MATHEMATICAL BIOSCIENCES
AND ENGINEERING
Volume **6**, Number **2**, April **2009**

doi:10.3934/mbe.2009.6.209

pp. **209–237**

MATHEMATICAL MODELLING OF TUBERCULOSIS EPIDEMICS

Juan Pablo Aparicio

School of Science and Technology, Universidad Metropolitana
San Juan 00928-1150, Puerto Rico
and
Instituto de Investigación en Energías no Convencionales
Universidad Nacional de Salta, 4400 Salta, Argentina

Carlos Castillo-Chavez

School of Human Evolution and Social Change
PO Box 872402 Tempe, AZ 85287-2402, USA
and
Mathematical, Computational and Modeling Sciences Center
PO Box 871904 Tempe, AZ 85287-1904, USA
and
School of Mathematical and Statistical Sciences, Arizona State University
PO BOX 871804 Tempe, AZ 85287-1804, USA
and
Santa Fe Institute
1399 Hyde Park Road Santa Fe, NM 87501, USA

Abstract. The strengths and limitations of using homogeneous mixing and heterogeneous mixing epidemic models are explored in the context of the transmission dynamics of tuberculosis. The focus is on three types of models: a standard incidence homogeneous mixing model, a non-homogeneous mixing model that incorporates 'household' contacts, and an age-structured model. The models are parameterized using demographic and epidemiological data and the patterns generated from these models are compared. Furthermore, the effects of population growth, stochasticity, clustering of contacts, and age structure on disease dynamics are explored. This framework is used to asses the possible causes for the observed historical decline of tuberculosis notifications.

1. **Introduction.** Tuberculosis (TB) is an ancient disease that still supports huge levels of prevalence across the world. The urbanization process that exploded after the Industrial Revolution has played a fundamental role in the observed patterns of TB spread in industrialized nations [37, 38, 39]. TB is a disease with slow dynamics and consequently, TB epidemics must be studied and assessed over extremely long windows in time. Transmission takes place primarily within small clusters of acquaintances. Thus TB patterns are driven by processes whose dynamics take place over distinct temporal, spatial and organizational scales. TB in the US is in the declining phase of an epidemic that peaked around the middle of the nineteenth

2000 *Mathematics Subject Classification.* Primary: 92D25, 92D30; Secondary: 92C60.

Key words and phrases. tuberculosis, non-autonomous systems, stochastic models, demography.

JPA is a member of the CONICET.

century. Full understanding of the actual course of TB dynamics in the US requires an appropriate description of historical epidemic patterns. Such description requires a sound modeling approach of epidemic processes that involve hundreds of millions of individuals over centuries. Simple standard compartmental models with some modifications turn out to be quite useful in the study of the long-term dynamics of TB, particularly when the age structure of a population is included. Naturally, the study of all processes associated with transmission do not need to be incorporated in detail. The selection of a model is (or it should be) intimately connected to the question. For example, a study of the competition dynamics between strains might benefit from the incorporation of a detailed description of the network of contacts between individuals or groups. Individual based models (which follow every individual in a population) may provide a most useful tinkering tool in the study of the role of social networks on TB transmission.

Different models are used to study the long-term dynamics of tuberculosis in this manuscript. The list includes simple compartmental models with standard incidence (extensions of the models in [3, 6]), aggregated cluster models (extensions of the models in [1, 53]), and age structured models. We focus our attention on the use of demographic and epidemiological data to parameterize these models as they are used to capture the patterns of TB over long time horizons. Our manuscript concludes with the use of these models in some preliminary applications, a discussion on the possible reasons behind TB decline in the US, and a discussion of challenges associated with TB dynamics. In some sense, the work in this manuscript complements and expands the work that we published in a comprehensive review that appeared in this Journal [21].

1.1. **Basic tuberculosis epidemiology.** Tuberculosis is mostly transmitted through the air by persons coughing with pulmonary tuberculosis. The probability of transmission per contact, per relevant unit of time is in general quite low [9, 22, 43]. Individuals at high risk of infection include those who are frequently exposed for long period of times to infectious individuals.

Infected individuals may remain asymptomatic over their entire lives (latent TB). Active-TB (the clinical disease) can develop into pulmonary and extrapulmonary forms. Extrapulmonary TB is common in children while pulmonary TB is frequent in adults. *Mycobacterium tuberculosis*, the causal agent of the disease, is transmitted almost exclusively via pulmonary cases (exceptions could include laryngeal TB). Cases arising within five years after infection are classified as primary tuberculosis while cases arising after five years from first infection are known as secondary tuberculosis. Endogenous reactivation cases (exacerbation of an old infection) or exogenous tuberculosis cases (the result of TB activation due to reinfection) are also classified as secondary tuberculosis. This distinction is standard but somewhat arbitrary.

The number of new cases of active TB decline almost exponentially when view as a function of age of infection [56] (see Fig. 6). In a ten year study (reviewed in [56]) it was noted that nearly 60% of the new cases arose during the first year following infection while the cumulative number of cases generated over the first five years after infection accounted for nearly 95% of the total cases observed. If this exponential decline in progression risk were to be maintained over the life time of individuals in the population then the contribution of endogenous reactivation to progression would be small, less than the 5%. However, increases in the risk of endogenous reactivation in the elderly have been observed. Reactivation in the

elderly is due to multiple reasons including those responsible for immune system depression.

There are differences in infectiousness. Pulmonary tuberculosis cases that are smear positive are significantly more infectious than the smear negative (culture positive or culture negative cases) [51, 56].

Heterogeneity comes from many directions including from the variability in infectious periods. However, it is known that most of the secondary infections generated by a source case do take place within the first months following TB activation[8]. That is, these cases are generated at a time when the disease may not be apparent. Case fatality among untreated pulmonary TB cases is around 50%. Recovered individuals, naturally or from treatment, may develop active TB again, a phenomenon known as TB relapse.

2. **Aggregated models.**

2.1. **Epidemiological states.** In order to capture key relevant complexities in the study of the transmission dynamics of TB, we consider seven classes (Fig. 1). Uninfected individuals, the U class; First-infected individuals, members of the high-risk latent class E, that is, individuals who are assumed to be asymptomatic and non-infectious but capable of "quickly" progressing to the clinical disease (or active-*TB*) at the per-capita rate k; Individuals who do not progress to the active *TB* class quickly enough from the class E are moved to the low-risk latent class L at the per capita rate α. The use of high- and low-risk latent classes captures in a simple way the observed patterns of fast TB progression that characterize a small fraction of recently infected individuals (see Section 4.2 below). Individuals in the low-risk class L progress to active TB at the per-capita rate k_L where they are assumed to be susceptible to reinfection. Re-infection may accelerate progression to active TB. Therefore we move re-infected individuals to a new high-risk class E^* where they progress to active TB class at the per-capita rate k^*. Individuals can escape progression (returning to the class L) at the same per capita rate α as individuals in the class E. Previous infections with *M. tuberculosis* afford some protection that may translate into reduced susceptibility to reinfection and/or reduced risk of progression to the class E^*, of high-risk latent TB ($k^* < k$). New cases of active TB are classified as pulmonary with probability q, collected in the A_p class. Extrapulmonary cases (A_e) are assumed to be not infectious. Recovered individuals (class R) may develop active TB again (TB-relapse) at the per capita rate k_{Rp}.

Uninfected individuals are recruited at the time dependent rate $B(t)$. Non-TB related mortality is μ while d and d_e represent pulmonary and extrapulmonary tuberculosis mortalities respectively. Therefore removal rates in the pulmonary and extrapulmonary active TB compartments are given by $\gamma = (\mu + d + r)$ and $\gamma_e = (\mu + d_e + r_e)$ where r and r_e are the corresponding recovery rates.

2.2. **Cluster models.** The use of a different unit, *generalized household*, in modeling transmission has proved useful [1, 53]. The *generalized household* (or in short the *cluster*) of an individual is by definition the group of his/her close and/or frequent contacts. Members of a generalized household include close "associates" such as house-mates, co-workers and/or classmates. Most of TB infections are generated in active clusters, that is, in a cluster or generalized household that includes an

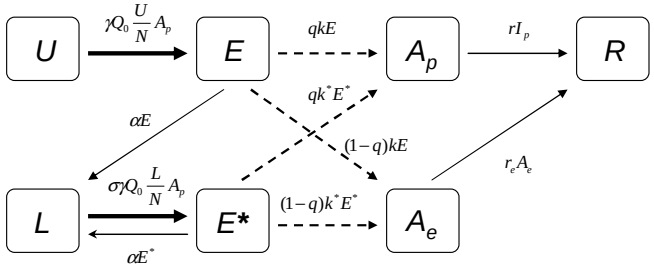

FIGURE 1. Transfer Diagram. Individuals are recruited into the uninfected class U; moved into the high-risk latent class E after the infection from where they progress to active TB. A fraction q develops pulmonary TB (A_p) while the complementary fraction 1-q develops extrapulmonary TB (A_e). E individuals move to the low-risk latent class L or to A_p or A_e. L individuals may become reinfected and move to the high-risk class E^* from there they may develop active TB, or may return to the low-risk class L. Active cases if they recover move to the R class. Transmission is represented in the above transfer diagram using thick arrows and progression with dashed arrows. Progression from low-risk latent class L to A_p (at the rate $qk_L L$) and to A_e (at the rate $(1-q)k_L L$) are not shown. Infection and reinfection rates (see expressions (15) and (16)) are also displayed.

infectious individual. It is possible to build deterministic models that account extremely well for within cluster transmission. Here, we briefly review the main ideas behind a modeling approach that we introduced in previous works (for details see [1, 53]).

Infectious individuals are rare in developed countries. Hence, it is reasonable to assume that there is only one infectious individual per generalized household. That is, we assume that new cases of active TB will appear after the recovery (or death) of the source case, and that the probability of belonging to two or more different active cluster at the same time is negligible. This approximation may be invalid for countries, or historical periods, where higher incidence levels are observed.

Active generalized households are by definition households that include a single infectious individual. $N_c(t)$ denotes the number of individuals in active clusters, the active 'cluster population' at risk, from close contacts, of acquiring TB infections. β is used to denote the per susceptible mean risk of infection, the risk that comes just for being a member of an active generalized household (close contacts). When infectious individuals recover or die (at the per-capita rate γ), the risk of infection becomes null in the corresponding generalized household. In such a case a group

of size $n+1$ (or just n if the infectious individual dies) is deactivated, that is, group members are brought back to the 'general population' of size $N_{nc}(t)$, the population of individuals who are not members of an active generalized households, where n denotes the mean generalized household size. $N_{nc}(t)-$individuals can only acquire a TB (latent) infection via random (casual) contacts with members of the overall population. We use the subscript 1 to identify non-infectious epidemiological classes in the population of active clusters (where transmission is extremely likely to occur), and subscript 2 for the corresponding classes in the general population (where transmission is possible through casual, random and rare contacts). The risk of acquiring TB infection from individuals in the general population is not zero but very low. Latently-infected individuals do progress to the class of pulmonary active TB at the rate $q(kE_2 + k^*E_2^* + k_L L_2 + k_{Rp}R_2)$. With the 'birth' of a new infectious individual, a cluster of size n is moved, from the general population to the population of active clusters. Extrapulmonary cases are not explicitly tracked in the cluster formation processes but we account for them here any way. Therefore we set $N_i \equiv U_i + E_i + L_i + E_i^* + R_i$, $i = 1, 2$, $N_c \equiv N_1 + I$, and $N_{nc} \equiv N_2 + A_e$. Using the above definitions, assumptions and notation leads to the following non-linear system of equations for the transmission dynamics of TB in clusters and (possible) via random, rare, casual contacts (see [1, 53])

$$\frac{dU_1}{dt} = -(\beta+\gamma)U_1 + F(U_1, U_2), \tag{1}$$

$$\frac{dE_1}{dt} = \beta U_1 - (\alpha+\gamma)E_1 + F(E_1, E_2), \tag{2}$$

$$\frac{dL_1}{dt} = \alpha(E_1 + E_1^*) - (\sigma\beta+\gamma)L_1 + F(L_1, L_2), \tag{3}$$

$$\frac{E_1^*}{dt} = \sigma\beta L_1 - (\alpha+\gamma)E_1^* + F(E_1^*, E_2^*), \tag{4}$$

$$\frac{dR_1}{dt} = -\gamma R_1 + F(R_1, R_2), \tag{5}$$

$$\frac{dA_p}{dt} = q(kE_2 + k^*E_2^* + k_L L_2 + k_{Rp}R_2) - \gamma A_p, \tag{6}$$

$$\frac{dU_2}{dt} = B(t) - \mu U_2 + \gamma U_1 - F(U_1, U_2), \tag{7}$$

$$\frac{dE_2}{dt} = -(\mu+\alpha+k)E_2 + \gamma E_1 - F(E_1, E_2), \tag{8}$$

$$\frac{dL_2}{dt} = \alpha(E_2 + E_2^*) + \gamma L_1 - (\mu+k_L)L_2 - F(L_1, L_2), \tag{9}$$

$$\frac{dE_2^*}{dt} = \gamma E_1^* - (k^*+\mu+\alpha)E_2^* - F(E_1^*, E_2^*), \tag{10}$$

$$\frac{dR_2}{dt} = rA_p + r_n A_e + \gamma R_1 - (\mu+k_{Rp})R_2 - F(R_1, R_2), \tag{11}$$

$$\frac{dA_e}{dt} = (1-q)(kE_2 + k^*E_2^* + k_L L_2 + k_{Rp}R_2) - \gamma_e A_e, \tag{12}$$

where $F(X_1, X_2)$ is a short notation for the expression

$$F(X_1, X_2) = q(kE_2 + k^*E_2^* + k_L L_2 + k_{Rp}R_2)n(\rho X_1/N_1 + (1-\rho)X_2/N_2). \tag{13}$$

Most new cases of active TB arise just a few years after infection or re-infection. Therefore only a small fraction of the members of an activated generalized household, generated by the 'birth' of a new (source) case, may also belong to the active generalized household where the source case was infected, overlap is possible. This overlap between successive active households is modelled using the clustering coefficient ρ in (13).

The *contact number* is defined as the average number of infections caused by one active case (during the whole infectious period), when all contacts are with uninfected individuals. When the clustering coefficient is zero, the contact number obtained from the cluster (generalized household) model is

$$Q_0 = \frac{\xi n}{1+\xi} \tag{14}$$

where $\xi = \beta/\gamma$ is the average (non-dimensional) risk of infection over the mean infectious period $1/\gamma$. This number is not the basic reproductive ratio because only a fraction of the infections will become actively infectious cases. When the clustering coefficient is greater than zero, a more useful definition of the contact number may be worded as "the number of secondary infections caused by one infectious individual in the second generation [1, 7]". In the generalized household case, the contact number is effectively reduced by the effect of repeated contacts within the same group of individuals.

Incorporation of casual infections (reinfections) is straightforward [1]. We assume that infectious individuals also produce infections from casual contacts (the result of random encounters) with individuals that do not belong to his/her generalized household. The probability of casual infections among susceptible individuals who are in an active generalized households is neglected. The rate of casual infections produced is $m\beta A_p \frac{S_2}{N_2}$ while the reinfection (casual) rate is $m\sigma\beta A_p \frac{L_2}{N_2}$. These terms move individuals from the S_2 to the E_2 and from the L_2 to the E_2^* classes, respectively. The non-dimensional parameter m is related to the mean size of the network of close casual contacts of a typical infectious individual [1] while σ accounts for the possible protection conferred to an individual from previous infections.

2.3. **Standard compartmental models.** Our definition of contact number requires the generation of Q_0 secondary infections by an average infectious individual placed in a totally susceptible population. If only the fraction U/N of his/her contacts is with uninfected individuals then the number of (first) infections produced is given by Q_0U/N and the infection rate per unit of time is obtained by dividing this number by the infectious period $1/\gamma$. Explicitly, when we start from a simpler standard incidence epidemiological model, we get

$$\gamma Q_0 \frac{U}{N} A_p. \tag{15}$$

Similarly, the re-infection rate is given by

$$\sigma\gamma Q_0 \frac{L}{N} A_p \tag{16}$$

where σ models the possible protection conferred by previous infections ($\sigma \leq 1$).

A simpler aggregate deterministic model for TB transmission that includes vital dynamics is given by the following system (see also Fig. 1)

$$\frac{dU}{dt} = B(t) - \mu U - \gamma Q_0 \frac{U}{N} A_p, \tag{17}$$

$$\frac{dE}{dt} = \gamma Q_0 \frac{U}{N} A_p - (k + \mu + \alpha)E, \tag{18}$$

$$\frac{dL}{dt} = \alpha(E + E^*) - (\mu + k_L)L - \sigma\gamma Q_0 \frac{L}{N} A_p, \tag{19}$$

$$\frac{dE^*}{dt} = \sigma\gamma Q_0 \frac{L}{N} A_p - (k^* + \mu + \alpha)E^*, \tag{20}$$

$$\frac{dA_p}{dt} = q(kE + k^*E^* + k_L L + k_{Rp}R) - \gamma A_p, \tag{21}$$

$$\frac{dA_e}{dt} = (1 - q)(kE + k^*E^* + k_L L + k_{Rp}R) - \gamma_e A_e, \tag{22}$$

$$\frac{dR}{dt} = rA_p + r_e A_e - (\mu + k_{Rp})R, \tag{23}$$

where $\gamma = \mu + d + r$ and $\gamma_e = \mu + d_e + r_e$. In this simplified version recruitment brings individuals only to the uninfected class. Models with immigration must include the arrival of infected individuals (see also 4.1 below). A secondary case is expected to share some contacts with the source case (from where he/she caught the infection), hence the rate of first infections is in general lower than (15) while the rate of reinfection is typically higher than (16). These discrepancies are "corrected" with the (phenomenological) clustering parameter ρ (see expression (13)), a factor that is 'lost' in the above simpler model.

3. **Age structured models.** Age plays an important role in tuberculosis dynamics for several reasons including the fact that most cases of active tuberculosis in children are non infectious [30]. Furthermore, TB mortality among children less than one year of age was extremely high before the chemotherapy era. In addition, active tuberculosis in developed countries has been seen or assumed to be as the result of the endogenous reactivation of old infections among primarily middle-aged or older individuals.

An age structured version of Model (17) is obtained straightforwardly (see for example [17]). In this version, the evolution of the dynamical variables (here only for the uninfected population) is given by

$$\frac{\partial}{\partial t}u(t,a) + \frac{\partial}{\partial a}u(t,a) = \int_0^\infty \gamma(t,a')Q_0(t,a')a_p(t,a')\frac{u(t,a)}{N_a(t)}da' \tag{24}$$

where $\gamma(t,a)Q_0(t,a)$ accounts for possible variations of infectiousness and pattern of mixing with age. Variables like $u(t,a)$ or $a_p(t,a)$ represent the *densities* of individuals. Thus the number of uninfected individual with ages between a and $a + \Delta a$ is given by $\int_a^{a+\Delta a} u(t,b)db$. The total uninfected population used in the aggregated models (given by Equations(17)-(23)) can be recovered by integrating the age structured model over all ages $\left(U(t) = \int_0^\infty u(t,a)da\right)$ under the assumption that the parameters are age-independent. The problem is completely specified with the inclusion of an initial age distribution (for example $u(t_0,a)$ that includes an initial completely uninfected population) and the renewal condition $u(t,0) = \int_0^\infty b(t,a)N(t,a)da$, where $b(t,a)$ is the age-specific per capita birth rate and $N(t,a)$ is the population of individuals of age a at time t. Numerical solutions can be computed, for example, through the use of a simple forward schemes like:

$$U(t + \Delta t, a + \Delta t) = s(t,a)U(t,a) - \Delta t\gamma Q_0 I\frac{U(t,a)}{N} \tag{25}$$

$$U(t,0) = b(t)N(t)\Delta t \tag{26}$$

where $U(t,a) = \int_a^{a+\Delta t} u(t,a')da'$, $b(t)$ is the time-dependent per capita birth rate, $N(t)$ is the total population, and $s(t,a)$ is the survivorship function, that is, the function that gives the fraction of the population of age a surviving after a period of time Δt. In our numerical studies, we assumed that neither γ nor Q_0 are age-dependent. In our simulations of the age-structured model the population is divided in $A/\Delta t$ age classes, where A is the maximum age considered. After an initial age distribution is specified, the numerical scheme outlined in (25)-(26) is used to follow the time evolution of each cohort.

4. **Available data and parameterization.**

4.1. **Demography.** For the entire United States birth and death rates estimates are available since 1900. For Massachusetts the data goes back to 1850 (U.S. Bureau of the Census, 1908, 1975). The observed total mortality rate (per person, per year) may be well represented by the smooth function of the calendar year t (see Fig. 2, left panel; see also [3]),

$$\mu_{TOT}(t) = \mu_f - \frac{\mu_f - \mu_i}{1 + exp((t - t_\mu)/\Delta_\mu)}, \tag{27}$$

with $\mu_i = 0.021yr^{-1}$, $\mu_f = 0.00887yr^{-1}$, $t_\mu = 1910yr$, and $\Delta_\mu = 16yr$. Parameters μ_i and μ_f are asymptotic values while t_μ, and Δ_μ determine when and how abrupt the transition is.

Life tables for United States based on census data are available (see for example [10]). There are also estimated life tables for the period 1850-1900 (see for example [40]). Life tables provide an estimate of the probability of death within one year (customarily denoted as q_x with x the age class) for different age classes and for different years. The survivorship function is estimated as $s(t,a) = 1 - q(t,a)\Delta t$. The values of $q(t,a)$ used in our simulations are obtained through the linear interpolation of the q_x values, which are provided in the life tables (see Fig. 2 right panel).

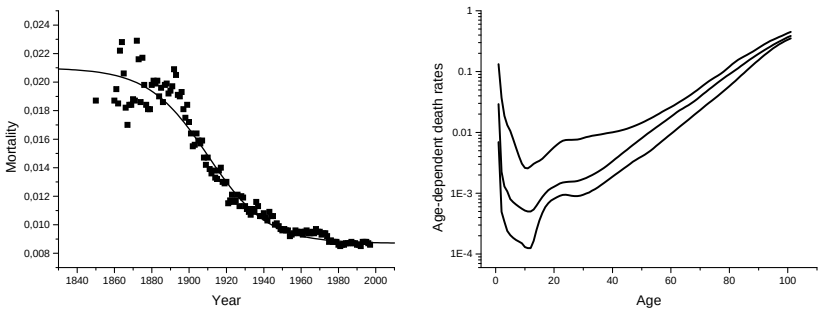

FIGURE 2. Left. Observed total mortality rates (squares) and their approximations given by expression (27) (continuous line). Right. Age-dependent probabilities of dying within one year (q_x) from period life tables for the United States based in the decennial census data from 1900, 1950, and 2000.

We computed the total mortality rate (per year, per person) with the aid of an age structured model. Results are displayed in Figure 3 together with the reported values. Agreement between model solutions and observed values is very good after 1900. The reason for the discrepancy in years before 1900 comes from the fact that we used the life tables values for 1900 for all years prior to 1900. In other words, we underestimate mortality patterns.

Since tuberculosis is mostly an urban disease, we only consider the dynamics of urban populations. The proportion of individuals living in urban populations over time for the U.S. and for the state of Massachusetts are available ([59], see also Fig. 4). The patterns observed in urban data are well captured by the sigmoid shaped function (see also [3])

$$P_U(t) = P_{Uf} + \frac{P_{Ui} - P_{Uf}}{1 + exp[(t - t_P)/\Delta_P]}, \tag{28}$$

which is also used for back-extrapolation. The parameters P_{Ui} and P_{Uf} are asymptotic values of the proportion P_U while t_P and Δ_P determine when and how abrupt the transition is.

The difference between the birth-death population growth and observed urban population growth is considered as immigration. At the beginning of industrialization in the United States, most immigration came from rural areas, essentially free of tuberculosis [37]. However, during the last century most of the immigrants have arrived from countries with higher TB incidence. Models should be modified to include the influx of new individuals in the infected classes in a model that attempts to capture the trends over the recent decades. Here, the role of arriving infected immigrants is ignored.

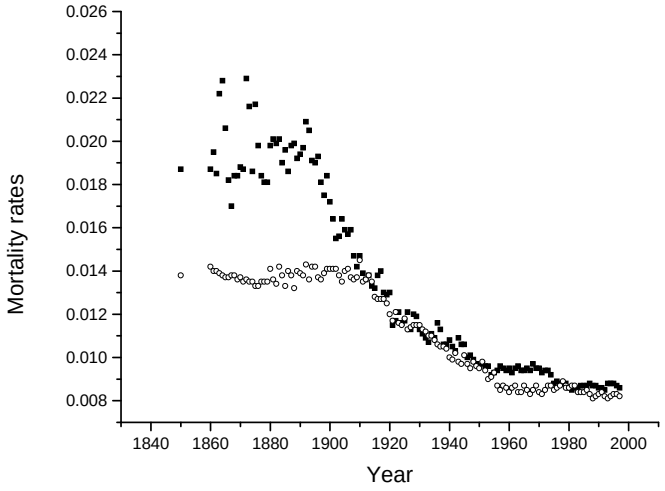

FIGURE 3. Observed mortality rates for United States (solid squares) and those obtained with the age-structured model (open circles).

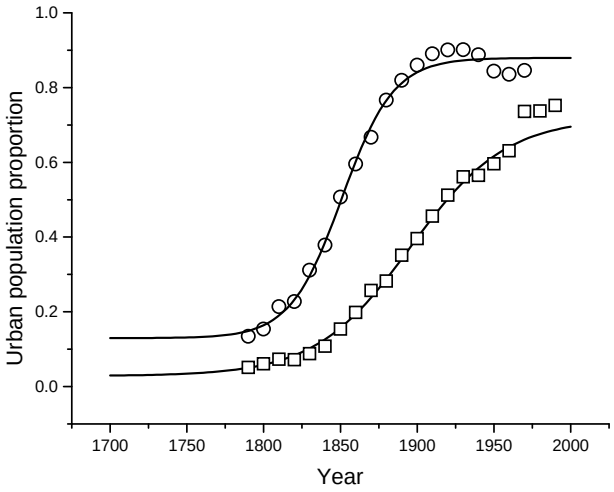

FIGURE 4. Proportion of urban populations for United States (squares) and the state of Massachusetts (circles). In both cases we fit the data to the sigmoid shaped function (28). Parameter values obtained from best least squares fit are: $P_i = 0.028236$, $P_f = 0.71835$, $t_{PU} = 1895.6$, $\Delta_P = 30.998$ for US, and $P_i = 0.12965$, $P_f = 0.87949$, $t_P = 1851.0$, $\Delta_P = 16.817$ for Massachusetts.

4.2. **Epidemiological parameters.** Estimating parameters in epidemiology is not straightforward. Here, we make use of parameter estimates collected from the relevant literature.

4.2.1. *Proportion of infectious and non infectious cases.* Studies show that between 50% to 85% of new TB cases are the result of progression from the high-risk latent class to the active TB class and are pulmonary. Furthermore, between 50% to 100% of the new cases that are the result of progression from the low-risk latent class to the TB active class are also pulmonary (see Blower et al., 1995, and references therein). Data from Massachusetts show that approximately 70% of the new TB cases can be classified as pulmonary over the last few years [23, 24, 25, 26]. Hence, we assume that on the average 70% ($q = 0.7$) of TB progressions lead to pulmonary cases.

4.2.2. *Contact number.* The contact number is the total number of secondary infections caused by one infectious case placed in an uninfected population. The transmission rate is defined as the average number of infections produced by a typical infectious individual during one year. Estimates of this rate are in the range 10-15 ([56] sec. 6.2.2). The average infectious period has been estimated in the range 0.5 to 2 years [8, 12, 33]. Using these estimates results derived estimates for the contact number Q_0 in the range 7 to 30. These levels of variability are not surprising. In fact, several facts suggest that variability in the number of secondary

cases generated through contacts with a source case is indeed high. Epidemiological studies [41, 49] have identified infectious individuals capable of generating more than 200 infections in periods as short a few months. Most outbreaks are in some sense exceptional, that is, the result of the effectiveness a few super-spreaders.

Sputum-positive individuals are more infectious than-sputum negative individuals. In fact, the study of Rose et al. [51] has shown that their difference in infectivity is at least 50%. Additional studies support the possibility of an even greater difference ([56] sec. 6.2.2). Here, we pre select a reasonable average value. The impact of variation on this last average value is taken into account indirectly when we consider a range of values for Q_0.

4.2.3. *Infectious period length.* Most of the secondary infections generated by a source case occur within his/her first few months as infectious [8]. Hence, fixing the mean infectious period at 0.5 years is reasonable. This choice, as it turns out, does not play a significant role in our results. In standard models the infectious period length determines the number of secondary cases produced by infectious individuals. Within our modeling framework, however, this number is varied independently as we vary the values of Q_0. It is possible to incorporate an explicit dependence on mean infectious period $1/\gamma$ using an expression like Expression (14).

4.2.4. *TB-induced mortality and recovery rates.* Records of pulmonary tuberculosis mortality have been regularly kept in the USA since 1850 (US Bureau of the Census, 1975). Reliable data on the incidence of active TB (pulmonary and extrapulmonary) have been collected for all United States since 1952 but data on incidence of pulmonary tuberculosis in Massachusetts have been available since 1915 [31].

Concern with pulmonary tuberculosis derives from the fact that there is a high case fatality associated with it. Before the availability of treatment almost 2/3 of the active TB sputum-positive cases and about 50% of the total pulmonary cases, died within 5 years after the onset of TB [56]. The effective mean infectious period $(1/\gamma)$ account for all cases of mortality. Hence, $\gamma \equiv \mu + d + r$ where μ is the natural per-capita mortality rate, d is the per-capita pulmonary tuberculosis induced death rate, and r is the per-capita recovery rate. The TB-induced mortality d is estimated from the formula $d \equiv p(\gamma - \mu)$ and the recovery rate from the expression $r = (1 - p)(\gamma - \mu)$ where p is the case fatality. We set $p = 0.5$ for $t < 1950$ (that is before the chemotherapy era) and use linear interpolation between this value and the actual value of 0.07 for $t \in [1950,\ 2000]$. The values of the general, non TB related mortality rate μ used in these calculations are obtained by subtracting the contribution to the total mortality by TB as modelled in Expression (27) (for age-structured and aggregated models).

Age structure data do exist. The case fatality of untreated extrapulmonary tuberculosis in children aged less than one year is above 50%, but decrease rapidly with age [30]. Unfortunately, mortality records are scarce. Some modelers have set the contributions to the mortality rate from extrapulmonary TB equal to the general (non TB related) mortality [60, 61], while others have simply disregarded extrapulmonary TB in their models. We take the contributions of extrapulmonary tuberculosis mortality (d_e) to be somewhere between the general mortality rate μ and the pulmonary TB mortality rate d. The recovery rate is estimated through the expression $r_e = \gamma_e - \mu - d_e$ where $1/\gamma_e$, the mean residence time in the extrapulmonary TB class, is taken to be equal to the mean infectious period $1/\gamma$ (see section below).

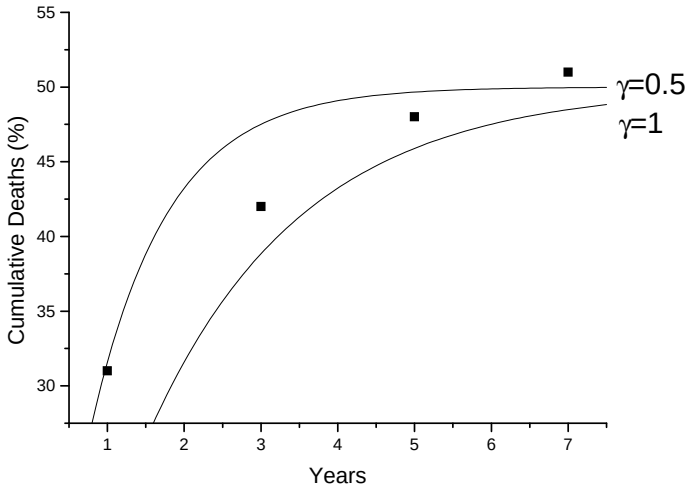

FIGURE 5. Observed deaths (in percentages) from pulmonary TB in Denmark (from Styblo 1991). Continuous lines represent model predictions (Case Fatality=$0.5(1-exp(-rt))$ with $d = r = \gamma$). We take $\gamma = 0.5$ and 1.

4.2.5. *Risk of progression to active TB.* The number of new cases of active TB exhibit approximate exponential decline as the age of infection increases [56]. In fact, about 60% of the new cases are the result of progression after the first year following infection, 95% of the cases surfacing within the first five years after the start of infection.

The decline in risk of progression as a function of age of infection is captured by setting $k + \alpha + \mu = 0.8363yr^{-1}$ (Fig. 6). A cohort of newly infected individuals ($t = 0$) exit according to the function as $E(t) = E_0 exp(-0.8363t)$ and therefore, the generation of new cases per unit of time evolves according to the expression $k(t)E(t)$ which decline almost exponentially because $k(t)$ is almost constant. Since progression and mortality rates are time dependent, α is also time dependent. Fortunately, setting $\alpha = 0.8363$ a constant, is enough to capture the observed patterns of TB progression as k is varied over a wide range under the constraint $\mu + k \ll \alpha$.

The value of the phenomenological parameter α facilitates the reproduction of the observed exponential decline in the risk of progression. Nevertheless, the function $k(t)$, must still be estimated from other data. Long term studies have shown that only a small fraction (between 5% to 10% in developed countries) of individuals infected with TB will ever develop active tuberculosis. Estimates of this fraction using a standard incidence model result in a function of progression rates from where the values of k can be estimated as it is outlined in the following heuristic argument.

If $E(t)$ denotes the population of recently first-time infected individuals then the fraction of individuals who progress to active TB from E is estimated as $f_E = \frac{k}{k+\alpha+\mu}$. The fraction of individuals who survive this latency period and do not

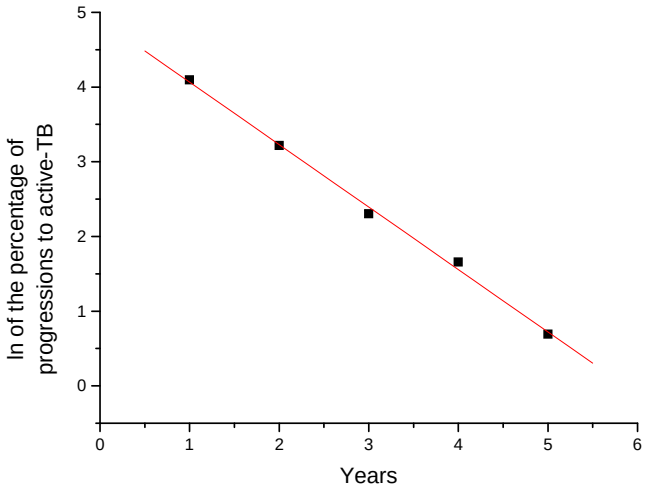

FIGURE 6. Percentage of individuals progressing to active TB after infection (log scale, taken from [56]). Continuous line is the least squares linear fit (slope=-0.8363). Because $\alpha \gg \mu + k$ we considered $\alpha = 0.8363$ for all t.

progress to active TB is $\frac{\alpha}{k+\alpha+\mu}$. This last group of individuals belong to the low-risk latent class L. If we disregard reinfection then the fraction of this population that develops active TB is estimated by $\frac{k_L}{k_L+\mu}$. Hence, roughly speaking, the fraction of the initial population that develops active TB from the class L is $f_L = \frac{\alpha}{k+\alpha+\mu}\frac{k_L}{k_L+\mu}$. Since in our model the effects are additive, we can roughly say that the fraction of individuals who develop active TB from both latent stages is

$$f = f_E + f_L = \frac{k}{k+\alpha+\mu} + \frac{\alpha}{k+\alpha+\mu}\,\frac{k_L}{k_L+\mu}. \tag{29}$$

We use this rough approximation as a non dimensional measure of the risk of progression to active TB. Actually, the relative contributions from the high-risk and low-risk latent classes to the active TB class is unknown. However, it is an accepted view that the latter plays a significant role only in developed countries with low incidence. In general we set $f = pf_E + (1-p)f_L$ with $p \geq 0.8$. The values of k and k_L are obtained from the values of f_E and f_L in each case. In other words, we take

$$k = pf\frac{(\alpha+\mu)}{(1-pf)}, \tag{30}$$

$$k_L = (1-p)\frac{(f/F_L)\mu}{(1-(1-p)f/F_L)}, \tag{31}$$

where $F_L = \alpha/(\alpha+k+\mu)$.

Our model of progression from the high-risk latent class E to the active TB class account for both, primary TB and endogenous reactivation, at least during the first ten years following infection, that is, the period for which observations have been recorded [56]. Also including progressions from the low-risk latent class

($k_L > 0$) may phenomenologically account for possible deviations from the predicted exponential decline in risk of progression for ages of infections greater than ten years. Certainly, in countries where infections with HIV are extremely likely, progression from low-risk cannot be ignored.

Possible protection conferred by previous infections with *M. tuberculosis* is modelled through the lowering of the risk of infection (through the parameter σ) and/or the lowering of the risk of progression to active TB (though the parameter σ_p).

The rate of TB relapse is set at the constant value $k_{Rp} = 0.00008yr^{-1}$ which results in a contribution of about 5% of the total incidence, the magnitude of today's observed contribution in Massachusetts [23, 24, 25, 26].

4.2.6. *Historical variation of epidemiological parameters.* In our model formulation two key epidemiological parameters determine the transmission-progression dynamics: The contact number Q_0 and the risk of progression f. The first controls the transmission process while the second controls the progression process. In Section 5.1 we show that at least one of these parameters must vary in time if the goal is to fit historical data. We use historical TB incidence and mortality data to estimate time-dependent consistent values for Q_0 and f.

Incidence of active TB for the United States and the state of Massachusetts has been recorded since 1953. There exist records on pulmonary TB mortality dating back to 1850. We model the incidence of active TB as proportional to pulmonary TB mortality before 1953. Data suggest that pulmonary TB mortality and the incidence of active TB can be approximated via a 2:1 relationship before treatment was available [31]. Hence, the assumption that pulmonary TB represents 70% of the number of total cases corresponds to an incidence of active tuberculosis of about 2.875 times the rate of mortality tied to pulmonary TB. We, therefore, estimate the incidence of active TB to be 2.875 times the tuberculosis mortality rates prior to 1953. Estimated and observed incidence rates are log transformed and the results are fitted to polynomials functions in order to obtain smooth functions that capture the trends (see the Appendix). From incidence trends (time series) we cannot obtain simultaneously the values of Q_0 and f. We assume that one of them is estimated independently and use the data (series) to estimate the other. For example if f is kept constant, the values of $Q_0(t)$ are obtained in the following way. Before 1845 we fix both parameter values in such a way that the simulated incidence of active pulmonary TB (that is the incidence obtained with the model) match (with some small error) the value of the smooth function (what we call $Inc(t)$) representing the estimated/observed incidence of pulmonary TB at $t = 1845$. After $t = 1845$ we use the time series to set the new values of $Q_0(t)$ in the following way. First we compute the relative error

$$\varepsilon = \frac{Simulated\ incidence(t) - Inc(t)}{Inc(t)}$$

and update Q_0 using the expression

$$Q_0(t + dt) = (1 - \varepsilon)Q_0(t)$$

after each step of the numerical integration scheme. The simulated incidence (per 10^5 population, per year) of active TB is estimated from the models as $(kE+k^*E^*+k_LL+k_{Rp}R)10^5/N_{tot}$, where N_{tot} is the total population (note that in the model N represents the urban population). The relationship to the age structured model comes from the obvious definitions of the variables E, E^*, L, and R as the integrals

of the corresponding densities over all ages. The above simple scheme produces excellent results (see Fig. 7). The functions $Inc(t)$ used in each case, USA and Massachusetts, are described in the Appendix.

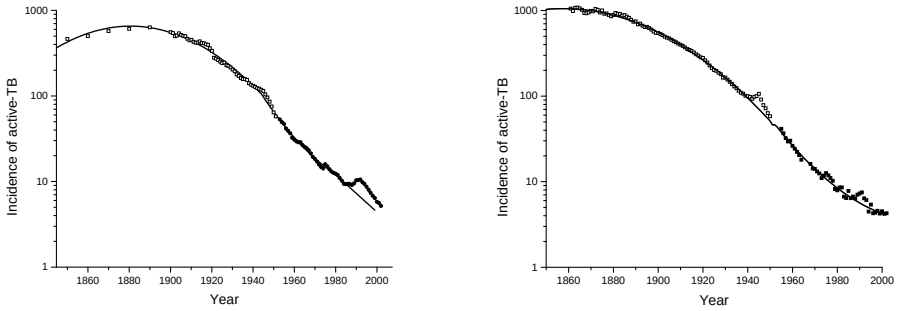

FIGURE 7. Observed incidence of active TB (all forms, solid squares, per year and per 100,000 population), estimated incidence of active TB (estimated as 2.875 times TB mortality rates, open squares), for the United States (left) and Massachusetts (right). The simulated incidences obtained from model (17-23) are shown in continuous lines.

5. **Some results and applications.**

5.1. **Vital-dynamics driven epidemics and the effect of initial conditions.**

In this section, we study TB dynamics when the epidemiological parameters f and Q_0 remain constant in time. Results obtained with the standard model (17-23) and the cluster model (1-12) are qualitatively similar. After a small peak (not always observed) there is a large epidemic spanning the last two centuries. Figure 8 plots the simulated (using the standard model) and the estimated (as 2.875 times the observed TB mortality) solutions. The simulated time evolution of the epidemics during the period of available data (1850-present) does not depend on the initial conditions. In fact the effect of the initial conditions virtually vanish after about 75 simulated years.

Vital dynamics driven epidemics with constant values of f and Q_0 may explain part of the observed decline in tuberculosis rates (like in [12]). However they cannot explain neither the timing nor the magnitude of the observed decline. According to model results the only way one can explain the trends in TB decreases are from reductions in transmission and/or progression over time.

In the aggregated model different initial conditions correspond to differences in the time of introduction of TB. Fortunately, the differences in the time of infection introduction are negligible after a few decades (also in the age-structured model). However, the establishment of an stationary age profile may take over a century since individuals in our model may live up to age 100. Furthermore, since the birth rate and the age dependent mortality rates change with time a stationary age profile is never reached. In order to minimize the effects of the initial age structure we proceeded as follow. We used birth and age-dependent mortalities

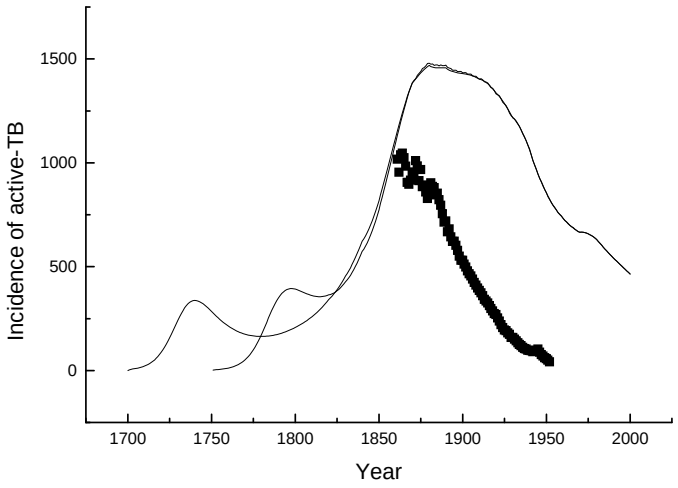

FIGURE 8. Vital dynamics driven epidemics. In one case epidemic start in 1700 while in the other in 1750. By 1850 both solutions are almost identical.

values for 1850 and run the simulations for 200 simulated years. We compute the age structure at the end of each simulation and use it as the initial condition for the next simulation. We carry out this process until variations at the end of consecutive simulations become negligible. We then use this "last" age structure distribution as the initial age distribution for all epidemiological simulations.

In the simulations, initial time is set at simulated year 1700 in order to minimize the role of the initial conditions.

5.2. **Stochastic effects.** Current Massachusetts values for the incidence of active TB are about 270 new cases per year. The expected stochastic fluctuations are of the order of $\sqrt{270} \sim 16$, that is, they are relatively small. In fact it can be shown that in this case the dynamics is almost deterministic [2]. Stochastic effects are likely to have been significant at the beginning of the epidemic. We perform standard stochastic simulations using the standard homogeneous mixing model. Monte Carlo simulations are carried out in a standard way, that is, the right-hand terms of model (17-23) are considered as transition rates of a stochastic process (see for example [50]).

Since quasi-exponential population growth was common during the nineteenth century, we consider initial infection-free populations of size 1000 and 10000 undergoing exponential growth (with parameter 0.03476 yr^{-1}) for 150 simulated years. In stochastic and deterministic simulations births are computed to match this population growth.

At the population level, stochasticity plays a role only for small populations. In Figure 9 we plot stochastic simulations together with the deterministic solutions when the initial population is between 10^3 and 10^4. Eventually stochastic realizations converge to the deterministic solutions.

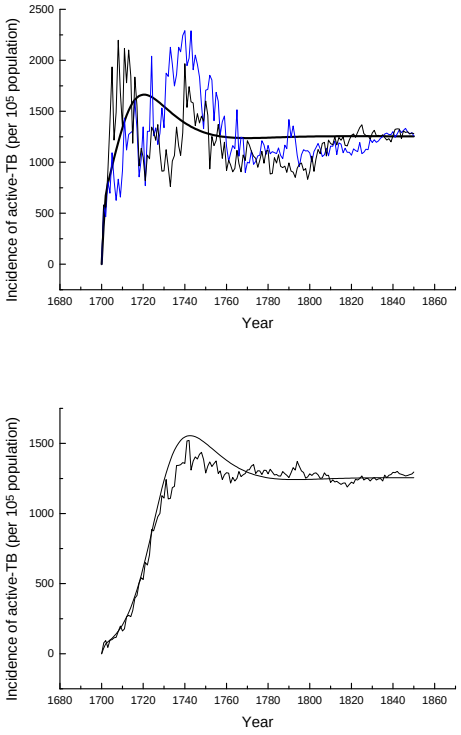

FIGURE 9. Stochastic realizations together with the deterministic solution for an initial population of 1000 (top) and 10000 (down). Realizations converge to the deterministic solutions as the population increase.

The likelihood of an epidemics is determined with the basic reproductive number $\mathcal{R}_0$ (see Table 1). In our framework an average infectious individual placed in an infection-free population will produce Q_0 secondary infections (by definition of Q_0) before he/she dies or recovers. From this number only the fraction f will become active TB cases and only the fraction q of them will develop pulmonary TB. The estimated basic reproductive number is therefore approximated by $\mathcal{R}_0 \simeq qQ_0f$. Numerical simulations show that this approximation is an excellent estimate for the model threshold (17-23). For $\mathcal{R}_0 > 1$ an epidemic is very likely to take place irrespective of population size.

The minimum population size for disease persistence is known as the critical community size. Tuberculosis dynamics are inherently slow, hence population growth must be factored in. We estimate the likelihood of tuberculosis establishment starting from an infection-free population of size N_0 undergoing population growth. Because at the beginning of the epidemic an average infectious individual will produce Q_0 secondary infections, our initial condition includes $N_0 - Q_0$ uninfected individuals and Q_0 recently infected (belonging to the E class). We compute the

TABLE 1. Frequency and time to extinction for different community sizes. Statistics were obtained from 1000 realizations. Mean values for time to extinction and the corresponding standard deviations measured in years.

N_0	Frequency of Extinction	Time to extinction (SD) (years)
500	0.584	6.53 (7.7)
1000	0.594	5.97 (5.9)
10000	0.573	5.27 (4.9)

probability of extinction as the frequency of the realizations for which the infected population $(E+E^*+I)$ become extinct before 100 simulated years, for different values of N_0. We also compute the time to extinction. The results of these simulations are remarkably similar for a wide range of community sizes. In Table 1 we collect the frequency of extinction and average times to extinction from 1000 stochastic simulations for different values of N_0. The values of $Q_0 = 10$ and $f = 0.22$ are used in all simulations. For these values $\mathcal{R}_0 = qQ_0f = 1.54$ and the probability of TB establishment turn out to be quite high, about 40% of the cases considered.

5.3. **Clustering.** Solutions for the incidence of active TB and TB-mortality obtained with the cluster (1-12) and homogeneous mixing models (17-23) turn out to be quite similar (see left panel in Fig. 10). We take the per-capita transmission constant through time. Specifically, Q_0 is assumed to be constant in the standard incidence model and n, β, γ constant in the cluster model. The value of n was obtained from the relation [1] $Q_0 = n(1-\rho)\frac{\beta}{\beta+(1-\rho)\gamma}$ with $\beta = \gamma$ and clustering coefficient $\rho = 0.5$. The initial values of f require the matching of an incidence close to the estimated value for 1845. The values obtained (for f) are model-dependent: 0.2 for standard incidence model and 0.475 for the cluster model. The difference comes from the fact that cluster models disregard progression to active TB in active clusters and at the epidemic peak a large fraction of the population is in active clusters. Therefore the contribution to active TB comes from a smaller population in cluster models than in standard incidence models (17-23). As the incidence rate reaches the low values observed during the past fifty years, the values of f converge to a value of around 0.1, a value close to actual estimates.

The most important difference between homogeneous mixing and cluster models comes from the relative contributions of reinfections and first infections to the generation of active TB cases. The standard incidence model does not account for the clustering of contacts and the number of new reinfections and first-infections are proportional to the populations σL and U, respectively. Cluster models, instead, take into account the clustering (of the contacts) through the clustering coefficient ρ. Therefore an infected individual has a higher probability of re-infection than a never infected individual to be infected.

Figure 10 shows the time evolution of the ratio of reinfections to first-infections from simulations generated by both models. The homogeneous mixing model gives a ratio below one after 1939 reaching the value of 0.30 by 2000. The cluster model generates a significantly higher value for this ratio. For $\rho = 0.5$ reinfections accounts for 55% of the total by year 2000.

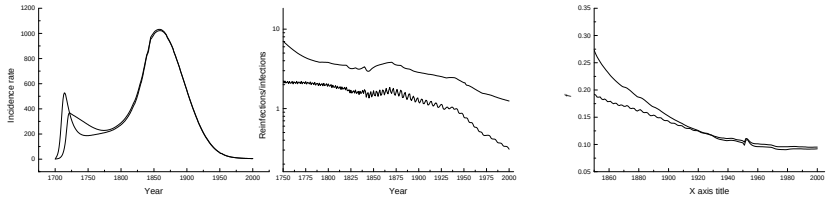

FIGURE 10. (Left panel) Model solutions obtained via the mass-action and cluster models. (Center panel) Relative contributions of re-infections to first infections via the mass-action and cluster models. The values of $f(t)$ obtained in each case are shown in the right panel. Parameter values are: $n = 40$, $\rho = 0.5$, $\beta = \gamma$ for the cluster model and $Q_0 = 15$ for the homogenous mixing model. In both cases $\sigma = 1$. We use the data from Massachusetts.

A natural question is whether or not clustering plays a significant role in the competitive interactions of *M. tuberculosis* strains. The results could differ significantly from those obtained through the use of homogeneous mixing models.

5.4. **Age structure.** The assumption of age-independent parameters, reduce the age-structured model to an aggregated model equivalent to the one given by the System (17-23). Integrating over all ages leads, for example, to the following equation for the uninfected population

$$\frac{dU(t)}{dt} = \int_0^\infty \left[\frac{\partial}{\partial t}u(t,a) + \frac{\partial}{\partial a}u(t,a)\right] da = B(t) - \gamma Q_0 A_p(t) U(t)/N(t) - \mu(t)U(t), \tag{32}$$

where $A_p(t) = \int_0^\infty a_p(a,t)da = A_p(t)$, $U(t) = \int_0^\infty u(t,a)da$ and we used the renewal condition $u(t,0) = \int_0^\infty b(t,a)N(t,a)da \equiv B(t)$.

This is not our case as many parameters are age-dependent. Figure 11 collects the survival curves for the United States obtained from different time periods. We are not plotting the actual observed survivorship functions but rather the curves generated from the use of age dependent mortality rates under the assumption that each cohort experiences the same mortality rates during all posterior times. Figure 11 also shows the survivorship curves generated for the age-independent mortality rates $\mu_{tot}(1900)$, $\mu_{tot}(1950)$, and $\mu_{tot}(2000)$. Here, $\mu_{tot}(t)$ is the total mortality rate (per year, per person) computed from Expression (27). The age independent mortality (also known as type II mortality) used in aggregated models like (17-23) produce long tailed survival curves not present in the realistic curves obtained with the age dependent mortalities (which are close to a type I mortality, $\mu(a) = 0$ for $a \leq A$, ∞ otherwise). In our model we assume that the survivorship function $(1 - \mu(a)dt)$ is zero for $a \geq A$ =100 years. From Figure 11 we see that this assumption is a good approximation. We also see that in all cases, type II mortality generates a significant proportion of surviving individuals older than A. As a consequence, the infectious population in homogeneous mixing models will produce a significative higher number of infections coming from this unrealistic

long tail. Therefore, even under the assumption of an age-independent mortality rate, the force of infection in homogeneous mixing models results in the generation of a higher number of infections than those generated by the age-structured model. In other words to obtain the same incidence of active TB a larger value for Q_0 is needed in age-structured models than in homogeneous mixing models. Furthermore, a key difference that derives from the use of aggregated models versus age-structured models comes from the fact that children rarely develop infectious tuberculosis. We used an age-dependent probability of progressing to pulmonary tuberculosis $q(a)$ that is set to zero for $a \leq a_{min}$ but equal to 0.7 otherwise. We set $a_{min} = 8$ years.

In Figure 12 (left panel) we compare the time evolution of the incidence of active TB obtained with the aggregated homogeneous mixing model (17-23) and the age-structured model. In all cases we incorporate vital-dynamics in epidemics driven with constant (epidemiological) parameters. The age-structured model is used to explore two cases. In one, the probability of developing pulmonary TB is age-independent, that is $q(a) = 0.7$ for all a. In the other we take $q(a) = 0$, for $a \leq a_{min} = 8$ years and $q(a) = 0.7$ otherwise.

In the first case the solutions of the homogeneous mixing and age-structured models are similar, with the latter producing lower values for the incidence of active TB, a consequence of the difference in mortality types. The inclusion of an age-dependent probability of developing pulmonary tuberculosis leads to significant quantitative differences.

In Figure 12 (right panel) we compare the numerical solutions generated with the age-structured model for different values of Q_0 with $a_{min} = 8yr$ with the solution of the homogeneous mixing model with $Q_0 = 12$. This Q_0 value was chosen because it produces a value for the incidence of active tuberculosis that matches the estimated

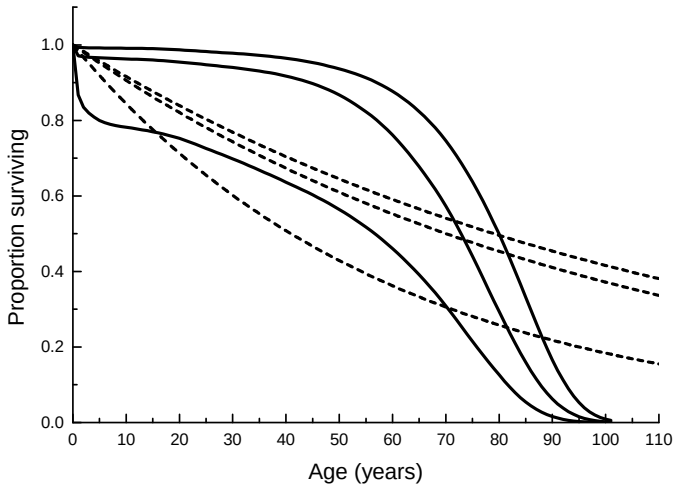

FIGURE 11. Proportion surviving an ideal cohort born in 1900, 1950 and 2000 that experiences age-dependent mortalities corresponding to the periods in which they were born for their entire life. Same curves obtained with the age-independent mortalities $\mu_{tot}(t = 1900)$, $\mu_{tot}(t = 1950)$, and $\mu_{tot}(t = 2000)$ (dashed lines).

value for Massachusetts in 1850 for $f = 0.2$ (constant through time). When the same value for Q_0 is used in the age-structured model a significant lower incidence of active TB is generated. If a comparable value for the incidence of active TB is desired, the value for Q_0 comes to about 20 (see Fig. 12). The homogeneous mixing model is likely to underestimate Q_0.

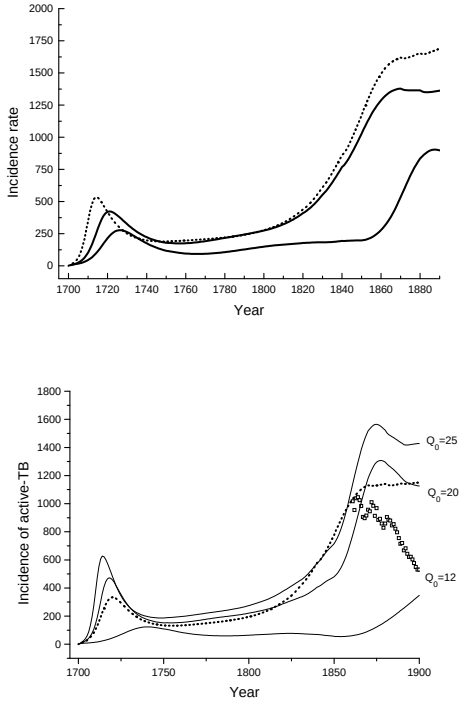

FIGURE 12. Incidence of active TB (rate per 10^5 population) obtained using the mass action model (17-23, doted line) and the age-structured model. Age-dependency of the probability to develop pulmonary TB has a significant impact in the disease dynamics (left panel). In all cases $Q_0 = 15$ and $f = 0.2$. Right. Solutions of the age-structured model for different values of Q_0 (12, 20 and 25) are compared with the solution of the mass action model (17-23) with $Q_0 = 12$. In all cases $f = 0.2$. Open squares are estimates of the incidence of active TB for Massachusetts.

5.5. **Possible causes for the historical decline of tuberculosis.** Several hypotheses have been proposed to explain the long-term decline in tuberculosis rates over the past century. These causes include:

1. Purely dynamical reasons [12, 37, 38, 39].
2. A reduction in transmission. Transmission could have been reduced as a consequence of public health measures (isolation of active cases in sanatoria, use

of antibiotic treatment) and the improvement of living and working conditions (ventilation and reduced overcrowding).
3. A reduction in progression. This explanation is also tied in to the theory of improving of living conditions [44] and the hypothesis of host-parasite coevolution [54, 55].

Next, we discuss the likelihood of these different possibilities under the results generated by our modeling framework.

It has been suggested that the observed decline in TB is the result of the expected natural course of this epidemic [12]. However, as it is shown in Section 5.1, vital-dynamics driven epidemics cannot account for the timing or the rate of the observed decline even when reinfection is excluded. These results strongly suggest the possibility that transmission and/or progression parameters have been declining in time.

The hypothesis of reduction in transmission as the cause behind the decline of tuberculosis rates is by far the more popular. In our framework reductions in the transmission correspond to reductions in the contact number Q_0.

However, Q_0 (see expression (14)) is not an independent factor, in fact it is a function of the mean infectious period length $(1/\gamma)$, the average risk of infection per case per susceptible (β), and the size of the network of close contacts of a typical active TB case (n and m) among others. Therefore reductions in Q_0 may come from:

1. a decrease in the mean infectious period $(1/\gamma)$, which can be achieved by the removal of infectious individuals soon after diagnosis (isolation or treatment);
2. a decrease in the environmental risk of infection (β), the result, for example of improving ventilation in workplaces;
3. a decrease in the number of contacts per individual, the result of reductions in mean household size, or reductions in crowdedness.

Public health measures like the sanatorium movement[29] or the widespread use of antibiotics after the fifties, have reduced the effective mean infectious period. However it should be noticed that the number of infections do not increase linearly with mean infectious period. For the simple case of exponentially distributed periods, Q_0 is given by (14), while for the more realistic case of fix periods, we have that $Q_0 = n(1 - e^{-\beta/\gamma})$. That is, most of the infections occur during the first months of infectiousness. This is consistent with the fact that ill individuals are naturally isolated from the rest of the population. So reductions in the average infectious period may not have been significant enough to explain TB decline.

Reductions in transmission are studied under a null hypothesis of a constant risk of progression, that is, a constant value for $f(t)$ is taken.[1] In developed countries it is estimated that between 5% to 10% of the infected individuals develop active tuberculosis. We start from the conservative upper value of $f = 0.1$ and compute the values of $Q_0(t)$ for which the simulated incidence matches the observed/estimated incidence as described in 4.2.6 (we also imposed the restriction that $\varepsilon < 0.01$). We use United States data and the age-structured model with a minimum age for developing pulmonary tuberculosis equal to eight years. Two cases are considered: the progressions coming from the high risk class is 100% of the total ($p = 1$ in (30)), or it is only 80% ($p = 0.8$). Maximum and minimum values for $Q_0(t)$ are displayed

[1] We used expression (29) for an estimation of the risk f although this expression was obtained from the model without age-structure (17-23).

in Table 2. In both cases the maximum values obtained are unrealistically large. Minimum values occurs at the end of the simulations and are in the right ball park.

TABLE 2. Maximum and minimum values of $Q_0(t)$ obtained for $f = 0.1$. We assume that a proportion p of the infected individuals progress from the high risk class, while the proportion $1 - p$ from the low-risk latent class L. In both cases maximum values are achieved at $t \approx 1860$.

p	Q_{0max}	Q_{0min}
1	180	15
0.8	453	10

Reductions in progression rates are studied in a similar fashion. We considered two values for Q_0: 10 and 20. In each case, the two values of $p = 1$ and $p = 0.8$ are used. Maximum and minimum values obtained with our model are collected in Table 3. The case $Q_0 = 20$ corresponds to the usual assumption of ten new infections caused per infectious individual per year with an infectious period of two years [33, 56]. For $Q_0 = 20$ the risk of progression to active TB has to decline from about 40% to about 7%. For $Q_0 = 10$ the maximum values rise to about 65%, a high (and perhaps unrealistic) value. Relatively small variations in f may indeed explain most of the observed decline.

TABLE 3. Maximum and minimum values of $f(t)$ obtained for $Q_0 = 10$ and $Q_0 = 20$. We assume that a proportion p of the infected individuals progress from the high risk class, while the proportion $1 - p$ from the low-risk latent class L. In both cases maximum values are achieved at $t \approx 1860$.

p	Q_0	f_{max}	f_{min}
20	1	0.33	0.073
20	0.8	0.41	0.065
10	1	0.54	0.146
10	0.8	0.65	0.126

6. **Challenges and future work.** Most models for the transmission dynamics of tuberculosis consider constant parameters and constant population size. These assumptions implicitly restrict the study of TB dynamics over small windows in time. TB dynamics are slow and, consequently, TB epidemics unfold over decades. Describing TB epidemics must account not only for population growth but also for the variation in time of parameters. Here (see also [3, 4]) we show that demography and changes in contact and epidemiological parameters have had a significant influence on disease dynamics.

The first conclusion we draw is that understanding of the time evolution of TB epidemics cannot ignore the study of its dynamics over long temporal scales.

Different population level models are introduced to carry out a comparative study of the dynamics of tuberculosis. Some aspects of tuberculosis dynamics including strain competition, probably would benefit from the use of individual based models that account for the structure of the social contact networks and possibly intra-host processes. Individual based models, however, cannot be used to capture epidemics evolving over centuries and over large spatial (countries) scales in a tractable way. However, individual based models may be used to estimate parameters used in simpler population level models.

There are many future promising research directions linked to TB that should be pursued over long-term time horizons and large spatial scales. We discuss some possibilities.

Competition and evolution of drug resistance. Previous research has explored *Mycobacterium tuberculosis* strain competition [13, 20, 34] under the assumption of homogeneous mixing. Inhomogeneous mixing is likely to play a strong role on competitive interactions. The evolution of drug resistance is tied in to intra-host competition where inhomogeneous mixing plays a key role. In Section 5.3 we show that the clustering of the contacts implies that recovered individuals have a higher risk of re-infection than susceptibles (never infected) of becoming infected. Further insights into the dynamics of strain competition may benefit from the explicit incorporation of networks of contacts. Individual based models may be useful in disentangling social network effects. The evolution of drug resistance on the other hand, demands attention to intra-host processes, including within host competition. These individual-level processes may be embedded within individual based models, or modeled independently. Whatever the approach, the effect of intra-host processes should be computed and some "average" and variance included in individual based models.

Superspreading. Molecular epidemiology studies show that, in general, an active pulmonary case produces few secondary infections [11]. This observation is compatible with the reproductive number estimates of around one in developed countries [4, 61]. Here, values of Q_0 in the range 10 to 25 are used. Classic epidemiological studies have documented the existence of infectious individuals that have generated hundreds of secondary infections (see [49] and references therein). It is unknown how common these super-spreaders are and the extent of their influence over the course of an epidemic. The role of super-spreaders should be explored in population level models such as those used here. New insights may be gained from studying the impact of super-spreaders in individual based models.

Age related problems. The probability of developing pulmonary tuberculosis is age-dependent and the incorporation of this fact has a dramatic effect on disease dynamics. Patterns of contacts in social networks are age dependent and thus age impacts transmission rates. Inhomogeneous mixing may be modelled using mixing matrices- *Who Acquires Infection from Whom* [14, 15, 16, 18, 19, 52]. Here, we only considered the most conspicuous age-dependency, namely age-dependent progression rates: children do not develop pulmonary tuberculosis. Important omitted characteristics include age of infection (but see [60, 61]). Long latency periods and strong age-dependence on risk of developing active tuberculosis support the view that age of infection must play a significant role on TB disease dynamics. We incorporated this effect through the use of a high- and a low-risk latent class, but there are other approaches [36, 58].

Host-parasite coevolution. Increases in human resistance, combined with a decrease in *M. tuberculosis* virulence must be tied in to reductions on average progression rates. With the beginning of the Industrial Revolution, tuberculosis reached epidemic levels. The growth of the cities with high levels of crowdedness and bad living and working conditions were big contributors. At this time TB accounted for about one quarter of all deaths. High levels of TB mortality suggest the possibility that selection played a role during this window in time. By the middle of the twentieth century TB was out of the top ten causes of death. Did natural selection produce a shift in average population progression rates over a short window in time? The estimation of the magnitude of the contribution of selection would require a genetic-epidemiological model (see for example [5]). The influence of genetic heterogeneity has been the subject of some research [46] but the study of the potential coevolution of *M. tuberculosis* with its human hosts has received much less attention.

7. Epilogue.

Tuberculosis dynamics has been the subject of a considerable body of theoretical (mathematical) work ([21] and references therein). Researchers have been interested in the influence of latency periods, existence of multiple steady states, effects of clustering, role of population heterogeneity, or the role of reinfection [1, 35, 36]. Some researchers have used mathematical models in the context of specific applied problems including the quantification of the potential effect of control strategies [6, 32, 47]. It is in this case that the use of more detailed models including realistic parameterizations become more relevant.

There are two classes of parameters of interest: demographic and epidemiological. Variations in each have important consequences. Because TB dynamics is slow in general, population growth cannot be disregarded. Population growth is the result of births, deaths and immigration. While in general birth and age-specific death rates are available, the modelling of immigration is complex. In developed countries, immigration usually brings an important influx of infected individuals. These form of recruitment (births and immigration) alter the population and epidemiological age profiles. The transmission dynamics of tuberculosis are tied to age profile because demographic and epidemiological factors are age-dependent. Aggregated models produce substantially higher values for TB incidence than age-structured models, a consequences of the long tail "age distribution" generated by type II mortality (observed mortality is closer to a type I mortality). Type I mortality translates in the existence of a maximal age. Therefore when age-structured models are parameterized (using observed age-dependent mortalities) the simulated incidence obtained is lower than the incidence generated by corresponding aggregated model, for the same parameter values.

Detailed models requires realistic parameterization. Hence they are too complex to study analytically. However, they are extremely useful in the design and evaluation of control strategies, estimation of the contribution of reinfection, or the evaluation of strain competitive outcomes in social networks.

Reductions in transmission are usually credited to public health measures. Yet, we show that reductions in progression rates are mostly influenced by socioeconomic improvements. Thus, what is the relative importance of improved socioeconomic conditions versus targeted medical interventions in battling infectious diseases like TB [27, 42, 44, 45, 57]?

Tuberculosis is an opportunistic disease. Infected individuals with weakened immune systems are at significant risk of developing clinical TB disease (active TB). High tuberculosis prevalence is therefore observed in individuals with AIDS. Malnutrition, alcoholism, drug abuse, concurrence of other infectious diseases, and psychological stress decrease immune response levels. We have argued [3] that the effect of improved living conditions may be related to increases in the average ability of the immune system for containing the proliferation of *M. tuberculosis* in infected hosts. In other words, we have argued that improved living conditions result in reductions in the risk of TB progression.

Our results strongly suggest that reductions in progression rates may account for a significant proportion of the decreases in TB. In addition, our conclusions cannot be conclusive when it comes down to the impact of reductions in transmission on TB declines. It can be argued that the social changes experienced over the last century (urbanization) may have increased the likelihood of transmission per case. Although working conditions have substantially improved since the beginning of the Industrial Revolution, changes in places of work have often accelerated disease spread (reduction of air volume per occupant or the use of ventilation systems which re-circulate the air[48]). Household size has also decreased over the last century while the average size of the network of close and frequent contacts, may have not changed that much. Increasing levels of school attendance can easily compensate for decreases in household size. Mass public transportation and globalization have increased the level of heterogeneity of casual contact rates and, consequently, the likelihood of super-spreader events.

Acknowledgments. This research has been partially funded by NSF, NSA, and Sloan Grants to the Mathematical and Theoretical Biology Institute. JPA acknowledges support from CONICET.

Appendix. The piecewise smooth functions used in our models to represent the estimated/observed incidence of active tuberculosis for the whole United States and for the state of Massachusetts are explicitly formulated below. In each case we chose few simple polynomials to carry out the fit. Other choices may lead to a better fit, but given the high uncertainty on estimated values, choice becomes irrelevant. Our choice produces excellent results (see Fig. 7).

USA data. Between 1845 and 1944, and between 1944 and 1979, we use a polynomial of degree two which produces the best fit (in the least squares sense) to the natural logarithm of the 2.875 times mortality data. For $t > 1979$ we use a polynomial of degree one. The fitted function $Inc(t)$ obtained is

$$Inc(t) = \begin{cases} exp(-1599.38143 + 1.70781t - 0.000454057t^2) & 1845 < t < 1944 \\ exp(2667.46 - 2.65571t + 0.0006615t^2) & 1944 < t < 1979 \\ exp(101.683531 - 0.0501t) & 1979 < t < 2000 \end{cases}$$

Massachusetts data. Between 1845 and 1953, and after 1953 we approximated the incidence by polynomials of degree two. These two curves do not intersect and are connected with a horizontal line. The function $Inc(t)$ obtained is:

$$Inc(t) = \begin{cases} exp(-1214.31417 + 1.31483t - 0.000353917t^2) & 1845 < t < 1950 \\ exp(3.834937) & 1950 < t < 1952.325 \\ exp(2327.06362 - 2.30263t + 0.00056991t^2) & 1952.325 < t < 2000 \end{cases}$$

REFERENCES

[1] J. P. Aparicio, A. F. Capurro and C. Castillo-Chavez, *Transmission and dynamics of tuberculosis on generalized households,* J. Theor. Biol., **206** (2000), 327–341, doi:10.1006/jtbi.2000.2129.

[2] J. P. Aparicio and H. G. Solari, *Population dynamics: A Poisson approximation and its relation to the Langevin process*, Phys. Rev. Lett. **86** (2001), 4183–4186.

[3] J. P. Aparicio, A. F. Capurro and C. Castillo-Chavez, *Markers of disease evolution: The case of tuberculosis,* J. Theor. Biol., **215** (2002), 227–237.

[4] J. P. Aparicio, A. F. Capurro and C. Castillo-Chavez, *Long-term dynamics and re-emergence of tuberculosis,* in "Mathematical Approaches for Emerging and Reemerging Infectious Diseases: An Introduction" (Minneapolis, MN, 1999), 351–360, IMA Vol. Math. Appl., 125, Springer, New York, 2002.

[5] J. P. Aparicio, H. G. Solari and N. Bonino, *Competition and coexistence in host-parasite systems: The myxomatosis case*, Population Ecology, **46** (2004), 71-85.

[6] J. P. Aparicio and J. C. Hernández, *Preventive treatment of tuberculosis through contact tracing*, Mathematical studies on human disease dynamics, 17–29, Contemp. Math., **410**, Amer. Math. Soc., Providence, RI, (2006).

[7] J. P. Aparicio and M. Pascual, *Building epidemiological models from R_0: An implicit treatment of transmission in networks*, Proceedings of the Royal Society: Biological Sciences, **274** (2007), 505–512.

[8] D. S. Barnes, "The Making of a Social Disease: Tuberculosis in the Nineteenth-Century France," University of California Press, 1995, 5–13.

[9] T. Baxter, *Low infectivity of tuberculosis*, The Lancet, **342** (1993), 371.

[10] F. C. Bell and M. L. Miller, "Life Tables for the United States Social Security Area 1900-2100," Actuarial Study No. 120, Social Security Administration, Pub. No. 11-11536, 2005.

[11] W. R. Bishai, N. M. Graham, S. Harrington, D. S. Pope, N. Hooper, J. Astemborski, L. Sheely, D. Vlahov, G. E. Glass and R. E. Chaisson, *Molecular and geographic patterns of tuberculosis transmission after 15 years of directly observed therapy*, J. Am. Med. Assoc., **280** (1998), 1679-1684.

[12] S. M. Blower , A. R. McLean, T. C. Porco, P. M. Small, P. C. Hopwell, M. A. Sanchez and A. R. Moss, *The intrinsic transmission dynamics of tuberculosis epidemics*, Nature Medicine, **1** (1995), 815–821.

[13] S. Blower, K. Koelle and T. Lietman, *Antibiotic resistance-to treat or not to treat?*, Nature Medicine, **5** (1999), 358–359.

[14] S. P. Blythe, C. Castillo-Chavez, J. Palmer and M. Cheng, *Towards unified theory of mixing and pair formation*, Math. Biosci., **107** (1991), 379–405.

[15] S. P. Blythe, C. Castillo-Chavez and G. Casella, *Empirical methods for the estimation of the mixing probabilities for socially structured populations from a single survey sample*, Mathematical Population Studies, **3** (1992), 199–225.

[16] S. P. Blythe, S. Busenberg and C. Castillo-Chavez, *Affinity in paired event probability*, Math. Biosci., **128** (1995), 265–284.

[17] F. Brauer and C.Castillo-Chavez, "Mathematical Models in Population Biology and Epidemiology," Texts in Applied Mathematics, 40, Springer-Verlag, New York, 2001.

[18] S. Busenberg and C. Castillo-Chavez, *A general solution of the problem of mixing subpopulations, and its application to risk-and age-structured epidemic models for the spread of AIDS*, IMA J. of Mathematics Applied in Med. and Biol., **8** (1991), 1–29.

[19] C. Castillo-Chavez, S-F Shyu, G. Rubin, and D. Umbauch, *On the estimation problem of mixing/pair formation matrices with applications to models for sexually-transmitted diseases*, in "AIDS Epidemiology: Methodology Issues", (eds. K. Dietz, V. T. Farewell, N. P. Jewell), 384–402, Birkhauser, Boston-Basel-Berlin, 1992.

[20] C. Castillo-Chavez and Z. Feng, *To treat or not to treat: The case of tuberculosis,* J. Math. Biol., **35** (1997), 629–656.

[21] C. Castillo-Chavez and B. Song, *Dynamical models of tuberculosis and applications*, Mathematical Biosciences and Engineering, **1** (2004), 361–404.

[22] Centers for Disease Control and Prevention, *Exposure of passengers and flight crew to Mycobacterium tuberculosis on commercial aircraft, 1992-1995*, *MMWR*, **44** (1995), 137–140.

[23] Centers for Disease Control and Prevention, *Reported Tuberculosis in the United States, 1996*, (1997) Available from www.cdc.gov/nchstp/tb/.

[24] Centers for Disease Control and Prevention, *Reported Tuberculosis in the United States, 1997,* (1998), Available from www.cdc.gov/nchstp/tb/.
[25] Centers for Disease Control and Prevention, *Reported Tuberculosis in the United States, 1998,* (1999) Available from www.cdc.gov/nchstp/tb/.
[26] Centers for Disease Control and Prevention, *Reported Tuberculosis in the United States, 1999,* (2000) Available from www.cdc.gov/nchstp/tb/.
[27] J. Colgrove, *The McKeown thesis: A historical controversy and its enduring influence*, Am. J. Public Health., **92** (2002), 722–725.
[28] T. M. Daniel, "Captain of Death: The Story of Tuberculosis," 37–39, University of Rochester Press, 1997.
[29] A. L. Davis, *History of the sanatorium movement*, In "Tuberculosis" (eds. W. N. Rom, and S. M. Garay), Little, Brown and co, 1996.
[30] P. R. Donald and N. Beyers, *Tuberculosis in childhood*, in "Clinical Tuberculosis" (ed. P.D.O. Davies), Chapman and Hall Medical, 1998.
[31] G. Drolet, *Present trend of case fatality rates in tuberculosis,* Am. Rev. of Tuberc., **37** (1938), 125–151.
[32] C. Dye, G. P. Garnett, K. Sleeman and B. G. Williams, *Prospects for worldwide tuberculosis control under the WHO DOTS strategy*, The Lancet, **352** (1998), 1886–1891.
[33] C. Dye, S. Scheele, P. Dolin, V. Pathania and M. C. Raviglione, *Global burden of tuberculosis estimated incidence, prevalence, and mortality by country*, JAMA, **282** (1999), 677–686.
[34] C. Dye and B. G. Williams, *Criteria for the control of drug-resistant tuberculosis*, Proc. Natl. Acad. Sci. USA, **97** (2000), 8180–8185.
[35] Z. Feng, C. Castillo-Chavez and A. F. Capurro, *A model for tuberculosis with exogenous reinfection*, Theor. Pop. Biol., **57** (2000), 235–247.
[36] Z. Feng, W. Huang and C. Castillo-Chavez, *On the role of variable latent periods in mathematical models for tuberculosis*, Journal of Dynamics and Differential Equations, **13** (2001), 425–452.
[37] E. R. N. Grigg, *The arcana of Tuberculosis*, Am. Rev. Tuberculosis and Pulmonary Diseases, **78** (1958), 151–172.
[38] E. R. N. Grigg, *The arcana of Tuberculosis III*, Am. Rev. Tuberculosis and Pulmonary Diseases, **78** (1958), 426-453.
[39] E. R. N. Grigg, *The arcana of Tuberculosis IV,* Am. Rev. Tuberculosis and Pulmonary Diseases, **78** (1958), 583–608.
[40] M. R. Haines, *Estimated Life Tables for the United States*, 1850–1900, Historical Paper n^o 59, National Bureau of Economic Research, 1994.
[41] E. Lincoln, *Epidemics of tuberculosis*, Adv. Tuberc. Res., **14** (1965), 157–201.
[42] B. G. Link and J. C. Phelan, *McKeown and the idea that social conditions are fundamental causes of disease*, Am. J. Public Health, **92** (2002), 730–732.
[43] J. W. McFarland, C. Hickman, M. T. Osterholm and K. L. MacDonald, *Exposure to Mycobacterium tuberculosis during air travel*, The Lancet, **342** (1993), 112–113.
[44] T. McKewon and R. G. Record, *Reasons for the decline of mortality in England and Wales during the nineteenth century*, Population Studies, **16** (1962), 94–122.
[45] T. McKewon, "The Origins of Human Disease," Basil Blackwell, 1987.
[46] B. M. Murphy, B. H. Singer, S. Anderson and D. Kirschner, *Comparing epidemic tuberculosis in demographically distinct heterogeneous populations*, John A. Jacquez memorial volume, Math. Biosci., **180** (2002), 161-185.
[47] C. J. L. Murray and J. A. Salomon, *Modeling the impact of global control strategies,* Proc. Natl. Acad. Sci. USA, **95** (1998), 13881–13886.
[48] E. A. Nardell, J. Keegan, S. A. Cheney and S. C. Etkind, *Theoretical limits of protection achievable by building ventilation*, Am. Rev. Respir. Dis., **144** (1991), 302–306.
[49] J. Raffalli, K. Sepkowitz and D. Armstrong, *Community-based outbreaks of tuberculosis*, Arch. Intern. Med., **156** (1996), 1053–1060.
[50] E. Renshaw, "Modelling Biological Populations in Space and Time," Cambridge Studies in Mathematical Biology, 11, Cambridge University Press, Cambridge, 1991.
[51] D. E. Rose, G. O. Zerbe, S. O. Lantz and W. C. Bailey, *Establishing priority during investigation of tuberculosis contacts,* Am. Rev. Respir. Dis., **119** (1979), 603–609.
[52] G. Rubin, D. Umback, S-F. Shyu and C. Castillo-Chavez, *Application of capture-recapture methodology to estimation of size of population at risk of AIDS and/or other sexually-transmitted diseases,* Statistics in Medicine, **11** (1992), 1533–1549.

[53] B. Song, C. Castillo-Chavez and J. P. Aparicio, *Tuberculosis models with fast and slow dynamics: the role of close and casual contacts,* Jacquez memorial volume, Math. Biosci., **180** (2002), 187–205.
[54] W. W. Stead, *Genetics and resistance to tuberculosis,* Annals of Internal Medicine, **116** (1992), 937–941.
[55] W. W. Stead and J. H. Bates, *Geographic and evolutionary epidemiology of tuberculosis*, in "Tuberculosis," Little, Brown and co. (eds. Rom W.N. and Garay, S.M.), (1996), 77–84.
[56] K. Styblo, "Epidemiology of Tuberculosis," Selected Papers, **24**, Royal Netherlands Tuberculosis Association, The Hague, 1991.
[57] S. Szreter, *Rethinking McKeown: The relationship between public health and social change,* Am. J. Public Health, **92** (2002), 725–729.
[58] H. Thieme and C. Castillo-Chavez, *How may infection-age-dependent infectivity affect the dynamics of HIV/AIDS?*, SIAM J. Applied Math., **53** (1993), 1447–1479.
[59] U. S. Bureau of the Census, "Historical Statistics of the United States: Colonial Times to 1970," Washington DC: Government Printing Office, 1975.
[60] E. Vynnycky and P. E. M. Fine, *The natural History of Tuberculosis: the implications of age-dependent risks of disease and the role of reinfection,* Epidemiol. Infect., **119** (1997), 183–201.
[61] E. Vynnycky and P. E. M. Fine, *Interpreting the decline in tuberculosis: The role of secular trends in effective contact,* International Journal of Epidemiology, **28** (1999), 327–334.

Received February 6, 2008. Accepted November 2, 2008.

E-mail address: juan.p.aparicio@gmail.com
E-mail address: ccchavez@asu.edu