Epidemics 18 (2017) 81–91

Contents lists available at ScienceDirect

Epidemics

journal homepage: www.elsevier.com/locate/epidemics

Comparison and validation of two computational models of Chagas disease: A thirty year perspective from Venezuela

Sarah M. Bartsch[a,b], Jennifer K. Peterson[c], Daniel L. Hertenstein[a,b], Laura Skrip[d], Martial Ndeffo-Mbah[d], Alison P. Galvani[d], Andrew P. Dobson[c], Bruce Y. Lee[a,b,*]

[a] Public Health Computational and Operations Research (PHICOR), Johns Hopkins Bloomberg School of Public Health, United States
[b] Global Obesity Prevention Center, Johns Hopkins University, United States
[c] Department of Ecology and Evolutionary Biology, Princeton University, United States
[d] Center for Infectious Disease Modeling and Analysis (CIDMA), Yale School of Public Health, United States

ARTICLE INFO

Article history:
Received 17 December 2016
Received in revised form 7 February 2017
Accepted 7 February 2017

Keywords:
Chagas disease
Trypanosoma cruzi
Model
Simulation
Model comparison

ABSTRACT

Background: Mathematical models can help aid public health responses to Chagas disease. Models are typically developed to fulfill a particular need, and comparing outputs from different models addressing the same question can help identify the strengths and weaknesses of the models in answering particular questions, such as those for achieving the 2020 goals for Chagas disease.
Methods: Using two separately developed models (PHICOR/CIDMA model and Princeton model), we simulated dynamics for domestic transmission of *Trypanosoma cruzi* (*T. cruzi*). We compared how well the models targeted the last 9 years and last 19 years of the 1968–1998 historical seroprevalence data from Venezuela.
Results: Both models were able to generate the *T. cruzi* seroprevalence for the next time period within reason to the historical data. The PHICOR/CIDMA model estimates of the total population seroprevalence more closely followed the trends seen in the historic data, while the Princeton model estimates of the age-specific seroprevalence more closely followed historic trends when simulating over 9 years. Additionally, results from both models overestimated *T. cruzi* seroprevalence among younger age groups, while underestimating the seroprevalence of *T. cruzi* in older age groups.
Conclusion: The PHICOR/CIDMA and Princeton models differ in level of detail and included features, yet both were able to generate the historical changes in *T. cruzi* seroprevalence in Venezuela over 9 and 19-year time periods. Our model comparison has demonstrated that different model structures can be useful in evaluating disease transmission dynamics and intervention strategies.

1. Introduction

Chagas disease, caused by the protozoan parasite *Trypanosoma cruzi*, is one of the world's most important neglected tropical diseases (NTDs). It infects approximately 6–7 million people worldwide (World Health Organization, 2016) and results in an estimated $627.46 million in healthcare costs and $7.19 billion in societal costs annually (Lee et al., 2013). Given its substantial burden Chagas is one of the ten NTDs targeted for control or elimination by 2020, with one of the London Declaration's stated goals for being 100% certified interruption of domestic transmission in Latin America (Tarleton et al., 2014). Historically, control of Chagas disease has focused on vector control. This can be achieved directly by vector reduction using insecticides or indirectly through housing modifications.

Mathematical models are simplifications of real life that are developed to address a particular need or question (Garnett et al., 2011). Model development must balance the actual complexity of biological systems with the simplifying assumptions that ensure computational tractability (Lee, 2008). Additionally, models are not a one size fits all. The applicability of different models to answer specific research and public health questions lies in appropriateness and flexibilities of specific methodologies employed. Thus, assessing and comparing mathematical models and determining if they capture relevant features of reality for a particular application is fundamental to optimal model design (St-Pierre, 2016). While model assessments and comparisons have been conducted

* Corresponding author at: Johns Hopkins Bloomberg School of Public Health, 615 N. Wolfe Street, Room W3501, Baltimore, MD 21205, United States.
E-mail address: brucelee@jhu.edu (B.Y. Lee).

http://dx.doi.org/10.1016/j.epidem.2017.02.004

in other fields/pathogens (notably human immunodeficiency virus (Hontelez et al., 2013; Eaton et al., 2012)), little has been done in the realm of NTDs (Hollingsworth et al., 2015).

In this study, we parameterize two differently structured, independently developed, Chagas disease transmission models to evaluate the same research question using the same input/baseline data. We compare model results, and discuss possible causes of differences. Comparing outputs from different models addressing the same question can help identify the strengths and weaknesses of the models to answer particular questions. For example, one model may be best at answering policy questions related to disease prevalence and control in humans, while another may be better suited to answer questions about ecology and vector control. Model comparison can also help us gain understanding on how data informs parameter estimation and impacts output. Understanding model strengths and weaknesses can aid various decision makers in knowing which model is best apt to answer questions and in interrupting model results, which can be helpful in achieving the 2020 goals.

2. Methods

We independently developed two *T. cruzi* transmission models (described below). The comparison consisted of simulating the transmission of *T. cruzi* in the domestic setting in the two models and comparing the resulting seroprevalence between the models and to the historical seroprevalence data. Table 1 shows key input parameter values and sources for both models, while Fig. 1 provides an outline of each model.

2.1. PHICOR/CIDMA model

This model was developed by a team at Johns Hopkins Bloomberg School of Public Health and the Center for Infectious Disease Modeling and Analysis (CIDMA) at Yale School of Public Health. It was originally developed to answer questions about vector control on *T. cruzi* transmission (i.e., measuring new acute Chagas cases) and the role of non-human hosts on a larger scale than previous models, and has three general age categories to explore potential target populations for interventions. Developed in Python (Python Software Foundation, Wilmington, DE), this compartmental simulation model represented vector and host populations involved in *T. cruzi* transmission and included triatomines, human hosts, and non-human hosts (i.e., dogs) and vector-borne transmission among these populations in the domestic habitat (Fig. 1). The model ran in monthly time steps (i.e., t = 1 month or 30 days), chosen due to the long disease course of Chagas, and simulated a 41-year period. During each time step, epidemiological and clinical rates defined transitions between model compartments, stratified by the different vector and host populations. Vectoral transmission in this model was governed by the force of infection.

Triatomine bugs could be susceptible (not infected with *T. cruzi* and able to become infected) or infectious (infected with *T. cruzi* and able to transmit to vertebrae hosts upon biting). Upon biting an infectious host (human and viable non-human), a susceptible bug had probabilities of becoming infected with *T. cruzi*, depending on the disease state of the host. The number of triatomine bugs (N_V = 475,972) in the model was determined from the carrying capacity, or the number of bugs sustainable in the habitat, which was assumed to be 50 bugs per person (consistent with previous work (Peterson et al., 2015)). The following formulas describe the susceptible and infectious states for triatomine bugs:

$$\frac{dS_V}{dt} = b_V - \gamma_V S_V - d_V S_V$$

$$\frac{dI_V}{dt} = \gamma_V S_V - d_V I_V,$$

where b_v is the number of bug births, d_v is the triatomine death rate, and γ_V is the force of infection. The number of bug births is determined by the birth rate, carrying capacity, and total number of triatomines by the following formula:

$$b_V = birthrate * N_V * \frac{carrying\ capacity - N_V}{carrying\ capacity}.$$

The following formula determine the force of infection (Υ_v):

$$\gamma_V = \beta \left[\frac{p_D \theta_D I_D + p_H \left(\theta_A A_H + \theta_I (I_H + \theta_C) \right)}{p_H N_H + p_D N_D} \right],$$

where β represents the triatomine biting rate, Θ is the probability of transmission (or infectivity), N_H, and N_D the number in the human and dog populations, respectively; p_H and p_D, and describe the vector feeding preferences for humans and dogs, respectively.

The human population (N_H) consisted of 10,000 persons at the start of the simulation and was comprised of three age groups, *i*: (0–19 years old, 20–39 years old, and 40 years and older following historical age-specific demographic data from the World Population Prospects (United Nations, 2015)). The human population is divided into four states: susceptible (S_H, not infected with *T. cruzi* and able to become infected), acute stage Chagas disease (A_H, infected with *T. cruzi* and able to transmit, exhibit mild and nonspecific symptoms, and person has microscopically detectable parasitemia), indeterminate stage Chagas disease (I_H, asymptomatically infected with *T. cruzi* and able to transmit), and symptomatic chronic stage Chagas disease (C_H, infected with *T. cruzi*, able to transmit, and show symptoms of chronic disease such as cardiomyopathy and/or megaviscera). Thus, a person in any of the three Chagas disease states are considered positive. Upon the bite of an infectious triatomine, a susceptible human had a probability of becoming infected with *T. cruzi*, based on the force of infection (γ_H), and once infectious, persons were remained infectious in absence of treatment (i.e., once seropositive, always seropositive, with no decay). Those in the acute and symptomatic chronic states of disease had probabilities of Chagas-related mortalities. These states and the transmission between them are described by the following four equations:

$$\frac{dS_{Hi}}{dt} = b_{Hi} - \gamma_H S_{Hi} - d_H S_{Hi}$$

$$\frac{dA_{Hi}}{dt} = \gamma_H S_{Hi} - \pi_H A_{Hi} - (d_H + \mu_{HA}) A_{Hi}$$

$$\frac{dI_{Hi}}{dt} = \pi_H A_{Hi} - \lambda_H I_{Hi} - d_H I_{Hi}$$

$$\frac{dC_{Hi}}{dt} = \lambda_H I_{Hi} - (d_H + \mu_{HC}) C_{Hi}$$

where b_H is the number of people entering each age group (i.e., number of births or number of persons aging (United Nations, 2015)), d_H is the human death rate from all causes, μ_{HA} is the probability of Chagas related mortality in the acute phase of disease, and μ_{HC} is the probability of Chagas related mortality in the chronic phase. Two variables, π_H *and* λ_H, describe the rate of movement from the acute phase to the indeterminate phase and the indeterminate phase to the chronic phase, respectively.