Proceedings of the 2017 Winter Simulation Conference
W. K. V. Chan, A. D'Ambrogio, G. Zacharewicz, N. Mustafee, G. Wainer, and E. Page, eds.

AN AGENT-BASED MODEL TO INVESTIGATE BEHAVIOR IMPACTS ON VECTOR-BORNE DISEASE SPREAD

Anna Paula Galvão Scheidegger
Amarnath Banerjee

Industrial and Systems Engineering Department
Texas A&M University
3131 TAMU
College Station, Texas 77843-3131, USA

ABSTRACT

This study aims to use agent-based simulation as a tool to illustrate the importance of human behavior in the dynamics of vector-borne disease spread. For this, a baseline compartmental model was developed and, based on it, four different scenarios considering human behavior were proposed: two assuming the whole population adopts the same behavior and two assuming each individual has his/her own behavior. Paired t-test was used to compare the proposed scenarios with the baseline, based on two output responses from the simulation experiments: total number of infected people and duration of the epidemic. Results from the data analysis indicate that behavior is an important factor and, as such, it must be further investigated and included in infectious-disease spread models to obtain more accurate results. As a final remark, we presented possible explanations to why human behavior has been neglected in many epidemiological models up to now.

1 INTRODUCTION

Infectious diseases are a major common problem among countries throughout the world. Due to globalization and large-scale travels, infectious diseases can quickly spread to any part of the globe. The problem has a greater impact on developing countries, where coupled with maternal causes (e.g. congenital disorder) and nutritional deficiencies, it accounted for 52% of all deaths in 2015, compared to only 7% in developed countries (WHO 2016). Despite the relatively low overall number (around 17% of worldwide deaths are caused by infectious diseases), number of deaths by itself is not a proper measure of the burden of infectious diseases. Although some infectious diseases do not lead to a high number of deaths, they can lead to large epidemics that result in loss of healthy years of life, higher proportion of years of life lost due to deaths at younger ages, morbidity, economic losses and overload to the health systems (WHO 2015). The Ebola outbreak in 2014/2015, for example, not only led to a high number of cases and deaths, but it also spread tension throughout the world and overwhelmed the healthcare systems of the affected countries in Africa (Chan 2014, Walker et al. 2015, Wenham 2017).

According to the WHO (2015), infectious disease transmission is influenced by multiple socioeconomic, environmental and ecological factors, such as population density and movement, climate change, increasing antimicrobial resistance, urbanization, land-use changes, and human behavior. The Zika outbreak in 2015/2016 showed how little is known about some of the infectious diseases and how different the epidemic outcomes can be in different countries (Chang et al. 2016, Johansson et al. 2016). The measles outbreak in California in the end of 2014 and the recurrence of yellow fever in Brazil in the beginning of 2017 raised the hypothesis that human behavior can be one of the major factors in the spread or containment of an epidemic (Majumder et al. 2015, Zipprich et al. 2015).

On the one hand, this context means that infectious diseases prevention and control are a very complex area of investigation. On the other hand, it also means that most of the cases and the number of deaths are preventable, as indicated by Brandeau (2008). Therefore, modeling and simulation are particularly useful for this complex context where someone needs to analyze "what-if" scenarios without actually implementing the changes in the real world. As a result, simulation has been increasingly applied to examine disease spread (Scheidegger and Banerjee 2017b).

Despite the advancements observed in the area and the growing number of publications, this research field still requires a lot of attention and investments. Especially, there are some factors, such as human behavior, agents' heterogeneity and interdependencies among the parameters, which have been frequently neglected in the simulation models and possibly prevents further progress in the area (Funk et al. 2015). As exception among many studies, Wang et al. (2015) and Verelst, Willem and Beutels (2016) affirmed that it is clearly important to include human behavior in infectious disease spread models. When contrasting these authors' opinion to the academic reality a simple question arises: Is human behavior indeed a significant factor in the spread of infectious diseases? If the answer to this question is yes, why has this factor been neglected by simulation specialists?

To provide a direction for these questions, we consider a simple compartmentalized agent-based simulation (ABS) model. The purpose is to illustrate how the spread of infectious diseases, more specifically a vector-borne disease, can be affected by human behavior. In order to show this, we consider five different scenarios: (i) a baseline SEIR-SEI (susceptible, exposed, infectious and recovered – susceptible, exposed and infectious) model, where no behavior is considered, (ii) a modified model where it is assumed that the whole population adopts the same behavior immediately after the epidemic reaches some specific threshold, (iii) a modified version of the second model, where it is assumed that the population changes its behavior only after some time the threshold has been reached, (iv) a modified model where it is assumed that each individual changes his/her particular behavior immediately after the epidemic reaches some specific threshold, and, (v) a modified version of the fourth model, where it is assumed that the behavior change occurs at different times for each individual. After running the experiments, we used paired t-test to evaluate the results.

In this section, we presented a general background of the topic, the motivation for the problem being investigated and our goal. Section 2 presents a brief overview of studies conducted in the area, while Section 3 describes the conceptual and computational models and the experimental design. Section 4 discusses the data analysis methods and provides the results of the experiments. Finally, Section 5 brings our final considerations, the limitations of this study, and suggestions for future work.

2 RESEARCH IN THE FIELD

In order to prevent, prepare and control epidemics, knowledge of transmission dynamics of infectious diseases is essential (Aleman, Wibisono, and Schwartz 2009). The field of epidemiology studies these aspects, and its first applications can be traced back to Daniel Bernoulli (Dietz and Heesterbeek 2002). Researchers have investigated how infectious diseases spread through human and animal populations in order to implement cost-effective preventive and control measures (Kreuger and Osgood 2015).

Initially, infectious diseases were solely modeled through a set of differential equations, called mathematical epidemiology models. The development of these models are mainly credited to the work of Ross (1908, 1911), Kermack and McKendrick (1927), and Macdonald (1952, 1957). These models are also known as Ross-Macdonald models. The Ross-Macdonald work consists in placing individuals in different health compartments and applying differential equations to determine the net flow to each compartment. To allow for analytical solution, especially in times where computer technology was in its early development, the model involved a single homogeneous population and a set of simplified concepts. Up to now it serves as the starting point for most disease spread research (Scheidegger and Banerjee 2017a). According to Sanchez and Sanchez (2015), various compartment classifications can be used, but the most traditional one is the so-called SEIR (susceptible, exposed, infectious and recovered).

Susceptible refers to the set of individuals who have not contracted the disease yet and, thus, are susceptible to the infection. Exposed are the individuals who have already been infected by the pathogen but cannot transmit the disease yet. Infectious refers to people who are infected and capable of spreading the disease to susceptible individuals. Finally, recovered refers to individuals who have recovered from the disease and, most of the times, are immune to new infections.

Usually disease dynamics is a complex nonlinear process that involves heterogeneous individuals and environments, and a diverse range of interconnected socioeconomic, behavioral, biological and environmental factors (Bandeau 2008, Lima et al. 2014). As a result of these characteristics, and despite the success of mathematical epidemiology models, Sanchez and Sanchez (2015), and Kreuger and Osgood (2015) indicated the need for using other approaches. The authors mentioned that this is particularly required in systems with small and/or heterogeneous populations, where aggregate and perfectly mixed approximation may not be accurate, and in complex systems with reciprocal causality, which is typically the case of infectious diseases. As expected, in the beginning of 1980's, with the advances in computer technology, the research field evolved and several computational/simulation models have been developed to study the spread of infectious diseases. Unlike mathematical models, simulation models allow for integration of data from different sources and at different levels, which makes it a suitable tool for studying complex systems (Li et al. 2016).

Bisset et al. (2009) cited three major computational approaches that have been applied to investigate infectious disease transmission, namely: (i) equation-based simulation or system dynamics simulation, (ii) ABS, and (iii) network simulation. Sanchez and Sanchez (2015) added that discrete-event simulation models are also applied to the study of disease transmission. However, as it will be discussed later, discrete-event epidemiology models are not so common and they are usually applied to the planning of healthcare systems and to the investigation of logistics and supply chain of preventive measures or treatments, such as vaccines and drugs.

According to Bobashev et al. (2007), choosing the most appropriate simulation method to model disease spread is not an easy task. As any method, each of these approaches has its own drawback. First of all, as a requirement of computational epidemiology models in general, we can cite the need to have access to accurate data in order to make good predictions. If parameter values are inaccurate or if the model is oversimplified, the discrepancies between the real-world outcomes and the model results may be very high. With respect to each method, on the positive side system dynamics simulation and network simulation are usually insightful and fast-processing tools, while on the negative side they frequently assume population is homogeneous and fully mixed and, thus, they do not allow for investigation of the impacts on different age groups or behaviors and they give little time varying information about epidemics. These are discussed in detail in Paleshi et al. (2011) and Bisset et al. (2009). In contrast, a relatively new approach, called agent-based simulation (ABS), gives a high level of modeling flexibility and it also allows for higher model fidelity by including heterogeneous populations and their multiple behaviors and interactions (Bisset et al. 2011). However, this method can be very time-consuming and requires considerable computational power, which can make it difficult to use for on-time real-world decision making. In order to take advantage of the best of each method, researchers have started to adopt hybrid simulation models that combine two or more simulation methods as an alternative approach to epidemiology modeling.

As examples of simulation studies in the field of epidemiology, we can cite: the work of Díaz, Akhavan-Tabatabaei, and Mura (2016) where an equation-based compartmental simulation model was developed to investigate the effects of vaccination strategies on the transmission of Human Papillomavirus (HPV) infection and cervical cancer; Lima et al. (2014), who developed a simulation framework to facilitate simulation modeling of dengue fever; Brandeau (2008) provided a discussion on three epidemiology studies they developed involving Human Immunodeficiency Virus (HIV) prevention and treatment, contact tracing in diseases such as tuberculosis and gonorrhea, and hepatitis B prevention and control; Yakob and Clements (2013) where they constructed a compartmental SEIR model with data

from the literature and fitted the model to empirical data from the Chikungunya outbreak in Reunion Island in 2005-2006; Zarei and Smith (2011) where they modified the predator-prey Lotka-Volterra equations to allow the analysis of subjective and objective parameters that affect disease spread in Susceptible-Infectious (SI) models; and, Bobashev et al. (2007) where they proposed a hybrid model that started with agent-based approach and when the number of infected individuals in each city was large enough, the respective city would be investigated by an equation-based approach. According to Sanchez and Sanchez (2015), the areas of investigation range from creating a structure for disease modeling, to modeling specific epidemics, to fitting models to empirical data.

To summarize this discussion, we performed a search for the term "disease" in the Winter Simulation Conference (WSC) archive, one of the main discussion forum on modeling and simulation. This search led to 371 results between 1968 and 2016, out of the total 9,601 publications (according to the ACM digital library). However, we realized that many studies published in the proceedings of this conference would not use the word "disease" in their title, as an example, some would use the name of the disease only. So, to get more information on what has been published about disease spread in the conference, we performed another search in the archive between 2007 and 2016. In this search, we did not use any keyword. Rather, we read the papers titles to find studies related to disease modeling and simulation. This second search resulted in the selection of 68 papers that included either chronic or infectious diseases or also changes in health habits and health care planning related to the study of diseases. After reading the abstract of those papers, we concluded that 49 out of the 68 papers were about infectious diseases. After reading these 49 papers, we were able to identify the simulation method being used in each one of them.

3 METHODOLOGY

3.1 Conceptual Model

3.1.1 Baseline Model Description

The system refers to infectious disease transmitted by mosquitoes, such as Dengue fever, yellow fever, Chikungunya and Zika. The transmission dynamics of these mosquito-borne diseases are very similar (WHO 2014). Therefore, in this study we are not going to differentiate between them.

In vector-borne disease we have three main types of agents: (i) the pathogen, that can be a virus or bacteria, for example; (ii) the vector, in this case a mosquito such as *Aedes aegypti* or *Aedes albopictus*; and, (iii) the final host, a human in this case. The baseline model can be represented by a SEIR-SEI (susceptible, exposed, infectious and recovered – susceptible, exposed and infectious), where the SEIR model is used to represent humans and the SEI model is used to represent the mosquito vector. The life-cycle of the pathogen can be described in four steps: (i) the pathogen is passed from an infectious mosquito (Mi) to a susceptible host (Hs) when the mosquito feeds from human blood; (ii) the pathogen infects the host (exposed host, He), who cannot transmit the disease yet to another mosquito. After the latent period, the pathogen reaches sufficiently high densities in the host blood (infectious host, Hi) to infect another susceptible mosquito (Ms). So, whenever a susceptible mosquito feeds from an infectious host (iii) the susceptible mosquito inoculates the pathogen (exposed mosquito, Me). Similar to the host, the mosquito cannot transmit the disease immediately when feeding from other susceptible hosts. However, (iv) after the latent period, the pathogen develops in the mosquito (infectious mosquito, Mi) to a point that it is in the salivary glands and ready to be transmitted during a subsequent bite on a susceptible host. After a recovery period, the host is recovered (Hr) and immune to the pathogen (this is not true for all mosquito-borne diseases, but it is the case for Chikungunya and same strain of Dengue fever, for example).

3.1.2 Model Assumptions

The main assumptions adopted in this work are: (i) human population is closed to birth, migration and death, so it is kept constant; (ii) probability of severe cases is not taken into account and 100% of infected individuals are symptomatic; (iii) there is only one host (humans); (iv) hosts become immune to infection after recovery; (v) the ratio of mosquitoes to humans is constant, so, mosquito births are set to balance deaths and to keep the mosquito population constant; (vi) there is only one mosquito species; (vii) mosquito bites are distributed randomly among hosts in the environment, i.e., distance or any other factor is not taken into account when modeling mosquito biting; (viii) all the parameters are constant over time and age and gender independent; (ix) the pathogen's lifecycle is not taken into account; and, (x) temperature and other climate data are not taken into account.

We recognize that such assumptions lead to an oversimplified model in comparison to the reality. However, the goal of this study is not to precisely predict the outcomes of an epidemic, rather we aim to explore the importance of human behavior in disease spread models and to serve as a point of departure for the elaboration of further detailed and more realistic models.

3.1.3 Proposed Scenarios

From the baseline model, some modifications are proposed in order to evaluate four different scenarios. Table 1 summarizes the scenarios and the proposed modifications.

Table 1: Scenarios to be run in the simulation model and their respective modifications.

Scenario	Modification description	Host level	Parameter associated
1	When the number of infected people reaches a specific threshold, the whole population adopts the same cautious behavior that reduces the population's probability of being exposed to the disease.	Population	Population's cautious factor
2	The population changes its behavior the same way as in the previous scenario, but only after some time the infected threshold has been reached. The change occurs in the whole population at the same time.	Population	Population's time to change behavior
3	When the number of infected people reaches a specific threshold, the individuals adopt cautious behavior (take preventive and control measures) that varies among them. This behavior reduces the individual's probability of being exposed to the disease.	Individual	Individuals' cautious factor
4	The individuals change their behavior the same way as in the previous scenario, but only after some time the infected threshold has been reached. The change occurs at different time for each individual.	Individual	Individuals' time to change behavior

3.1.4 Input Data and State Chart

The input parameters adopted in this work are shown in Table 2. The values of the parameters refer to the *Aedes aegypti* mosquito population and the Chikungunya disease, and these are based on the work of Dumont, Chiroleu, and Domerg (2008); Moulay, Aziz-Alaoui, and Cadivel (2011); and Yakob and Clements (2013). For disease specific parameters, such as daily mosquito latent rate and daily human

latent rate, we opted for not testing different values at high levels once the aforementioned authors agreed on the range of these parameters. Figure 1 shows the state chart of the mosquito and human populations.

Table 2: List of input parameters of the simulation model.

#	**Input parameters**	**Low level**	**High level**
1	Mosquito population size	100	1,000
2	Initial number of infectious mosquitoes	1	50
3	Daily mosquito latent rate	Uniform (0.333, 0.500)	
4	Daily mosquito mortality rate	Uniform (0.025, 0.05)	
5	Daily mosquito to human infect rate	Uniform(0.14,0.25)	
6	Human population size	500	2,000
7	Initial number of infectious humans	0	50
8	Daily human latent rate	Uniform (0.2, 0.500)	
9	Daily human recovery rate	Uniform (0.143, 0.25)	
10	Daily human to mosquito infect rate	Uniform (0.3, 0.475)	
11	Include same behavior for population [0, 1]	0	1
12	Include behavior for individual [0, 1]	0	1
13	Percent of infectious individuals to trigger cautious behavior [%]	0.05	0.10
14	Population cautious factor [%]	Uniform(0.8, 0.9)	Uniform(0.6, 0.9)
15	Individual cautious factor [%]	Uniform(0.8, 0.9)	Uniform(0.6, 0.9)
16	Include time to switch behavior [0, 1]	0	1
17	Population time to switch behavior [days]	Uniform (1,3)	Uniform (1,7)
18	Individual time to switch behavior [days]	Uniform (1,3)	Uniform (1,7)

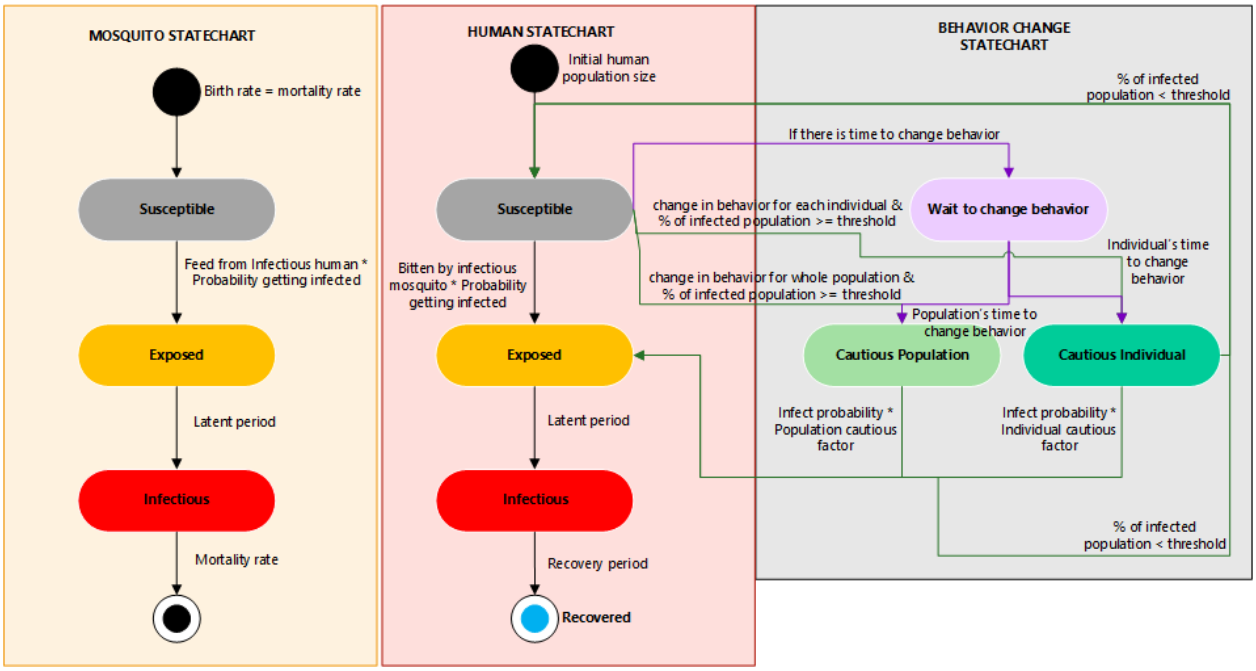

Figure 1: State chart of the agents.

3.2 Computational Model

As previously mentioned, several simulation methods can be applied in disease spread context and the choice of the method depends on several factors, such as level of detail desired, computational power and information available. At one end, there are ABS models, where heterogeneous individuals are described in interaction with each other and with the infectious agent. At the other end, there are equation-based models where individuals are considered perfectly mixed and have the same average characteristics. Although our model adopts several assumptions that make it simple, using equation-based simulation to model human behavior is usually a more complex task than using ABS. Therefore, even considering a perfectly mixed population, we opted to develop an agent-based model where individuals change their state to represent the change in behavior based on some predefined information threshold. The predefined information threshold is based on the percentage of infected individuals in the population and, thus, the information is considered prevalence-based. Despite the computational power required, the choice for the ABS method is also justified once the final goal is to develop more detailed models in future studies. Due to similarities between some mosquito-borne diseases, the model can be easily altered to represent other diseases, by changing the parameter values and/or adding another transition between recovered and susceptible states.

AnyLogic® (8.0.5. University version) was chosen as the simulation tool to build the model. The computational model involves three main agents: mosquito, human, and environment (main agent), where mosquitoes and humans live. Humans and mosquitoes are initiated according to their respective population size and their respective number of infectious individuals (as shown in Table 2). The time step used in the model was days. At each time step either humans or mosquitoes could change their state. Mosquitoes and humans were randomly distributed in the environment and since the contact between mosquitoes and humans was also randomly assigned, the distance between the agents had no importance in the model. The environment and the simulation model will be demonstrated during the presentation at the conference. They have been removed from the paper due to page limitation.

To verify the simulation model, we ran it in the interactive model with deterministic values and we added buttons to allow us to include and remove each of the proposed scenarios (population behavior, individual behavior and time to change behavior). We elaborated and followed a test protocol where each possible situation was tested. As an example, if population behavior was included in the model, but the percentage of infected individuals was smaller than the specified threshold, humans should stay in the susceptible state or move to the exposed state after being bitten by an infectious mosquito, but they should not move to any of the behavior states. Another example was if individual behavior and time were included in the model, individuals should change their behavior at different times, and so on.

3.3 Experimental Design

The baseline simulation model has a total of 10 input parameters. Compared to the baseline, the population behavior model and the individual behavior model has 2 extra varying parameters each, namely percentage of infected individuals to trigger cautious behavior and population/individual cautious factor. Finally, compared to both previous models, the inclusion of time to change population or individual behavior adds 1 extra parameter in each case. If we opted to run all possible combination of parameter values, we would have to run a large number of experiments, without having any information about the importance of behavior to the model or not. A considerable amount of computational power is required to run the ABS models. Therefore, instead of using a full factorial design, we opted for running the baseline and each of the four scenarios for two different situations: one with all parameters in the low level and another one with all parameters in the high level, resulting in a total of 10 iterations. The model was run for 2 years (730 days), which was long enough for the outbreak to be over in all the runs. Twenty replications were performed for each scenario and situation, leading to an experiment that involved a total

number of runs (N) = 200. We considered two output responses: total number of infected people and duration of the epidemic in days to statistically compare the results.

4 RESULTS AND DATA ANALYSIS

The main question being investigated in this paper is whether human behavior is a significant factor in the spread of infectious diseases. To answer this question we performed paired t-test on each of the scenarios against the baseline. The paired t-test was carried out using Minitab® and it was based on the two output responses from the simulation experiments. The results of the paired t-test comparing the baseline with each of the possible scenarios are presented in Table 3 – Table 6, followed by a brief description. The negative values are due to the paired t-test that tests the mean of pairwise differences. In this case, the pairwise difference is the difference between the baseline model and one of the scenarios being tested.

Table 3: Comparison between baseline and scenario 1 (population behavior).

	Total number of infected people		Duration of the epidemic (days)	
	Low Level	High Level	Low Level	High Level
95% Confidence Interval	(3.29, 18.31)	(12.58, 27.22)	(-51.08, -11.92)	(-30.90, -13.30)
T-value	3.010	5.690	-3.370	-5.260
P-value	**0.007**	**0.000**	**0.003**	**0.000**

Table 4: Comparison between baseline and scenario 2 (population behavior and time to change behavior).

	Total number of infected people		Duration of the epidemic (days)	
	Low Level	High Level	Low Level	High Level
95% Confidence Interval	(-69.10, 23.50)	(1.10, 18.70)	(-22.50, 32.10)	(-21.22, -2.78)
T-value	-1.030	2.350	0.370	-2.720
P-value	0.316	**0.029**	0.717	**0.013**

Table 5: Comparison between baseline and scenario 3 (individual behavior).

	Total number of infected people		Duration of the epidemic (days)	
	Low Level	High Level	Low Level	High Level
95% Confidence Interval	(-66.10, 28.80)	(15.80, 29.10)	(-70.4, -20.40)	(-44.94, -19.06)
T-value	-0.820	7.070	-3.810	-5.180
P-value	0.421	**0.000**	**0.001**	**0.000**

Table 6: Comparison between baseline and scenario 4 (individual behavior and time to change behavior).

	Total number of infected people		Duration of the epidemic (days)	
	Low Level	High Level	Low Level	High Level
95% Confidence Interval	(-66.70, 31.40)	(8.49, 19.81)	(-44.10, 15.60)	(-25.86, -2.04)
T-value	-0.750	5.240	-1.000	-2.450
P-value	0.461	**0.000**	0.330	**0.024**

According to the p-values, we can see that at the parameters' high level all the scenarios led to statistically different results, when compared to the baseline, at 95% level of confidence. Part of this

difference can be possibly explained by the fact that in the high level it was assumed that either the population or the individuals adopted a more cautious behavior (cautious factor follows Uniform distribution, with minimum of 0.6 and maximum of 0.9). This fact does not necessarily imply that behavior is important once we are directly reducing the probability of infections, but it is a possible indicator and, hence, it shows the need for a further detailed investigation. On the other hand, in the case of scenario 2 and 4, the time is also larger (it follows Uniform, with minimum of 1 day and maximum of 7 days), which could lead to the percentage of infected people exceeding the threshold and triggering behavior, but also going down the threshold before the behavior even changes. Despite this possibility, the results are still statistically different from the baseline, reinforcing the importance for further analyzing the impacts of human behavior on disease spread.

Another interesting conclusion from the analysis is that by immediately changing the behavior of the whole population, all the results were statistically different from the baseline, regardless of the level of the parameters. However, for individual behavior, only three were statistically different. Based on this, we would like to raise the attention of researchers to the need and value of validating disease spread models with empirical data. In some cases, adding more details may not lead to as accurate results as considering aggregate information. However, to rigorously decide this trade-off between level of detail/information and computational power required, it is paramount to perform model validation.

Another point is that, except for one case, all results were statistically the same as the baseline when considering the response of total number of infected people at low level. Therefore, in cases where the number of infected people is more important than the epidemic duration, it may not be worth to include human behavior information in the model. Of course, this also depends on the size of behavior change.

As a general conclusion, we say that human behavior can indeed alter the results of a disease spread model, even in simple cases. Therefore, it is important to carefully analyze the information and validate the model before opting to ignore this parameter. As the differences in the results between low and high level indicate, it is also important to perform a sensitivity analysis to examine the variation of the results according to the parameter values. A first step would be elaborating an experimental design, based on full or fractional factorial, and performing ANOVA on the results to determine the significance of each factor.

5 CONCLUDING REMARKS

This study used ABS as a tool to highlight the impacts that human behavior may have in the results of disease spread models and, consequently, in epidemics in the real world. ABS allows capturing heterogeneity of populations and including different rules and behavior according to some attributes. However, the method requires more computational power. So, it is important to perform trade-off analysis between model accuracy and time. A model that is very accurate but takes days to run usually has little to no value for decision making in the real world.

Despite the impacts of human behavior in the course of an outbreak, many disease spread models still ignore this factor. We have some hypothesis about why this may happen as an attempt to answer the second question raised in this paper. First, data collection on disease spread is usually difficult. When we think about data collection about human behavior during an outbreak, the difficulty is even greater. Therefore, lack of data on behavior may be the first reason for not including it in simulation models. Second, to incorporate behavior in simulation, it is necessary to have a more detailed model that makes use of agents, which requires a lot of processing power, especially when the agent population is large. Third, many of the researchers in epidemiology are either epidemiologists, entomologists or statisticians/mathematicians, with a lack of background in human engineering or human factors. Lately this scenario has been changing with a greater involvement of engineers and human factors specialists in the modeling of disease spread. However, incentives to the development of multidisciplinary work are still needed to bridge this gap.

We recognize that validation and sensitivity analysis are important steps of simulation modeling and, as such, these are some of the limitations of this work, along with the simplicity of the model. However,

the focus of this work was not to develop a model for outbreak prediction. Instead, we wanted to illustrate that ignoring human behavior in disease modeling may lead to misrepresentative results. As the results have shown, human behavior is an important factor that shall be investigated by researchers and simulation specialists when modeling disease spread. Even in simple cases, small changes in human behavior may lead to statistically different results. On the other hand, as the results have also shown, sometimes human behavior may be more accurately modeled at the aggregate level, i.e., considering the average of the whole population instead of individuals' behavior. The results of the model may also depend on the output response being analyzed, the level of information aggregation adopted and the value of the parameters. Therefore, this work reinforces the importance of paying attention to the following topics when building a simulation model: (i) defining the simulation goal and the output response of interest beforehand; (ii) validating the model using empirical data in order to ensure that the parameter values adequately represent the reality; and, (iii) deciding the trade-off between level of information and computational power required, once adding more details not always lead to information gain. Although the model developed here is simple, the results align with what is known in this research field, which indicates that the model is a suitable tool for exploratory research.

As future work, we intend to perform similar analysis on different rules for behavior inclusion, such as change in behavior based on number of infected individuals within a specific distance or based on the number of infected individuals in a social network (emotional proximity). Another proposal for future work is to include more details, such as number of symptomatic cases and vertical transmission among agents, and to analyze how much information is gained with the inclusion of these new parameters.

ACKNOWLEDGMENTS

The authors would like to thank CAPES (Coordenação de Aperfeicoamento de Pessoal de Nível Superior) for the financial support to this work through the Science without Borders program.

REFERENCES

Aleman, D. M., T. G. Wibisono, and B. Schwartz. 2009. "Accounting for Individual Behaviors in a Pandemic Disease Spread Model". In *Proceedings of the 2009 Winter Simulation Conference*, edited by M. D. Rossetti, R. R. Hill, B. Johansson, A. Dunkin, and R. G. Ingalls, 1977–1985. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Bisset, K. R., X. Feng, M. Marathe, and S. Yardi. 2009. "Modeling Interaction between Individuals, Social Networks and Public Policy to Support Public Health Epidemiology". In *Proceedings of the 2009 Winter Simulation Conference*, edited by M. D. Rossetti, R. R. Hill, B. Johansson, A. Dunkin, and R. G. Ingalls, 2020–2031. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Bisset, K., J. Chen, C. J. Kuhlman, V. S. Kumar, and M. V. Marathe. 2011. "Interaction-Based HPC Modeling of Social, Biological, and Economic Contagions over Large Networks". In *Proceedings of the 2011 Winter Simulation Conference*, edited by S. Jain, R.R. Creasey, J. Himmelspach, K.P. White, and M. Fu, 2938–2952. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Bobashev, G. V., D. M. Goedecke, F. Yu, and J. M. Epstein. 2007. "A Hybrid Epidemic Model: Combining the Advantages of Agent-Based and Equation-Based Approaches". In *Proceedings of the 2007 Winter Simulation Conference*, edited by S. G. Henderson, B. Biller, M.-H. Hsieh, J. Shortle, J. D. Tew, and R. R. Barton, 1532–1537. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Brandeau, M. L. 2008. "Infectious Disease Control Policy: A Role for Simulation". In *Proceedings of the 2008 Winter Simulation Conference*, edited by S. J. Mason, R. R. Hill, L. Mönch, O. Rose, T.

Jefferson, and J. W. Fowler, 1578–1582. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Chan, M. 2014. "Ebola virus disease in West Africa - no early end to the outbreak". *New England Journal of Medicine* 371(13): 1183-1185.

Chang, C., K. Ortiz, A. Ansari, and M. E. Gershwin. 2016. "The Zika outbreak of the 21st century". *Journal of autoimmunity* 68:1-13.

Díaz, D. A., R. Akhavan-Tabatabaei, and I. Mura. 2016. "A Compartmentalized Simulation Model for Evaluation of HPV Vaccination Policies in Colombia". In *Proceedings of the 2016 Winter Simulation Conference*, edited by T. M. K. Roeder, P. I. Frazier, R. Szechtman, E. Zhou, T. Huschka, and S. E. Chick, 1994–2005. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Dietz, K., and J. A. P. Heesterbeek. 2002. "Daniel Bernoulli's Epidemiological Model Revisited". *Mathematical Biosciences* 180:1–21

Dumont, Y., F. Chiroleu, and C. Domerg. 2008. "On a Temporal Model for the Chikungunya Disease: Modeling, Theory and Numerics. *Mathematical Biosciences* 213:80–91.

Funk, S., S. Bansal, C. T. Bauch, K. T. Eames, W. J. Edmunds, A. P. Galvani, and P. Klepac. 2015. "Nine challenges in incorporating the dynamics of behaviour in infectious diseases models". *Epidemics* 10:21-25.

Johansson, M.A., L. Mier-y-Teran-Romero, J. Reefhuis, S. M. Gilboa, and S. L. Hills. 2016. "Zika and the risk of microcephaly". *New England Journal of Medicine* 375(1):1-4.

Kermack, W. O., and A. G. McKendrick. 1927. "A Contribution to the Mathematical Theory of Epidemics". In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* 700–721. London: The Royal Society.

Kreuger, K., and N. Osgood. 2015. "Particle Filtering Using Agent-Based Transmission Models". In *Proceedings of the 2015 Winter Simulation Conference*, edited by L. Yilmaz, W. K. V. Chan, I. Moon, T. M. K. Roeder, C. Macal, and M. D. Rossetti, 737–747. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Li, Y., M. A. Lawley, D. S. Siscovick, D. Zhang, and J. A. Pagán. 2016. "Peer Reviewed: Agent-Based Modeling of Chronic Diseases: A Narrative Review and Future Research Directions". *Preventing Chronic Disease*, 13:E69.

Lima, T., T. Carneiro, L. Silva, R. Lana, C. Codeço, I. Reis, R. Maretto, L. Santos, A. Monteiro, L. Medeiros, and F. Coelho. 2014. "A Framework for Modeling and Simulating Aedes aegypti and Dengue Fever Dynamics". In *Proceedings of the 2014 Winter Simulation Conference*, edited by A. Tolk, S. Y. Diallo, I. O. Ryzhov, L. Yilmaz, S. Buckley, and J. A. Miller, 1481–1492. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Macdonald, G. 1952. "The Analysis of Equilibrium in Malaria". *Tropical Diseases Bulletin* 49: 813–829.

Macdonald, G. 1957. *The Epidemiology and Control of Malaria*. 1st ed. London: Oxford University.

Majumder, M.S., E. L. Cohn, S. R. Mekaru, J. E. Huston, and J. S. Brownstein. 2015. "Substandard vaccination compliance and the 2015 measles outbreak". *JAMA pediatrics* 169(5):494-495.

Moulay, D., M. A. Aziz-Alaoui, and M. Cadivel. 2011. "The Chikungunya Disease: Modeling, Vector and Transmission Global Dynamics". *Mathematical Biosciences* 229:50–63.

Paleshi, A., G. W. Evans, S. S. Heragu, and K. S. Moghaddam. 2011. "Simulation of Mitigation Strategies for a Pandemic Influenza". In *Proceedings of the 2011 Winter Simulation Conference*, edited by S. Jain, R.R. Creasey, J. Himmelspach, K.P. White, and M. Fu, 1340–1348. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Ross, R. 1908. *Report on the Prevention of Malaria in Mauritius*. 1st ed. London: Waterlow.

Ross, R. 1911. "Some Quantitative Studies in Epidemiology". *Nature* 87:466–467.

Sanchez, P. J., and S. M. Sanchez. 2015. "A Scalable Discrete Event Stochastic Agent-Based Model of Infectious Disease Propagation". In *Proceedings of the 2015 Winter Simulation Conference*, edited by

L. Yilmaz, W. K. V. Chan, I. Moon, T. M. K. Roeder, C. Macal, and M. D. Rossetti, 151–158. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Scheidegger, A. P. G., and A. Banerjee. 2017a. "A Mathematical Approach for Classifying Input Parameters for Infectious Disease Spread". In *Proceedings of the 2017 Industrial and Systems Engineering Conference*, edited by K. Coperich, E. Cudney, and H. Nembhard. Norcross: Institute of Industrial and Systems Engineers.

Scheidegger, A. P. G., and A. Banerjee. 2017b. "Bridging the Gap between Human Factors and Epidemiological Models: Recommendations for Future Research". In *Proceedings of the 2017 Applied Human Factors and Ergonomics Conference*. Berlin: Springer. (In Press)

Verelst, F., L. Willem, and P. Beutels. 2016. "Behavioural Change Models for Infectious Disease Transmission: A Systematic Review (2010–2015)". *Journal of the Royal Society Interface* 13: 20160820.

Walker, P. G., M. T. White, J. T. Griffin, A. Reynolds, N. M. Ferguson, and A. C. Ghani. 2015. "Malaria morbidity and mortality in Ebola-affected countries caused by decreased health-care capacity, and the potential effect of mitigation strategies: a modelling analysis". *The Lancet Infectious Diseases* 15(7): 825-832.

Wang, Z., M. A. Andrews, Z. X. Wu, L. Wang, and C. T. Bauch. 2015. "Coupled disease–behavior dynamics on complex networks: A review". *Physics of life reviews* 15:1-29.

Wenham, C. 2017. "What we have learnt about the World Health Organization from the Ebola outbreak." *Phil. Trans. R. Soc. B* 372(1721): 20160307.

WHO (World Health Organization). 2014. "A global brief on vector-borne diseases". WHO Press, Geneva, Switzerland.

WHO (World Health Organization). 2015. "Health in 2015: from MDGs, Millennium Development Goals to SDGs, Sustainable Development Goals". WHO Press, Geneva, Switzerland.

WHO (World Health Organization). 2016. The Top 10 Causes of Death. http://www.who.int/mediacentre/factsheets/fs310/en/index1.html.

Yakob, L., and A. C. Clements. 2013. "A Mathematical Model of Chikungunya Dynamics and Control: The Major Epidemic on Réunion Island". *PloS one* 8:e57448.

Zarei, M., and E. Smith. 2011. "Epidemiology from the Human Perspective, Using Subjectively Perceived Probability of Contagion". In *Proceedings of the 2011 IIE Annual Conference*, edited by A. Johnson and J. Miller, 1–8. Norcross: Institute of Industrial and Systems Engineers.

Zipprich, J., K. Winter, J. Hacker, D. Xia, J. Watt, and K. Harriman. 2015. "Measles outbreak-California, December 2014-February 2015". *MMWR Morb Mortal Wkly Rep* 64(6):153-154.

AUTHOR BIOGRAPHIES

ANNA PAULA GALVÃO SCHEIDEGGER is a PhD student in the Industrial and Systems Engineering department at Texas A&M University. She holds a master's degree in industrial engineering from Universidade Federal de Itajubá, Brazil. Her research interests include modeling and simulation and its application to disease and disaster management. Her e-mail address is apscheidegger@tamu.edu.

AMARNATH BANERJEE is a Professor and Corrie and Jim Furber '64 Faculty Fellow of Industrial and Systems Engineering at Texas A&M University. He received his Ph.D. in Industrial Engineering and Operations Research from the University of Illinois at Chicago, and BS in Computer Science from Birla Institute of Technology and Science, Pilani, India. His research interests are in modeling, simulation and visualization, with applications in manufacturing, health care, and information systems. His email address is banerjee@tamu.edu.