Calibrating and Validating Large Scale Agent-Based Simulations of Populations

Douglas Roberts, Philip C. Cooley, Diane Wagener, Geoffrey Frank
RTI International
Research Triangle Park, North Carolina
droberts@rti.org, pcc@rti.org, dwagener@rti.org, gaf@rti.org

ABSTRACT

Modeling the "human terrain" is important for effective training of non-kinetic missions in urban terrain. It is equally important for planning responses to epidemics of infectious diseases, where combinations of vaccines, antiviral drugs, and social distancing are the tools to be employed. Social distancing methods include quarantine; closing schools, workplaces, and community centers; and travel restrictions. Social distancing methods are relevant to both non-kinetic missions and responses to epidemics.

This paper describes the development of behavior models for large-scale agent-based models to support planning for responses to potential epidemics such as avian flu (Cooley et al., 2008). In these models, each person in an urban area is represented as an agent. Models have been constructed for several municipal areas, including Chicago, Illinois (Ferguson, Longini); Portland, Oregon (Eubanks); Pittsburgh, Pennsylvania (Cooley): and the Research Triangle Park, North Carolina (Cooley). These behavior models must represent the interaction patterns of the people in the area being modeled. This requires fusing data from multiple sources about the social networks and movement patterns of the population. This process starts with an allocation of agents to the households in the area that is consistent with available demographic information including age, gender, and socio-economic status. Algorithms assign these agents to their schools, workplaces, shopping malls, and hospitals in ways that are consistent with available source data. Additional algorithms estimate daily travel patterns and interaction times at these locations based on these assignments and additional source data on travel times.

ABOUT THE AUTHORS

Douglas Roberts is a senior computer scientist in the Research Computing Division at RTI International[1]. He is involved with the design, implementation, and use of large-scale distributed agent based models of social network systems. Recently he participated in a study to determine effective intervention strategies in the event of a national outbreak of pandemic influenza. Prior to coming to RTI, he was part of the team that developed TRANSIMS, a transportation system ABM, and EpiSims, an epidemiological simulation ABM.

Diane K. Wagener is the Principal Investigator for the Informatics team at RTI International on the Modeling of Infectious Disease Agent Study (MIDAS) project. She is a former public health advisor to the Centers for Disease Control and Prevention and the U.S. Assistant Secretary of Health and a genetic epidemiologist. She leads other projects analyzing genomics and proteomics associated with immune responses to vaccinations.

Philip C. Cooley is Assistant Director for Bioinformatics in the Research Computing Division at RTI International. He is a PI with Dr. Wagener for the Informatics team on the MIDAS project. He leads and develops infectious disease simulations and holds patents on software packages for survey data capture and statistical processing.

Geoffrey A. Frank is a Principal Scientist in the Technology Assisted Learning Center at RTI International. Recently, Dr. Frank has been developing an ontology to assist in finding relevant models and data sources for epidemiological models and simulations.

---

[1] RTI International is a trade name of Research Triangle Institute.

Calibrating and Validating Large Scale Agent-Based Simulations of Populations

Douglas Roberts, Philip C. Cooley, Diane Wagener, Geoffrey Frank
RTI International
Research Triangle Park, North Carolina
droberts@rti.org, pcc@rti.org, dwagener@rti.org, gaf@rti.org

INTRODUCTION

Modeling the "human terrain" is important for effective training of non-kinetic missions in urban terrain. The success of these missions often depends upon how a large and diverse civilian population will respond to different Courses of Action (COA). Their response is dependent upon their normal patterns of behavior and how those behaviors change in response to changes in the environment and to the interventions that are part of the COA.

This paper describes how agent-based models (ABM) can be used for planning responses to epidemics of infectious diseases, where combinations of vaccines, antiviral drugs, and social distancing are employed.

AGENT-BASED MODELS OF EPIDEMICS

Large-scale agent-based models have been used to support planning for responses to potential epidemics such as avian flu (Cooley et al., 2008). In these models, each person in an urban area is represented as an agent. Models have been constructed for several municipal areas, including Chicago, Illinois (Ferguson, 2006; Germann et al., 2006); Portland, Oregon (Eubank, 2005); Pittsburgh, Pennsylvania (Cooley et al., 2008): and the Research Triangle Park, North Carolina (Cooley et al., 2008).

Agent Models of Civilian Populations

The behavior models represent the interaction patterns of the people in the area being modeled. This requires fusing data from multiple sources about the social networks and movement patterns of the population. This process starts with an allocation of agents to the households in the area that is consistent with available demographic information including age, gender, and socio-economic status. Algorithms assign these agents to their schools, workplaces, shopping malls, and hospitals in ways that are consistent with available source data. Additional algorithms estimate daily travel patterns and interaction times at these locations based on these assignments and additional source data on travel times.

Social Network Models

A key tradeoff when constructing a model involves the level and nature of the population disaggregation. The social-network model disaggregates the population into subgroups according to certain characteristics (e.g., age, gender, religion, socio-economic status) and provides behaviors for each of these subgroups. The subgroup behavior models include appropriate models of the daily activities of that population subgroup. For example, infants typically stay home with a family member, school-age children go to school, adults go to work, and the aged stay at home or in group quarters. These agents interact with other agents that regularly share the same locations for some period of each day. Table 1 (Glass et al., 2005) shows how disaggregating the population by age, using empirical data for infection rates, and modeling the different social interaction patterns of the different age groups changes the overall infection rates. This information is critical for targeting interventions to specific demographic categories.

Table 1: Effects of Separate Demographic Models

	To Children	To Teenagers	To Adults	To Seniors	Total From
From Children	21.4	3.0	17.4	1.6	43.4
From Teenagers	2.4	10.4	8.5	0.7	21.9
From Adults	4.6	3.1	22.4	1.8	31.8
From Seniors	0.2	0.1	0.8	1.7	2.8
Total To	28.6	16.6	49.0	5.7	

Social network data are combined with transportation data to further specify the behavior of agents. Agents are initially assigned to households in a manner consistent with census data. The transportation data are used in combination with census data to determine where the agents work or go to school. Where the patterns of movement are more diffuse, a gravity model is used that prioritizes the destinations based on a metric that weighs the basic attractiveness of a destination against the distance to be traveled to reach the destination.

An ABM of a population must also characterize how the normal pattern of behavior of agents is disturbed by environmental effects. For epidemics, the environmental effect that is modeled is the spread of a disease and the changes in behavior of individuals when they get sick. Thus the model includes as parameters disease state–transition rates and likelihoods of contacts between infected and susceptible agents at the same location.

An ABM of a population must accurately represent the response of the population to potential interventions by authorities. Targeting interventions to specific subgroups is a critical element of an effective strategy for combating an epidemic. The social network model must be disaggregated to a level that distinguishes the effects of targeted interventions. At the same time, data sources are needed to substantiate the different subgroup models.

Some workplaces, such as hotels and hospitals, have special significance for epidemic models. Hotels are a primary location for the mixing of travelers. Hospitals are significant because their workers are more likely to come in contact with infected agents.

Disease Models for ABMs

A primary element of an agent-based epidemic model is a representation of the natural history of the disease in individuals in the model, whom we refer to as agents. One common approach is to represent the history in terms of a set of transitions between states such as the Susceptible–Infectious–Removed (SIR) model. The initial conditions of an epidemic model place all of the agents into either a *susceptible* state or an immune state (which is a special case of the *removed* state). Upon contact with agents in the infectious state, susceptible agents (S) transition to the infectious state (I) with a specified probability. After a number of days, the agents transition to the *removed* state (R).

Figure 1 shows a refined version of the SIR model used for simulations. To obtain mortality statistics as outputs from the simulation, the removed state has been divided into two substates: dead and immune. Because agents in the removed state are either dead or immune to the pathogen, they remain in state R for the rest of the time period of interest. Similarly, the infectious state has been divided into symptomatic and asymptomatic agents. Asymptomatic agents do not change their behaviors when they become infected and can impact the effectiveness of surveillance and reactive interventions. A significant step in a model's calibration process is to demonstrate that the probabilities for these transitions in the simulation are consistent with historical data.

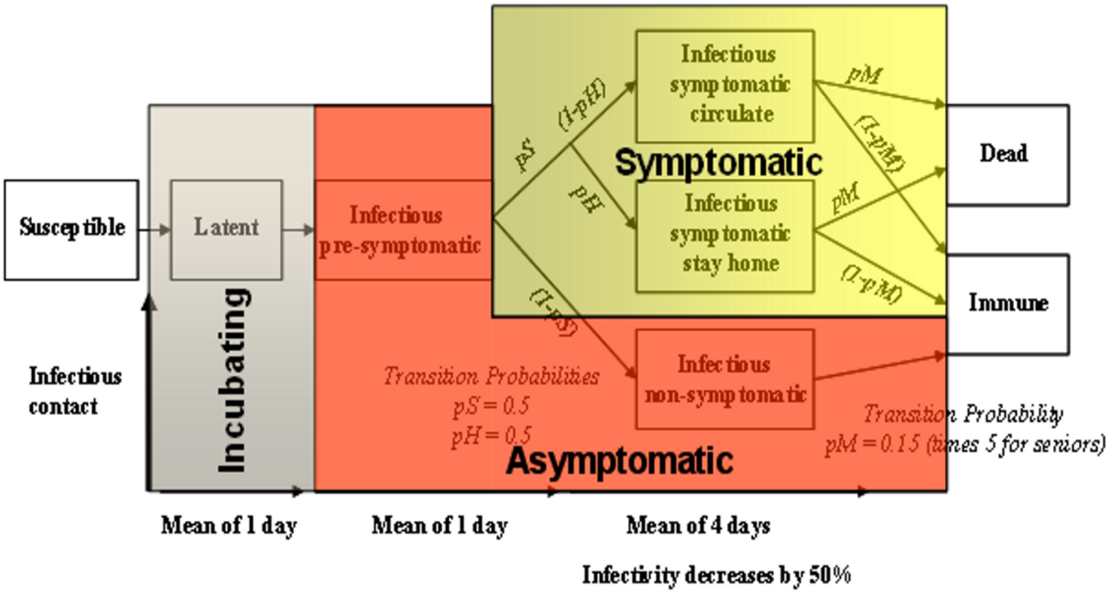

Figure 1: Influenza Disease States and Transition Rates

Several factors determine the rate of disease spread in a large scale epidemiological ABM. The parameters which characterize the disease virulence and infectiousness, the social network activity patterns, and assumptions of individual person behavior in times of disease outbreak are a few of the more important ones. When validating model output, or comparing one model's results with another's, it is essential that the modeling parameters and assumptions are consistent. Small differences in even a single parameter can lead to large differences in predictions of disease attack rates, mortality rates, etc.

All of these parameters contribute to the characteristic basic reproductive number of an infectious disease, named R_0. R_0 is defined as the number of secondary cases which one case would produce in a completely susceptible population (Deitz, 1993).

Modeling the Effects of Interventions

The primary goal for the development of these models is to assist decision makers in assessing the expected consequences of possible Courses Of Action (COA) or interventions through a series of “what if” studies. Development of these “what if” studies involves modeling a potential set of COA and ensuring that the models will distinguish the impact of the different COA. Typical interventions include vaccinations; treatment of the symptomatic population via antiviral medications; and social distancing. Social distancing methods include quarantine; closing schools, workplaces, and community centers; and travel restrictions. Social distancing methods are relevant to both non-kinetic missions and responses to epidemics. If interventions are directed at particular social networks (e.g., prophylactic treatment of children in a school), then the model must provide separate behaviors for the social group of interest.

Figure 2 (Longini et al., 2005) shows how different combinations of interventions for seven COA result in very different utilization of resources. In this case, the critical resource is the number of anti-viral treatments that are available at the time of the epidemic. The combination of school closures, other forms of social distancing, and targeted antiviral treatment is slightly more effective than the next best alternative COA.

Synthetic Populations

Many large scale human social network ABMs populate their simulations with synthetic populations. Using the population synthesizer provided by TRANSIMS (Hobeika, 2009), a synthetic population can be produced that is statistically indistinguishable from the real population for the region of interest, down to the block group level, which is the lowest level of detail provided by the US Census. Synthetic populations produced by the TRANSIMS synthesizer

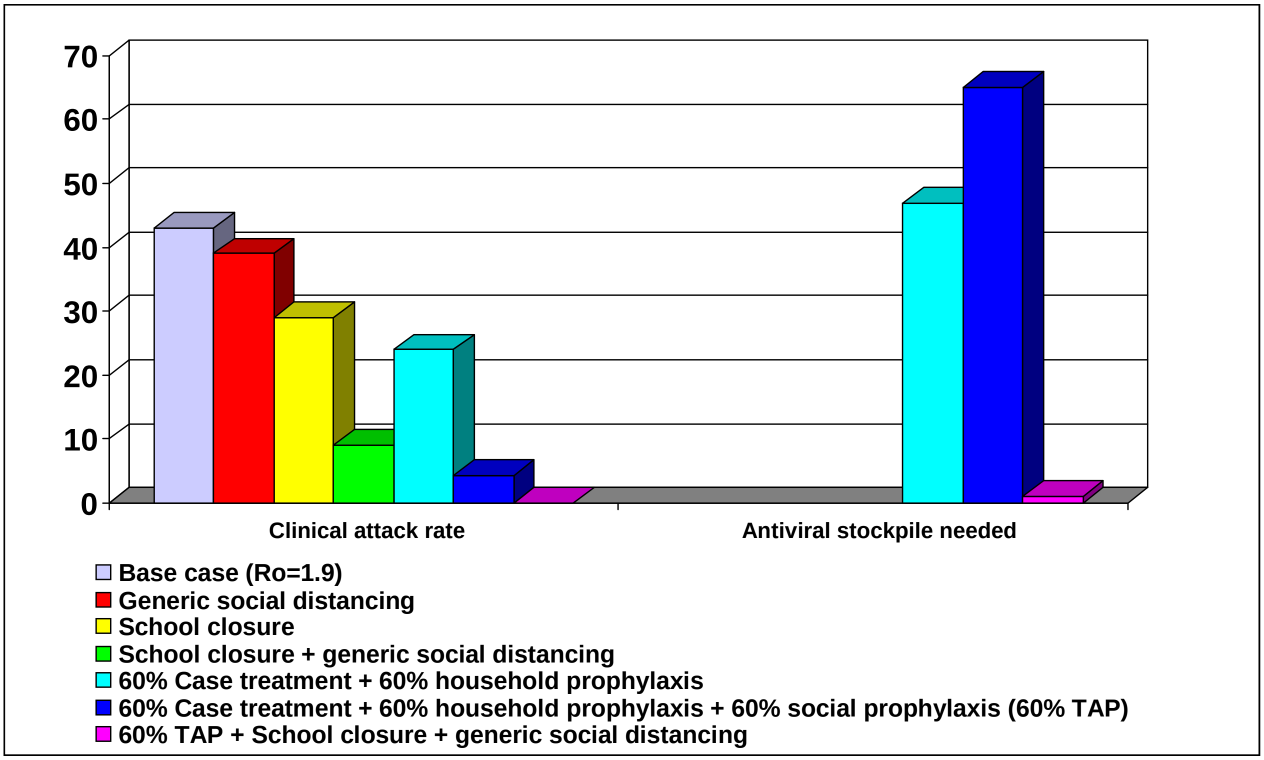

Figure 2: Effect of Different Combinations of Interventions on Consumption of Resources

provide demographic information that can be essential for validating simulation results for scenarios where an individual's characteristics are a factor regarding his response to the phenomenon being simulated. This data includes household members, household income, household member age, gender, ethnicity, education level, language spoken, occupation, travel time to work, military service, and many other variables (Wheaton, 2009).

CALIBRATION AND VALIDATION OF LARGE POPULATION MODELS

Calibration of a Model

To an extent, calibrating establishes the validity of the internal workings of the model and its results. Calibration is often used to establish the feasibility of the underlying computational model and the data (Hatchett, 2007) by showing that the model can generate results that match empirical data collected from historical events. Finding the empirical data that can be matched by simulation outputs and reflects the expected predictive simulation inputs is a challenge. Figure 3 shows simulation results for various intervention strategies for an influenza outbreak in Chicago, Il., using disease characteristics of the 1918 H1N1 influenza.

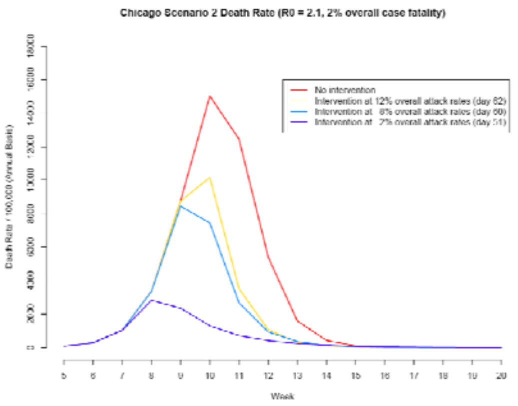

Figure 3. 1918 H1N1 Simulation Results for Modern-Day Chicago.

Tuning a model to fit detailed real data is a multi-step, often iterative, process in which the model's processes are altered so that the model's predictions come to fit, with reasonable tolerance, a set of detailed real data. Calibrating a model may require multiple simulation runs over a range of parameter values. In some cases, calibration results may require alterations to the fundamental programming, procedures, algorithms, or rules in the computational model.

Some models like EpiCast (Studt, 2007) take R_0 as an input, and vary parameters such as social network activity patterns to achieve calibration. Other models like EpiSims (Barrett, 2005) compute the value of R_0 for a given disease scenario, and vary the disease infectiousness, virulence, and incubation parameters to achieve calibration.

Validation of a Model

Schlesinger and his co-authors (Schlesinger et al., 1979) define validation as "substantiation that a computerized model within its domain of applicability possesses a satisfactory range of accuracy consistent with the intended application of the model." Zeigler (Zeigler, 1985) describes three forms of validation:

- Replicative validity: the simulation matches externally available data that has been generated by the modeled system (retrodiction).
- Predictive validity: the model matches data that can be acquired from the modeled system.
- Structural validity—the model reflects observed behavior and matches the process inherent to the process to produce the behavior. (Zeigler, 1984)

Calibration and Validation Challenges

Calibrating and validating simulations of the spread of infectious disease through large populations presents a significant challenge. Comparing simulation results to documented historical disease outbreak events is challenging for several reasons:

- Present day social network behavioral patterns are significantly different than historical patterns, and data to reproduce historical mobility and contact patterns usually does not exist. Social contact patterns are one of the primary factors which define the rate at which a disease can spread, and social distancing interventions have proven to be valuable in controlling an epidemic.

- Often, large scale simulations are run for hypothetical pandemic scenarios, such as H5N1 avian influenza. This disease has not yet migrated into the human population and still retained dangerous virulence and infectiousness characteristics. Simulations involving hypothetical disease characteristics have no real-world counterparts against which to validate.

- Human social network behavior changes in response to the presence of pandemic disease outbreak (Epstein, 2009), and it is difficult to capture those behaviors within a simulation.

ABM CALIBRATION METHODS

In this paper, we describe two aspects of calibration of a model:

- The alignment of the data driving the model with empirical information about the target population and its behaviors
- Tuning the model by adjusting parameters so that the output of the model matches historical data.

Alignment of Population Models with Demographic Data

The issue of how to validate civilian response to this scenario is complex.

One approach is to treat all working members of the population as equals, and then perform a parameter sweep simulating the impact on disease spread of varying the percentage of randomly-selected working members of the population who ignore the "stay at home" directive. Another, more detailed approach, will take population demographics into account.

For this scenario, we will use two different households to illustrate how civilian behavior can be a major factor regarding simulation results validation. The purpose of describing the two households is to demonstrate how household demographics can be a major factor in civilian behavior in times of pandemic disease outbreak.

Household 1

- Two household members, adult, married, no children, both with advanced university degrees.
- Age of male: 34
- Age of female: 32
- Household income: $167,000 per year.

Household 2

- The head of the household is a single unmarried woman, age 20, high school graduate.
- The woman has one daughter, age 4 living with her at home.
- Household income: $17,000 per year.

Calibration of the Model with Historical Data

Some useful data can be mined from existing historical data, like the 1918 "Spanish Flu" outbreak (Gottfredson, 2007). However, the social network mobility patterns in 1918 were sufficiently different from modern day patterns that it is not possible to completely validate model results using historical data.

ABM VALIDATION METHODS

Validation Using Multiple Models

One approach to validating large-scale social network ABMs is to run multiple, different models on the same problem. In fact, this approach is the only way to validate model results for problems of a certain class, e.g. simulations of pandemic disease outbreak in modern social networks. For such simulations, there is no historical data against which model results can be validated. This is the approach that was used in a recent study that was conducted to quantify the effects of various intervention studies in the event of a pandemic influenza outbreak in the United States. H5N1 influenza was the strain selected for the study. Since this strain has not yet migrated into the human population, it was necessary to make assumptions about its virulence and infectiousness. For the purpose of the study, it was assumed that the simulated H5N1 virus would have the same general characteristics as the 1918 strain.

To compare the three models, an urban region approximately the size of Chicago, ~8.6 million people, was selected (Halloran, 2007). One of the three models used was produced by a collaboration of investigators at the University of Washington and Fred Hutchinson Cancer Research Center in Seattle and the Los Alamos National Laboratories. This was the "EpiCast" model. A second model was produced by a collaboration of investigators at Imperial College and the University of Pittsburgh. This was the "Imperial" model. The third model was provided by Virginia Bioinformatics Institute of the Virginia Polytechnical Institute and State University. This was the "EpiSims" model.

As can be seen from Figure 4, the three models sometimes produced significantly different results for several analysis scenarios. The validation process used on this study was to examine the differences in assumptions and granularity implemented by the three models for the various scenarios that were run. Once the differences were identified, and demonstrated to be the source of variance in results from the other models, that part of the study was considered to have been validated.

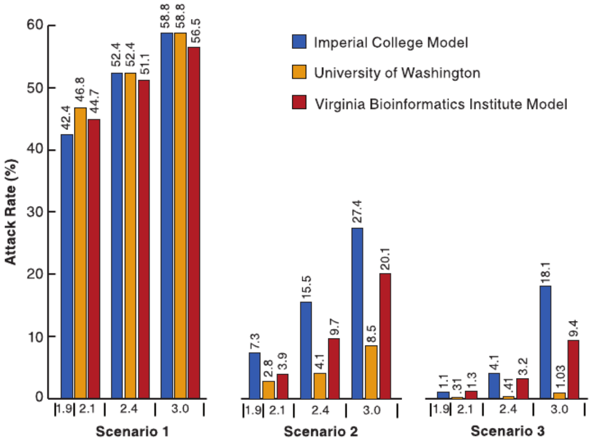

Figure 4. Results From Three Different Models

CONCLUSIONS

By their nature, large agent based social network simulations are difficult to validate. The models are often used to predict an outcome for a physical system that does not yet exist, and therefore no "ground truth" data is available against which to compare the simulation results. Validation options in these cases typically involve usage of multiple models and rigorously examining all modeling assumptions.

Validating large social network models against historical data is problematic because social network interaction patterns have changed significantly throughout recorded history, and fine-grained population mobility data is only now becoming available. As computing power approaches the petascale and large social network ABMs grow to utilize the additional computational resources, model validation will play an increasingly important role in analyzing large, complex systems.

ACKNOWLEDGEMENTS

We thank the National Institute of General Medical Sciences MIDAS Program for research funding (Grant #5U01GM70698-5).

REFERENCES

Barrett, C., Eubank, S. Smith, J. March, 2005. "If Smallpox Strikes Portland... ." Scientific American.

Cooley P, Ganapathi L, Ghneim G, et al. Using influenza-like illness data to reconstruct an influenza outbreak. Mathematical and Computer Modelling. 2008;48(5,6):929–939.

Deitz, 1993. "The estimation of the basic reproductive number for infectious diseases." Statistical Methods in Medical Research, Vol. 2, No. 1, 23-41 (1993)

DOI: 10.1177/096228029300200103

Epstein, J., Parker, J., Cummings, D., Hammond, R. "Coupled Contagion Dynamics of Fear and Disease: Mathematical and Computational Explorations". Mathematical and Computational

Explorations. PLoS ONE 3(12): e3955. doi:10.1371/journal.pone.0003955.

Eubank, S. December 2005. "Network Based Models of Infectious Disease Spread." Japanese Journal of Infectious Diseases 58(6):S9-13.

Ferguson N, Cummings DAT, Cauchemez S, et al. Strategies for containing an emerging influenza pandemic in Southeast Asia. Nature. 2005;437:209–214.

Ferguson N, Cummings DAT, Fraser C, et al. Strategies for mitigating an influenza pandemic. Nature. 2006;442:448–452.

Germann TC, Kadau K, Longini IM Jr., et al. Mitigation strategies for pandemic influenza in the United States. PNAS. 2006;103(15):5935–5940.

Glass Local mitigation strategies for pandemic influenza. NISAC, SAND Number: 2005-7955J

Gottfredson, 2007

Halloran ME, Ferguson NM, Eubank S, et al. Modeling targeted layered containment of an influenza pandemic in the USA. *PNAS*. 2008;105(12):4639–4644.

Hatchett R, Mecher CE, Lipsitch M. May 2007. Public health interventions and epidemic intensity during the 1918 influenza pandemic. NCI PubMed. PMID: 17416679

Hobeika, 2009. May, 2009. "TRANSIMS - the next generation planning/simulation model." Transactions of the Wessex Institute. 10.2495/UT040151.

Longini I Jr, Nizam A, Xu S, et al. Containing pandemic influenza at the source. Science. 2005;309:1083–1087.

Schlesinger, S., Crosbie, R.E., Gagne, R.E., Innis, G.S., Lalwani, C.S., Loch, J., Sylvester, R.J., Wright, R.D., Kheir, N., and Bartos, D 1979, Terminology for Model Credibility, Simulation; 34 (3) 103-104

Studt, July, 2007. “Modeling in the life sciences pushes into new areas: life science modeling and simulation have been in use for decades, but new areas continue to be developed that promise to accelerate research.(TEST, MEASUREMENT & ANALYSIS)(EpiCast and hardware-based artificial immune system)”. Advantage Business Media.

Wheaton, W.D., Cajka, J.C., Chasteen, B.M., Wagener, D.K., Cooley, P.C., Ganapathi, L., Roberts, D.J., and Allpress, J.L. (2009). Synthesized Population Databases: A US Geospatial Database for Agent-Based Models. RTI Press publication No. MR-0010-0905. Research Triangle Park, NC: RTI International.

B.P. Ziegler, Theory of Modeling and Simulation, Krieger, 1985, Malabar.