Original Research Article

Medical Decision Making

Medical Decision Making
2021, Vol. 41(6) 641–652

DOI: 10.1177/0272989X211007842
journals.sagepub.com/home/mdm
SAGE

Representing Tuberculosis Transmission with Complex Contagion: An Agent-Based Simulation Modeling Approach

Erin D. Zwick, Caitlin S. Pepperell, and Oguzhan Alagoz

Objective. A recent study reported a tuberculosis (TB) outbreak in which, among newly infected individuals, exposure to additional active infections was associated with a higher probability of developing active disease. Referred to as *complex contagion*, multiple reexposures to TB within a short period after initial infection is hypothesized to confer a greater likelihood of developing active infection in 1 y. The purpose of this article is to develop and validate an agent-based simulation model (ABM) to study the effect of complex contagion on population-level TB transmission dynamics. **Methods.** We built an ABM of a TB epidemic using data from a series of outbreaks recorded in the 20th century in Saskatchewan, Canada. We fit 3 dynamical schemes: base, with no complex contagion; additive, in which each reexposure confers an independent risk of activated infection; and threshold, in which a small number of reexposures confers a low risk and a high number of reexposures confers a high risk of activation. **Results.** We find that the base model fits the mortality and incidence output targets best, followed by the threshold and then the additive models. The threshold model fits the incidence better than the base model does but overestimates mortality. All 3 models produce qualitatively realistic epidemic curves. **Conclusion.** We find that complex contagion qualitatively changes the trajectory of a TB epidemic, although data from a high-incidence setting are reproduced better with the base model. Results from this model demonstrate the feasibility of using ABM to capture nuances in TB transmission.

Keywords
agent-based modeling, complex contagion, infectious disease, modeling, simulation, tuberculosis

Date received: January 22, 2021; accepted: March 12, 2021

Worldwide, tuberculosis (TB) is the leading cause of death due to a single infection.[1] With a global goal of eliminating TB by 2050[2,3] and rising threats of multidrug-resistant and extremely drug-resistant TB infection,[1,4] it is more urgent than ever to elucidate why TB persists as a threat to global health.

Complex transmission dynamics make TB difficult to control. One complexity of transmission, referred to as *complex contagion*, is the phenomenon by which multiple reexposures to active TB disease during early latency increase the risk of progression to active infection[5] (Figure 1). Although previous work has looked at the effect of reinfection occurring 5 or more years after initial infection,[6–8] the dynamical impact of numerous exposures within 1 y of initial infection is unknown.

The population-level implications of complex contagion are difficult to study using traditional epidemiologic methods. Previous studies using comprehensive contact tracing and diagnostic testing revealed complex contagion in TB epidemiological data.[5,9] It is not feasible to apply these techniques at a large scale, and the population-level impacts of complex contagion are not well understood. Disease transmission models are a cost-effective way to understand complex transmission

Corresponding Author
Oguzhan Alagoz, Department of Industrial and Systems Engineering, University of Wisconsin-Madison, 3242 Mechanical Engineering Building, 1513 University Avenue, Madison, WI 53706, USA (alagoz@engr.wisc.edu)

dynamics when relevant population-level data are difficult to collect. In particular, agent-based simulation models (ABMs) are flexible tools for inferring population-level effects from individual-level dynamics.[10,11]

One study found that within the first year after infection, each additional exposure to active TB increased the odds of progressing to active disease by a factor of 1.11 (95% confidence interval [CI], 1.06–1.16).[5] Results from a stochastic model fit to a small TB outbreak suggest complex contagion worsens an epidemic.[12] Those authors found that a model with complex contagion fit better than a model without it.[12] Building on this work, we develop an ABM to represent a TB epidemic in a large population over a longer time period while incorporating complex contagion transmission dynamics. The purpose of this study is to demonstrate the feasibility of modeling a TB epidemic with complex contagion dynamics by fitting versions of the model with and without complex contagion to data.

We used data from the First Nations and Métis populations of Saskatchewan, Canada, collected during a series of severe TB epidemics during the 20th century. This study population is of particular interest because TB was a leading contributor to mortality in Canada,[13] with disproportionate effects on these populations.[14–16] The epidemics studied here are well documented, with data comprising decades-long time series, surveys, and medical records.[14–16] Modeling data from these epidemics can help elucidate the role of complex contagion in a high-incidence setting during a prolonged outbreak.

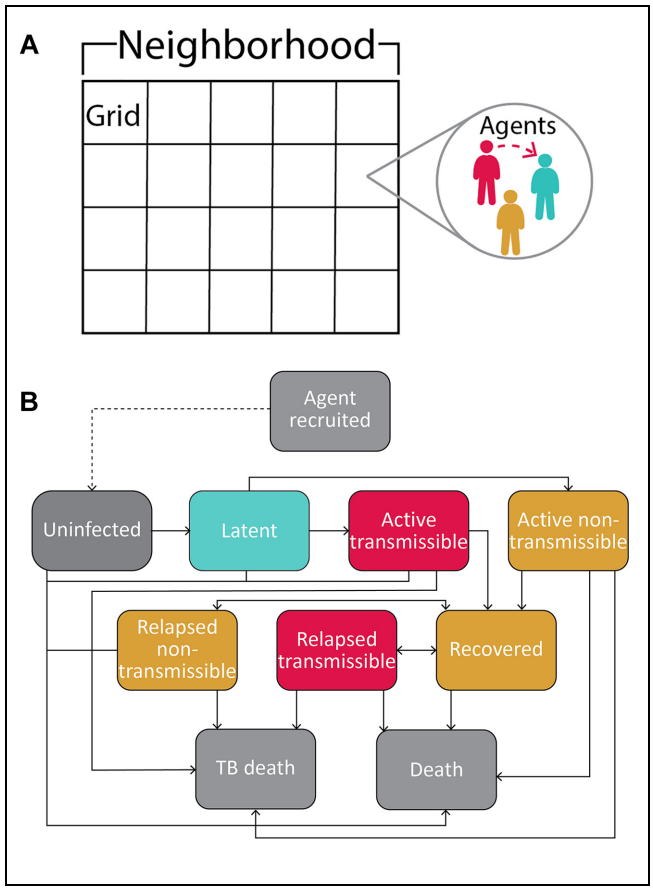

Figure 1 Model schematics. (A) Data structure for 1 neighborhood in the model. (B) Disease state transition diagram.

Methods

Study Population

We developed our ABM using archival epidemiologic and census data from the First Nations and Métis communities of Saskatchewan, Canada. Population and data sources have been described in Pepperell et al.[17]

Overview of the Model

We developed an ABM that simulates TB transmission on a network of connected communities. An ABM is an extension of traditional discrete-event simulation in which agents in the model have unique attributes and interact with each other and the environment, and updates occur in discrete time steps. TB infections can be protracted,[18–20] taking years or even decades after an initial infection for a person to develop active disease. In addition, untreated active disease lasts on average 2 to 3 y until death or self-cure.[20] To capture this wide range of latent and active disease times, we have modeled minimum time steps of 1 mo. The model geography is composed of neighborhoods representing 1 community each from our study population and agents represent

Department of Population Health Sciences, University of Wisconsin-Madison, Madison, WI, USA (EDZ); Department of Medicine and Department of Medical Microbiology and Immunology, University of Wisconsin-Madison, Madison, WI, USA (CSP); Department of Industrial and Systems Engineering, University of Wisconsin-Madison, Madison, WI, USA, PhD (OA). The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article. The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Financial support for this study was provided in part by grants from University of Wisconsin-Madison Graduate School, NIH (NIH/T15-LM007359 and NIAID R01AI113287), and the National Center for Research Resources (CTSA 2UL1TR000427). The funding agreement ensured the authors' independence in designing the study, interpreting the data, writing, and publishing the report.

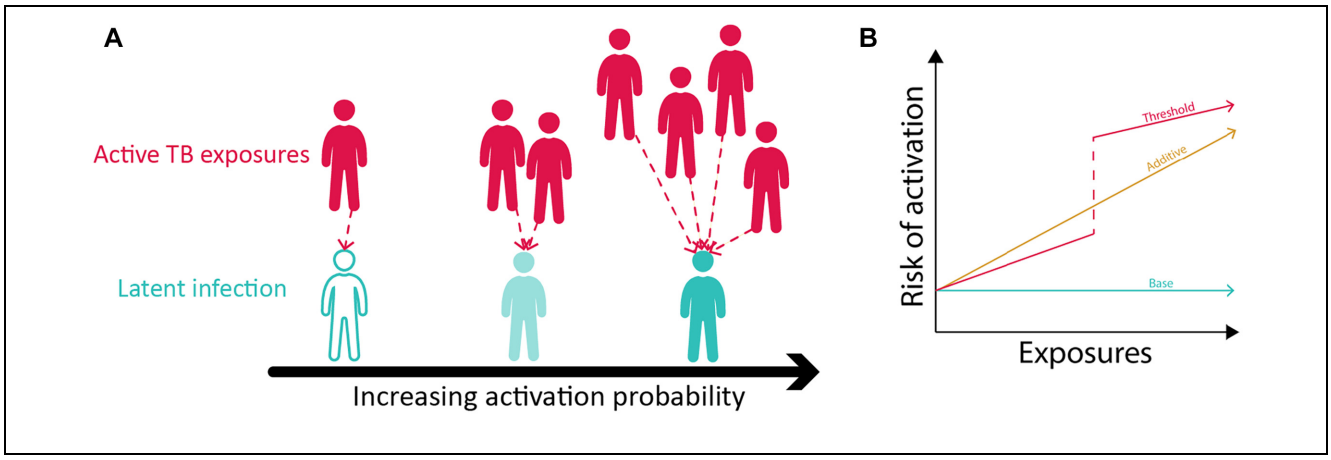

Figure 2 Diagrams of complex contagion. (A) Schematic of the relationship between exposures to active tuberculosis during latency and the increasing risk of developing active infection. (B) Comparison of the risk of activation during latency as a function of exposures during the first year of latency in the base (blue), additive (orange), and threshold (pink) models.

individuals. Agents move between grids within each neighborhood, and contacts occur between agents within a grid space and every surrounding grid (Figure 1). Agents move once per time step and can move between neighborhoods, returning to their home neighborhood at the start of each time step. The model allows for population growth, which we refer to as recruitment. Recruitment can occur through migration or birth, although for the purposes of modeling the study data, recruitment is entirely due to new births because migration in the study population was minimal during the study period.[14]

We used this ABM to evaluate 3 different TB transmission dynamics: the standard transmission model, in which additional exposure to active TB during early latency does not confer additional risk of activation (base model); additive complex contagion, in which each additional exposure during early latency confers an independent risk of activation (additive model); and threshold complex contagion, in which each additional exposure during early latency confers a small additional risk of activation up to some threshold of exposures, after which the risk jumps significantly (threshold model). Figure 2 presents a schematic of the relationship between exposure and risk in all 3 models.

Agents and the Environment

Agents are assigned a probability of changing neighborhoods during initialization. Moving probability is drawn from an exponential distribution such that, on average, agents will visit another neighborhood at least once per year of model time. Within the neighborhood, agents have a uniform probability of ending up on any grid space. Once assigned a grid space, agents mix homogeneously with each other within the grid space they are on as well as on any surrounding grid space (up to 9 grids total). Grids are designated as contaminated at time step t if an infectious agent visits that grid. Agents on the contaminated grid and surrounding grids are then assigned exposures equivalent to the number of infectious agents on those grid spaces. For this model, each neighborhood is initialized with enough grids such that, on average, each grid holds 4 agents at the end of the warmup period.

At the beginning of each time step, all agents return to their home neighborhood. Agents then move throughout the geography in a 2-step process. First, a Bernoulli trial determines if an agent will move outside of their neighborhood. If the agent moves, the new neighborhood is determined with uniform probability from non-home neighborhoods. Then, the agent is assigned a grid with uniform probability from the neighborhood they are staying in or traveling to. At the beginning of each year, defined as a time step that is a multiple of 12, each neighborhood shrinks because of non-TB deaths and then grows through new births. The non-TB death rate is set to the inverse of life expectancy (Table 1: model parameters). The neighborhood grows according to birth rates for the study population. All newly recruited agents are uninfected.

Disease State Transitions

Each agent has a dynamic TB disease state modeled as a discrete-time Markov chain with transitions drawn from exponential distributions based on previous models[12,21] (Figure 1). Agents can be in Uninfected (U), Latent (L),

Table 1 Input Parameters

Symbol	Parameter	Mean Value	Distribution	Range	Source
β	Transmission rate per contact per month	0.013[B] 0.006[A] 0.005[T]	Uniform	[0.001, 0.06]	Calibrated
$p_{transmissible}$	Probability of developing transmissible TB for active cases	0.54[B] 0.74[A] 0.68[T]	Uniform	[0.25, 0.85]	Calibrated
p_{fast}	Probability of developing active TB within 1 y of initial infection	0.3	—	—	21,22
v_{fast}	Progression rate to TB, fast (<1 y)	0.384	Exponential	—	22
v_{slow}	Progression rate to TB, slow (>1 y)	0.00438 (corresponds to 5%–10% progression in 20 y)	Exponential	—	21,22
c	Time to self-cure	36 mo	Exponential	—	20
μ_{TB}	Time to TB death	24 mo	Exponential	—	20
ω	Relapse rate	3% per year	Exponential	—	21
t_{rate}	Treatment rate	—	Piecewise linear function (see supplemental material)	[0.005, 0.52]	23,24
$t_{completion}$	Treatment completion probability	0.6	—	—	23,24
t_{length}	Treatment length	12 mo	Gamma(12,1)	—	23
$p_{prophylaxis}$	Probability of receiving treatment during latency each month	0.27[B] 0.08[A] 0.08[T]	Uniform	—	15
Λ	Recruitment rate per year per 1000 individuals	50	—	—	14,24
$1/\mu$	Life expectancy (1/death rate)	60 y	—	—	24
m	Move probability	0.15	Exponential	—	14,16

B, base model; A, additive model; T, threshold model.

Active–Transmissible (AT), Active–Nontransmissible (AN), Recovered (R), Relapsed (RL), TB Death (TD), or Natural Death (ND) states. Agents are only infectious in the AT state and susceptible to infection in the U and R states. Agents in the Latent (L) state have been exposed but are not infectious. Once infected, the agent is subjected to a Bernoulli trial to determine if they will experience a fast or slow progression to active infection with probability of fast progression (p_{fast}). The progression rate is drawn from an exponential distribution for both fast and slow activations. For fast activations, progression takes at most 12 mo (v_{fast}), and for slow activations, it takes more than 12 mo, with only 5% to 10% of infections activating within 20 y of initial infection (v_{slow}). The activation time for each latently infected agent is stored in an activation clock and decreases at each time step. Once the activation clock reaches 0, the agent will move to either the AT or AN state. See Table 1 for a further explanation of parameter values. Agents in either active state (AT or AN) can self-cure or receive treatment and enter the R state or experience death due to TB (TD). Recovered agents can experience RL or become independently reinfected, reentering either active state (AT or AN). Agents in any alive state can die of non-TB causes.

Exposures

We evaluated 3 types of exposure models: 1) a base model with no complex contagion, 2) an additive complex contagion, and 3) threshold complex contagion. In all 3 models, the exposures of the uninfected agents are counted at each time step. An agent is considered exposed if an actively infected, transmissible agent is on the same grid or an adjacent grid at that time step. At the subsequent time step, the agent is subjected to a Bernoulli trial with probability β that exposure leads to latent infection. The trial is repeated once for each exposure that agent accrued over the previous time step.

For the complex contagion models, exposures during latency are also counted. In the additive model, exposures during latency are counted at each time step, and each results in a Bernoulli trial to determine if a latent exposure will become an active infection with probability β. In the threshold model, exposures during latency are counted over the first 12 mo since the initial infection, the time period determined by Ackley et al.[12] to be of most importance in complex contagion. Exposures are treated as in the additive model until the threshold exposure level is reached. At that time, the agent becomes twice as likely to have a successful exposure[12] (i.e., a Bernoulli trial with probability $2*\beta$), and any additional exposures over the threshold are subjected to Bernoulli trials with probability β. If one of these exposures is successful, that new infection is determined to be a fast activation (i.e., <12 mo in latency) or slow activation (>12 mo) by a Bernoulli trial with probability p_{fast} of becoming a fast activation. Then, a new length of latency is drawn from an exponential distribution with rate v_{fast} for fast activations and v_{slow} for slow activations. If this length of latency is less than the current time left on the agent's activation clock, the time on the clock is reduced to the lower value. See Supplemental Figure S1 for a flow diagram of how latent exposures are modeled under complex contagion.

Treatment

We model 3 treatment introductions coinciding with changes in treatment policy in the study population over the modeled time period. The first 2 introductions reflect the discovery of TB-treating drugs: streptomycin and para-aminosalicylic acid (PAS; 1944–1946) and isoniazid (1952).[14] Streptomycin was not very effective until combined with PAS,[25] so we average the years of discovery to introduce treatment in 1945. The second introduction occurs in 1952, with the discovery of isoniazid and more effective therapy.[26] Finally, we introduce treatment for latent infections likely to become active, in 1966.[15] Infections likely to become active were typically identified in the study population by mass chest x-ray screenings, and risk level was determined based on the results of the chest x-ray.[14,15] In the model, we do not include mass screenings, so any latently infected agent "likely to become active" is eligible to start receiving treatment. We have used "within 24 months of activation" (i.e., 24 time steps or fewer on the activation clock) as a proxy for an infection likely to become active based on TB treatment guidelines.[27,28] Any latently infected agent with 24 or fewer time steps remaining on their activation clock has a probability $p_{prophylaxis}$ of starting treatment at each time step.

The treatment rate (t_{rate}) was calibrated during the testing phase of model building. The length of treatment and probability of successfully completing treatment ($t_{completion}$) are based on previous analyses of the study data.[29] After treatment is introduced in the model, any active agent has some probability of entering treatment, determined by a Bernoulli trial. If the agent is selected for treatment, a Bernoulli trial determines if the agent will successfully complete treatment. If treatment will be completed, the duration is a Gamma random variable with a mean of 12 mo (shape = 12, scale = 1). If the agent ceases treatment before cure, the duration is a Gamma random variable with a mean of 9 mo (shape = 9, scale = 1). Once an agent completes treatment, she or

Table 2 Fitting Procedure

Phase	Variables Fitted	Target Output	Sets Tested
Calibration	Transmission rate (β) Transmissible TB probability (f) Threshold value	Mortality rate (1934–1944)	780^{B} 780^{A} 2340^{T}
Testing	Treatment rate (t_{rate}) Treatment completion ($t_{completion}$)	Mortality rate (1945–1958) Treatment rate (1945–1955)	672^{B} 96^{A} 216^{T}
Validation	—	Incidence rate (1965–1975)	28^{B} 4^{A} 9^{T}

B, base model; A, additive model; T, threshold model.

he moves to the R state. If the agent fails treatment, she or he returns to the active (AT or AN) state she or he came from for the remainder of the duration of the original time in active state. Agents are assigned a new treatment completion probability so that, if selected again, they might successfully complete treatment. Latently infected agents who successfully complete treatment return to the U state.

Input Parameters

The model was parameterized from 3 sources. The first data source includes archival epidemiologic and census data, as described above. Second, published clinical and epidemiologic literature pertaining to TB infection was used for clinical parameters. Third, other TB transmission models were used for parameters pertaining to disease state transitions (Table 1: model parameters).

The archival data set includes TB sanatorium admission and discharge records (described in Zwick and Pepperell[29] and Pepperell et al.[17]), incidence time series (1952–1975),[23] and mortality time series (1935–1958).[23] We parameterized the length of treatment and proportion of successfully completed treatments based on analyses by Zwick and Pepperell.[29] We used mortality time series for calibration and testing and incidence time series for validation output targets. Census data were used to estimate birth and non-TB death rates.[24] Clinical and epidemiologic literature was used to estimate the proportion of agents that activated within 1 y, TB-specific mortality rate, self-cure rate, and relapse rate. Disease state transitions are assumed to be exponentially distributed.[8,21,22,30]

Model Fitting

Target outputs for model fitting were determined by dividing study data into 3 subsets for calibration, testing, and validation based on the time period. Calibration was performed using mortality data from 1934–1944, testing was performed using mortality data from 1945–1959, and validation was performed using incidence data from 1965–1975. None of the parameters that were estimated during calibration were changed during the testing and validation phases, except for the following: treatment rate was increased during the testing phase and treatment rate and treatment completion probability were increased during the validation phase. In addition, we introduced treatment for latent infections starting in 1966 during the validation phase. Treatment rates were determined from previous analyses of medical records.[29] See Supplemental Table S1 for the changes in parameter values through these different phases.

Calibration

Because of the novel contact structure in this model, there are no data available to directly estimate the probability of infection per exposure (β). We also lacked data to directly estimate the proportion of active infections that are transmissible ($p_{transmissible}$). To fit these parameters, we developed a calibration process similar to Codella et al.[31] For β, we tested values ranging from 0.001 to 0.06 incremented by 0.001, and for $p_{transmissible}$, we tested values ranging from 0.25 to 0.85 incremented by 0.05, generating 780 parameter combinations for the base and additive models. For the threshold model, we also calibrated the threshold number of exposures, at 5, 10, and 15 exposures per year for a total of 2340 calibration sets (Table 2: fitting procedure). We simulated each parameter set using1000 replications to sufficiently account for the stochasticity of the model (Supplemental Figure S2).

We determined the best-fitting parameter sets by comparing the simulated mortality rate to the observed

mortality rate from 1934–1944. We calculated the 95% CI around the observed mortality rate, generating an envelope around our target output.[32] We counted the number of times simulated mortality rate violated those confidence bounds (i.e., envelope violations). Then, we calculated the mean square error (MSE) to determine the distance between the simulated mortality rate and observed mortality rate. We define the MSE for parameter set P as

$$MSE_P = \sum_{t=1}^{n} \left(y_t - \frac{\sum_{s=1}^{1000} \hat{y}_{t,P,s}}{s} \right)^2$$

for each time point t and each simulation run s. We retained all parameter sets p that fell within the lowest 10% MSE and had 2 or fewer envelope violations. The set of best-fitting parameter sets, P^{TOP}, was used for all subsequent phases of fitting and analysis.

Testing

We tested 3 different treatment models against mortality rates[23] from 1945–1958 using the distance measurement defined above. For the testing phase, we fit treatment rate per month (t_{rate}) and treatment completion probability ($t_{completion}$). We tested 3 treatment structures (1 introduction in 1952, 2 introductions in 1945 and 1952, and 2 introductions in 1945 and 1952) with linearly increasing treatment intensity over time. For each model, we tested treatment probabilities ranging from 0.04% to 50% per agent per month and treatment completion probabilities of 50%, 55%, and 60% based on data from literature.[23] See Table 2 for fitting sets from each model and the supplemental material for treatment rate functions.

Verification and Validation

We followed verification and validation techniques similar to Codella et al.[31] and Kleijnen.[33] Our model is written in C++ using Microsoft Visual Studio IDE 2017. The model has an agent-level and time step–level debugging output, which was used to determine whether agent behavior matched the intended agent logic (supplemental material).

We validated our model in 2 phases[31]: face validation, calibration, testing, validation against withheld data, and sensitivity analysis. Face validation, an ongoing process, involved checking that the agent logic, environment, agent interactions, disease state transitions, and parameters produced a sufficiently accurate representation of a real TB epidemic given the simplifying assumptions necessary to create the model. One of the authors, an infectious disease physician and TB researcher, provided expertise on TB transmission and control, helping ensure that agents' interactions and TB transmission occurred with sufficient accuracy.

Validation against withheld data was performed by comparing modeled incidence rates against observed incidence in the study population between 1965 and 1975. Incidence was not used as a performance metric at any other point in the development of our model. Sensitivity analysis was performed to examine the effect of changes in input parameters on model results.

Validation against Withheld Data

We withheld all data from 1959–1975 during the building, calibration, and testing phases of model fitting. This includes incidence time series data starting in 1952 through 1975, a subset of which was used as a validation set. The only structural changes to the model after testing are capping the treatment probability at 80%, increasing the treatment completion probability to 90% after 1958, and introducing treatment for agents with latent infections. We take these values from the medical histories of the study population.[15] Treatment for latent infections is introduced in the model at the time step corresponding to January 1965. Treatment probability per agent per time step for latently infected agents likely to become active ($p_{prophylaxis}$) is 0.27 in the base model and 0.08 in the additive and threshold models (Table 1: model parameters).

Sensitivity Analysis

We performed a deterministic 1-way sensitivity analysis to look for a significant change in MSE in modeled output. We varied each parameter within 50% of its fitted value and ran the model holding the rest of the parameter values constant. We considered any MSE from the sensitivity run that exceeded the maximal accepted MSE from P^{TOP} to be a significant change in MSE.

Results

Initial Conditions

We initialized with 3786 agents across 37 neighborhoods with 5 active transmissible infected agents randomly distributed across geographic space. We found that the model enters a steady state of infections just after the 600th time step (50 y of model time), so we set a warm-up period of 660 time steps (55 y). Each parameter set was run for 1000 replications to sufficiently account for the stochasticity of the model (Supplemental Figure S2).

Calibration

We retained 28 sets for P^{TOP} from the base, 4 from the additive, and 9 from the threshold models (Table 2: fitting procedure; Supplemental Figure S3). The base model calibration resulted in an average probability of infection per exposure per month (β) of 1.3% and an average proportion of transmissible active infections ($p_{transmissible}$) of 54%. Both additive and threshold model calibrations tended toward a lower β, with averages of 0.6% and 0.5%, respectively. The additive model calibration resulted in an average $p_{transmissible}$ of 74%, whereas the threshold model averaged 68%. The threshold model calibrated to an average threshold of 10 exposures per year. All parameter results are shown in Tables 1 and 2.

Testing and Treatment Structure

We found that 2 treatment introductions for active infections (1945 and 1952) with increasing intensity resulted in the best fit to mortality data from 1945–1958. We fit a piecewise linear function such that the rate of treatment administered to active cases per month (t_{rate}) increased from 0.04% to 3.5% between 1945 and 1951 and then jumped to 7.2% in 1952, increasing to 36% in 1958. This treatment function was preferred in all 3 versions of the model. The base model had a 55% treatment completion probability ($t_{completion}$), whereas the additive and threshold models had 50% $t_{completion}$.

The treatment probability per latently infected agent per month ($p_{prophylaxis}$) was determined to be 0.27 in the base model and 0.08 in both the additive and threshold models to match the incidence rates during the validation period (Figure 3, left side). We found that significantly more latently infected agents required treatment to achieve the same incidence level in the base model as compared with both complex contagion models (Figure 3, right side).

Validation against Withheld Data

We found that the threshold model had the lowest MSE, followed by the additive and then the base model, fit to 1965–1975 incidence data. However, this seems to be an artifact of how high the modeled 1965 incidence is compared with the observed incidence (Figure 4). The threshold model underestimates the true incidence for most of the validation period, whereas both the base and additive models seem to provide better fit after 1965. Indeed, when we recalculated the MSE excluding 1965, we found that base has the best fit, followed by the additive, then the threshold models. We also found that the base model fits mortality targets from calibration and testing the best (Figure 5). In addition, we found that the additive and threshold models produced fewer latent infections and required less treatment of latent infections to achieve a similar epidemic decline than the base model.

Sensitivity Analysis

We found that all 3 models were most sensitive to changes in birth rate and the proportion of fast activations. The threshold model was particularly sensitive to movement between neighborhoods, whereas the base and additive were not. We found that the models were fairly robust to changes in fast activation rate, recovery rate, relapse rate, slow activation rate, and treatment length. Life expectancy and TB death rate had variable effects on incidence and TB mortality across the 3 models. The full sensitivity results are presented in Supplemental Figures S4 and S5.

Discussion

In this study, we developed and fit an ABM that models TB transmission and control under 3 dynamical schemes. We also estimated the transmission rate and proportion of transmissible TB cases in a high-incidence setting. Together, these provide a framework for future models of TB transmission that may help epidemiologists and infectious disease experts understand the effect of underlying transmission dynamics on the scope of a TB epidemic.

We showed that all 3 versions of the model produce qualitatively realistic epidemic curves (Figure 4). Each version of the model reproduces the relationship between mortality and incidence in settings before/without and after/with antibiotic treatment from canonical works by Styblo[34] and Grigg.[19] Although the base model provides the best fit to mortality data, we found that the threshold model fits the incidence better from 1952 onward (Figure 4), despite overestimating mortality in previous years. This is because more infections activate sooner in the complex contagion versions of the model, resulting in more deaths early on, and a smaller pool of latent infections remain after the initial epidemic subsides. The burden of latent infections affects the reemerging epidemic: once the initial epidemic is brought under control, the incidence starts climbing again around 1959 (Figure 4). This secondary epidemic is most apparent in the base model, which produced a steep uptick in incidence starting in 1959, whereas both complex contagion models produced incidence curves that appear to plateau rather than start a secondary epidemic (Figure 4). We also found that all 3 models fit a high proportion of fast

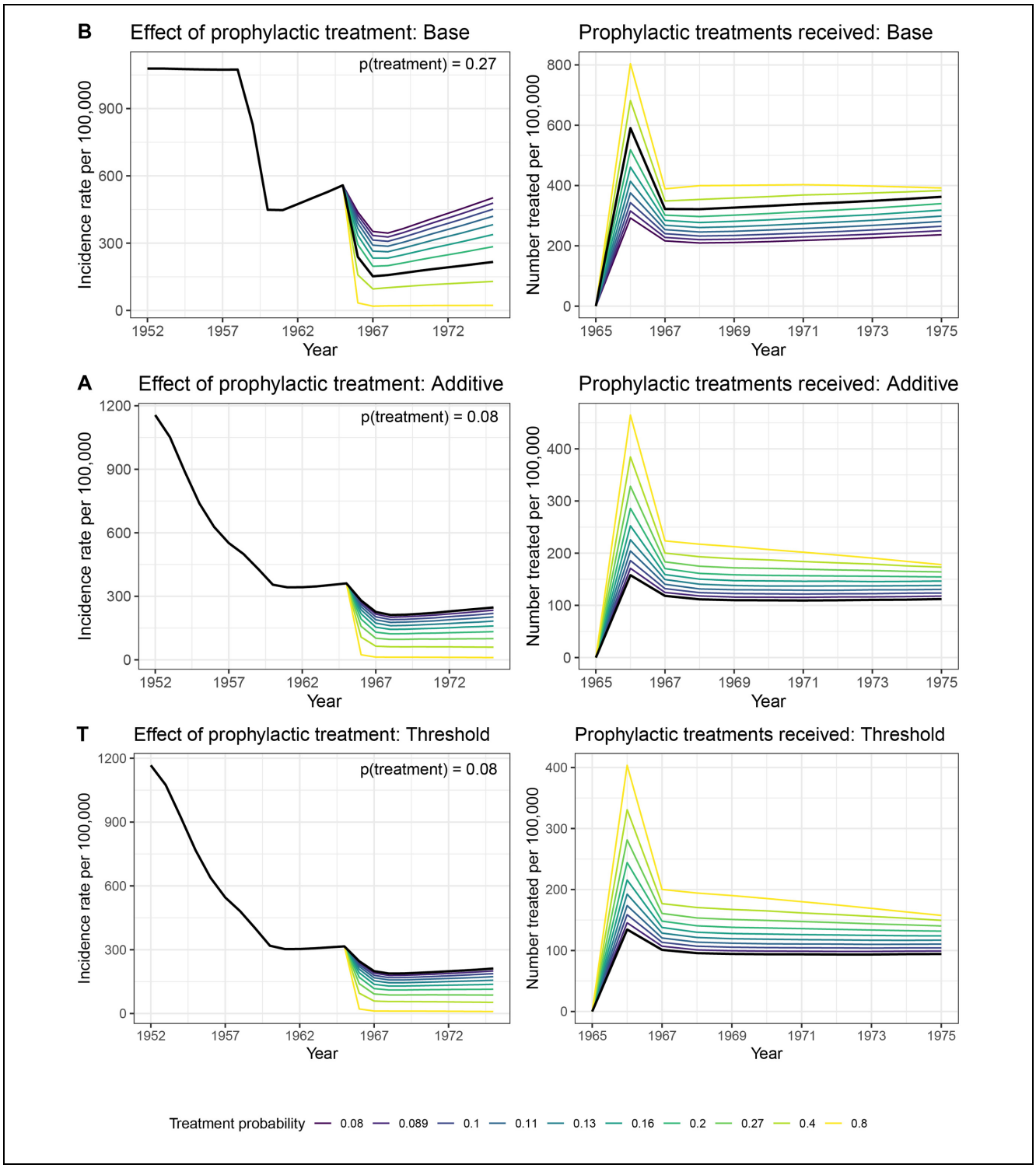

Figure 3 (L) The effect of treatment of latent tuberculosis infections on incidence rates for the base (B), additive (A), and threshold (T) models under varying treatment probabilities. (R) Latent infection treatment rate per year across treatment probabilities. The treatment probability is per latently infected agent per month. The modeled values are bolded in black: 0.27 for the base model and 0.08 for the additive and threshold models.

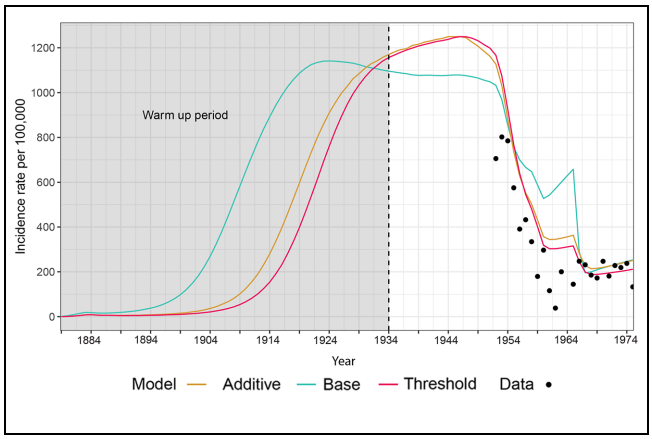

Figure 4 Modeled tuberculosis incidence rate for base (blue), additive (orange), and threshold (pink) models averaged over all best-fitting sets. Points represent observed incidence rates.

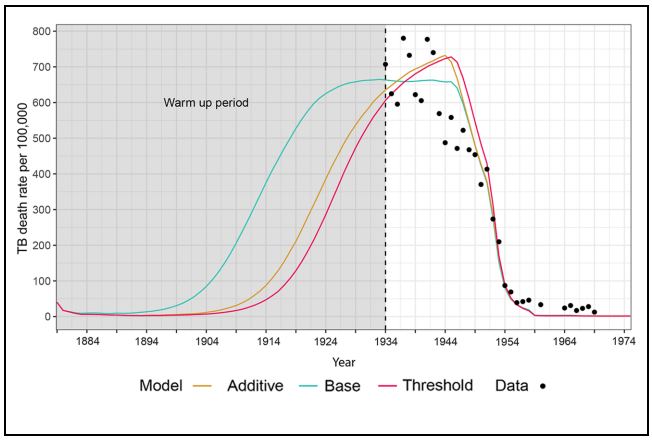

Figure 5 Modeled tuberculosis (TB)–specific mortality rate for base (blue), additive (orange), and threshold (pink) models averaged over all best-fitting sets. Points represent observed TB mortality rates.

activations as compared with other TB models,[12,21,35] although our estimates did not exceed the reported values in the medical literature.[18] We expect a high proportion of fast activations under complex contagion[5,12] and hypothesize that some of the dynamical impacts of complex contagion are captured in this parameter.

There are few ABMs of TB transmission dynamics fit to data. There are models using theoretical populations to test the effect of close versus casual contacts,[10] determine the role of age structure in transmission,[36] or predict the emergence of drug resistance.[37] However, none of these are fit to data, nor do they consider complex transmission dynamics. Tuite et al.[38] fit a model to data in a similar high-incidence setting in Northern Canada; however, their model only captures a 10-y period. TB epidemics have slow transmission dynamics and long doubling times[21]; thus, to fully capture a TB epidemic, we modeled almost 100 y and fit more than 40 y to data. To our knowledge, there is only 1 previous transmission model that has been used to study complex contagion: the study by Ackley et al.,[12] which was used as a starting point for our model. Like Ackley et al., we found that of the 2 complex contagion schemes, threshold fits better than additive and base models to incidence; however, the base model had a better fit to mortality during calibration and testing.

One of the most striking differences in the complex contagion models compared with the base model is the burden of latent infections. This is evident in both the reemerging epidemic, around 1959, as slow progressing latent infections start to activate and increase incidence rates (Figure 4), and in the differential effects of treatment for latent infections across models (Figure 3). To achieve similar incidence rates in the base model as compared with either complex contagion model, the probability that a latently infected agent starts treatment each month ($p_{prophylaxis}$) is more than 3 times higher in the base model at 0.27 as compared with 0.08 in both complex contagion models. In addition, latent agents receive treatment in the base model at or more than 4 times the rate as the complex contagion models in 1966, the peak of latent treatment in the model (600 per 100,000 population in the base model versus 125–150 per 100,000 population in the additive and threshold models, respectively; Figure 3). At the same $p_{prophylaxis}$ value, the base model has a significantly higher incidence while still treating more latently infected agents than either complex contagion model (Figure 3). This suggests that, depending on dynamical context (complex contagion or no complex contagion), we can expect widely different burdens of latent TB infection in a population with widely different latent treatment rates required to prevent a secondary epidemic.

In addition to wide variation in the amount of underlying latent infections across models, we found that complex contagion produces a steeper epidemic curve (Figure 4). Both model behaviors have implications for treatment interventions. Because the epidemics take off more steeply under complex contagion, early active case detection and intervention would be crucial to bring the epidemic under control in complex contagion conditions. Fewer latent infections would reduce the amount of latent detection and treatment under the same circumstances, which is apparent in the lower latent treatment rates from the complex contagion models (Table 1). Whereas in the base model, we observed a longer epidemic start up and

many more latent infections, as evidenced by the increasing incidence rate starting around 1959 (Figure 4). This suggests that in the absence of complex contagion, focusing on treatment for latently infected individuals is important for completely stopping the outbreak.

We acknowledge that the study data are not granular enough to fully distinguish all 3 exposure models. For example, although we note a major difference in the burden of latent TB infection in both complex contagion models as compared with the base model, we do not have study data to compare these model results with the true burden of latent TB infection. Similarly, the active TB incidence and TB mortality time series used for fitting are not detailed enough to fully distinguish the 3 exposure models as in Ackley et al.,[12] which used contact tracing data. However, a successful fit of all 3 models to the data demonstrates the feasibility of incorporating complex contagion in dynamical models of TB transmission and suggests the potential importance of incorporating complex contagion transmission dynamics into future TB models.

Our model does have some additional limitations. Because of data limitations, we did not introduce treatment for latent infections until 1966, although medical histories[14,15] and model behavior suggest that latent treatment may have been introduced earlier. We also assume that treatment has a binary outcome: successful completion or failure to complete treatment. If an agent fails treatment, they return to their previous active disease state. In reality, individuals who do not complete treatment, receive only partial treatment, or have their treatment interrupted and later resumed have a higher risk of developing difficult-to-treat multidrug resistant TB,[4,39] which we do not model. However, while this is an area for further exploration, we do not find evidence that changes to treatment (Supplemental Figures S4 and S5) affected the output of our model significantly. The model could be easily adapted to include different treatment outcomes. Finally, although we include structure at the community level, we do not account for household structure or network movement within a neighborhood. This simplifying assumption was deemed appropriate to study the direct effects of complex contagion, although we acknowledge that contact structure plays an important role in TB transmission.[10,40]

We developed this model with an appropriate level of detail to capture individual-level interactions while still being generalized enough to avoid extreme assumptions or limitations because of data scarcity. To maintain the focus on population-level effects of transmission dynamics, we did not consider different treatment efficacy or effectiveness in this model. Further, we collapsed noninfectious pulmonary TB and extrapulmonary TB into 1 category, capturing differences in infectious TB only to reduce the number of parameters we had to fit.

All 3 versions of this model serve as a basis for future, more complex models of TB transmission and control. Future applications of the model may explore different formulations for complex contagion, including more complex risk functions, or investigate the applicability of complex contagion in lower-incidence settings.

ORCID iD

Oguzhan Alagoz https://orcid.org/0000-0002-5133-1382

Supplemental Material

Supplementary material for this article is available on the *Medical Decision Making* website at http://journals.sagepub.com/home/mdm.

References

1. World Health Organization. *Global Tuberculosis Report 2019*. Geneva (Switzerland): World Health Organization; 2019.
2. World Health Organization. *WHO: Global Tuberculosis Report 2018*. September 18, 2018. Available from: http://www.who.int/tb/publications/global_report/en/. Accessed February 15, 2019.
3. Dye C, Glaziou P, Floyd K, Raviglione M. Prospects for tuberculosis elimination. *Annu Rev Public Health*. 2013; 34(1):271–86.
4. Lange C, Chesov D, Heyckendorf J, Leung CC, Udwadia Z, Dheda K. Drug-resistant tuberculosis: an update on disease burden, diagnosis and treatment. *Respirology*. 2018; 23(7):656–73.
5. Lee RS, Proulx J-F, Menzies D, Behr MA. Progression to tuberculosis disease increases with multiple exposures. *Eur Respir J*. 2016;48(6):1682–9.
6. Cardona P-J. A dynamic reinfection hypothesis of latent tuberculosis infection. *Infection*. 2009;37(2):80.
7. Vynnycky E, Fine PE. Interpreting the decline in tuberculosis: the role of secular trends in effective contact. *Int J Epidemiol*. 1999;28(2):327–34.
8. Feng Z, Castillo-Chavez C, Capurro AF. A model for tuberculosis with exogenous reinfection. *Theor Popul Biol*. 2000;57(3):235–47.
9. Acuña-Villaorduña C, Jones-López EC, Fregona G, et al. Intensity of exposure to pulmonary tuberculosis determines risk of tuberculosis infection and disease. *Eur Respir J*. 2018;51(1).
10. Kasaie P, Andrews JR, Kelton WD, Dowdy DW. Timing of tuberculosis transmission and the impact of household contact tracing. an agent-based simulation model. *Am J Respir Crit Care Med*. 2014;189(7):845–52.

11. Chhatwal J, He T. Economic evaluations with agent-based modelling: an introduction. *Pharmacoeconomics*. 2015; 33(5):423–33.
12. Ackley SF, Lee RS, Worden L, et al. Multiple exposures, reinfection and risk of progression to active tuberculosis. *R Soc Open Sci*. 2019;6(3):180999.
13. Public Health Agency of Canada. Epidemiology of tuberculosis in Canada. In: *Canadian Tuberculosis Standards*. 7th ed. Ottawa: Public Health Agency of Canada; 2014. https://www.canada.ca/en/public-health/services/infectious-diseases/canadian-tuberculosis-standards-7th-edition/edition-13.html. Accessed January 29, 2020.
14. Lux MK. *Medicine That Walks*. Toronto (Canada): University of Toronto Press; 2001.
15. Wherrett GJ. *Miracle of the Empty Beds: History of Tuberculosis in Canada*. Toronto (Canada): University of Toronto Press; 1978.
16. Waldram JB, Herring DA, Young TK. *Aboriginal Health in Canada*. 2nd ed. Toronto (Canada): University of Toronto Press; 2006.
17. Pepperell C, Hoeppner VH, Lipatov M, Wobeser W, Schoolnik GK, Feldman MW. Bacterial genetic signatures of human social phenomena among *M. tuberculosis* from an Aboriginal Canadian population. *Mol Biol Evol*. 2010; 27(2):427–40.
18. Comstock GW. Epidemiology of tuberculosis. *Am Rev Respir Dis*. 1982;125(3P2):8–15.
19. Grigg E. The arcana of tuberculosis with a brief epidemiologic history of the disease in the U.S.A. *Am Rev Tuberc*. 1958;78(2):151–72.
20. Tiemersma EW, van der Werf MJ, Borgdorff MW, Williams BG, Nagelkerke NJD. Natural history of tuberculosis: duration and fatality of untreated pulmonary tuberculosis in HIV negative patients: a systematic review. *PLoS One*. 2011;6(4).
21. Blower SM, Mclean AR, Porco TC, et al. The intrinsic transmission dynamics of tuberculosis epidemics. *Nat Med*. 1995;1(8):815–21.
22. Ackley SF, Liu F, Porco TC, Pepperell CS. Modeling historical tuberculosis epidemics among Canadian First Nations: effects of malnutrition and genetic variation. *Peer J*. 2015;3:e1237.
23. Tuberculosis Morbidity and Mortality 1938-1963, National Health and Welfare Volume 1224. RG29 File 311-T7-7. Ottawa: Library and Archives Canada.
24. Library and Archives Canada. Indian Affairs annual reports, 1864-1990. March 19, 2013. http://www.bac-lac.gc.ca/eng/discover/aboriginal-heritage/first-nations/indian-affairs-annual-reports/Pages/introduction.aspx. Accessed March 5, 2020.
25. Hinshaw C, Feldman WH, Pfuetze KH. Treatment of tuberculosis with streptomycin: a summary of observations on one hundred cases. *J Am Med Assoc*. 1946;132(13): 778–82.
26. East African/British Medical Research Councils. Controlled clinical trial of four short-course (6-month) regimens of chemotherapy for treatment of pulmonary tuberculosis: third report. *Lancet*. 1974;304(7875):237–40.
27. Iseman MD. *A Clinician's Guide to Tuberculosis*. Philadelphia: Lippincott Williams & Wilkins; 2000.
28. Sterling TR, Njie G, Zenner D, et al. Guidelines for the treatment of latent tuberculosis infection: recommendations from the National Tuberculosis Controllers Association and CDC, 2020. *Am J Transplant*. 2020;20(4):1196–206.
29. Zwick ED, Pepperell CS. Tuberculosis sanatorium treatment at the advent of the chemotherapy era. *BMC Infect Dis*. 2020;20(1):831.
30. Aparicio JP, Castillo-Chávez C. Mathematical modelling of tuberculosis epidemics. *Math Biosci Eng*. 2009;6:209–37.
31. Codella J, Safdar N, Heffernan R, Alagoz O. An agent-based simulation model for *C. difficile* infection control. *Med Decis Mak Int J Soc Med Decis Mak*. 2015;35(2): 211–29.
32. Alagoz O, Ergun MA, Cevik M, et al. The University of Wisconsin breast cancer epidemiology simulation model: an update. *Med Decis Making*. 2018;38(1 suppl):99S–111S.
33. Kleijnen JPC. Verification and validation of simulation models. *Eur J Oper Res*. 1995;82:145–62.
34. Styblo K. The relationship between risk of tuberculosis infection and risk of developing infectious tuberculosis. *Bull Int Union Tuberc Lung Dis*. 1985;60:117–9.
35. Gomes MGM, Franco AO, Gomes MC, Medley GF. The reinfection threshold promotes variability in tuberculosis epidemiology and vaccine efficacy. *Proc R Soc Lond B Biol Sci*. 2004;271(1539):617–23.
36. Graciani Rodrigues CC, Espíndola AL, Penna TJP. An agent-based computational model for tuberculosis spreading on age-structured populations. *Phys Stat Mech Its Appl*. 2015;428:52–9.
37. de Espíndola AL, Bauch CT, Cabella BCT, Martinez AS. An agent-based computational model of the spread of tuberculosis. *J Stat Mech Theory Exp*. 2011;2011(05):P05003.
38. Tuite AR, Gallant V, Randell E, Bourgeois A-C, Greer AL. Stochastic agent-based modeling of tuberculosis in Canadian Indigenous communities. *BMC Public Health*. 2017;17:73.
39. Faustini A, Hall AJ, Perucci CA. Tuberculosis treatment outcomes in Europe: a systematic review. *Eur Respir J*. 2005;26(3):503–10.
40. Song B, Castillo-Chavez C, Aparicio JP. Tuberculosis models with fast and slow dynamics: the role of close and casual contacts. *Math Biosci*. 2002;180(1):187–205.