A Stochastic Model to Simulate the Spread of Leprosy in Juiz de Fora*

Vinícius Clemente Varella[1], Aline Mota Freitas Matos[2], Henrique Couto Teixeira[2], Angélica da Conceição Oliveira Coelho[3], Rodrigo Weber dos Santos[1], and Marcelo Lobosco[1]

[1] Graduate Program on Computational Modeling
[2] Graduate Program on Biological Sciences, PPGCBIO
[3] Graduate Program in Nursing
UFJF, Rua José Lourenço Kelmer, s/n, 36036-330, Juiz de Fora, MG, Brazil
vcv_varella@hotmail.com,alinemotafreitas@yahoo.com.br
{henrique.teixeira,angelica.fabri,rodrigo.weber,marcelo.lobosco}@ufjf.edu.br

Abstract. This work aims to simulate the spread of leprosy in Juiz de Fora using the SIR model and considering some of its pathological aspects. SIR models divide the studied population into compartments in relation to the disease, in which S, I and R compartments refer to the groups of susceptible, infected and recovered individuals, respectively. The model was solved computationally by a stochastic approach using the Gillespie algorithm. Then, the results obtained by the model were validated using the public health records database of Juiz de Fora.

Keywords: Leprosy, computational modelling, epidemiology, compartmental model, SIR model, Gillespie's algorithm, SSA algorithm

1 Introduction

Despite the decrease in number of leprosy cases in the world, some countries, like India, Indonesia and Brazil, still have difficulties in controlling this disease, which represents a big challenge to their public health systems [13]. In 2016, 22,710 new cases were registered as receiving standard multidrug therapy (MDT) in Brazil, with a registered prevalence rate of 1.08 per 10,000 population [13]. Therefore, Brazil has not yet eliminated leprosy as a public health problem: elimination is defined as the reduction of prevalence to a level below one case per 10,000 population [12].

Mathematical and computational tools can be useful to understand the spread of leprosy. A computational model is the implementation, using programming languages, of a mathematical model that describes how a system works. Then, simulations are made with the purpose of studying the behaviour of this system in the occurrence of distinct scenarios. The main contribution of this paper is

* The authors would like to thank UFJF, FAPEMIG, CAPES, CNPq and Minas Gerais Secretary of State for Health.

ICCS Camera Ready Version 2018
To cite this paper please use the final published version:
DOI: 10.1007/978-3-319-93713-7_51

the development of a computational model that simulates, with a reasonable degree of fidelity, the spread of leprosy in a Brazilian city, Juiz de Fora, over time, using for this purpose a compartmental model, SIR (Susceptible, Infected, Recovered), solved using a stochastic approach. The results obtained from the model are then validated through a comparison with historical data of the diagnosis of the disease.

Compartmental models are frequently used to simulate the spread of diseases. For example, a SIR model was used to simulate the transmission of H1N1 virus [5], Dengue virus [2] and Cholera [3]. A recent work [10] proposed a compartmental continuous-time model to describe leprosy dynamics in Brazil. Approximate Bayesian Computation was used to fit some parameters of the model, such as the transmission coefficients and the rate of detection, using for this purpose leprosy incidence data over the period of 2000 to 2010. Then, the model was validated on incidence data from 2011 to 2012. In this work, a much simpler model was used to describe the leprosy dynamics in Juiz de Fora. The number of parameters used in our model is reduced, which allow us to fit them manually to data. Also, other work [1] used four distinct approaches (linear mixed model, back-calculation approach, deterministic compartmental model and individual-based model) to forecast the new case detection rate of leprosy in four states in Brazil (Rio Grande do Norte, Ceará, Tocantins and Amazonas). In this work, we proposed the use of a simple compartmental model using a stochastic approach. A pre-defined structure of twelve compartments was used in other work [8] to represent health conditions with respect to leprosy, and flows from these compartments are calculated according to Markov transition rates. In this work, only three compartments were used.

The remain of this work is organized as follows. First, Section 2 presents a very short overview of the disease. Then, Section 3 presents the proposed model and a draft of its computational implementation using the Gillespie algorithm. Section 4 presents the results obtained and finally Section 5 presents our conclusions and plans for future works.

2 Leprosy

Leprosy is an infectious disease caused by the bacterium *Mycrobacterium leprae* that affects the skin and peripheral nerves, and can reach the eyes and internal tissues of the nose. The entry route of the bacterium into the body is not definitively known, but the skin and the respiratory route are most likely [9].

It is a highly infectious disease with low pathogenicity, that is, many people are infected, however, few get sick. It is estimated that between 90% to 95% of the human population is resistant to leprosy. Also, some people, when infected, may evolve to spontaneous cure.

Leprosy has cure, although in some cases it may leave patients with physical disabilities. Access to treatment is universal in countries where it occurs and is essential for disease control because, after the start of treatment, there is a fall in the bacillary load and the patient ceases to transmit the disease. However, its

ICCS Camera Ready Version 2018
To cite this paper please use the final published version:
DOI: 10.1007/978-3-319-93713-7_51

control is a challenge mainly due to the possibility of long periods of incubation of the bacterium and the frequent delays in its diagnosis. For this reason, the number of reported cases is much lower than the actual number of infected individuals. Treated patients may be infected again, and relapse of the disease may also occur. Leprosy deaths are not reported in the literature, although in untreated cases the disease may evolve to physical disabilities, loss of sensitivity and impairment of neural and muscular structure.

3 Methods

In order to model the spread of leprosy in Juiz de Fora, a SIR model was used. Then, a computational model that implements it was solved using a deterministic and a stochastic approach. The deterministic implementation solves the system of ODEs that describe the SIR model using the Python's package *SciPy*. This library has a package called "integrate". One of the functions available in this package is called "odeint", and it is used to solve numerically a system of first order ODEs in the form $\frac{dy}{dt} = f(y, t)$ using the LSODA function of *odepack* package in FORTRAN. The choice of the numerical method to be used is made automatically by the function based on the characteristics of the equations. The function uses an adaptive scheme for both the integration step and the convergence order. The function can solve the ODEs system using either the BDF (Backward Differentiation Formula) or the Adams method. BDF is used for stiff equations and the implicit Adams method is used otherwise.

For the stochastic implementation the system of equations was transformed into a set of equivalents stochastic processes that were implemented using the Gillespie's Algorithm [4]. The Python programming language was also used in the implementation.

The models consider a constant population along the period of 20 years. The population was fixed in 441,816 inhabitants, which was approximately the number of inhabitants living in Juiz de Fora in 1996, according to the census [6]. Data for adjusting and validating the model was obtained from SINAN (*Sistema de Informação de Agravos de Notificação*) database. It is mandatory to register all cases of the disease in Brazil in this database. For this study, it was available all cases recorded in Juiz de Fora from 1996 to 2015. Half of data was used to adjust the parameters of the model (data from 1996 to 2004). The other half of data was used to validate it. Only the records in which the patients live all period in the city were considered, i.e., if the patient moved to another city, the record was disregarded.

3.1 Modelling the Spread of Leprosy

For simulating the spread dynamics of leprosy, it was used the SIR mathematical model [7]. SIR divides the studied population in three compartments according to their state in relation to the disease studied: susceptible (S), infected (I) and recovered(R). This model is described mathematically in Equation 1.

ICCS Camera Ready Version 2018
To cite this paper please use the final published version:
DOI: 10.1007/978-3-319-93713-7_51

$$\frac{dS}{dt} = -\beta SI, \frac{dI}{dt} = \beta SI - \mu I, \frac{dR}{dt} = \mu I, \quad (1)$$

where β represents the infection rate and μ represents the recovery rate.

The susceptible compartment is characterized by all individuals who are susceptible to the contagion of the disease or those who have been contaminated but have not yet manifested the disease and have insufficient amount of bacilli to be a possible transmitter. The infected compartment is composed by all the infected individuals that can transmit the disease, i.e., those sick.

The infection rate, β, was chosen in a way to simulate the diffusion effect observed in spatial models. In this way, it has a similar behaviour to the solution of the heat equation, given by:

$$\phi(x,t) = \frac{1}{\sqrt{4\pi tk}} \exp(\frac{-x^2}{4kt}). \quad (2)$$

In this equation, k represents the thermal conductivity of the material. Since in this work we are considering only the temporal dimension, the spacial aspects of the heat equation were ignored ($x = 0$) to define the infection rate, β, i.e., only the term $\frac{1}{\sqrt{4\pi tk}}$ was considered.

The constant values of Equations 1 and 2 were manually adjusted in order to fit qualitatively the number of infected and recovered cases. For fitting purposes, it was considered that the number of reported cases is less than the number of existing cases. Also, in order to reproduce the oscillatory characteristic observed in the number of infected cases, the infection rate was multiplied by the term $(sin(\frac{\pi t}{28}) + 1)$. The sum by one in the trigonometric term prevents negative values. so, the values found for the parameters were the following: $\beta = \frac{1}{\sqrt{4\pi t(4*10^{11})}}(sin(\frac{\pi t}{28}) + 1)$ and $\mu = 0.025$.

There is a tiny possibility of relapse of the disease (about 0.77% for multibacillary cases and 1.07% for paucibacillary ones [11]). For this reason, and due to the lack of notifications of relapse in Juiz de Fora, the possibility of relapse was not considered in this work. In other words, after infected, an individual recovers from the disease, and is not susceptible again.

3.2 Gillespie Implementation

The first step to implement the Gillespie algorithm is to define the equivalent reactions, as follows: a) $S + I \rightarrow I + I$: a susceptible reacts with an infected, producing two infected; and b) $I \rightarrow R$: an infected recovers.

Algorithm 1 presents the implementation of the SIR model. The Gillespie model works with the probabilities of a reaction to occur, one reaction per iteration. Two values are drawn from the uniform distribution in the unit interval (lines 6 and 8). The first one is used to compute the time of the next reaction (line 7). The second value is used to choose which reaction will occur in this time interval (lines 9-11). For each reaction, it is computed the probability of its occurrence (line 5). All probabilities are summed (line 5), and this value is

ICCS Camera Ready Version 2018
To cite this paper please use the final published version:
DOI: 10.1007/978-3-319-93713-7_51

used in the computation of the interval of the next reaction (line 7), as well as in the choice of the reaction that will occur (line 8). The second value drawn is compared to the normalized probability of occurrence of each reaction; if the drawn value is into an interval associated to a reaction, populations affected by that reaction are updated accordingly and a new iteration starts (lines 9-11).

Algorithm 1 Gillespie implementation of the SIR model

1: **while** $t < t_{max}$ **do**
2: **if** i==0 **then**
3: break;
4: **end if**
5: R1 = β * s * i; R2 = μ * i; R = R1+R2;
6: ran = uniformly_distributed_random(0,1);
7: t_n = -log (ran)/R; t = t + t_n;
8: **if** uniformly_distributed_random(0,1) < $\frac{R1}{R}$ **then**
9: s = s-1; i = i+1;
10: **else**
11: i = i-1; r = r+1;
12: **end if**
13: **end while**

4 Results

Figure 1 compares the number of infected and recovered cases registered in Juiz de Fora between 1996 to 2015, the results of the deterministic model and some of the results obtained by the stochastic approach. It's possible to notice that the Gillespie's solutions are very close to the deterministic solution. These oscillations are expected because the SSA algorithm adds noise to the solution, which may represent bacterial seasonality, changes in the treatment of disease, and so on. Figure 2 shows the CDF (cumulative distribution function) graph using 10,000 executions of the Gillespie algorithm. The graph shows that the probability of leprosy to be eradicated in the city before 2,045 is 99.21%. Therefore, the model indicates that there is no great risk of an outbreak in Juiz de Fora, if the model assumptions are kept constant. Implicitly the rates used in the model capture all aspects related to the combat of leprosy in the city. This projected scenario is in accordance with the results obtained in other works [10, 1], that estimated that elimination of leprosy as a public health risk would require, on average, 44–45 years.

One aspect that should be highlighted is the occurrence of the so-called "hidden prevalence", i.e., cases not diagnosed and therefore not reported and registered in SINAN. The hidden prevalence occurs due to the unpreparedness or lack of knowledge about the symptoms of the disease by the health teams, so its diagnosis is made most often in advanced stages, after 5-7 years after infection. In this case, the reality may be very different from the SINAN numbers and, as consequence, our results and projections may be wrong.

ICCS Camera Ready Version 2018
To cite this paper please use the final published version:
DOI: 10.1007/978-3-319-93713-7_51

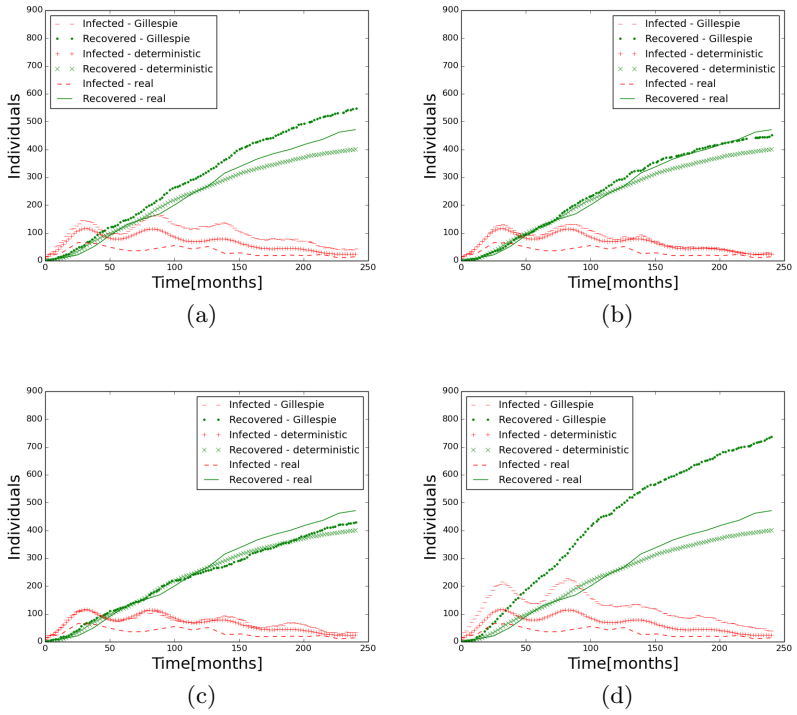

Fig. 1: Distinct results (a-d) obtained by the execution of the Gillespie algorithm. For comparison purposes, all figures also reproduces the deterministic result and the number of infected and recovered cases in Juiz de Fora.

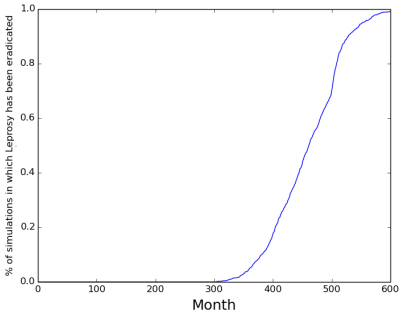

Fig. 2: CDF computed for 10,000 executions of the Gillespie algorithm. The graph presents the percentage of simulations in which leprosy has been eradicated. The probability of leprosy be eradicated after 600 months (starting in 1996) is 99,21%.

ICCS Camera Ready Version 2018
To cite this paper please use the final published version:
DOI: 10.1007/978-3-319-93713-7_51

5 Conclusion and Future Works

This work presented a mathematical-computational model to simulate the spread of leprosy in Juiz de Fora. The mathematical model was implemented using two approaches: a deterministic and a stochastic one. The deterministic approach used ODEs to model the spread of leprosy, while the stochastic one used the Gillespie's Stochastic Simulation Algorithm. The results of both approaches were qualitatively validated by the comparison to the historical number of diagnosis of leprosy in Juiz de Fora. Both the deterministic and the stochastic simulations obtained numbers of cases with the same order of magnitude of the registered cases, and the shapes of the curves were similar to the ones that describe the history of cases in Juiz de Fora. As future work, improvements can be made in the model to better fit its results to the historical series of leprosy cases. Also, the number of under-reporting cases needs to be verified. We plan to extend the model to a spatial domain using PDEs because the disease transmission occurs more frequently in places with vulnerability to health. Including this information in the model can be very useful and improve the quality of results.

References

1. Blok, D., Crump, R.E., Sundaresh, R.R., Ndeffo-Mbah, M., Galvani, A.A.P., Porco, T.C., de Vlas, S., Medley, G.F., Richardus, J.H.: Forecasting the new case detection rate of leprosy in four states of Brazil: A comparison of modelling approaches. Epidemics **18**, 92–100 (Mar 2017). https://doi.org/10.1016/j.epidem.2017.01.005
2. Esteva, L., Vargas, C.: Analysis of a dengue disease transmission model. Mathematical biosciences **150**(2), 131–151 (1998)
3. Fister, K.R., Gaff, H., Schaefer, E., Buford, G., Norris, B.: Investigating cholera using an SIR model with age-class structure and optimal control. Involve, a Journal of Mathematics **9**(1), 83–100 (2015)
4. Gillespie, D.T.: A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. J. Comput. Phys **22**(4), 403–434 (1976)
5. Hattaf, K., Yousfi, N.: Mathematical model of the influenza A (H1N1) infection. Advanced Studies in Biology **1**(8), 383–390 (2009)
6. IBGE: Juiz de Fora. http://www.cidades.ibge.gov.br/xtras/perfil.php?codmun=313670
7. Kermack, W.O., McKendrick, A.G.: A contribution to the mathematical theory of epidemics. Proc R Soc Lon Ser-A **115**(772), 700–721 (1927)
8. Meima, A., Smith, W.C.S., Van Oortmarssen, G.J., Richardus, J.H., Habbema, J.D.F.: The future incidence of leprosy: a scenario analysis. Bull World Health Organ. **82**(5), 373–380 (2004)
9. Rees, R., McDougall, A.: Airborne infection with mycobacterium leprae in mice. Journal of medical microbiology **10**(1), 63–68 (1977)
10. Smith, R.L.: Proposing a compartmental model for leprosy and parameterizing using regional incidence in Brazil. PLoS Negl Trop Dis. **10**(8), e0004925 (2016)
11. World Health Organization: Risk of relapse in leprosy. Tech. rep., The Leprosy Unit - Division of Control of Tropical Diseases (CTD/LEP/941 1994)
12. World Health Organization: Global strategy for the elimination of leprosy as a public health problem. Tech. rep., WHO (WHO/LEP/967 1996)
13. World Health Organization: Weekly Epidemiological Record. Tech. rep., WHO (vol 92, n 35, pp 510-520 2017)

ICCS Camera Ready Version 2018
To cite this paper please use the final published version:
DOI: 10.1007/978-3-319-93713-7_51