

An Agent Based Model for Simulating the Spread of Sexually Transmitted Infections

Grant Rutherford¹, Marcia R Friesen¹, Robert D McLeod¹

¹Electrical & Computer Engineering, University of Manitoba, Canada

Abstract

This work uses agent-based modelling (ABM) to simulate sexually transmitted infection (STIs) spread within a population of 1000 agents over a 10-year period, as a preliminary investigation of the suitability of ABM methodology to simulate STI spread. The work contrasts compartmentalized mathematical models that fail to account for individual agents, and ABMs commonly applied to simulate the spread of respiratory infections. The model was developed in C++ using the Boost 1.47.0 libraries for the normal distribution and OpenGL for visualization. Sixteen agent parameters interact individually and in combination to govern agent profiles and behaviours relative to infection probabilities. The simulation results provide qualitative comparisons of STI mitigation strategies, including the impact of condom use, promiscuity, the form of the friend network, and mandatory STI testing. Individual and population-wide impacts were explored, with individual risk being impacted much more dramatically by population-level behaviour changes as compared to individual behaviour changes.

Keywords: Agent based modelling, modelling and simulation, sexually transmitted diseases, social networks

Introduction

The objective of this work was to develop an agent-based model (ABM) to simulate the spread of sexually transmitted infections (STIs) within a population of interacting agents. As a preliminary application of the ABM methodology to STI spread, the focus of this work was to explore the inherent suitability and potential of the ABM method to this particular context.

Agent based modelling is becoming an effective tool in understanding infection spread and is particularly well suited to environments where the agents themselves and their interaction with one another are the principal vectors of infection spread. Agent-based models have emerged in the past decades as a complementary approach to the long history of differential equation-based models that require a macroscopic perspective of the population of interest [1].

Agent based modelling is ‘bottom-up’ systems modelling from the perspective of constituent parts. Systems are modelled as a collection of agents (in this case, people) imbued with properties: characteristics, behaviours (actions), and interactions that attempt to capture actual properties of individuals. In the most general context, agents are both adaptive as well as autonomous decision-making entities who are able to assess their situation, make decisions, compete with one another on the basis of a set of rules, and adapt future behaviours on the basis of past interactions. Agent properties may be conceived by the modeller or may be derived from actual data that reasonably describe agents’ behaviours – i.e. their

movements and their interactions with other agents. The modeller's task is to determine which data sources best govern agent profiles in a given ABM simulation [2][3].

The foundational premise and conceptual depth of ABM is that simple rules of individual behaviour will aggregate to illuminate or exhibit complex and emergent group-level phenomena that is not specifically encoded by the modeller [3][3]. This emergent behaviour may be counterintuitive or a complex behavioural whole that is greater than the sum of its parts. Furthermore, ABM provides a natural description of a system that can be calibrated and validated by subject matter experts, and is flexible enough to be tuned to high degrees of sensitivity in agent behaviours and interactions. ABMs are considered particularly applicable to situations where interactions are local and potentially complex, where agents are heterogeneous, where the phenomenon has inherent temporal aspects, and where agents are adaptive [2]-[5].

Much of the work in ABM to date has focussed on the simulation of contact-based infection spread associated with influenza-like illnesses and other respiratory infections, including coarse-scale community, country, and global models [6]-[13] to finer-scale community and institutional models [14]-[15].

Increasingly, there are considerable data that can be used to improve the fidelity of an ABM in representing real social networks in ways that are highly amenable to understanding the type of contacts (casual and behaviour-agnostic) that facilitate infection spread. These data may be generated for other purposes but are increasingly available toward secondary and tertiary applications within an emerging 'data culture'. These data that can be mined and analysed to generate agent contact and movement patterns include, for example, intelligent transportation systems (vehicular, public, air travel, etc), cellular service provider data, and a range of location-based services and technologies that may leverage Smartphones and other Bluetooth-enabled mobile devices.

Additionally, there is a large body of literature on modelling sexually transmitted infection spread based on equation-based monolithic approaches or compartmental mathematical models [15]-[16]-[17]. Monolithic analytical models use one governing equation to model the population, while compartmental tools reduce the population to a few key characteristics which are relevant to the infection under consideration. These models are limited to treating simplified scenarios and are not amenable to including data directly pertaining to each *individual* agent. In contrast, the strength of ABM lies in its detailed and naturalistic representation of agents and scenarios and the ability to directly integrate real data.

However, modelling of sexually transmitted infections presents a more significant challenge to ABMs than modelling of contact-based respiratory infections. Proximate contact and geographical location (important factors in respiratory infection spread) are a minor factor in STI transmission, while agent behaviour becomes a defining parameter, including the formation of relationships, and an agent's network of contacts through which relationships may be established. In addition the topic of sexually transmitted infections and sexual behaviour is a sensitive topic for most people, and accurate and full disclosure of activities in order to accurately characterize agents in the model is difficult to achieve.

This work derives novelty in applying the ABM methodology to the context of STI spread, which has traditionally been modelled by equation-based and compartmental methods. At the early stage of applying ABM methods to the simulation of STI spread, qualitative comparison to general known outcomes is a meaningful objective. Overall, the usefulness of ABM validation by qualitative comparison to known outcomes has been established by others [1].

There are known limitations stated at the end of the paper, and these limitations would need to be addressed thoroughly before reliable results could be extracted from the simulation. At this stage of the work, the focus is on exploring the potential and suitability of the ABM *method* to this context. Drawing on these findings, future work will refine the ABM and then emphasize and analyse the *results* of the simulations themselves in detail.

Methods

The Agent Based Model

This work developed an agent-based model to simulate the spread of sexually transmitted infections, or STIs, which are difficult to study directly. The study created an ABM of a population of sexually active people. A disease is introduced to this population from the external world, and its progress can be traced through the model. This approach allows the effects of various mitigative and control policies and behaviours to be easily analyzed. The use of a computer model allows thousands of populations to be tested in a short amount of time, so the possible effects of new policies and behaviours can be evaluated quickly and easily.

Within the model, each agent has both predetermined and fixed parameters which regulate their behaviour. Each agent can also be in a number of different states which affect the expression of their predetermined parameters. The movement of agents from a healthy state to an infected state simulates the spread of a sexually transmitted infection through the population.

Each simulation was run with a population of 1000 people for a period of 10 simulated years (3652 days) after the initial infection. These output included the proportion of the population which was infected, and records of the individuals along with their final infection state and the number of people to whom they passed the infection. In all cases, the original source of the infection is the “outside world”, and the infection enters the population through sexual contact between the population and the outside world.

The process of simulating a population involves two major steps. The population is generated and a period of one year is simulated without any chance of infection in order for the proportion of monogamous relationships to become stable. Following this initial calibration step, the simulation runs for exactly 10 years after the first agent in the population is infected. Each simulation is repeated approximately 5000 times with a new population generated each time. Results are accumulated and averaged to show the trends in the data.

The ABM framework was coded in C++ using the Boost 1.47.0 libraries for the normal distribution and OpenGL for visualization.

Model Setup

The first step of simulation is the generation of the agent population. Each individual agent in the simulation is randomly assigned unique values which will govern their behaviour. These values are assigned from a distribution which is characterized by global mean and standard deviation parameters. These global distribution parameters define the mean and variance of a parameter in a population, but each individual agent is assigned a specific and fixed value for each parameter upon generation, so that their profile is unique in relation to other agents in the simulation. Standard parameter values for the distributions are given in Table 1.

For both the baseline simulation scenario and to explore various intervention scenarios, the simulation is run multiple times from same initial model conditions in order to average the results. The population of agents is regenerated for each simulation run, so that the agents are different each time, but the distributions from which their parameter values are drawn are constant for a given scenario.

More specifically, each agent is randomly assigned to be male or female. Following this, each person is assigned a unique value for each of 15 unique parameters, which vary with a restricted normal distribution between 0 and 1 (Table 1). Each of these 15 parameters represents a probability, or an offset to a probability. A 16th parameter, “desiredFriends” is an integer and is determined by taking the floor of a normally distributed value with a mean of 2 and a standard deviation of 2, and re-rolling the value if it falls below 1. In this manner everyone has at least one friend. Where possible, parameter values were informed by the body of literature on the topic [18][19].

Friendship Model

The “friendship” model simulates connections between people. These connections may be interpreted as frequent contact: a close friend, a co-worker, etc. During the simulation, these links and networks will often be used to find partners for relationships and for sex. Realistically, a person would be expected to have many such connections, but this model only included the most relevant. For the sake of clarity, from now on these connections will be referred to as friends.

The “*cliqueFactor*” parameter has the ability to control the separation of people into tight groups of friends (high *cliqueFactor*) or a more connected but randomly shaped network (low *cliqueFactor*). A *cliqueFactor* of 0.7 was chosen as the standard value for the model, which gives a well-connected network with tight clusters of friends. The network evolves over time as relationships instantiated and later dissolve, while maintaining a small world flavour.

Table 1: Population parameters with standard values

Parameter	Type	Mean	Std. Dev.	Description
IsMale	50/50	NA	NA	Male or Female
baseSexAvailability (bSA)	Re-roll	M = 0.5 F = 0.03	M = 0.2 F = 0.05	Base chance of availability for sex outside of monogamy
baseSexSeeking (bSS)	Restricted	M = 0.1 F = 0.002	M = 0.1 F = 0.002	Base chance to be seeking sex if also available, outside of monogamy
baseCondomUse (bCU)	Restricted	0.9	0.5	Base chance of desiring a condom
AvailableMonogamy	Re-roll	0.08	0.04	Chance of being available for a new mono. relationship
SeekingMonogamy	Re-roll	0	0.01	Chance of seeking a new monogamous relationship
ExitMonogamy	Re-roll	0	0.001	Chance of terminating an existing mono. relationship
MonogamousSexRate	Re-roll	0.15	0.1	Chance of desiring sex during monogamy
monoSexAvailabilityDecrease	Dependent	1 x bSA	0.4 x bSA	Offset to bSA while monogamous
monoSexSeekingDecrease	Dependent	1 x bSS	0.3 x bSS	Offset to bSS while monogamous
MonoCondomDecrease	Restricted	0.5	0.5	Offset to bCU while monogamous
diseaseAvailabilityDecrease	Dependent	1.2 x bSA	0.8 x bSA	Offset to bSA while diseased
DiseaseSeekingDecrease	Dependent	1.5 x bSS	0.5 x bSS	Offset to bSS while diseased
DiseaseCondomIncrease	Restricted	0.8	0.5	Offset to bCU while diseased
TestingHealthy	Restricted	-0.1	0.05	Daily chance of getting tested while healthy
TestingSymptoms	Restricted	0.1	0.1	Daily chance of getting tested while symptomatic
DesiredFriends	Special	2	2	Desired number of friends

Simulation

The simulation of the infection spread through the population happens in discrete units of time, chosen to be one day per step. During each simulated day, the following steps are done in order. The simulation continues for 3652 days (10 years) after the first agent is infected.

While the model is running there are a number of relationship changes that can occur. These include:

Step 1 - Monogamy changes:

In this step, a list of people in monogamous relationships is generated, along with probabilities of relationships ending, new relationships desired, and new monogamous relationships formed within the population. When new relationships form, there is a 95% chance that a person will choose someone from within their friend network, and a 5% chance that a person will choose someone at random from the list of available people. Finding someone through the friend network uses a breadth-first search of up to 5 hops, with the closest available person of the opposite gender selected. If more than

one person is found at the same distance, then one is selected at random. If no suitable person is found using the friend network, then the seeker will default to selecting someone at random from the list of available people of the opposite gender. The above process is repeated until either the seeking list is depleted, or the list of available people is depleted of either gender. At this point, monogamous relationship matching is completed.

Step 2 – Determining Sexual Encounters:

- **Monogamous Sexual Encounters:** Each person in a monogamous relationship uses their `monogamousSexRate` variable in order to determine if they desire monogamous sex. This value is first modified by subtracting their `diseaseAvailabilityDecrease` parameter if they are confirmed to have the infection. If they are symptomatic for the infection, but have not been tested, then they subtract half of the value of this parameter. If the resulting value is less than zero, then a zero probability is used.
- **Non-monogamous sexual encounters:** This is a complex parameter impacted by sex availability and modulated by the agent's non or symptomatic state, but following the general process of matching people into monogamous relationships. "Cheating" within monogamous relationships is also accounted for.

Step 3 - Spreading the Infection:

While the model is running there are a number of factors that impact the spread of infection, including the `baseCondomUse` parameter modified by their monogamy status and their infection status (including varying probabilities for symptomatic but unconfirmed, vs. tested and confirmed). Once condom use has been determined, each encounter between a person who is uninfected and a person with the infection has a probability of infection transmission (modified by condom use) and a subsequent probability of becoming symptomatic if infected. Becoming symptomatic further impacts probability of testing and probability of behaviour modification.

Step 4 - Testing for Infection:

The final step in the simulated day is to test members of the population for the disease, accounting for persons (asymptomatic) who may submit to a spontaneous test, and the probability of symptomatic persons seeking testing.

Results and Discussion

When evaluating the simulation, the metric used was the infected proportion of the population after 10 years. Ten years was chosen in order to capture an average infection prevalence of around 50%, which allows ample room to both evaluate changes in outcomes through parameter changes. With the standard parameter values, the average infection prevalence after 10 years was 48.71% over 4732 trials.

Two types of analysis were performed on the model. The first set involved 4732 trials (simulations) using the standard parameters, and collecting data on the 4,732,000 simulated agents (1000 agents x 4732 trials) generated for these trials. By measuring correlations between agent parameters and infection outcomes, some trends regarding individual risk can be identified. The second type of analysis involved using the standard parameter values for all parameters but one, and varying a single parameter through a range of values. These trials

can be used to investigate the results of general policy or behaviour changes applied to the entire population.

Individual Risk:

Condom use: The most pronounced impacts related to condom use. Condom use while healthy and single was strongly related to an individual's risk of acquiring an infection (Fig. 1). Condom use while single and aware of being infected is strongly correlated to the likelihood of passing the infection to others (Fig. 2).

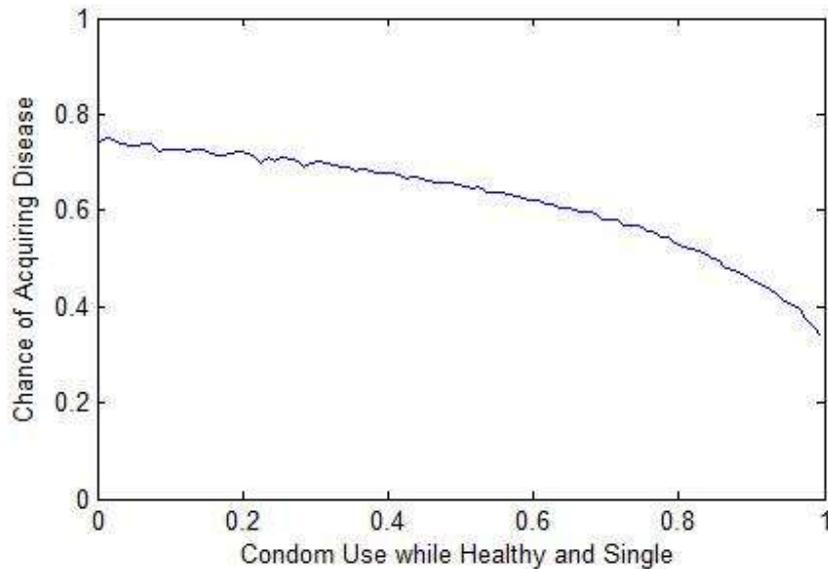


Figure 1: Individual Risk vs. Condom Use (Healthy & Single)

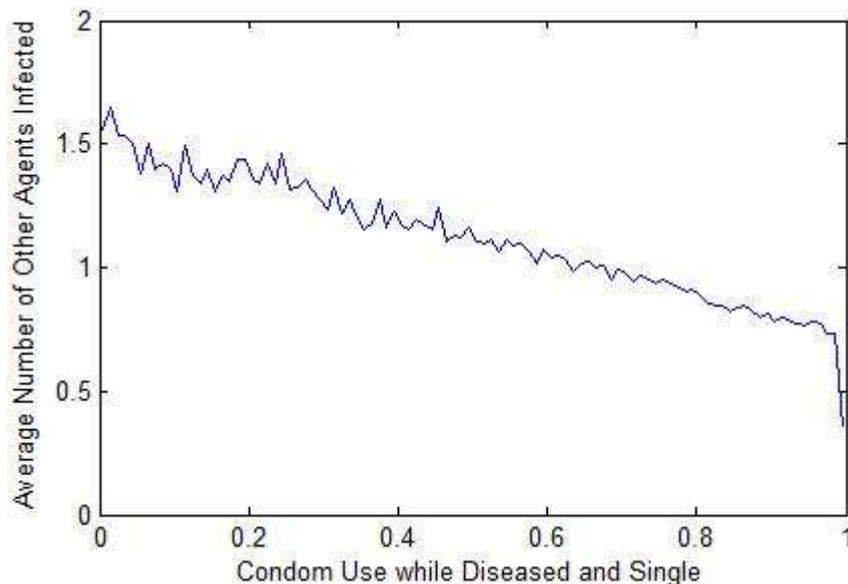


Figure 2: Infection Spread to Others vs. Condom Use (Infected & Single)

Sexual availability was correlated with individual risk of acquiring an infection. However, the risk of infection remains near 25% even for persons who are never available for sex, due to the possibility of acquiring infection through partners in monogamous relationships (accounting for a degree of ‘cheating’).

The risk of infection was only moderately affected by an agent’s average number of friends, with nine friends corresponding to an infection risk of 55.2% and having one friend corresponding to an infection risk of 46.5%. Since a large proportion of seeking occurs through the friend network, the proximity of a person to potential partners through the friend network will impact frequency of interaction.

Population Data:

In comparison with the individual data in the previous section, the population data was obtained by changing the standard model parameters, one by one, thereby of changing the properties and behaviour of the entire population. Four parameters were varied: condom use, promiscuity, time between mandatory testing, and cliqueness of the friend network. In total, 60 discrete values of each of the four noted parameters parameter were tested, and the simulation results were averaged over 50 repeated simulation runs at each value.

Condom Use: Global changes to condom use have a very significant effect on the disease prevalence (Fig. 3), with an infection prevalence of 97.4% and 13.9% with mean condom use of 0 and 1, respectively. Dramatic changes to condom have the potential to reduce the infection prevalence by more than a factor of 7.

Figure 3 shows a steep change in infection prevalence at a mean condom use between 60% and 80%, suggesting a threshold over which the ubiquitous use of condoms severely slows the spread of infection. Beyond this level, the benefits from higher condom attenuate.

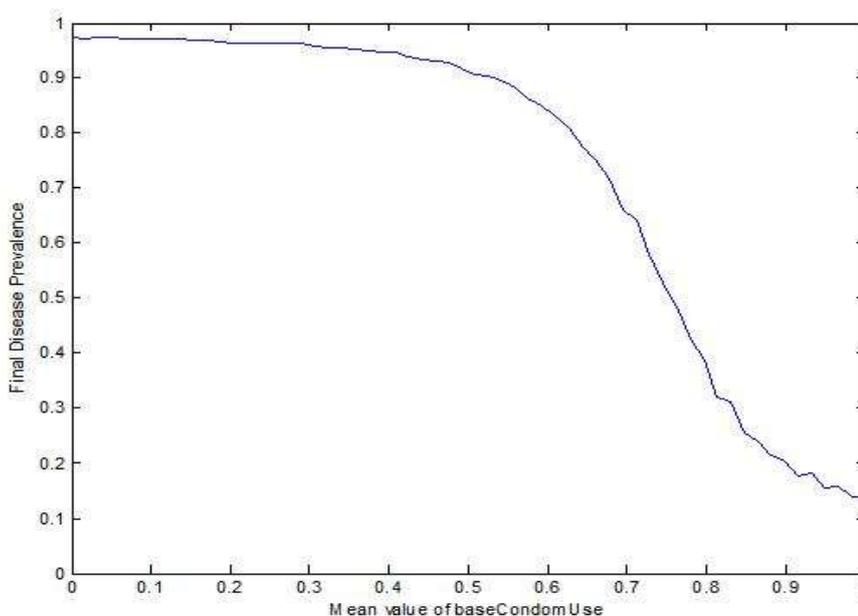


Figure 3: Infection Prevalence vs. Population Condom Use

Promiscuity: The impact of promiscuity was simulated by controlling availability for sex and sex-seeking characteristics. Population infection prevalence ranged from 48.7%, 13.6% and 88.2% with promiscuity parameters at standard, halved, and doubled values, respectively.

Mandatory Infection Testing: This parameter varied the length between mandatory infection testing of all individuals (including healthy, asymptomatic) from daily testing to testing every 14.7 years, in increments of 91 days. The infection prevalence decreased steadily as the test interval decreased. Annual testing resulted in population infection prevalence of 15.3%. With testing at 14.7 year intervals, the population infection prevalence dropped from 48.7% (the standard model) to 43.0%. Even in the impractical scenario of daily mandatory testing, an endemic level of infection remains in the population.

Friend Network: The impact of the form of the friend network was examined by altering the connectedness of the network. The results showed no meaningful correlation between the form of the friend network and infection prevalence within the population.

Conclusion

There are some known limitations to this agent-based model for STI spread. While the 16 parameters act both independently and in combination and represent numerous complex combinations and many unique agent profiles, they nonetheless do not capture the full range of agent profiles, agent behaviours, and infection transmission dynamics within a population. For example, this model only considers opposite-sex sexual encounters. Second, in this simulation, agents never leave the population and new agents do not enter the population. Agents remain in state and are immortal, representing a simplification of real movement of people into and out of networks. Furthermore, in this model, the agent behavioural probabilities (seeking, condom use, etc.) remain the same, regardless of whether the overall infection prevalence within the population is low or high. In reality, some behaviour modification is initiated when infection prevalence rises due to personal knowledge of risk and public health messages. Population norming becomes a feedback mechanism (e.g. public health messages toward hand hygiene and cough etiquette during pandemic influenza). In this model, a sexual encounter is equated with an infection transmission with some probability (determined by interacting agent parameters). In reality, the mapping of the sexual encounter to infection transmission is more complex, as disease epidemiology needs to be taken into account. It is known, for example, that the rate of transmission may be dependent on the stage of a disease, which is not accounted for in the current model.

Notwithstanding the limitations, the qualitative results of the model correspond well with known priorities in STI mitigation strategies. This finding supports the objective of determining the potential and suitability of the ABM methodology to the context of STI spread, and an STI-ABM's future potential as a tool in public health decision-making and policy. A person can significantly control their individual infection risk by using a condom and avoiding non-monogamous sex. This result is intuitive and supported by simulation. However, individual risk is not reduced as easily or dramatically through individual behaviour modification as it is through collective behavioural changes in the population which reduces collective risk. This mirrors the impact of vaccination for childhood diseases, where protection is only achieved through population-wide compliance. This result is also intuitive, considering the cascading effects within a network, where collective behaviour changes are amplified because individuals benefit from the changes to their own behaviour and also from the reduction in risk due to the change in behaviour of their peers. This makes

a strong case for public health policy promoting sexual health, such as the promotion of condom use and responsible sexual behaviour.

Mandatory testing for STIs can have a very strong effect in limiting their spread. Challenges including the lack of incentive for testing particularly if asymptomatic, the personal disincentives (e.g. potential embarrassment), and the stigma attached to STIs.

This work has presented an agent-based modelling framework for the simulation of sexually transmitted infection spread within a population of 1000 people, over a 10-year period, in order to explore the applicability of the ABM approach to STI modelling. While these results provide some preliminary support for the suitability of the ABM methodology to an STI application, all of these results need to be more robustly explored and verified by developing the ABM further to address the stated limitations. In general, the methodology is well suited to practitioners and educators, and lends itself to qualitative assessments of various mitigation and control measures related to STI spread. There is considerable refinement that may be easily undertaken as the ABM methodology explicitly facilitates dialog within an easily communicated lexicon.

An accompanying YouTube video illustrating the model and evolving network can be found at <http://www.youtube.com/watch?v=AqL8mBgns8U>

Funding

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Conflicts of Interests

None.

Corresponding Author

Marcia Friesen
Assistant Professor, Design Engineering
University of Manitoba, Canada
Email: marcia.friesen@ad.umanitoba.ca

References

1. Emrich S, Suslov V, Judex F. Fully agent based modelling of epidemic spread using Anylogic. Proc. EUROSIM 2007, 9-13 Sept. 2007, Ljubljana, Slovenia.
2. Bonabeau E. 2002. Agent-based modelling: Methods and techniques for simulating human systems. Proc Natl Acad Sci USA. 99(Suppl 3), 7280-87. <http://www.pnas.org/content/99/suppl.3/7280.full#xref-ref-3-1>. <http://dx.doi.org/10.1073/pnas.082080899>
3. Rand W, Rust RT. Agent-based modelling in marketing: Guidelines for rigor. International Journal of Research in Marketing 2011, http://ijrm.feb.uvt.nl/uploads/Forthcoming_D-10-00071_RandRust.pdf

4. Goldstone RL, Janssen MA. 2005. Computational models of collective behaviour. *Trends Cogn Sci.* 9, 424-30. <http://dx.doi.org/10.1016/j.tics.2005.07.009>
5. Hupert N, Xiong W, Mushlin A. 2008. The virtue of virtuality: The promise of agent-based epidemic modeling. *Transl Res.* 151(6), 273-74. <http://dx.doi.org/10.1016/j.trsl.2008.04.002>
6. Epstein JM. 2009. Modelling to contain pandemics. *Nature.* 460, 687. <http://dx.doi.org/10.1038/460687a>
7. Germann TC, Kadau K, Longini IM, Jr, et al. 2006. Mitigation strategies for pandemic influenza in the United States. *Proc Natl Acad Sci USA.* 103, 5935-40. <http://dx.doi.org/10.1073/pnas.0601266103>
8. Longini IM, Jr, Nizam A, Xu S, et al. 2005. Containing pandemic influenza at the source. *Science.* 309, 1083-87. <http://dx.doi.org/10.1126/science.1115717>
9. Ferguson NM, Cummings DAT, Fraser C, et al. 2006. Strategies for mitigating an influenza pandemic. *Nature.* 442, 448-52. <http://dx.doi.org/10.1038/nature04795>
10. Bobashev GV, Goedecke DM, Yu F, et al. A hybrid epidemic model: Combining the advantages of agent-based and equation-based approaches. *Proceedings of the 2007 Winter Simulation Conference 2007*;1532-1537
11. Merler S, Ajelli M, Jurman G, et al. Modeling influenza pandemic in Italy: An individual-based approach. *Proceedings of the 2007 intermediate conference of the Italian Statistical Society 2007*, http://www.sisstatistica.it/files/pdf/atti/SIS%202007%20Venezia%20intermedio_121-131.pdf
12. Skvortsov BAT, Connell RB, Dawson PD, et al. Epidemic modelling: Validation of agent-based simulation by using simple mathematical models. *MODSIM 2007 International Congress on Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand 2007*; 657-662.
13. Borkowski M, Podaima BW, McLeod RD. 2009. Epidemic modeling with discrete space scheduled walkers: Possible extensions to HIV/AIDS. *BMC Public Health.* 9(Suppl 1), S14. <http://dx.doi.org/10.1186/1471-2458-9-S1-S14>
14. Meng Y, Davies R, Hardy K, et al. 2010. An application of agent-based simulation to the management of hospital-acquired infection. *Journal of Simulation.* 4, 60-67. <http://dx.doi.org/10.1057/jos.2009.17>
15. Mukhi S, Laskowski M. Agent-based simulation of emergency departments with patient diversion. In *Electronic Healthcare, D. Weerasinghe, Ed. Berlin: Springer, 2009*:25-37.
16. Anderson RM, Geoffrey G. 2000. Mathematical models of the transmission and control of sexually transmitted diseases. *Sex Transm Dis.* 27, 636-43. <http://dx.doi.org/10.1097/00007435-200011000-00012>
17. Ferguson NM, Geoffrey GP. 2000. More realistic models of sexually transmitted disease transmission dynamics: Sexual partnership networks, pair models, and moment closure. *Sex Transm Dis.* 27, 600-09. <http://dx.doi.org/10.1097/00007435-200011000-00008>
18. Doherty IA, Shiboski S, Ellen J, et al. 2006. Sexual bridging socially and over time: A simulation model exploring the relative effects of mixing and concurrency on viral sexually transmitted infection transmission. *Sex Transm Dis.* 33, 368-73. <http://dx.doi.org/10.1097/01.olq.0000194586.66409.7a>
19. Azra GC, Geoffrey GP. 2000. Risks of acquiring and transmitting sexually transmitted diseases in sexual partner networks. *Sex Transm Dis.* 27, 587-97.
20. New York City Department of Health and Mental Hygiene. *Community Health Survey 2011*, <http://www.nyc.gov/html/doh/html/survey/survey.shtml>, accessed October 19, 2012.