Agent-Based Models to study tuberculosis in Nigeria

Marta Florido, Joan Llobera, and Isabel Medrano
(Dated: May 31, 2019)

Abstract. Tuberculosis (TB) is an infectious disease with a high prevalence in Nigeria. Its effect on the population has been modeled by adapting an agent-based model (ABM) designed for the study of TB in Ciutat Vella (Barcelona). After fitting the parameters, virtual experiments were performed increasing the notification rate, with the aim of exploring how a better education about TB of the population would improve the current situation. The results show that, in fact, ABM are useful tools to model reality, and the experiments suggest that by increasing the notification rate from current 16% to 25-30%, the prevalence decreases until eradication.

I. INTRODUCTION

Tuberculosis (TB) is an infectious disease caused by the bacteria *Mycobacterium tuberculosis* and spread through saliva droplets. It is one of the most extended diseases in the world: estimations say that around one third of the global population carries the bacteria. However, not all of them develop the illness. TB presents itself in two different forms: latent tuberculosis (infected) is asymptomatic and it is not contagious; people with tuberculosis disease (sick) present symptoms and can spread the bacteria [1].

Nigeria is the seventh country with higher TB burden and the second in Africa. It is estimated that over 400.000 people get infected by TB each year in Nigeria, and 155.000 people die from it. Massive urbanization with a rapidly growing population (at the current growth rate of 2.8%-3%, it is estimated that Nigeria's urban population will double in the next two decades [2]) have caused TB to be endemic in Nigeria. The notification rate is low (16-24%), and the problem is aggravated by the issues of drug-resistant TB and the HIV/AIDS epidemic, which increases the probability of developing the disease [3]. Gombe State is one of Nigeria's 36 districts. Its average TB prevalence over 10^5 people between 2012 and 2016 was of 363.6 people. Out of them, only 58.7 were diagnosed (16.14%) [4]. Despite the efforts made, TB incidence has remained constant in the last years, so it is necessary to develop new strategies and assess their effectiveness to decide which ones should be implemented.

The objective is to adapt an ABM designed to study TB at Ciutat Vella (Barcelona) [5] to the situation of Gombe State, changing both the parameters and the model itself. Once the model is fitted, some experiments will be performed to explore how the incidence of TB would change if some parameters were improved and people were more educated about TB.

II. MODEL DETAILS

A. ABM as a bottom-up approach

Classical epidemiology uses a top-down approach with SEI models [6]: the model describes the overview of the system without defining any subsystems. ABM (Agent-Based Models) take a different approach, a bottom-up one: they are built by describing the specific interactions between the agents and adjusting the parameters by comparing the average of several simulations of the whole system with the real global data [5].

These different approaches - top-down and bottom-up - meet different needs, the main difference being the size of the population to study. ABM requires creating individuals and simulating their interactions: when the populations are too big, computational times can become an insurmountable hurdle. Moreover, in this case, the individuals lose importance and only the averages are significant.

Nigeria has a population of over 190 million, while Gombe State's population is 3 million. While a bottom-up approach is impracticable to study TB in Nigeria, it can be used for Gombe, and it will show the heterogeneity of the system: chance plays a big role and different results can be obtained while using the same parameters. This heterogeneity is crucial in public health matters. Therefore, an ABM will be used to study TB in Gombe.

B. Description of the model

The basic entities of the system are agents. Each of them represents a person with specific properties and a certain status within the infection cycle.

Agents are placed in a two-dimensional grid, and every day, they move and increase their age. Healthy (susceptible plus recovered) people are an element of the grid (their characteristics won't be assigned until infection occurs), while infected, sick, under treatment and treated agents have certain characteristics, both mutual and specific to their category. The total population is assumed to be constant: when an agent dies, a new healthy person is automatically generated in the grid.

Healthy people can become infected if they meet a sick agent (actually, they must remain in contact with them for more than 6 hours a day, for a period of at least 2-3 months [1]). Then, an infected agent can develop the illness with a certain probability, which decreases with time. In fact, it will be assumed that if an agent stays infected for longer than 7 years, it will need to re-encounter a sick agent to be able to develop it [7]. Any infected agent can be re-infected after being re-exposed to a sick agent. A sick person can either: *(i)* die from TB, *(ii)* get diagnosed and receive treatment, or *(iii)* have a spontaneous recovery if it remains ill for longer than 3 years. If a sick person is smear positive (higher bacterial load), the probability of it infecting other agents will double [5].

The treatment should last for 6 months. However, people may leave treatment before finishing it. If it lasts less than

15 days, then the agent will become ill immediately, since the bacterial count will not have lower enough. After that, the longer it stays under treatment, the less their probability of relapse will be once it finishes the treatment and becomes treated (the probability will be assumed to be linear from 100% at 15 days to 1% at 6 months). During the first 15 days of treatment, an agent under treatment will be infectious.

A treated agent can either: *(i)* do not relapse in the following 2 years, so it is assumed to have recovered, *(ii)* get in contact with a sick person and become infected, or *(iii)* relapse and go back to being sick.

A schematic summary of the model can be found in figure 1. A more detailed description can be consulted at the original paper [5]. The focus of this description will be the implemented changes:

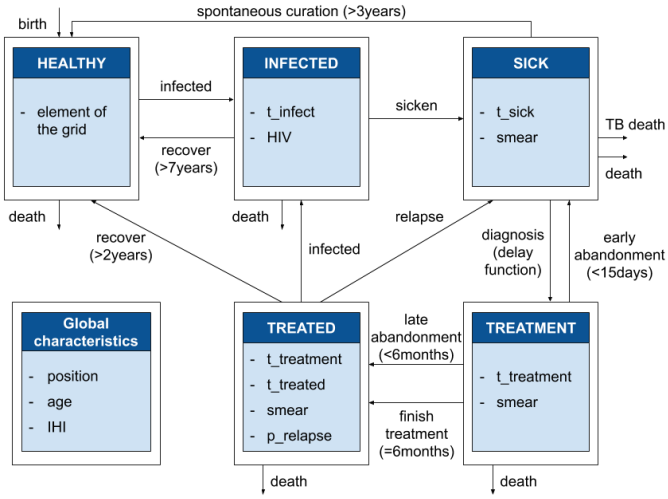

FIG. 1. Schematic structure of the model. The different states of the agents are shown together with the interactions between them and the specific characteristics in each phase. Global characteristics refer to those applied to all agents.

- **Simplifications**: it will be assumed that some characteristics of the agents implemented for Ciutat Vella will not affect the Gombe model: smoking, diabetes (lack of data) or being foreign (not relevant in this context). Moreover, the characteristics of the agents will be assigned from those of the population, since there is not enough data to create probability functions that link infected agents with infecting agents.

- **Acquired immunity**: in Ciutat Vella, since the prevalence is so low, it is assumed that an agent can not have TB twice in its life. In Gombe, the prevalence is much higher, so it is possible that an infected agent has already gone through the illness before. In this case, the infected agent will have a probability of developing the illness (*p_infect*) 7 times higher [8]. Moreover, this increased exposure has created some kind of immunity on 70% population, either because they have already gone through the disease, they have been infected before, or they have been exposed to it. This immunity factor decreases the probability of developing the illness by a factor of 0.1 [9].

  These two effects can be merged into two new parameters. It will be assumed that some percentage of the people (*p_immune_total*) will develop the sickness, with a probability multiplied by some factor (*f_immunity*$\geq$ 0.1) to take into account both of the effects explained before. Each agent will have a characteristic infection history index (*IHI*) that will determine whether the factor *f_immunity* should be applied or not.

- **Diagnosis process**: in Ciutat Vella, all TB cases are detected. In Gombe, only 16% [4] (*p_going_to_hospital*) of the sick people will go to the hospital, with an average delay of 90 days. Moreover, once the treatment has started, the agent is still contagious for 15 days. In Ciutat Vella, protective measures are taken so that they do not become in contact with susceptible people within that period; however, this does not happen in Gombe and they may spread the disease for another 15 days.

- **Spontaneous recovery**: in Gombe, some agents will be sick for longer than 3 years. Once this time has passed, the agents have an annual probability of recovering of 30% (*p_spontaneous_recovery*) [10].

III. EXPERIMENTAL DATA

The main sources of data for TB in Nigeria are the World Health Organization (WHO), which publishes annual reports on several epidemics, including specific reports for high burden countries. Experimental data about prevalence (average of 363.6/10^5 over five years, between 2012 and 2016) is taken from WHO [3]. Another important source is the National TB and Leprosy Control Programme (NTBLCP). The total amount of diagnosed people in Gombe (average of 58.7/10^5 between 2012 and 2016) is taken from them [4].

Moreover, during fieldwork in Gombe State between July and October 2018, data from 54 diagnosed people was obtained through personal interviews with patients from different health centers in the District, related to their age and gender, their occupation, their living arrangements, risk factors (HIV infection, smoking) or the time it took them to go to the hospital since the symptoms first appeared [11]. The histogram of this delay can be found in figure 2.

IV. RESULTS

A. Initial estimation

The initial estimation of parameters and initial conditions are based on bibliographic sources and will be used as a starting point for the simulations. Table I shows the initial estimation of the parameters with their source.

TABLE I. Initial parameter estimation (prior to fitting)

Parameter	Estimation	Source
Healthy	6290[(1)]	Estimation
Infected	3600	SEI models
Sick	36	Exp.Data [3]
Treatment	3[(2)]	Exp.Data [4]
Treated	71	Estimation
p_infect	0.0152207	[(3),(4)]
p_infected_male[(5)]	0.526	[4]
p_spontaneous_recover	0.3	[10]
p_going_to_hospital	0.16	[6]
p_smear (smear-positive)	0.3428	[12]
p_HIV (HIV-positive)	0.049	[4]

(1) Total population is assumed to be 10000
(2) Six people receive a 6-month treatment a year.
In a given moment, three people are being given treatment
(3) For smear-positive (for smear-negative, apply 0.5 factor) [5]
(4) Assuming 20 infections/sick/year and a mean density of 0.4 [7]
(5) The proportion of infected agents that are males

B. Delay function

The objective is to adjust the data of the time before diagnosis to a probability function, so that the likelihood of going to the hospital can be computed for every sick person, every day. There are two main constrictions: *(i)* agents take at least 20 days to go to the hospital and *(ii)* only 16% of sick agents will go to the hospital, therefore, the probability function must be zero between days 0 and 20, and its area must be 0.16.

The outliers were deleted and a script in R was used to adjust several functions (normal, log-normal, Poisson, negative binomial, exponential, gamma, logistic, geometric) with MLE and compare their goodness-of-fit applying Akaike's Information Criterion (AIC). The top-three best adjustments can be found in figure 2.

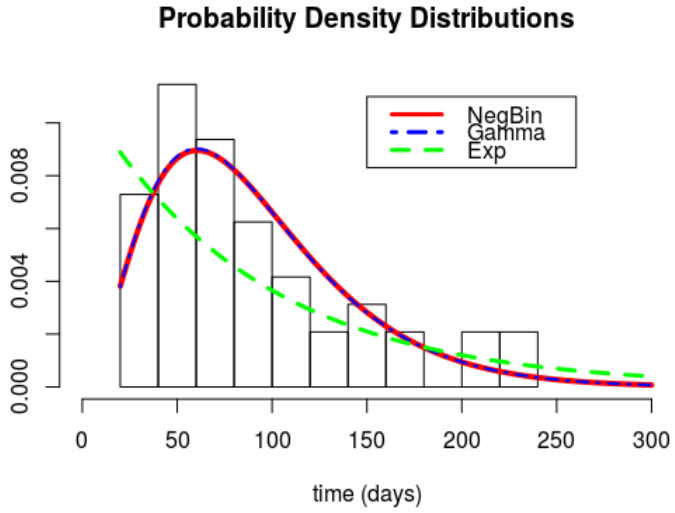

FIG. 2. Fitting of the delay function. PDF of top-three best adjusts: (1) negative binomial, (2) gamma and (3) exponential distributions.

The chosen distribution is the Negative Binomial, with PDF: $\frac{\Gamma(x+n)}{\Gamma(n)x!}p^n(1-p)^x$ with $n = 3.102$, $\mu = n\frac{1-p}{p} = 90.055$

C. Fitting parameters

The prevalence in Gombe State has remained constant over the last years, so the situation is assumed to be stationary. Therefore, the simulations must run long enough to achieve an equilibrium state. Once it is reached, the parameters and initial conditions are changed to improve the rates between the categories: for example, if the simulation yields too many people under treatment, the probability of going to the hospital may have to be reduced. This process is repeated until the equilibrium state of the simulation matches the experimental data. Then, the simulation can be considered to be representative of the situation in Gombe State.

The focus of the fitting has been the following parameters: *p_immune_total, f_immunity, p_go_to_hospital, p_spontaneous_rec, p_infect*. However, since the population considered is low, high precision cannot be achieved and the randomization has a huge effect. The following facts were observed:

(1) Once the number of sick agents went below 50, TB was eradicated. (2) Once the number of infected agents went above 3500, the number of sick agents went above 350.

Therefore, obtaining a prevalence around 36 and a total number of infected people around 3600 is impracticable with these parameters. The simulations achieved values close to the data but did not stay at them for long enough to be considered stable points.

After several sets of parameters tried, the result chosen was the one that: *(i)* had reached a stable state (changes ±10% around the average), *(ii)* got closer to the simulations and *(iii)* had parameters consistent with the bibliographic sources. The final estimations of the initial conditions are computed by taking the mean of the last 100 years of the mean of the 200-year simulations. Their initial and final estimations can be found on table II.

TABLE II. Parameter fitting
Initial and final estimations of parameters and initial conditions.

Parameter	Initial	Final
Healthy	6290	7812
Infected	3600	1996
Sick	36	125
Treatment	3	17
Treated	71	50
p_immune_total	0.7	0.82
f_immunity	0.2	0.1
p_infect	0.003805	0.0033484

There are some differences between the results obtained and the data from Gombe. Some improvements can be implemented to the simulations, that should be explored to keep increasing the veracity and accurateness of the model. Computational time made impossible to implement them at this stage.

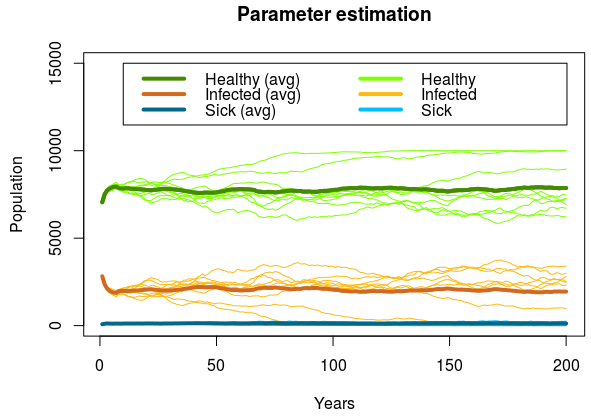

FIG. 3. Parameter estimation: reaching equilibrium with the first set of initial conditions and the final set of parameters. Nine simulations are plotted, with their average emphasized.

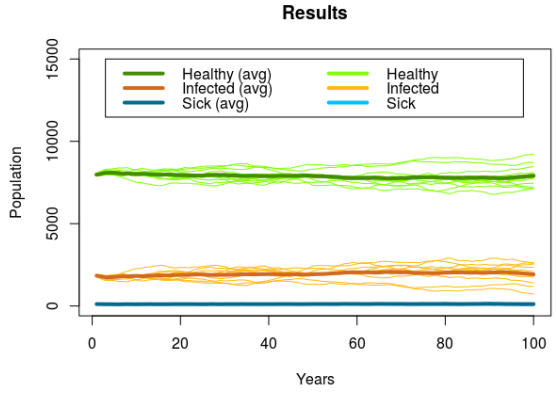

FIG. 4. Results 1: reaching equilibrium with the final set of initial conditions and set of parameters. Ten simulations are plotted, with their average emphasized.

One example is increasing the population: it would decrease the effect of randomization that makes that, with the same parameters, several different outcomes are obtained. The stability around the equilibrium points would be more significant, and the precision would increase. Moreover, the number of simulations when computing averages could be increased: the more simulations ran (more data to compute averages), the more certainty could be had over the results obtained.

V. EXPERIMENTS

The virtual experiment performed is to increase the notification rate (decreasing the delay time will not have a relevant effect if only a small proportion of the sick agents go to the hospital). In fact, educating the population about TB, its symptoms, treatment, and preventive measures, would imply an increase in the notification rate as well as a decrease of delay time, treatment abandonment rate, probability of infection (for example, agents under treatment during the first two weeks would no longer spread the disease...).

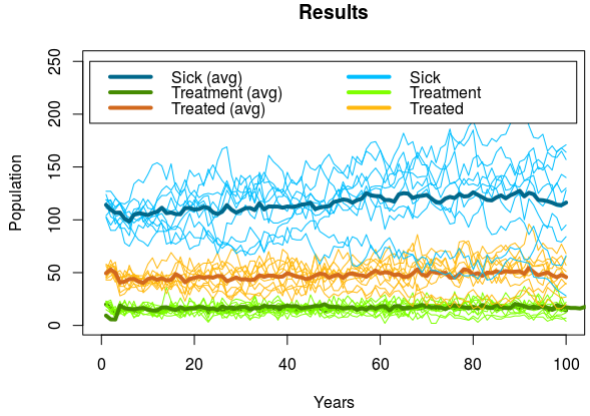

FIG. 5. Results 2: number of sick, treatment and treated agents.

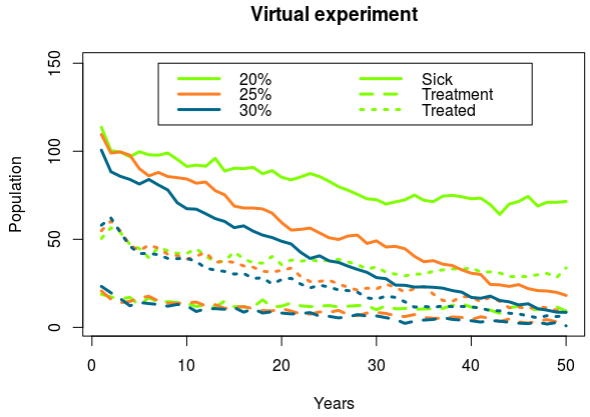

FIG. 6. Virtual experiments: increasing the notification rate from 16% to 20%, 25% and 30%. Eradication of TB achieved at 25%.

The results of increasing the notification rate from 16% to 20%, 25% and 30% can be found in figure 6. TB eradication is achieved in the second and third simulation.

VI. CONCLUSIONS

Models do not just describe reality, but also allow a way to experiment with it. They can be used to discover how some changes may affect the global situation without having to implement them in real life. In this case, ABMs have proven to be very useful to study the effect of TB in Gombe, and how the situation could be improved by the implementation of different strategies. In the present work, we have assessed the effectiveness of increasing the notification rate. Improvements in the model parameters and simulations can still be done in order to obtain results that are closer to reality.

VII. ACKNOWLEDGMENTS

We would like to thank Clara Prats and Nura M.R. Ahmad for their assessment and kind help throughout the work.

[1] Cardona P-J. *Patogenesis de la tuberculosis y otras micobacteriosis.* Enferm Infecc Microbiol Clin. 2018;36:38-46.

[2] Aliyu, A. A. et al (2017). *Urbanization, cities, and health: The challenges to Nigeria - A review*. Annals of African Medicine, 16(4), 149-158.

[3] Global tuberculosis report 2018. Geneva: World Health Organization; 2018. Nigeria TB country profile.

[4] Data from the National Tuberculosis and Leprosy Control Programme, NTBLCP.

[5] Prats et al (2016). *Individual-Based Modeling of Tuberculosis in a User-Friendly Interface: Understanding the Epidemiological Role of Population Heterogeneity in a City. Front.* Microbiol. 6:1564

[6] Ahmad, N. M. R. et al (2018). *Analyzing Policymaking for Tuberculosis Control in Nigeria.* Complexity, 2018, 1âĂŞ13

[7] Sloot, R. et al (2014). *Risk of Tuberculosis after Recent Exposure. A 10-Year Follow-up Study of Contacts in Amsterdam.* American Journal of Respiratory and Critical Care Medicine, 190(9), 1044-1052.

[8] : Uys P. et al (2015). *The Risk of Tuberculosis Reinfection Soon after Cure of a First Disease Episode Is Extremely High in a Hyperendemic Community.* PLoS ONE, 10(12), e0144487.

[9] Heimbeck, J. (1928). *Immunity to tuberculosis.* Archives of Internal Medicine, 41(3), 336.

[10] Tiemersma, E. W.et al (2011). *Natural History of Tuberculosis: Duration and Fatality of Untreated Pulmonary Tuberculosis in HIV Negative Patients: A Systematic Review.* PLoS ONE, 6(4), e17601.

[11] Fieldwork performed by N.M.R.Ahmad in Gombe State between July and October 2018. Data from 54 diagnosed people was obtained through personal interviews with patients from different health centers in the District.

[12] Obasanya, J. et al (2015). *Tuberculosis case detection in Nigeria, the unfinished agenda.* Tropical Medicine and International Health, 20(10), 1396-1402.