calculated by using equation (2) and the parameter estimates obtained by LHS. Therefore, the results of this analysis were 1,000 estimates of R_0 ; these results were plotted as a distribution function of R_0 . We then performed 1,000 simulations (using the LHS parameter values). For each simulation, first we calculated the endemic infection incidence rate (from our analytical results), then we simulated the epidemic, and finally we measured the length of the epidemic (which we operationally defined as the number of years the epidemic took to reach the point at which the simulated infection incidence rate approached (within one person year) the value of the analytically derived endemic infection incidence rate). In each of the 1,000 simulations only one epidemic occurred before the system reached equilibrium (that is, the equilibrium was always approached monotonically). The results from the uncertainty analysis of the epidemic length were plotted as a distribution function.

Acknowledgements

We thank Nelson Freimer, Bob May, Herb Hethcote, Carlos Castillo-Chavez and John Jacquez for useful comments during the course of this work, and George Cauthen for providing us with access to historical tuberculosis mortality data from New York City. We are grateful to the Robert Wood Johnson Foundation, The Royal Society, National Institutes on Drug Abuse (Grant No. 1R29DA08153) and National Institute of Allergy and Infectious Diseases (Grant No. AI33831 and No. K08AI01137-01) for financial support.

RECEIVED 17 JANUARY; ACCEPTED 12 JUNE 1995

- Long, E.R. The decline of tuberculosis, with special reference to its generalized form. *Bull. Hist. Med.* **8**, 819–843 (1940).
- Ransome, A. *Researches on Tuberculosis* 1–11 (Smith, Elder & Co., London, 1898).
- Grigg, E.R.N. The arcana of tuberculosis: With a brief epidemiologic history of the disease in the USA - Parts I, II & III. *Am. Rev. Tuberc. pulm. Dis.* **78**(2), 151–172, 426–453, 583–603 (1958).
- Styblo, K. in *Advances in Respiratory Medicine* (eds Flenley, D.C. & Retly, J.L.) 77–108 (Churchill Livingstone, Edinburgh, 1986).
- McKeown, T. & Record, R.G. Reasons for the decline of mortality in England and Wales during the nineteenth century. *Population Studies* **16**, 94–122 (1962).
- Newsholme, A. *The Prevention of Tuberculosis* 269–276 (Methuen & Co., London, 1908).
- Stead, W.W. Genetics and resistance to tuberculosis: Could resistance be enhanced by genetic engineering? *Ann. intern. Med.* **116**, 937–941 (1992).
- Comstock, G.W. Epidemiology of tuberculosis. *Am. Rev. respir. Dis.* **125**(2), 8–16 (1982).
- Lurie, M.B. *Resistance to Tuberculosis* 30–41 (Harvard Univ. Press, Cambridge, Massachusetts, 1964).
- Nardell, E., McInnis, B., Thomas, B. & Weidhass, S. Exogenous reinfection with tuberculosis in a shelter for the homeless. *New Engl. J. Med.* **315**, 1570–1575 (1986).
- Small, P.M. *et al.* Exogenous reinfection with multidrug-resistant *Mycobacterium tuberculosis* in patients with advanced HIV infection. *New Engl. J. Med.* **328**, 1137–1144 (1993).
- Macdonald, G. The analysis of equilibrium in malaria. *Trop. Dis. Bull.* **49**, 813–829 (1952).
- Anderson, R.M. & May, R.M. *Infectious Diseases of Humans: Dynamics & Control* 17–19 (Oxford Univ. Press, Oxford, 1991).
- Ryan, F. *The Forgotten Plague* (Little, Brown & Co., Boston, 1992).
- Comstock, G.W. & Cauthen, G.M. in *Tuberculosis: A Comprehensive International Approach* (eds Reichman, L.J. & Hershfield, E.S.) 23–47 (Marcel Dekker Inc., New York, 1993).
- Centers for Disease Control. Tuberculosis Statistics in the United States (1990).
- Centers for Disease Control. Tuberculosis Statistics in the United States (1992).
- Small, P.M. *et al.* The epidemiology of tuberculosis in San Francisco: A population-based study using conventional and molecular methods. *New Engl. J. Med.* **330**, 1703–1709 (1994).
- Alland, D. *et al.* Transmission of tuberculosis in New York City: An analysis by DNA fingerprinting and conventional epidemiologic methods. *New Engl. J. Med.* **330**, 1710–1716 (1994).
- Blower, S.M. & Medley, G. Epidemiology, HIV & drugs: Mathematical models & data. *Br. J. Addiction* **87**, 31–39 (1992).
- Blower, S.M. & McLean, A.R. Prophylactic vaccines, risk behavior change and the probability of eradicating HIV in San Francisco. *Science* **265**, 1451–1454 (1994).
- Waalder, H.T., Geser, A. & Andersen, S. The use of mathematical models in the study of the epidemiology of tuberculosis. *Am. J. publ. Health* **52**, 1002–1013 (1962).
- Waalder, H.T. Cost-benefit analyses of BCG vaccination under various epidemiological situations. *Bull. int. Union Tuberc.* **41**, 42–52 (1968).
- Waalder, H.T. & Piot M.A. The use of an epidemiological model for estimating the effectiveness of tuberculosis control measures. Sensitivity of the effectiveness of tuberculosis control measures to the coverage of the population. *Bull. World Health Org.* **41**, 75–93 (1969).
- May, R.M. Population biology of microparasitic infections. in *Mathematical Ecology: An Introduction* (eds Hallan, T.G. & Levin, S.W.) 405–442 (Springer-Verlag, New York, 1986).
- McLean, A.R. & Anderson, R.M. Measles in developing countries. Part I. Epidemiological patterns and parameters. *Epidemiol. Infect.* **100**, 111–33 (1988).
- McLean, A.R. & Anderson, R.M. Measles in developing countries. Part II. The predicted impact of mass vaccination. *Epidemiol. Infect.* **100**, 419–42 (1988).
- May, R.M. & Anderson, R.M. Parasite-host coevolution. *Parasitology* **100**, S89–S101 (1990).
- Blower, S.M., Hartel, D., Dowlatabadi, H., May, R.M. & Anderson, R.M. Drugs, sex & HIV: A mathematical model for New York City. *Phil. Trans. R. Soc. Ser. B* **321**, 171–187 (1991).
- Blower, S.M. & Dowlatabadi, H. Sensitivity & Uncertainty Analysis of complex models of disease transmission: An HIV model, as an example. *Int. Statist. Rev.* **62**, 229–243 (1994).
- Grzybowski, S. & Enarson, D.A. The fate of cases of pulmonary tuberculosis under various treatment programmes. *Bull. int. Union Tuberc.* **53**, 70–75 (1978).