ABSTRACT

Title of dissertation: AN AGENT-BASED MODELING APPROACH TO REDUCING PATHOGENIC TRANSMISSION IN MEDICAL FACILITIES AND COMMUNITY POPULATIONS

Sean Louis Barnes, Doctor of Philosophy, 2012

Dissertation directed by: Professor Bruce Golden
Robert H. Smith School of Business

The spread of infectious diseases is a significant and ongoing problem in human populations. In hospitals, the cost of patients acquiring infections causes many downstream effects, including longer lengths of stay for patients, higher costs, and unexpected fatalities. Outbreaks in community populations cause more significant problems because they stress the medical facilities that need to accommodate large numbers of infected patients, and they can lead to the closing of schools and businesses. In addition, epidemics often require logistical considerations such as where to locate clinics or how to optimize the distribution of vaccinations and food supplies.

Traditionally, mathematical modeling is used to explore transmission dynamics and evaluate potential infection control measures. This methodology, although simple to implement and computationally efficient, has several shortcomings that prevent it from adequately representing some of the most critical aspects of disease transmission. Specifically, mathematical modeling can only represent groups of in-

dividuals in a homogenous manner and cannot model how transmission is affected by the behavior of individuals and the structure of their interactions.

Agent-based modeling and social network analysis are two increasingly popular methods that are well-suited to modeling the spread of infectious diseases. Together, they can be used to model individuals with unique characteristics, behavior, and levels of interaction with other individuals. These advantages enable a more realistic representation of transmission dynamics and a much greater ability to provide insight to questions of interest for infection control practitioners.

This dissertation presents several agent-based models and network models of the transmission of infectious diseases at scales ranging from hospitals to networks of medical facilities and community populations. By employing these methods, we can explore how the behavior of individual healthcare workers and the structure of a network of patients or healthcare facilities can affect the rate and extent of hospital-acquired infections. After the transmission dynamics are properly characterized, we can then attempt to differentiate between different types of transmission and assess the effectiveness of infection control measures.

AN AGENT-BASED MODELING APPROACH TO REDUCING PATHOGENIC TRANSMISSION IN MEDICAL FACILITIES AND COMMUNITY POPULATIONS

by

Sean Louis Barnes

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2012

Advisory Committee:
Professor Bruce Golden, Chair
Professor Radu Balan
Professor Anthony Harris
Professor Paul Schonfeld
Professor Paul Smith
Professor Edward Wasil

Dedication

This dissertation is dedicated to the youth who march onward and upward. Pursue your dreams with everlasting fervor.

Acknowledgments

First and foremost, I thank my wife Nadya and our families for their enduring support, which enabled me to pursue this degree with the necessary level of dedication to be successful. I also thank my parents, sister, and grandparents for setting an example for me to pursue the highest level of education and supporting me through all of my endeavors. Finally, I thank all of the friends, co-workers, professors, and administrators that have provided even the slightest level of encouragement, support, and guidance throughout my education.

This dissertation would not have been possible without the enthusiasm, support, and direction of my advisor, Dr. Bruce Golden. I owe a debt of gratitude to Dr. Edward Wasil, who provided a great deal of valuable insight into the various projects we worked on together. We had many productive discussions, and his very careful editing has helped me communicate my research more clearly. I am also grateful to Dr. Paul Smith, Dr. Radu Balan, Dr. Paul Schonfeld, and Dr. Anthony Harris for serving on my dissertation committee.

I thank Dr. James Yorke for accepting me into the doctoral program and speaking with me on several occasions while I deliberated over the decision. I thank Dr. Konstantina Trivisa for her support as my program chair. I thank Alverda McCoy for helping me with all of my administrative needs over the years and hosting me on my initial visit to the University. I thank Dr. Harris again, along with Dr. JJ Furuno and Dr. Eli Perencevich for sharing their time and expertise during our collaborative work at the University of Maryland Medical Center. I thank Dr. Inbal

Yahav-Shenberger for initiating an enjoyable and productive collaboration on my final graduate research project and contributing a significant portion of Chapter 6. I thank Dr. Wolfgang Jank for helpful discussions related to factorial design and regression analysis. I thank the Department of Mathematics and the Graduate School for supporting my travel expenses to various conferences. I thank Dr. Aleksey Zimin for providing access to the Genome computing cluster at the University of Maryland, as well as his guidance along with Dr. Radu Balan during the Applied Mathematics and Scientific Computation project course. I thank fellow graduate students David Anderson, Stuart Price, and Max Scharrenbroich for their work on several research and course projects. I thank Dr. Chris Groer, Dr. Carter Price, and Dr. Damon Gulczynksi for their advice throughout different stages of my program.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ABM	Agent-Based Modeling
$ABMS$	Agent-Based Modeling and Simulation
CDC	Centers for Disease Control and Prevention
DES	Discrete Event Simulation
FDA	Functional Data Analysis
$fPCA$	Functional Principal Components Analysis
HCW	Healthcare Worker
$LTCF$	Long-Term Care Facility
$MRSA$	Methicillin-resistant *Staphylococcus aureus*
PCR	Polymerase-Chain Reaction
PCA	Principal Component Analysis
SD	Systems Dynamics
SIR	Susceptible-infected-recovered
SNA	Social Network Analysis
VRE	Vancomycin-resistant *Enterococcus*

Chapter 1

Introduction

1.1 Pathogenic Transmission in Medical Facilities and Community Populations

The spread of infectious diseases is a significant and ongoing problem in human populations. For centuries, societies have fallen victim to various diseases that have spread throughout their population. Some of the most infamous epidemics are the Black Death (14^{th} century) and the Great Influenza Pandemic (1918). More recently, HIV/AIDS, SARS, and multiple outbreaks of influenza have demonstrated how vulnerable our increasingly connected world is to the transmission of infectious diseases. In addition, the threat of bioterror attacks has grown, and the proper measures for differentiating attacks from naturally-occurring epidemics are currently not in place. Outbreaks in community populations–whether due to an infectious disease or a biotteror attack–cause significant problems because they stress the medical facilities that need to accommodate large numbers of infected patients, and they can lead to the closing of schools and businesses. In addition, epidemics often require logistical considerations such as where to locate clinics or how to optimize the distribution of vaccinations and food supplies. Historically, the best defense for combating disease outbreaks has been vaccination and quarantine. However, due to

resource constraints and economical considerations, alternative non-pharmaceutical solutions are desired.

The problem of preventing hospital-acquired infections is another well-publicized problem and study of the best measures to control them has been extensive [24]. These types of infections cause many downstream effects in hospitals, including longer lengths of stay for patients, higher costs, and unexpected fatalities on the order of almost 100,000 deaths each year in the U.S. alone [74]. Hospitals have become increasingly more vulnerable to various types of infection, most frequently to methicillin-resistant *Staphylococcus aureus* (MRSA), vancomycin-resistant *Enterococcus* (VRE), *Clostridium difficile* (C. diff), and Acinetobacter. These pathogens are often carried asymptomatically by patients, which leads to undetected transmission to other patients through the transiently colonized hands of healthcare workers.

Computer modeling and simulation are decision-aiding tools that can be leveraged to assist infection control professionals in developing strategies for reducing or preventing the transmission of infectious diseases [27]. Control measures are often expensive in terms of time, money, and resources. Computational models can help to alleviate some of the risk in implementing intervention strategies. Studies of this problem have not focused on modeling the interactions between the critical actors in a given scenario, whether it is patients, nurses, and physicians in a hospital, individuals in a community, or medical facilities in a particular region.

This dissertation focuses on developing agent-based models that use interactions at both the individual level and the facility level to better understand transmission dynamics and to evaluate strategies for intervention. By using these methods,

we can explore how the behavior of individual healthcare workers and the structure of a network of patients or healthcare facilities can affect the rate and extent of transmission. After the transmission dynamics are properly characterized, we can attempt to differentiate between different types of transmission and assess the effectiveness of infection control measures.

1.2 Agent-Based Modeling and Simulation

Historically, modeling the transmission of infectious diseases has focused on results derived from systems dynamics (SD) and discrete event simulation (DES) methods [30, 44, 51]. These techniques can provide valuable insight, and are ideally suited for many problems. Both methods typically focus on system-level behavior, but they differ in how the system is modeled and how time is simulated. SD models typically represent entities as continuous variables whose states change continuously with time, whereas DES models use individual components whose states only change at discrete moments in time. In either case, the goal is to aggregate the system behavior and draw conclusions on how the system evolves over time under internal and external forces, such as a change of an internal policy or a surge in the demand of a particular service.

Recently, a new modeling and simulation methodology has gained momentum with respect to healthcare applications. The methodology is most commonly known as agent-based, or individual-based, modeling (ABM). In contrast to SD and DES methodologies, ABM focuses on modeling individuals, interactions between indi-

viduals, and, in some cases, interactions with a physical or influential surrounding environment [53]. These activities can then be aggregated to simulate how a system behaves over time. The focus on individual agents and their interactions makes ABM an ideal tool for analyzing complex systems such as healthcare facilities and community populations, because there are many components to these systems and outcomes can be difficult to predict without adequate model representation. A comparison between the three modeling frameworks is shown in Table 1.1.

Agents can interact with each other in many ways. Interactions may occur in a spatial environment, which could be a simple one-dimensional circular network, a two- or three-dimensional Cartesian grid, or a specific geographic location or region. Interactions can also be aspatial and constrained by relational considerations, where agents only interact if there is an explicit connection between them, such as being a member of the same family, working together, or being cared for by the same healthcare worker. For these cases, physical space has no effect on the outcome of the interactions, and therefore it does not need to be modeled explicitly. Social network analysis (SNA) is often paired with ABM for problems in which the structure of the interaction network is not uniform, implying that each individual may interact with only a specific subset of the population. For these types of problems, agents are often represented by nodes in a network and interactions are represented by edges, which can be weighted to represent the frequency or type of interactions.

There are several advantages of ABM over SD and DES. First, ABM is a more realistic modeling approach for many problems, especially problems in which there are multiple types of actors that interact in different ways. In these cases, it is very

Table 1.1: Comparison of systems dynamics (SD), discrete event simulation (DES), and agent-based modeling (ABM) methodologies.

Attribute	**SD**	**DES**	**ABM**
Model Perspective	System level	System level	Individual level
Level of Realism	Low	Moderate	High
Model Flexibility	Low	Moderate	Very high
Time Domain	Continuous	Discrete	Discrete
Run Times	Very fast	Moderate	Slows with increasing system size and complexity
Model Inputs	Rate parameters and flow characteristics	Entity types, arrival times, queuing parameters, resource scheduling, process flows	Agent characteristics and interactions, environment specification
Model Outputs	Dynamics, steady state values, analytic expressions	Wait times, resource utilization, throughput	Unlimited
Software	Spreadsheet (e.g., Microsoft Excel), any programming language, mathematical software (e.g., MATLAB[1], Mathematica[2]	Object-oriented programming languages (e.g., C++, Python[3], Java), commercial simulation packages (e.g., Arena[4], AnyLogic[5], SIMUL8[6]), mathematical software (e.g., MATLAB, Mathematica)	Object-oriented programming language (e.g., C++, Python, Java), open source ABM software (e.g., NetLogo[7], Repast[8], MASON[9], Swarm[10])

[1] http://www.mathworks.com/products/matlab/
[2] http://www.wolfram.com/mathematica/
[3] http://www.python.org
[4] http://www.arenasimulation.com/
[5] http://www.xjtek.com/
[6] http://www.simul8.com/
[7] http://ccl.northwestern.edu/netlogo/
[8] http://repast.sourceforge.net/
[9] http://cs.gmu.edu/ eclab/projects/mason/
[10] http://www.swarm.org/index.php/Main_Page

straightforward to model these actors as agents that have distinct sets of behaviors and characteristics without making assumptions as to how the system would be affected by each type. Individuals are not represented in SD models. In DES models, individuals are explicitly modeled, but their states are a function of their status in some type of pre-defined system process. In addition, agent-based models facilitate detailed analysis of both individual- and system-level behavior, because metrics at each level can be updated with each interaction. Agent-based models are easier to explain than most SD and DES models because of their direct correlation to reality, which is an important factor in gaining the confidence of healthcare professionals and ultimately having an impact. Oftentimes, SD models are mathematical models that can become quite complicated. Usually, DES models are described by intricate flow diagrams. These abstractions can cause difficulty in explaining model concepts to healthcare professionals who are not trained in mathematical or computational modeling.

As in all methodologies, there are disadvantages to ABM as well. These models can become very complex when they begin to represent a high level of detail. When this happens, it becomes difficult to separate the actual effect of each input parameter in the model. In addition, agent-based models can become very computationally expensive, which requires excessively long computer run times for simulations. This problem has been alleviated to some degree by high-performance and parallel computing techniques, but it demands additional developmental resources that are not required by SD or DES models. Agent-based models face different challenges than SD and DES model concerning the underlying assumptions. SD and DES models

require assumptions about system-level parameters, whereas ABM requires assumptions about individual-level parameters and the nature of interactions. Some of these requirements can be satisfied easily, by collaborating with experts who have experience working in these environments. Other requirements can be more difficult to quantify. However, the advantage of ABM is that few assumptions need to be made about the system as a whole, because the system response is determined by the activities at the individual level.

Agents can have several characteristics that can be used to distinguish ABM from SD or DES models. These characteristics and their definitions are summarized in Table 1.2. There are other characteristics that are associated with agent-based models, but the set described in Table 1.2 consists of the most prevalent features. Agent-based models do not necessarily possess all of these characteristics, and, in some cases, the distinction between ABM and DES can be difficult.

In the following section, we review and evaluate a selected body of research that has applied agent-based modeling techniques to characterizing the transmission of infectious diseases in medical facilities and community populations. We review the agent-based modeling literature in this research area, briefly describe the core methods, summarize the key results, and identify best practices. We highlight areas where agent-based modeling and simulation fill a significant gap that has not been addressed by other methods. Finally, we provide some new questions that may be of interest to healthcare researchers and practictioners.

Table 1.2: Agent characteristics and definitions

Characteristic	**Definition**
Autonomy	Agents act independently of other agents
Heterogeneity	Agent characteristics and evolution of state are sufficiently different for all agents
Awareness	Agents can have varying levels of knowledge of the system state, ranging from ignorance to omniscience
Memory	An agent can remember its state and/or the state of the system at earlier points in time
Adaptation	Agents can change their behavior over time based on the current state of the system or prior experience
Goal Oriented	Agent actions are aimed at accomplishing an objective
Rationality	Agent actions are aimed toward their best interests
Interactivity	Agents can exchange information or resources with other agents
Reactivity	Agent state or behavior can change in reaction to the environment or changes in the behavior of other agents
Mobility	Agents can move within the environment

1.3 Epidemiological Models

The field of epidemiology is concerned with how disease spreads in a population. Studies that model the transmission of infectious diseases can focus on increasingly large scales, ranging from single units or wards to entire hospitals, communities, cities, and global pandemic scales. Vector-borne diseases, or those that are transmitted by way of an intermediate carrier, are the most commonly modeled types. Airborne, waterborne, food-borne, respiratory, and sexually transmitted diseases can be modeled as well. These diseases can be modeled with different degrees of specificity to include incubation periods and periods where a colonized individual is infectious but asymptomatic. Understanding how diseases are transmitted and determining the best ways to control transmission are critical to preventing excessive spread. Epidemic modeling is an ideal method for experimenting with various control strategies.

Historically, epidemiological models are mathematical, or compartmental, models that predict how proportions of the population in each state evolve over time. The susceptible-infected-recovered (SIR) model [48] is the most well-known of these models, which forms the foundation for many other compartmental models [4, 5, 10, 11, 19, 54, 69, 76]. The SIR model equations are shown in Equations 1.1, 1.2, and 1.3, where S, I, and R represent proportions of the population that are in the susceptible, infected, and recovered states, respectively. β and γ are the transmission and recovery rates. These equations can be integrated over time to generate population transmission dynamics. They provide interesting results for both deterministic

and stochastic scenarios. An important measure for this type of model is the basic reproduction number, R_0, which is the average number of secondary infections (i.e., transmissions) per primary case in an entirely susceptible population. R_0 is a key metric in predicting the extent to which an infection is likely to spread. If $R_0 > 1$, then an epidemic is likely to grow because, on average, each infected person transmits the disease to more than one other person. If $R_0 < 1$, then an epidemic is likely to become extinct at some point in the future. Some models use R_0 as an input to drive transmission within a population, whereas other models use a transmission rate or probability parameter and calculate the R_0 value for the population based on the number of primary and secondary infected individuals. These models have provided significant insight into the effects of certain parameters on transmission dynamics and the effectiveness of various infection control measures.

$$\frac{dS}{dt} = -\beta SI \tag{1.1}$$

$$\frac{dI}{dt} = \beta SI - \gamma I \tag{1.2}$$

$$\frac{dR}{dt} = \gamma I \tag{1.3}$$

Mathematical models have several assumptions and limitations that prevent them from producing more valuable results. The first key assumption is that populations modeled by mathematical equations are well-mixed, meaning that all individuals within the population interact with equal probability. For example, the SI term in the SIR model equations (e.g., Equations 1.1 and 1.2) shows that transmission is proportional to the interaction between all susceptible and infected individuals, rather than a specific subset. All individuals within each compartment are assumed

to be homogeneous, which prevents analysis of how mixed or extreme behavior, such as individuals that are more likely to spread a disease (i.e., superspreaders) or healthcare workers (HCWs) that are less likely to wash their hands, can affect transmission dynamics. It can be difficult to implement time-varying or conditional behavior in a mathematical model, and thus analysis of control measures is often performed by simply varying input parameters without modeling the interactions involved in a particular intervention. Some variables, such as handwashing probabilities and screening test return times, can be approximated with simple parameters. Others, such as isolation or the effect of staffing ratios are more difficult to model using differential equations. Several advances in mathematical modeling have been made in recent years to account for these limitations, but these results are only valid for particular scenarios [6].

Agent-based, or computational, disease spread models have extended the research established by mathematical models, and have addressed many of their limitations. They have reinforced many conclusions from mathematical and discrete event simulation models and have provided additional detail about the nature of transmission. The key advantage of agent-based models is that they simulate the interactions that serve as the primary mechanism for transmission and they are capable of implementing many infection control measures explicitly. In addition, agent-based models are more adept at simulating stochastic effects, which must be captured when modeling heterogeneous populations. As a result, they have contributed significantly to a better understanding of epidemics.

The articles reviewed in this section fall mainly into two categories: process-

oriented models of transmission and network models of transmission. Process-oriented models simulate agents that are moving through a series of stages before being removed from the population, much like the movement in a DES model. Network models simulate the spread of disease between individuals using relational connections as the primary mechanism of transmission. Intermediate carriers such as HCWs do not need to be modeled explicitly in this type of model because the connections among individuals are explicit, unlike in a process-oriented model. Exemplary methods and contributions for both types of models are summarized in the following subsections.

1.3.1 Process-Oriented Models of Transmission

Process-oriented models of transmission are the most natural application of ABM to epidemiology. A common example is a simulation of patient-to-patient transmission in a hospital, in which patients are admitted, visited by HCWs, and discharged (see Figure 1.1 for a sample patient flow diagram). What separates process-oriented models from traditional DES models is that there are often multiple patient and HCW types, and their behavior is often dynamic. Transmission typically occurs through HCWs, who spread an infection from one patient to another because they fail to wash their hands adequately. These types of transmissions, from an already infected (i.e., primary) patient to a newly infected (i.e., secondary) patient, are known as hospital-acquired, or nosocomial, infections. Hospital intensive care units (ICUs) are commonly modeled in an agent-based framework. ABM

is ideally suited for this type of model because populations are small and diverse, the patients are more susceptible to infection than in other hospital units, contacts between patients and HCWs are frequent and intimate, and stochastic effects are of considerable importance.

Within agent-based models, there is often an increased ability to track various simulation data and provide additional insight into the transmission dynamics. These models can evaluate the effectiveness of various infection control measures in order to offer recommendations of which measure or bundle of measures should be implemented. Typical control measures that are modeled include the hand washing behavior of HCWs, diagnostic screening of patients, isolation of infected patients, vaccination, and decolonization, in which colonized patients undergo a therapeutic process that negates their ability to infect others.

There is a set of pathogens that are commonly modeled in these types of simulations, and many of these are resistant to antibiotic treatments. MRSA and VRE are the most common of the modeled pathogens. Patients are colonized with MRSA and VRE prior to developing an infection. This scenario is particularly difficult because colonized patients are often asymptomatic and, therefore, they can only become identified by using active surveillance techniques such as diagnostic screening. Consequently, these patients can spread the pathogen to HCWs and ultimately other patients before any intervention is started. In addition, treatment for these resistant organisms is often difficult, therefore protecting patients from acquisition is the most effective approach for ensuring their safety.

The articles discussed in this section all follow a similar pattern for modeling

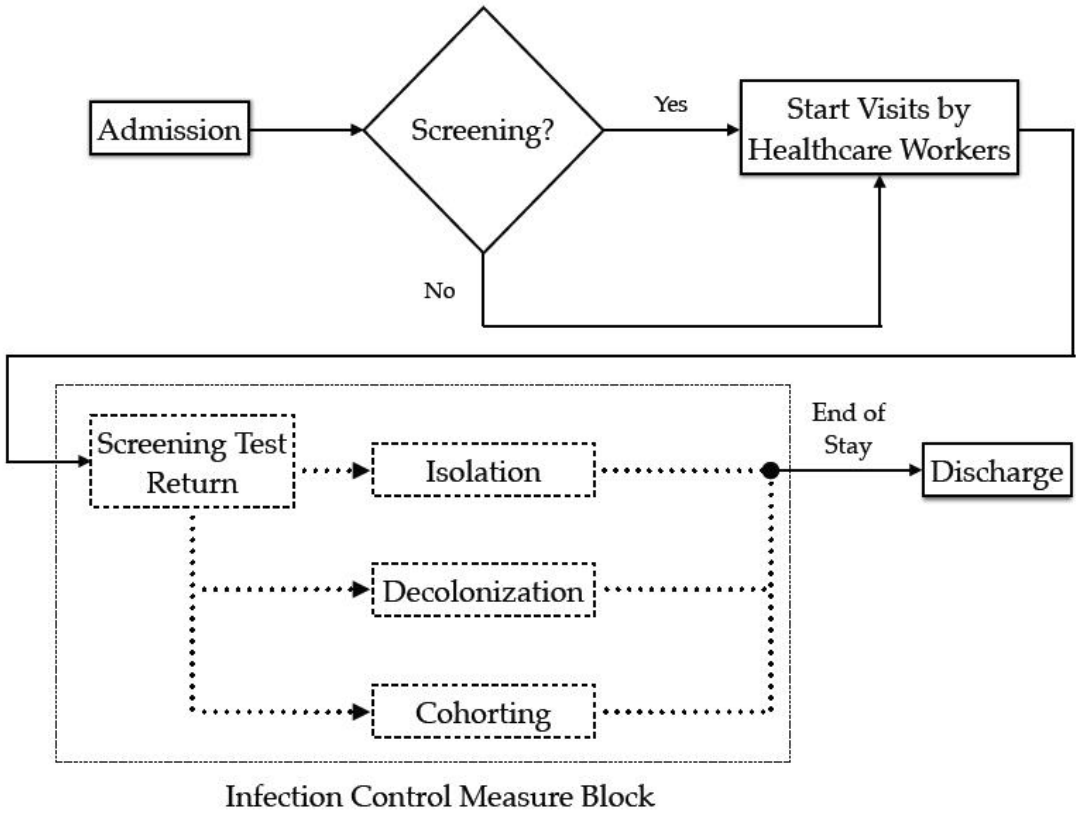


Figure 1.1: Sample patient flow diagram of a disease spread model in a hospital. Optional infection control measures are indicated by dashed lines and dotted arrows.

transmission of an infectious disease in a hospital. Many models incorporate resistant pathogens and assess the potential effectiveness of specific control measures. However, these models differ slightly in their specific sets of experimental parameters and the nature of their results. These differences are summarized in Table 1.3.

We begin with an agent-based model that highlights the specific level of detail afforded by ABM that is not possible using SD or DES methods. Hotchkiss et al. [42] tested the effects of several factors and infection control measures on transmission. The authors were able to demonstrate that early detection and subsequent isolation of infected patients, quick patient turnover, cohorting patients, and limiting the frequency of physician visits could all reduce the likelihood of a significant outbreak.

In addition to evaluating infection control measures and determining the most influential external factors, ABM can provide more realistic transmission dynamics by introducing more model complexity. This additional detail not only increases the relevancy and strength of simulation results, but it can be used to address the concerns of healthcare professionals who are skeptical of the model validity. Ong et al. [66] implemented a spatially explicit agent-based model of influenza transmission in a hospital unit that represented several types of HCWs, including physicians, nurses, health attendants, clerks, and cleaners. Ambulant and non-ambulant patients were also modeled. Transmission of influenza is airborne, therefore, additional model considerations must be taken into account because agents can transmit the disease without coming into direct contact with each other. Results directly related to transmission were limited, but the model generated a reasonable distribution of contacts

Table 1.3: Summary of process-oriented model parameters

Model	D'Agata	Hotchkiss	Meng	Ong	Temime
Software Platform	MATLAB	Mathematica	AnyLogic	PathoSim[1]	Java
Spatial representation	No	Yes	Yes	Yes	Yes
HCW Types	1	2	0	5	3
HCW shifts	Yes	Yes	No	Yes	Yes
Variable hand hygiene	*	*	No	No	Yes
Hygiene efficacy	*	*	No	No	Yes
Patient screening	No	Yes	Yes	No	No
Patient isolation	No	Yes	Yes	No	No
Decolonization	Yes	No	Yes	No	No
Variable staffing ratios	No	Yes	No	Yes	Yes
Patient cohorting	Yes	Yes	No	No	Yes
Variable transmissibility	No	Yes	Yes	Yes	Yes
Patients colonized or infected on admission	No	Yes	Yes	No	Yes
Variable # of patient visits	Yes	Yes	No	Yes	Yes
Variable patient lengths of stay	Yes	Yes	Yes	No	No
Bacterial load	Yes	No	No	No	No
Antibiotic resistance	Yes	No	No	No	No
Visitors	No	No	No	Yes	No

[1] http://www.ross-scientific.com/products.htm

* Variable hand hygiene and hygiene efficacy were not implemented explicitly, but transient HCW colonization times were variable and followed an exponential distribution

between all agents in the unit. This distribution could be used as a basis for modeling transmission between any pair of agents in future work. The model by Meng et al. [57] incorporated multiple routes of transmission and variable transmission rates between patients, but it produced limited results related to actual dynamics. Temime et al. [82] describe an agent-based model of pathogenic transmission in a hospital, but the implications of these results are more appropriately addressed in the following section on network models of transmission.

Agent-based models can be used to address questions related to antibiotic resistance, which is an important issue in disease control. Several models address the implications of antibiotic resistance at the microbiological level (they are not covered in this dissertation). D'Agata et al. [25] focused on the effects of antibiotic resistance on transmission and the competition between resistant pathogens in a hospital. Single patient and HCW types were modeled explicitly, with each having eligible states of being susceptible to or colonized with resistant and/or non-resistant pathogens. The authors demonstrated that initializing decolonization treatments on patients quickly and for shorter durations can eliminate both resistant and non-resistant strains from the population. The authors were able to develop a corresponding differential equation model that facilitated model validation and additional analysis.

1.3.2 Network Models of Transmission

Network models of transmission provide a different perspective for analyzing the spread of infectious diseases. They are an abstract representation of the physical

interactions that can lead to transmission, in contrast with the direct correlation to reality afforded by process-oriented transmission models. The edges, or connections, between nodes in the network can represent a relationship between patients in a hospital or between individuals in a community. The structure or distribution of these connections has a significant effect on transmission dynamics [47, 71], and different network structures can be designed to represent various scenarios. Whereas process-oriented models can identify the best interventions, network models can provide insight as to where those interventions should be directed, such as targeting individuals for vaccinations or closing schools or hospital wards.

There are several common structures for interaction networks. They are characterized by the frequency distribution of node connections in the network, also known as the degree distribution [1]. Examples of each network type and the corresponding degree distributions are shown in Figure 1.2. Regular networks have all nodes with the same degree. Edges can be structured (e.g., nodes are connected to their nearest k neighbors) or randomly distributed to other nodes in the network. Random networks are generated by assigning an equal probability to each potential edge in the network. Each edge is then chosen at random based on the specified probability, which forms a network in which the degree distribution follows a binomial model. Small-world networks [85] fall somewhere in between regular and random networks by re-wiring (i.e., redirecting an edge from one node to another at random) a certain proportion of edges in a structured, regular network. Nodes in this type of network are still highly clustered, but disease could spread more quickly because there are shortcuts to other, highly susceptible parts of the network. A

random network is a special case of a small-world network that has re-wired all of its edges. A special class of networks are exponential networks, whose degree distribution follows a negative exponential trend. These networks have been found to be the most realistic structure for social interaction networks [6]. Finally, a scale-free network has a power law degree distribution, in that there are a few nodes with a large number of connections and many nodes with a small number of connections. These networks are called scale-free because the average distance between any two nodes increases very slowly as the number of nodes increases. Disease transmission through these types of networks is likely to find the highly connected nodes quickly. However, transmission to the remaining population is likely to take much longer because there are fewer paths to nodes on the periphery of the network.

Mathematical models inherently assume that populations are well-mixed, which means each individual has an equal probability of interacting with all other individuals. In general, this assumption corresponds most closely to a regular network, whether the connections are structured or random. For certain applications, this configuration could be appropriate if each individual has approximately the same number of social contacts (e.g., child daycare). In other cases, a small-world, exponential, or scale-free network is a more accurate representation because there can be individuals who have connections to several population subgroups. These highly connected individuals often have the greatest effect on transmission. Neglecting to model them explicitly can have a significant effect on the results that are ultimately communicated to healthcare organizations.

The first set of models investigates how direct transmission can occur between

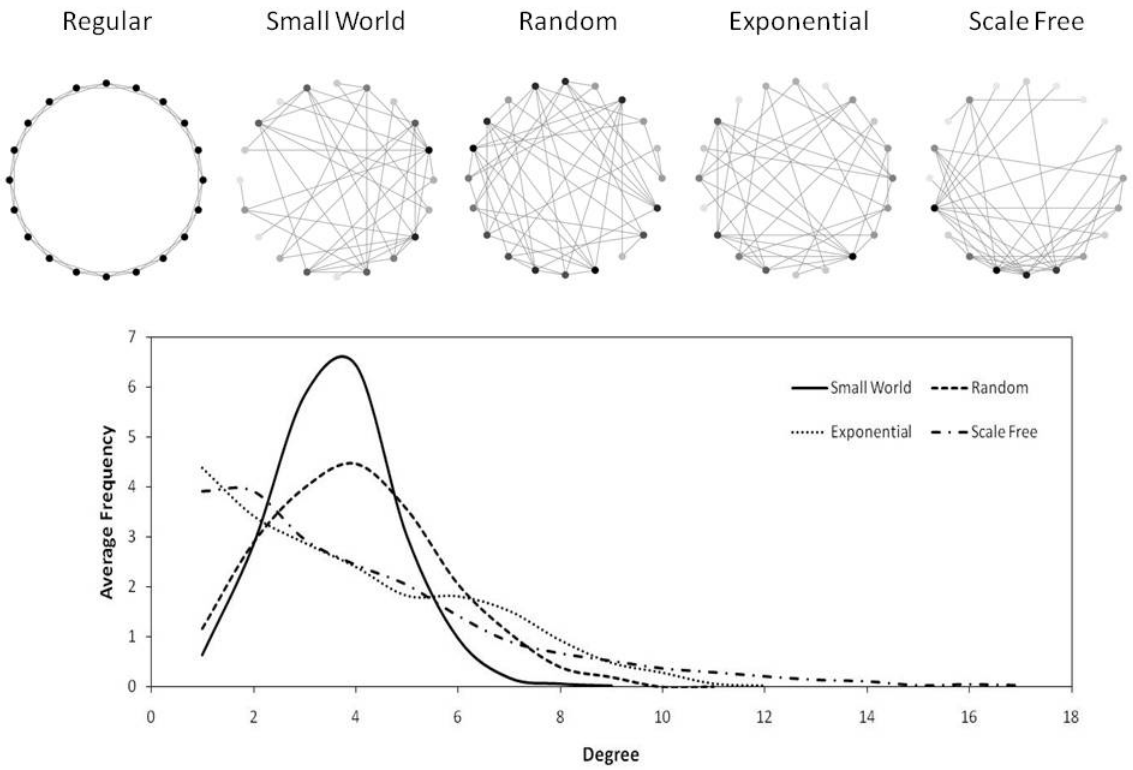


Figure 1.2: Sample network instances for regular (degree = 4), small world (initial degree = 4, re-wiring probability $p = 0.5$), random (edge probability = 0.2105), exponential (mean degree = 4) and scale free (generated using the Barabási-Albert model for preferential attachment [1]) structures. Darker shaded nodes have a higher degree relative to the other nodes and lighter shaded nodes have a relatively lower degree. The bottom plot shows the mean degree distribution for each network type with 20 nodes and approximately 40 edges, averaged over 100 samples. The regular network degree distribution (not shown in the graph) has a constant degree distribution, with all 20 nodes having a constant degree of 4. As the network structure changes from regular to scale free, the degree distribution becomes more skewed, with several nodes being highly connected and the remaining nodes having relatively few connections.

individuals in a general population. These models typically consist of a homogeneous population of agents with essentially no individual characteristics other than their infection status. However, heterogeneity enters the model because the degree of each node in the network is not constant. The goal of these models is to characterize how the structure of the network affects the rate and extent of transmission, and to further demonstrate the limitations of homogeneous models.

Bansal, Grenfell, and Meyers [6] presented strong evidence that homogeneous mixing models such as the SIR model do not accurately predict epidemics for realistic contact networks. The authors were able to demonstrate that several empirical contact networks could all be approximated by computer-generated networks with exponential degree distributions. Homogeneous mixing models, although reasonably accurate for characterizing transmission dynamics on regular, random networks, did not accurately predict epidemics on the more heterogeneous exponential and scale-free networks.

Christley [17] focused on identifying the most susceptible individuals, or those most likely to become infected in the event of an outbreak, in random and small-world networks. This type of analysis is very useful because the results could be used in developing strategies for targeting individuals for vaccination, isolation, or quarantine. The author experimented with various measures of node centrality (i.e., measures of a node's importance in a network), and determined that the degree of a node was as good an indication of an individual's risk of infection as more complicated measures that would require more information to compute. Eubank [28] proposed several local and global measures of network structure that could

have significant implications for transmission of infectious diseases.

The next set of network models contain multiple types of agents that interact, whether they are explicitly or implicitly represented in the model. These studies focused on the structure of the network. In addition, they characterized how interactions between different agent types affect transmission. The degree of heterogeneity in these network models facilitates analysis of the relative effect of each type of HCW. These models can account for HCW behavior as well, and bring consideration to other potential aspects of transmission such as HCW-to-HCW transmission and patient sharing. These types of interactions are not often considered, but can lead to increased levels of transmission in certain circumstances.

Temime et al. [81] constructed a model of a hospital ICU with three types of HCWs that visit patients. Two types of HCWs are assigned to specific groups, or cohorts, of patients, whereas the third type visits all patients (see Figure 1.3). The assigned HCWs represent nurses and physicians, and the third type (the peripatetic HCW), represents someone who could potentially come into contact with any patient, such as a nursing assistant or respiratory therapist. The model demonstrated the threat posed by the latter type, and presented results that a single, non-compliant peripatetic HCW could cause the same level of transmission as if all HCWs were moderately non-compliant (i.e., 19-23%). These effects become even more significant when HCW-to-HCW transmission occurs.

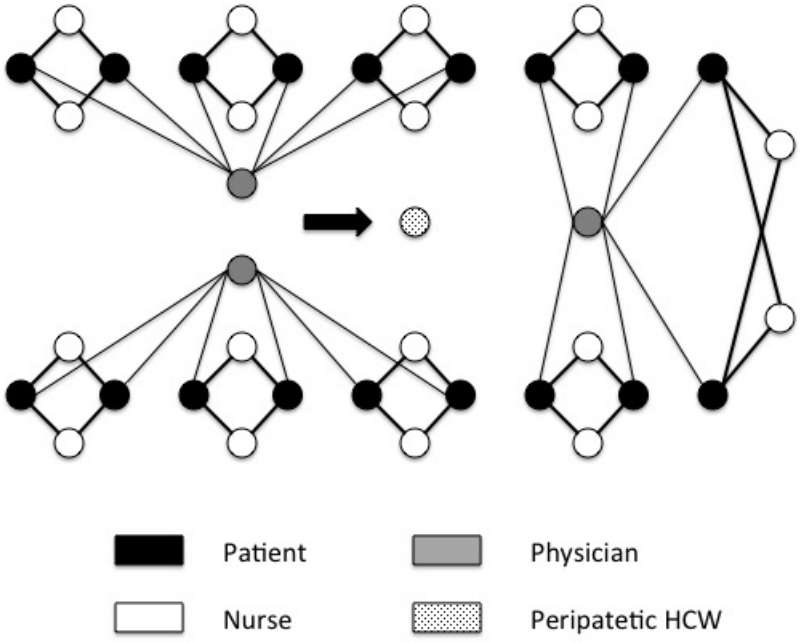


Figure 1.3: This figure shows the network of contacts in the ICU in the Temime model. There are 18 patients and three types of HCWs: two profiles of HCWs, corresponding to nurses and physicians, are assigned to subgroups of patients and one peripatetic-type HCW visits every patient once each day.

1.3.3 Pandemic Models

Pandemic modeling is another research area in which agent-based modeling has been used successfully. Diseases such as malaria, SARS, smallpox, and various strains of influenza have been modeled. Control measures are most often concerned with logistics, such as distributing vaccinations, locating community clinics, delivering emergency rations and medical supplies, and evaluating the effects of social distancing (e.g., school closings). The initial work in applying agent-based modeling techniques to pandemic scenarios was done primarily by Carley et al. [15] and Cummings et al. [23], which both demonstrated the value of agent-based models in generating pandemic dynamics and evaluating response strategies. Since this initial work, agent-based pandemic models have begun to incorporate massive data sets that reflect detailed demographic, social, transportation, and even climate characteristics of a particular geographic region. Although many of the characteristics and advantages of these models are applicable to smaller scales, this research area is beyond the scope of this dissertation. Therefore, we do not provide additional coverage of this topic.

1.4 Outline of Dissertation

The next five chapters in this dissertation describe applications of agent-based modeling and social network analysis to characterizing the spread of infectious diseases in medical facilities and community populations. In Chapter 2, we present an agent-based model of MRSA transmission in a hospital and use the model to explore

the effects of infection control measures and healthcare worker behavior. In Chapter 3, we apply a full 2^k factorial design to the output of the model developed in Chapter 2 to evaluate the effectiveness of the hand-hygiene compliance of healthcare workers and hospital staffing ratios. In Chapter 4, we examine how patient movement through a network of hospitals and long-term care facilities affects the long-term MRSA prevalence levels in each type of facility. Chapter 5 explores how the structure of the patient network and healthcare worker behavior affects transmission in a hospital. Chapter 6 compares the transmission dynamics for an epidemic to those of a bioterror incident, and explores diagnostic methods for differentiating between the two scenarios. Chapter 7 provides a brief conclusion to the dissertation.

Chapter 2

MRSA Transmission Reduction using Agent-Based Modeling and Simulation

2.1 Overview

Patients and healthcare workers (HCWs) frequently interact, creating the opportunity for the transmission of infectious diseases. If someone becomes colonized with methicillin-resistant Staphylococcus aureus (MRSA) or another pathogen, the bacteria could spread by way of HCWs to many others within the hospital population. As a result, many patients fall victim to hospital-acquired infections because a HCW carrying the bacteria made contact with a susceptible area of the patient's body. It is estimated by the Committee to Reduce Infection Deaths that infections acquired in US hospitals lead to over 100,000 deaths per year and an additional $30.5B in hospital costs [74]. Over 300,000 (of 2 million) infection cases involved MRSA, with close to 20,000 of those cases resulting in fatalities.

Some experts agree that hospital-acquired infections are almost entirely preventable [55], given a committed and capable healthcare institution. However, studies have shown that such measures have proven difficult to implement and enforce, due to both HCW non-compliance and cost considerations. These control measures require money and resources in terms of materials, dedicated personnel, and addi-

tional bed capacity. In some documented cases, they have proven to be cost-effective when preventing even a small number of infections [69].

Typically, the first measure taken by hospitals is promoting awareness aimed at improving hand-hygiene compliance of HCWs. Hospitals begin to enforce the use of alcohol-based hand disinfectants, gloves, and gowns that increase the efficacy of each hand-hygiene activity, as washing hands improperly may not remove a sufficient amount of bacteria. The next effort usually involves active surveillance, where patients are screened for MRSA, at admission and/or with some frequency during their stay. This policy allows for the detection of colonized, but asymptomatic, patients so that measures can be taken to prevent further transmission.

Among these additional measures are patient isolation, decolonization, and improvement of HCW-to-patient ratios. Patient isolation confines a detected colonized or infected patient to a single room. The decolonization process involves a regimen aimed at reducing or removing the presence of bacteria on the skin of a patient, which is done typically through the use of antibiotics and chlorhexidine bathing. This process reduces the probability that an HCW will acquire the bacteria after a visit. Increasing HCW-to-patient ratios decreases the range of transmission, as patients can only transmit the bacteria to others who share their caregivers.

There is some disagreement about whether hand washing is the ultimate solution to infection control. Some results have shown that compliance, if raised to high enough levels, could prevent transmission almost entirely [88]. Others have shown that hand washing is not a sufficient measure and that additional measures must be taken to reduce transmission to acceptable levels [10]. There is also some uncer-

tainty concerning the primary source of transmission. In certain scenarios, nurses may be more likely to transmit MRSA to patients, whereas in others, physicians may be the more likely source. Our research effort seeks to validate the most effective measures for minimizing MRSA transmission, and then explore novel questions related to infection control. The aim of this chapter is to describe an agent-based model and its use in studying MRSA transmission dynamics within a hospital.

2.2 Methodology

Historically, problems in epidemiology were solved by using case-control studies, cohort studies, and randomized controlled trials. In infection control, quasi-experimental studies are carried out to determine the effectiveness of various interventions, but unfortunately, only a few cluster randomized trials have been performed [38]. An expansion in methodology led to the use of mathematical modeling and simulation [4, 5, 10, 11, 18, 19, 48, 54, 69, 76] to investigate the spread of MRSA within hospitals. These computational models allow researchers to evaluate potential solutions in a virtual environment in order to help hospital administrators determine the best infection control policy. However, these models have limitations, as they are driven by the macroscopic behavior of the system. Even when properly calibrated, mathematical models lack realism because they fail to account for the low-level interactions that drive the system. These interactions are naturally represented by agent-based modeling [53], which we use to investigate this problem.

Agent-based modeling and simulation (ABMS) is a powerful technique that

seeks to generate emergent characteristics from simple, rule-based individual actions. One goal of ABMS is to generate macroscopic trends that are not necessarily anticipated from the definition of microscopic behavior. Building an agent-based model involves specifying the characteristics and behavior of the agents. The agents have a state, the ability to change that state with time, and the ability to interact in some environment. Complexity can be added to agent-based models by introducing learning so that agents can change their behavior over time. We use ABMS to define agents in a hospital, specifically patients, nurses, physicians, and visitors, who interact with each other. The interactions between these agents are the mechanism by which transmission occurs in the hospital. There are few ABMS models in this field. Temime et al. [81] have shown–albeit for a very specific scenario–that ABMS is capable of providing insight to new and relevant questions. Our model complements this work by focusing on the interactions between patients and HCWs and seeks to provide a better understanding to MRSA transmission in a hospital setting.

In order to account for different outcomes in a hospital, stochastic effects are required. We use Monte Carlo methods in the design of our model. ABMS can become computationally expensive quickly, as interactions among many agents is simulated. Running many Monte Carlo replications take a long time to execute. Consequently, an important capability of the model is to be able to execute serially or in parallel, so that more demanding test cases can be simulated in a reasonable amount of time.

2.3 Conceptual Model

Each agent in the model is defined by its characteristics and behavior. This type of modeling is supported best by object-oriented programming, in which object classes are defined with inherent characteristics and functionality. The simulation model is developed in Python, a dynamic object-oriented programming language [68]. In addition to basic Python, the SimPy, Parallel Python, NumPy, and SciPy modules are critical for building the model. SimPy has process oriented, discrete event simulation classes and methods that are used to develop the simulation architecture. Parallel Python is used to implement a capability so that Monte Carlo simulation replications can be executed simultaneously on multi-processor machines. NumPy provides a multi-dimensional array functionality that features many useful operations similar to MATLAB. SciPy is a module used for scientific computation tasks and provides random number generation functions.

The agents in the simulation are represented as processes, including patients, HCWs, and visitors. The hospital has single and double rooms, a staff of HCWs, and an infection control policy. The only modeled interactions are between patients and HCWs, and patients and their visitors. Interactions between HCWs are not modeled because there is insufficient data to support a significant contribution to transmission from such interactions. Simulation parameters can be input directly into the model or specified by spreadsheet. The primary transmission-related parameters are summarized in Table 2.1.

Agents in the simulation are generated by a source agent that varies in its

Table 2.1: MRSA transmission factors.

Parameter Name
Performance Related
· Hand-hygiene compliance
· Hand-hygiene efficacy*
· HCW-to-patient ratios
· Patient screening test return times
External Factors
· Transmissibility
- Patient to HCW
- HCW to patient
- Visitor to patient
· Length of stay
· Number of daily contacts
· Proportion of colonized patients admitted
· Proportion of high-risk patients admitted
· Number of visitors per day

* Defined as the probability of a hand-hygiene event during which the HCW successfully removes the bacteria

operation depending on the type of agent being generated. Patients are generated continually and are housed in a waiting room to replace discharged patients so that the hospital remains fully occupied. Patients can be admitted in a susceptible or colonized state, depending on the admission rate of colonized patients. This particular parameter can be adjusted to reflect the proportion of colonized patients that may be readmitted to the hospital or transferred from other hospitals. Colonized patients carry MRSA on their skin, but they are asymptomatic and therefore require screening tests to be detected. Colonized patients can also develop infections during their stay, but they can be detected visually and precautionary action can be taken while the patients state is confirmed with a test. Both colonized and infected patients can transmit MRSA to HCWs. HCWs are generated at the beginning of

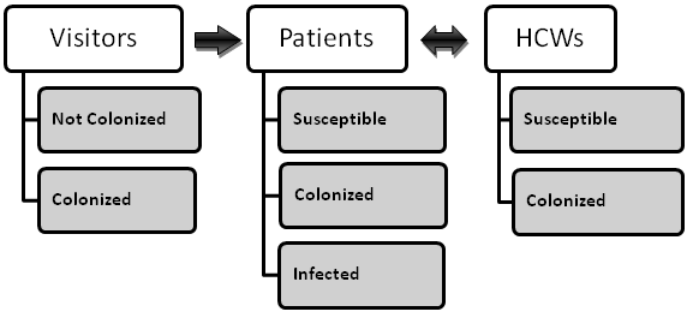


Figure 2.1: Agent interactions and state transitions.

the simulation as specified by input parameters. Each HCW is initialized at the uncolonized state. A fixed number of visitors are generated each day, each visiting a single patient in the hospital at random. The colonized state of the visitors is determined by choosing a random number and comparing it to the parameter-specified visitor colonization rate. Agent interactions and state transitions are summarized in Figure 2.1.

When patients are admitted to the hospital, they can be screened for MRSA if it is specified by the hospital infection control policy. Patient length of stay and the required number of visits are specified by the user and are defined after the patient is admitted. These parameters can be constant values for all patients or they can be randomly generated and vary for each patient. Patients are visited each day by nurse and physician agents, and sometimes visitor agents. Each HCW is assigned a fixed-size cohort of patients at random as the patients enter the hospital. The size of the cohort is specified by parameters that define the nurse-to-patient and physician-to-patient ratios. Nurse and physician cohorts are assigned at random and,

therefore, two patients that share a nurse do not necessarily also share a physician. As patients are discharged from the hospital, HCWs acquire new patients for their cohort. A ratio is defined to represent the proportion of visits by nurses. Physicians perform the remainder of visits. For each patient visit, a random number is drawn to determine whether the visit will be carried out by the assigned nurse or physician. The transmission of MRSA can occur in three ways:

1. A colonized patient transmits the bacteria to a susceptible HCW,
2. A transiently colonized HCW transmits the bacteria to a susceptible patient,
3. A colonized visitor transmits the bacteria to a susceptible patient.

The transmission of MRSA between agents is determined stochastically, based on the risk level of the patient and the behavior of the HCWs who visit the patient. Patients who have previously been colonized are at a higher risk to acquire MRSA again. Our model represents two risk levels (low and high) to account for this factor. HCWs that visit high risk, colonized, or infected patients are also more likely to wash their hands because they are aware of the state of the patient. Our model allows for different handwashing probabilities based on the risk level of the patient. Nurses and physicians have the opportunity to wash their hands after the visit. The probability that HCWs move back to the uncolonized state is based on the probability that they wash their hands and the efficacy of the act of handwashing.

A patient remains in the colonized state until the patient develops an infection or completes a decolonization regimen. A patient who clears MRSA through

decolonization is no longer infectious, but is immediately susceptible to becoming colonized again by a HCW or visitor. The colonized state of the patient can only be determined by a screening test whereas the infected state is determined by a physical examination and positive culture. If active surveillance is enforced, the patient may be screened periodically for colonization. Screening test return times can be adjusted to reflect those in practice, such as traditional culture testing or more advanced polymerase chain reaction testing, which is more accurate and yields test results more quickly. The patient can begin the decolonization or treatment process for an infection only after the test results are returned. If a patient develops an infection, the patients stay is extended for treatment; otherwise, that patient is released at the end of the original length of stay and another patient is allowed to enter the hospital.

The decolonization process is modeled in a simple manner. A parameter defines the MRSA clearance time for both colonization and infection. After this time passes, the patient moves back to the susceptible state. While a patient agent is undergoing the decolonization process, it can still transmit MRSA to the HCW, unless the patient is isolated. A patient can only be isolated in a single room and, as a consequence, there must be a single room in the hospital that does not contain a colonized or infected patient. Susceptible patients who occupy a single room can be swapped out to accommodate a patient who needs to be isolated. If there are no single rooms available, the patient who needs to be isolated must wait until a patient holding a single room is discharged or decolonized. HCWs who visit isolated patients wear gloves and gowns so as to not acquire MRSA from the isolated patient.

An isolated patient cannot transmit MRSA to other patients in the hospital.

2.4 Parallel Computing

Agent-based models can take substantial run times because of their explicit representation of interactions. Consequently, executing many Monte Carlo replications can become computationally prohibitive, if done serially. However, each replication is independent from the others, so it is advantageous to run as many simulations as possible in parallel and aggregate the results. In order to assess the effectiveness of parallel computing, two scenarios were tested: a small case with many replications and a large case with a small number of replications. These two scenarios were run on the Genome cluster at the University of Maryland. We used 32 processors and 128 GB RAM. The results of the comparison are given in Tables 2.2 and 2.3.

As shown in Tables 2.2 and 2.3, multiple processors provide the dramatic speedup, by a factor almost equal to the number of processors for a smaller number of processors. As the number of processors increases, there is some degradation in speedup due to the extraction of results from a larger number of processors, as indicated by the total job time, which is the sum of processing time across all processors. However, even with this degradation, the run, or wall-clock, times are faster with more processors, so an advantage remains.

For larger cases, it is clear that simulations are more difficult to run quickly, as single replications are computationally expensive. The degradation in speedup is

Table 2.2: Parameters and results of running a small case on Genome cluster.

Small Case	**N**	**Job Time** (sec)	**Run Time** (sec)	**Speedup**
· 100 days, 250 replications	1	747	747	–
· 10 single, double rooms	2	752	377	1.98
· 10 nurses, 5 physicians	4	746	188	3.97
· 10-day length of stay	8	752	96	7.78
· 5 daily contacts	16	761	50	14.94
· No infection control measures	32	941	33	22.64

N: number of processors.

more apparent because each processor loses efficiency by running fewer replications. However, the benefit of parallelization is greater in this case, because running larger numbers of replications in serial would take prohibitive amounts of time. Overall, parallelization is quite valuable. Even with the degradation in speedup as the number of processors increases, the run times continue to decrease. In addition, there is no penalty for parallelization because the replications are independent and, therefore, no accuracy is lost.

Many of the experiments described in the following sections are similar to the small case, in that single replications have relatively short execution times. However, when parameter variation is introduced, in conjunction with the number of replications required to produce stable results for each run, total job times begin to grow quickly. Therefore, these experiments require parallel processing to complete in reasonable amounts of time.

Table 2.3: Parameters and results of running a large case on Genome cluster.

Large Case	**N**	**Job Time** (min)	**Run Time** (min)	**Speedup**
· 500 days, 25 replications	1	136.9	136.9	–
· 50 single, 150 double rooms	2	138.4	71.84	1.91
· 50 nurses, 20 physicians	4	136.1	37.91	3.61
· 10-day length of stay	8	133.7	21.10	6.49
· 5 daily contacts	16	141.3	11.88	11.52
· All infection control measures	25	182.3	8.96	15.28

N: number of processors. Only 25 cases were run so that the serial case could run relatively quickly

2.5 Results

2.5.1 Infection Control Measures

In order to evaluate the infection control measures, a baseline case was defined, as specified in Table 2.4. Infection control measures were enforced separately from the baseline case to assess their effectiveness. Isolation and decolonization depend on patient screening to detect colonized patients, so in these cases these measures were used in conjunction with patient screening. HCW-to-patient ratios were improved without any additional control measure. The results of this experiment are summarized in Table 2.5.

As expected, a HCW-to-patient ratio of one nearly eliminates transmission, as colonized patients who are admitted are unable to transmit MRSA to other patients by way of HCWs. Maintaining this level in practice, however, is unlikely, especially outside of intensive care units. Even as the HCW-to-patient ratio decreases to 1:2, transmission increases dramatically. It turns out that HCW-to-patient ratios less than 1:3 offer little improvement on transmission from the baseline case. It is

Table 2.4: Baseline case parameters.

Baseline case
· 100 days, 250 replications
· 5 HCWs (Single type)
· 10 single, 10 double rooms
· 5% of patients admitted are colonized with MRSA
· 5 daily contacts per patient
· 5-day average patient length of stay
· 50% hand-hygiene compliance, 80% efficacy
· Proportion of high-risk patients admitted
· No interventions

Table 2.5: Summary of infection control measure performance.

Mean statistic	**Baseline**	**Isolation**	**Decolonization**	**HCW-to-Patient Ratio**	
				1:1	1:2
Patients colonized	51.46	39.56	45.42	**34.79**	40.65
Colonized patients admitted	36.50	34.48	34.76	**33.85**	33.89
Number of secondary cases	14.96	5.08	10.66	**0.94**	6.75
Colonized patient days (%)	6.49	5.66	5.72	**5.14**	5.64
Ratio of primary to secondary cases	0.41	0.15	0.31	**0.03**	0.20

Bold numbers indicate best performance.

evident that patient isolation is a better control measure than decolonization and is likely the best overall measure due to the difficulty in achieving a 1:1 HCW-to-patient ratio. Decolonization suffers greatly from the long time it actually takes for a patient to clear colonization. Therefore, patients can still spread MRSA throughout the hospital if they are not isolated. The performance of these infection control measures appears consistent with the results reported in the literature.

2.5.2 Nurses vs. Physicians

Now that ABMS has demonstrated results consistent with the literature, simulation experiments can be performed to provide insight into questions relevant to hospitals. For example, which type of HCW, nurse or physician, is responsible for the most transmissions? This particular question has been difficult for epidemiologists to address. What effect do nurses and physicians who practice poor hygiene have on transmission? ABMS is well-suited to answer these questions, as it is straightforward to track the number of colonizations directly attributable to nurses and physicians. Under what circumstances does a hospital that practices high hand-hygiene compliance and additional infection control measures become susceptible to an outbreak? Unless otherwise noted, each experiment was simulated for 100 days and was repeated for 100 simulation replications.

The question of who is responsible for more transmissions is important because hospitals want to focus educational programs where they would have the most significant impact. Nurses and physicians have different degrees of interaction with

patients and, therefore, would require a different approach to reduce transmission. Typically, nurses see patients much more often, and the literature shows that they are more likely to wash their hands. Physicians see many more patients, but less frequently. These contrasting service patterns make it difficult to predict the primary source of transmission.

2.5.2.1 General Ward Settings

Three experiments were conducted in a 50-patient hospital with 10 nurses over a 100-day period. All experiments varied the proportion of patient visits from nurses and measured the proportion of patients colonized by nurses. In the first experiment, the number of physicians was varied from 1 to 5, with equal hand-hygiene compliance among the different HCWs, to determine the effects of nurse-to-physician ratios on transmissions. The results of this experiment are summarized in Figure 2.2. In two additional experiments, we examined the effects of physician hand-hygiene compliance on the proportion of colonizations by nurses. Physician compliance was varied up to the compliance of nurses, which was 50% and 80% in the second and third experiments, respectively.

In the first experiment, it is clear that the nurse-to-physician ratio does not have a significant impact on the transition point where nurses colonize more patients than physicians. For this scenario, whichever class receives the majority of patient visits is likely to account for the majority of transmissions. In the second and third experiments, the difference in hand-hygiene compliance greatly affects the

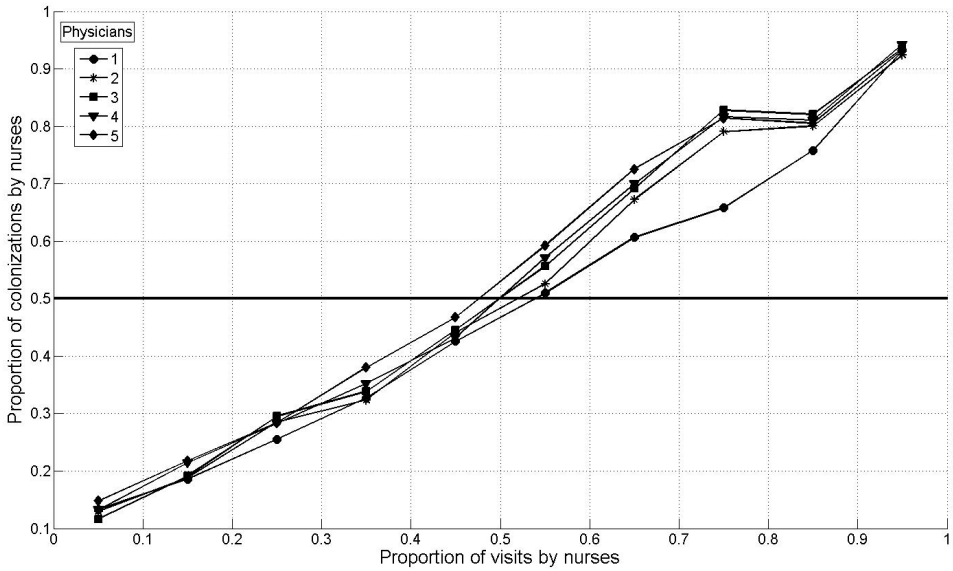


Figure 2.2: Proportion of nurse colonizations in a general ward, varying the number of physicians in a 50-patient hospital ward with 10 nurses.

transition point and shifts it further to the right as the difference grows, as shown in Figures 2.3 and 2.4. In practice, nurses typically account for 80-90% of contacts with patients. Therefore, our results indicate that nurses account for the majority of colonizations in a general ward type setting. These results suggest that hospitals should be aware of which HCW type sees patients more often and specifically target education programs towards that group.

2.5.2.2 Intensive Care Unit Settings

Transmission dynamics could differ in this respect depending on the type of unit where the patients and HCWs reside. To explore this concept, the previous experiments were run in a simulated intensive care unit (ICU) that houses 20 patients

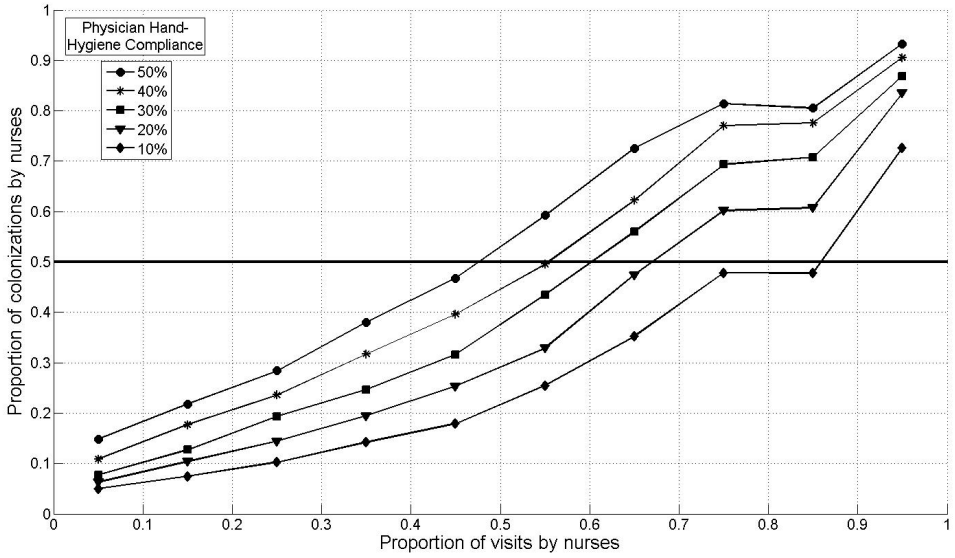


Figure 2.3: Proportion of nurse colonizations in a general ward, varying physician hand-hygiene compliance in a 50-patient hospital ward with 10 nurses—50% nurse hand-hygiene compliance.

in single rooms with 10 nurses and two physicians. In this scenario, nurses are much more likely to transmit the pathogen to the two patients they visit, but are unable to transfer to any other patients in the ICU. Physicians present a greater risk because they are capable of transferring MRSA to as many as 10 patients, whereas each nurse only directly affect 2 patients. The results of this experiment for low, moderate, and high hand-hygiene compliance values for nurses are shown in Figures 2.5 and 2.6, which differ in the proportion of visits to patients by nurses.

In general, the results in Figures 2.5 and 2.6 indicate that when physicians practice good hand-hygiene, nurses become responsible for most of the colonizations. Only at very low compliance (30% or less) do physicians colonize more patients. However, these results are somewhat misleading because they imply that physicians

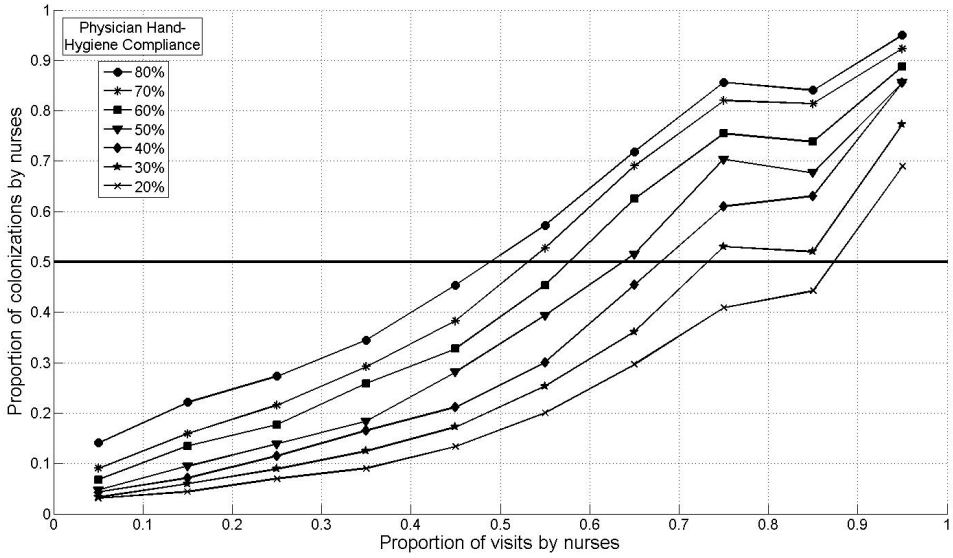


Figure 2.4: Proportion of nurse colonizations in a general ward, varying physician hand-hygiene compliance in a 50-patient hospital ward with 10 nurses—80% nurse hand-hygiene compliance.

do not pose as great a danger to patients in an ICU. In fact, physicians pose perhaps a greater threat to patients because they can transfer MRSA from one patient cohort to another whereas nurses can only colonize the remaining susceptible patients in their respective cohorts. In this particular case, nurses simply colonized the patients sooner than physicians because of the higher frequency of visits. To demonstrate this observation, we simulated the same ICU with a 1:1 nurse-to-patient ratio, so that only physicians could transmit MRSA to other patients. There is a small degree of transmission by nurses, mainly due to nurses becoming transiently colonized by their current patient and transmitting the bacteria to a new patient in the same room. The results of this experiment are shown in Figure 2.7, which shows that physicians account for almost all transmission in the ICU, except at extremely high

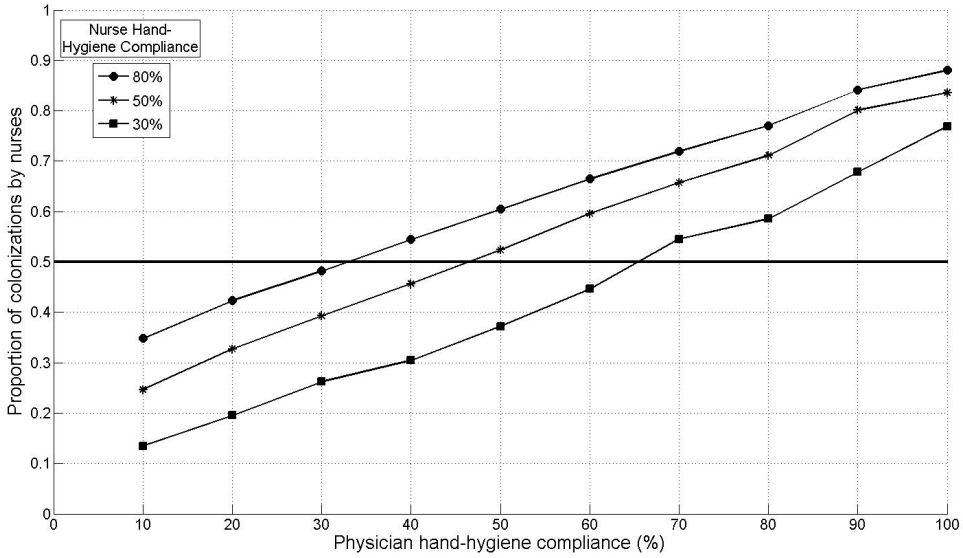


Figure 2.5: Proportion of nurse colonizations in an intensive care unit with a 1:2 nurse-to-patient ratio, as a function of physician hand-hygiene compliance with nurses performing 80% of the visits.

values of hand-hygiene compliance. We point out that the number of colonizations by physicians remained roughly the same between the two nurse-to-patient ratios (1:2 versus 1:1), whereas nurse colonizations decreased dramatically. This trend indicates that physicians, given enough time, are likely to transmit the bacteria to patients regardless of nurse behavior.

In fact, physician colonizations in this type of setting are nearly independent of nurse activity because physicians visit a larger number of patients, but less frequently than nurses. Nurse colonizations are dependent on physician activity because physicians have the capability of introducing MRSA into various patient cohorts. In effect, there is positive feedback in the system each time a physician is colonized by a patient in one cohort and spreads it to another cohort. These effects

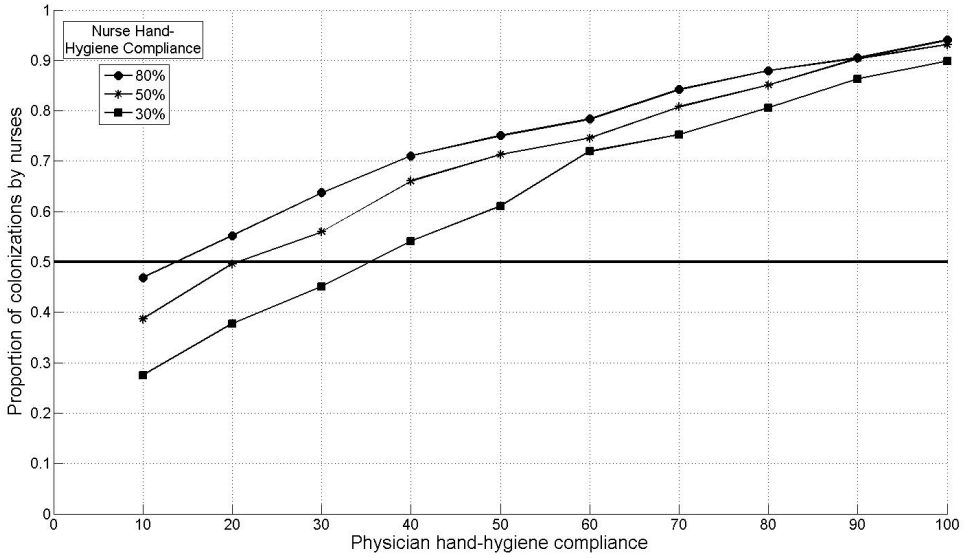


Figure 2.6: Proportion of nurse colonizations in an intensive care unit with a 1:2 nurse-to-patient ratio, as a function of physician hand-hygiene compliance with nurses performing 90% of the visits.

are shown in Figures 2.8 and 2.9, where there is little variation in the number of colonizations by physicians for a specific physician hand-hygiene compliance, but sizeable increases in nurse colonizations as the nurse-to-patient ratio decreases.

We have seen the effect of varying the behavior of nurses and physicians as a whole, but it is also worthwhile to consider the effects of rogue HCWs. A rogue HCW is less compliant with respect to handwashing than the rest of the medical staff. In separate cases, the effect of a rogue behavior was examined for a nurse and physician in an ICU where the rest of the staff practices 80% compliance. The results of these two experiments are summarized in Figures 2.10 and 2.11. These figures illustrate the threat that physicians pose to patients in an ICU setting, as rogue physician behavior can lead to a sizeable increase in the number of colonizations.

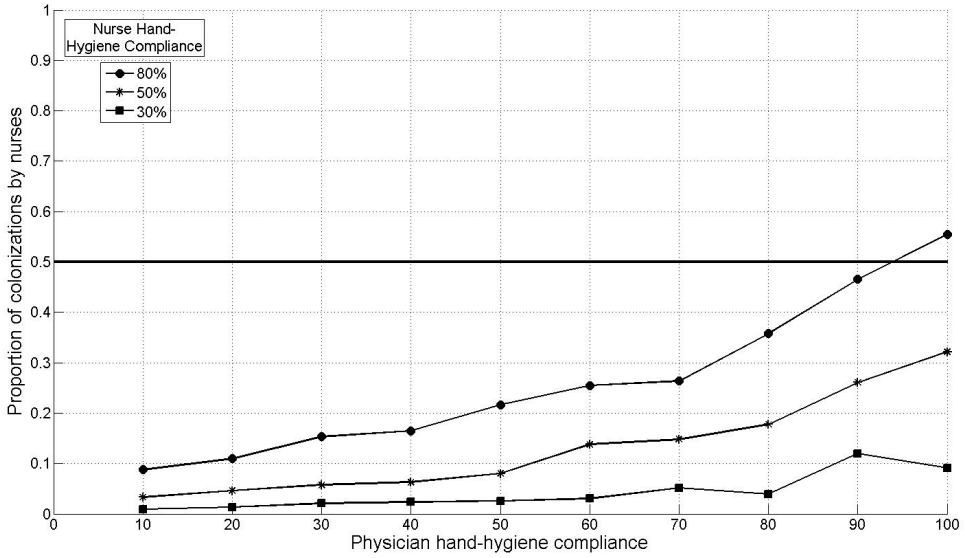


Figure 2.7: Proportion of nurse colonizations in an intensive care unit with a 1:1 nurse-to-patient ratio, as a function of physician hand-hygiene compliance with nurses performing 90% of the visits.

Rogue nurse behavior has minimal effect, even at the lowest compliance values, as they are only able to colonize a small number of patients (two in this case) in their respective cohorts. We see from these results that rogue behavior is an important issue when the number of HCWs is small. However, in cases where the number of HCWs is large, rogue effects are less evident.

Based on our results, hospital administrators may be interested in the relative benefit of increasing HCW-to-patient ratios, especially when compared to the reduction in transmissions attributable to higher hand-hygiene compliance. This issue is only relevant to nurses because physicians earn much higher salaries and typically would not be hired only to reduce MRSA transmission rates. A comparison of these benefits is shown for a similar ICU, where the HCWs all practice the same level of

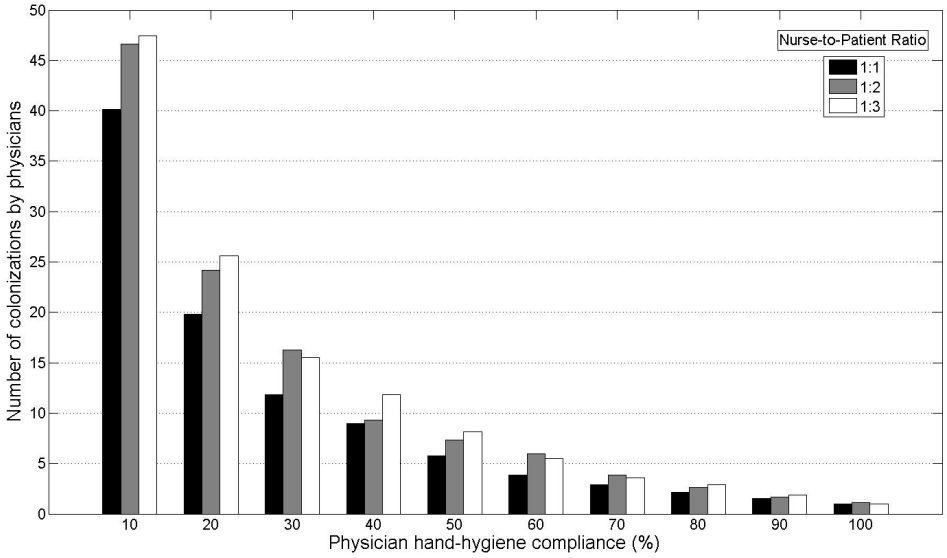


Figure 2.8: Colonizations by physicians in an intensive care unit, as a function of physician hand-hygiene compliance and nurse-to-patient ratios.

hand-hygiene compliance and nurses visit the patient 90% of the time. The results of this experiment are summarized in Figure 12.

The key observation from Figure 12 is obtained by comparing the reduction in the number of colonizations by nurses as a result of increasing the nurse-to-patient ratio with the effect of increasing hand-hygiene compliance. Clearly, there is great benefit in increasing the hand-hygiene compliance of nurses already on staff. A comparable reduction in the nurse colonizations attributable to increasing the hand-hygiene compliance of one nurse from 30% to 80% would require hiring an additional 9 nurses at 30% compliance.

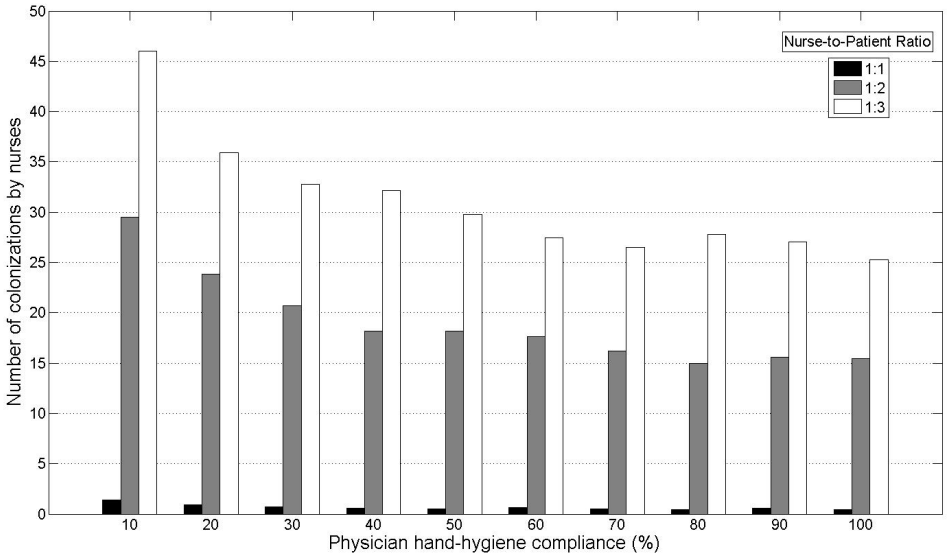


Figure 2.9: Colonizations by nurses in an intensive care unit, as a function of physician hand-hygiene compliance and nurse-to-patient ratios.

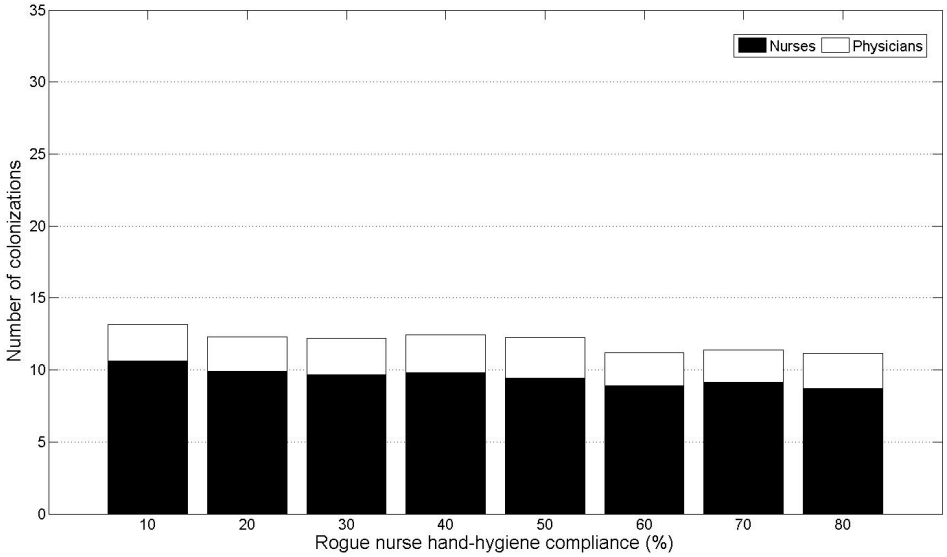


Figure 2.10: Colonizations by nurses and physicians in an intensive care unit, as a function of rogue nurse hand-hygiene compliance.

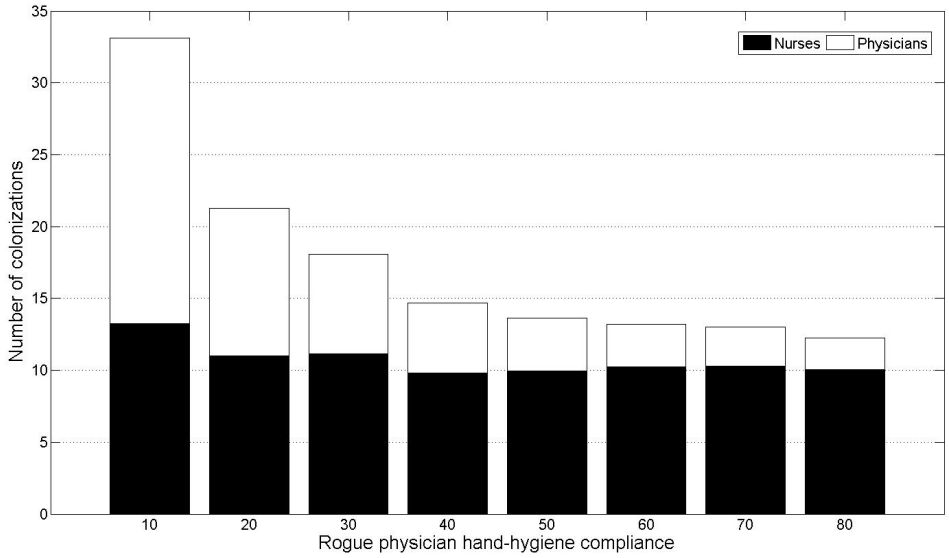


Figure 2.11: Colonizations by nurses and physicians in an intensive care unit, as a function of rogue physician hand-hygiene compliance.

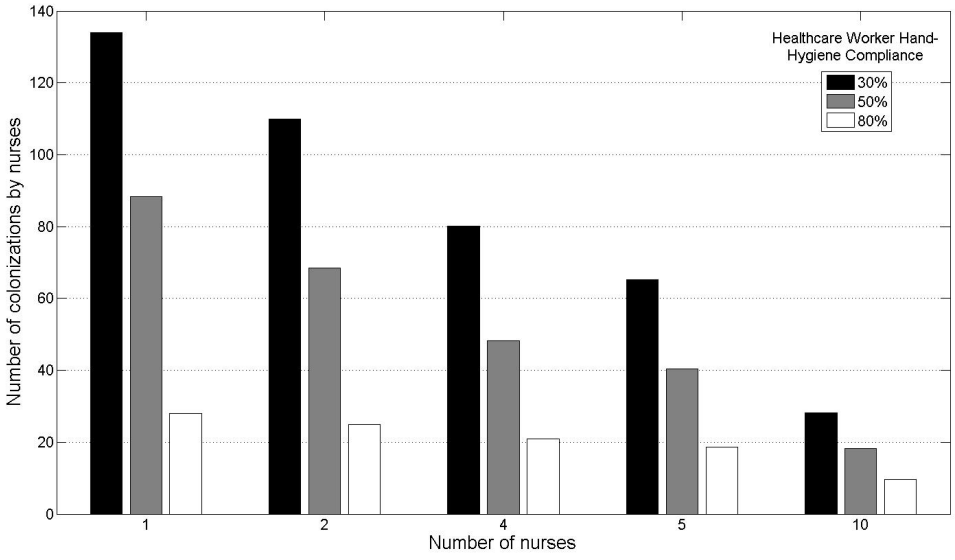


Figure 2.12: Number of patients colonized by nurses in an intensive care unit, as a function of nurse-to-patient ratio and healthcare worker hand-hygiene compliance.

2.5.2.3 Susceptibility of High-Performance Hospitals

Some hospitals have already made significant progress in reducing the incidence of MRSA transmission. In order to investigate the degree of susceptibility of these high-performance hospitals to MRSA outbreaks (ratio of secondary to primary MRSA cases greater than 1), we consider a 100 patient hospital with 20 nurses and 10 physicians. The HCWs comply with hand washing 70% of the time, but practice no additional infection control measures. Even with such a high compliance rate, this hospital is susceptible to an outbreak in the following cases:

- 10 daily contacts or more between HCWs and each patient,
- 20 day or more average patient length of stay,
- Transmissibility greater than 0.15,
- Hand-hygiene efficacy less than 0.6,
- 200 or more visitors per day at 2% transmission rate.

Suppose we increase the performance of the hospital by employing patient screening on admission with one-day test result return times, patient isolation, and decolonization. It would appear that this hospital would not be susceptible to MRSA outbreaks due to the effort put forth to prevent and control infection. For the most part, this assessment is accurate. Moderate changes in several transmission factors, such as hand-hygiene efficacy, number of daily contacts, proportion of colonized admitted patients, screening test return times, and patient length of stay, do not greatly affect transmission. Only two cases appear to lead the system to an outbreak:

(1) a highly transmissible pathogen (greater than 0.28 for this case) and (2) a high visitor rate (greater than 200 per day at 2% transmission rate). In the first case, a highly transmissible pathogen is able to transfer between patients and HCWs even at high compliance values. The second case can lead to an outbreak due to the small-world effect because visitor introductions create new pockets of colonization that allow transmission to occur in different areas of the hospital.

2.5.2.4 Summary

We have presented the results from many experiments that examined the balance between nurse and physician colonizations as well as the performance of the hospital as a system. We have seen that transmission dynamics vary between different hospital units and that nurses and physicians can have significantly different effects on patients depending on their behavior. The greatest threat appears to be from physicians. They have the capability of colonizing patients directly and they can introduce MRSA into a nurse cohort that leaves the remaining patients susceptible to colonization. Hand-hygiene compliance is a critical factor for infection control. In particular, increasing the compliance of the hospital can reduce transmission much more effectively than hiring additional workers. At the same time, hand-hygiene compliance is not the only solution to preventing MRSA transmission. Outbreaks can still occur even when hospitals practice high compliance. Hospitals need to be aware of this observation, and should pursue additional control measures.

2.6 Conclusions

Agent-based modeling and simulation is a powerful tool for analyzing complex systems. When applied to epidemiological problems, it is straightforward to represent individuals and the interactions among them to model the transmission dynamics of diseases and the effectiveness of various infection control measures. Parallel processing is a valuable capability, as agent-based models are typically computationally intensive, requiring a large number of computer cycles.

Reducing MRSA transmission in hospitals involves three key approaches, based on our results:

1. Minimize the size of patient cohorts,

2. Screen for and isolate colonized and infected patients,

3. Decrease the likelihood of transmission between patients and HCWs by minimizing the number of visits and enforcing the use of gloves and gowns with colonized and infected patients.

Frequent transmission occurs when the patient population is well-mixed, which occurs when many patients share the same HCW and allows for transmission to take place across the cohort. Isolating patients and maintaining favorable HCW-to-patient ratios can serve to segment the patient population, so that colonized patients are less likely to transfer the bacteria to others. HCWs can reduce the probability of transmission by practicing proper hand-hygiene. Hospitals can help by minimizing patient length of stay and the number of daily contacts.

In future work, we would like to investigate the effects of modeling time-varying hand-hygiene compliance on the transmission dynamics of MRSA in a hospital. All experiments so far have assumed that the hand-hygiene compliance of HCWs is fixed. This restriction implies that HCWs are not able to make adjustments to their compliance based on the recent history of colonizations within the hospital. This type of study would provide insight into the value of informing HCWs about MRSA colonizations.

Chapter 3

Comparing the Relative Effectiveness of Hospital Infection Control Measures using Factorial Design

3.1 Overview

Infections with MRSA are a major problem among hospitalized patients [21, 60]. Interventions have been proposed to decrease the transmission of MRSA in high-risk healthcare settings such as the intensive care unit [62, 86]. Two key factors that have been identified as effective infection control measures are improving the hand-hygiene compliance of healthcare workers and increasing the ratio of nurses to patients in the hospital.

Hand-hygiene compliance is a cornerstone of infection control. Many studies have attempted to demonstrate that improving hand-hygiene compliance leads to fewer hospital-acquired infections. Pittet et al. [67] demonstrated that the introduction of alcohol-based hand disinfectant, which likely caused an increase in hand-hygiene compliance, reduced the incidence of MRSA infections. However, there are no experimental studies that have assessed the effect of improving hand-hygiene compliance on MRSA colonization acquisition rates.

Improving nurse-to-patient ratios has also been advocated as a potentially effective intervention at reducing the transmission of antibiotic-resistant bacteria.

This intervention should lead to fewer MRSA acquisitions by limiting the number of contacts between nurses and patients, thereby decreasing the opportunity to spread MRSA to multiple patients. Fridkin et al. [32] reported that the reduction of the nursing staff below a critical level may contribute to an increase in catheter-associated bloodstream infections. However, there are no experimental studies that have assessed the effect of changing the nurse-to-patient ratio on MRSA acquisition rates.

Hospitals and their resident epidemiologists often have limited resources and must choose between several potentially viable interventions aimed at decreasing patient-to-patient transmission of antibiotic-resistant bacteria such as MRSA. Mathematical modeling and simulation can be used to assess the potential benefits of different interventions when experimental trials have not been performed or cannot be performed due to ethical considerations [3, 32, 56]. There are several studies that have used mathematical modeling and computer simulation to evaluate the effectiveness of one or both of these factors in reducing the incidence of hospital-acquired infections [4, 9, 10, 36, 69, 78]. These studies have compared the effectiveness of bundled interventions on MRSA acquisition for a particular hospital configuration, but have provided little or no consideration as to when, or in what cases, one intervention is better than another. In addition, these studies have not fully characterized how changing the level of one factor affects the effectiveness of the other factors. These interaction effects are a strong determinant of how much transmission can be reduced by improving a limited set of control measures. In this chapter, we compare the relative effectiveness of increasing the hand-hygiene compliance of nurses and in-

creasing the nurse-to-patient ratio in reducing the transmission of hospital-acquired infections. This comparison should help inform hospitals and hospital epidemiologists that are considering these interventions.

3.2 Methodology

We use a stochastic, agent-based model developed by Barnes, Golden, and Wasil [7] to simulate a 20-bed ICU and the transmission of MRSA among patients through direct contact with healthcare workers. The agent-based formulation of the model facilitates the separation of behavior among patients, nurses, and physicians, so that their individual effects on the entire system can be evaluated. In addition, within each agent class, each individual can have unique characteristics and behavior, which provides more flexibility to analyze the system in detail.

3.2.1 Model Assumptions and Parameter Estimates

Our model simulates patients entering the ICU, occupying a single room, and being discharged after their stays are completed. Each patient is visited a constant total number of times each day, but the relative proportion of visits by nurses and physicians can vary. During these visits, patients are susceptible to acquiring MRSA from transiently colonized healthcare workers. After each visit to a patient, a healthcare worker washes his or her hands with a given probability and a specified efficacy of removing MRSA. A healthcare worker removes any MRSA bacteria that may have been transiently acquired from a colonized patient if the handwashing

event is successful, and he or she does not pose an immediate threat to transmitting MRSA to other patients. If handwashing is omitted or the event is unsuccessful, the healthcare worker can potentially transmit MRSA to other patients during subsequent visits.

In order to assess the relative effectiveness of improving hand-hygiene compliance among nurses and increasing the nurse-to-patient ratio, we assumed that both the number of physicians and their hand-hygiene behavior was constant (i.e., only parameters related to nurses were varied). Two physicians were responsible for 10 patients each in the ICU. Physician compliance was held constant at 65% throughout the experiments. The other key parameters for our model are summarized in Table 3.1. Given a set of input parameters, we simulate the interactions among patients and healthcare workers over a specified time period, and produce statistics on the number of transmissions to patients by nurses and physicians.

3.2.2 Factorial Design Methodology

A full 2^k factorial design is applied to compare hand-hygiene compliance of nurses and nurse-to-patient ratios in reducing MRSA transmission. A 2^k factorial design specifies two levels for each of the k factors used in the experiment: a plus-level represents the factor value that has the better effect on the response and a minus-level that has a less desirable effect. Simulations are then conducted using the parameter values specified by each design point to generate a sample of the system response at that level. A full design simulates all factor-level combinations,

Table 3.1: Factorial design model parameters and values.

Parameter	Value	References
Number of physicians	2 (1:10 ratio)	
Physician hand-hygiene compliance	0.65	[22, 26]
Hand hygiene efficacy	0.95 (low)	[34, 65]
Proportion of admitted patients positive with MRSA	0.10	[41, 75, 83]
Transmission probability from patient to healthcare worker	0.20	
Transmission probability from healthcare worker to patient	0.05	
Patient length of stay	logN(0.693,1.1646)*	[45, 2]
Visits per day per patient	48	[70, 49]
Proportion of visits by nurses	0.90	

* logN(α, β) represents the lognormal distribution with scale parameter α and shift parameter β. The parameters α and β were determined to generate a distribution with a mean length of stay of 3.94 days and a median length of stay of 2 days.

Table 3.2: Sample factorial design.

Design Point	Factor A	Factor B	Response
1	-	-	R_1
2	-	+	R_2
3	+	-	R_3
4	+	+	R_4

whereas a partial design would use only a subset of the design points to evaluate the effects of the various factors. The two input factors for this study are the hand-hygiene compliance of nurses and the nurse-to-patient ratio in the ICU. The system response is the number of MRSA acquisitions in the ICU over the simulation period. A sample factorial design is shown in Table 3.2.

There are two primary results derived from a factorial experiment. The first result is the set of main effects, each of which represents the average effect on the

response by increasing one factor from its minus-level to its plus-level. Main effects, denoted by e_k for some factor k, can have positive or negative values, which represent increasing or decreasing the response of the system as the factor level changes from minus to plus. Main effects that are close to zero suggest the factor has little to no effect on the system. The second result is the set of interaction effects between all unique subsets of two or more factors. Interaction effects, denoted by $e_{k_1 k_2}$ for factors k_1 and k_2, characterize how changing one factor affects the ability of the other factors in the chosen subset to influence the response. Small interaction effects imply that improving multiple factors has nearly the same effect as adding together the benefits of improving each factor by itself, which is the maximum potential improvement. When interaction effects are large, other factors are less likely to have a significant effect on the response when one factor is at its plus-level. Interaction effects can be difficult to interpret, because they only provide a relative measure for how dependent one factor is on other factors. Equations to calculate the main effect of two factors (A and B) and the interaction effect between them are shown in Equations 3.1, 3.2, and 3.3.

$$e_A = \frac{-R_1 - R_2 + R_3 + R_4}{2} \tag{3.1}$$

$$e_B = \frac{-R_1 + R_2 - R_3 + R_4}{2} \tag{3.2}$$

$$e_{AB} = \frac{R_1 - R_2 - R_3 + R_4}{2} \tag{3.3}$$

We examine whether increasing the hand-hygiene compliance of nurses from 0% to 100% in various increments causes a larger reduction in MRSA acquisition rates than increasing the nurse-to-patient ratio from 1:4 to 1:1 in various increments. Rather than determine a single measure of effectiveness for each measure, we iteratively apply a 2×2 factorial design across the entire parameter space. Beginning at 0% compliance and a nurse-to-patient ratio of 1:4, we compute the mean main effects and mean interaction effect across 25 Monte Carlo replications over selected factor-level combinations, using the number of MRSA acquisitions as the response. We use increments of 5%, 10%, 15%, 20%, and 25% between the plus- and minus-levels for the hand-hygiene compliance of nurses and increases of one (e.g., 1:4 to 1:3) and two levels (e.g., 1:4 to 1:2) for the nurse-to-patient ratio. By computing the main and interaction effects across all selected factor-level combinations, we are able to gain a better understanding of where in the parameter space each factor has an advantage.

There is one issue that arises when comparing interaction effects across the entire set of iterative factorial experiments. Normally, interaction effects are computed over a set of k factors in a single experimental design. In this case, interaction effects between all subsets of factors are comparable because they are computed using the same set of system responses, and therefore, are relative to the same scale. However, when system responses are generated over a range of parameter values, interaction effects are computed on different scales, which undermines the comparisons. This issue does not affect main effects, which are absolute measures of the effects of factors on the response and can be compared across the parameter space

with no loss of generality.

In order to address this issue concerning interaction effects and facilitate comparisons across the entire parameter space, we introduce the concept of the maximum interaction effect. The maximum interaction effect represents the largest possible interaction effect for a given set of factors which, for a 2×2 design, occurs when improving one factor has the same benefit as improving the other factor or both factors. In terms of the sample design in Table 2, the maximum interaction effect occurs when $R_2 = R_3 = R_4$. We can insert this expression into Equation 3.3 and then normalize all interaction effects using the maximum interaction effect for each case to facilitate meaningful comparisons. These two expressions are shown in Equation 3.4 and Equation 3.5.

$$e_{AB}^{max} = \frac{R_1 - R_2 - R_3 + R_4}{2} = \frac{R_1 - R_4}{2} \tag{3.4}$$

$$\hat{e}_{AB} = \frac{e_{AB}}{e_{AB}^{max}} = \frac{R_1 - R_2 - R_3 + R_4}{R_1 - R_4} \tag{3.5}$$

3.3 Results

The results of the factorial design experiments are summarized in Figures 3.1 to 3.5. Figure 3.1 gives the response values at each factor-level combination, shown as data series for each nurse-to-patient ratio level. Figures 3.2 to 3.5 compare the relative effectiveness of the two factors for four different changes in the nurse-to-patient ratio and highlight the interaction between the two factors. There is one

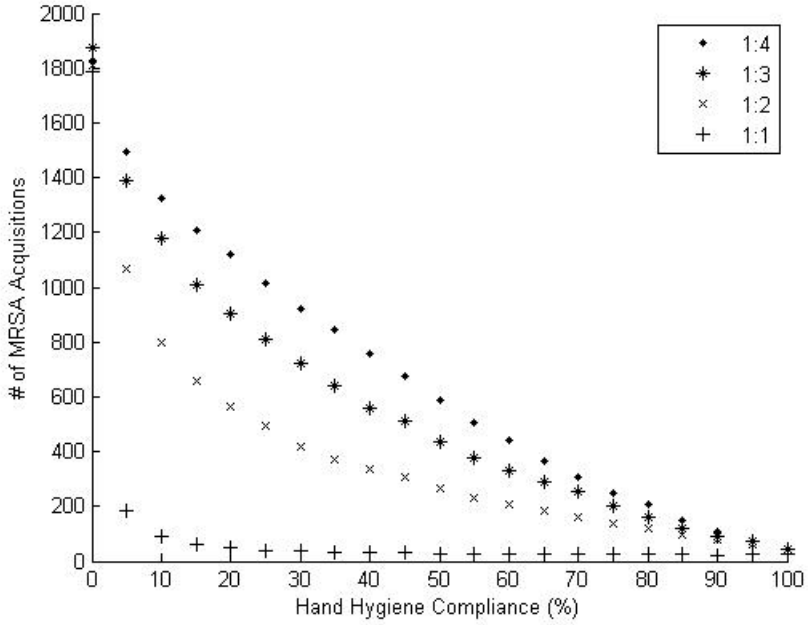


Figure 3.1: Summary of response values at all factor-level combinations of hand-hygiene compliance and nurse-to-patient ratio.

figure for each of the investigated changes in nurse-to-patient ratio, that is, from 1:4 to 1:3 (Figure 3.2), 1:3 to 1:2 (Figure 3.3), 1:2 to 1:1 (Figure 3.4), and 1:4 to 1:2 (Figure 3.5). Each series of data points in the figures on the left contains values that represent the difference between the mean main effect of increasing the hand hygiene of nurses and the mean main effect of the corresponding change in the nurse-to-patient ratio. A positive value indicates that the improvement in nurse-to-patient ratio resulted in more prevented MRSA acquisitions than the corresponding increase in hand-hygiene compliance. A negative value indicates the improvement in hand hygiene led to a better result. A line is drawn at zero to separate cases where nurse-to-patient ratio performed better from cases when hand-hygiene compliance resulted in more effective infection control.

Each series contains these difference values for a specified increment in hand-hygiene compliance. For example, each series labeled 5% contains difference values that use the x axis compliance as the minus-level and a hand-hygiene compliance level 5% higher as the plus-level to compute the main effect. Series labeled 25% use the x axis compliance as the minus-level and a compliance that is 25% higher as the plus-level. Each series only extends to a compliance value that has 100% as its maximum plus-level (i.e., 5% terminates at 95%, 10% at 90%, etc.).

Each data series in the figures on the right represents the set of normalized interaction effects over the entire range of compliance values for a particular increase in the nurse-to-patient ratio. Similar to the difference value figures, each data point represents the normalized interaction effect computed using the baseline compliance level indicated by the x axis, a plus-level specified by the series compliance increment, and the increase in nurse-to-patient ratio for that particular figure. Normalized interaction effects have a maximum value of one, which represents the case where the interaction effect is actually maximized. This value is not likely to occur in any experiment, unless one factor has no effect on the response. Interaction effects can also be negative, which represents cases in which improving one factor actually improves the effect of the other factor on the response. However, these cases only occur when compliance increases from very low levels.

A two-sided, two-sample t-test with unequal and unknown variances was applied to each main effect difference value. A two-sided, single-sample t-test with unknown variance was applied to each normalized interaction effect. We conducted these tests at the $\alpha = 0.05$ level. Equations for the test statistics and degrees of

Table 3.3: Statistical test equations.

Experimental Output	**Test Statistic***	**Degrees of Freedom**
Main Effect Difference Value	$T = \frac{\bar{e}_A - \bar{e}_B}{\sqrt{s_A^2/n + s_B^2/n}}$	$\nu = \frac{(s_A^2/n + s_B^2/n)^2}{\frac{\left(s_A^2/n\right)^2}{n-1} + \frac{\left(s_B^2/n\right)^2}{n-1}}$
Normalized Interaction Effect	$T = \frac{\bar{\hat{e}}_{AB}}{\sqrt{\frac{\left(\hat{s}_{AB}^2/n\right)^2}{n-1}}}$	$\nu = n - 1$

* Parameters denoted with a bar above represent sample means. Parameters of the form s^2 represent sample variances. n is the number of samples (i.e., simulation replications).

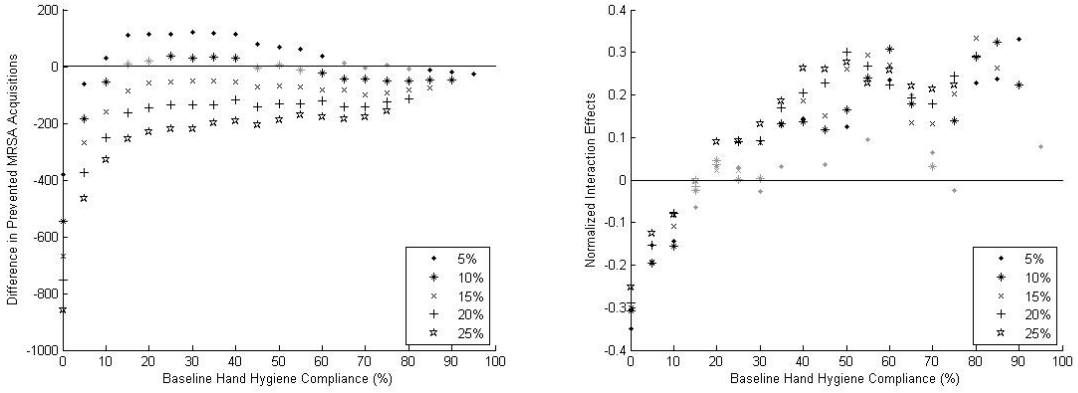


Figure 3.2: Summary of difference values between hand-hygiene compliance main effects and the main effect of increasing the nurse-to-patient ratio from 1:4 to 1:3 (left) and the corresponding normalized interaction effects (right). Non-significant results are shaded in gray ($\alpha = 0.05$).

freedom are summarized in Table 3.3. Non-significant difference values cannot differentiate between the impact of hand-hygiene compliance and the nurse-to-patient ratio. Non-significant interaction effects suggest that improving both factors can achieve the maximum possible benefit. Statistical significance is reflected in Figure 3.2 through Figure 3.5, in the form of shaded values. Significant results are shaded in black, whereas non-significant results are shaded in gray and typically fall close to zero.

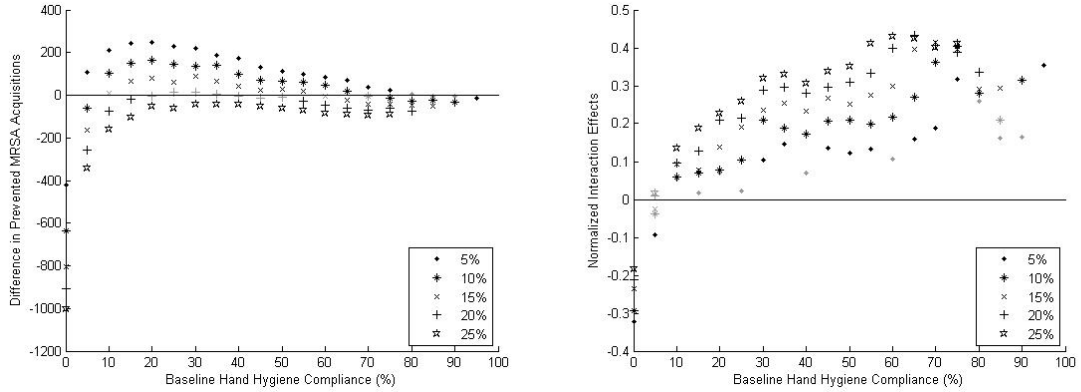


Figure 3.3: Summary of difference values between hand-hygiene compliance main effects and the main effect of increasing the nurse-to-patient ratio from 1:3 to 1:2 (left) and the corresponding normalized interaction effects (right). Non-significant results are shaded in gray ($\alpha = 0.05$).

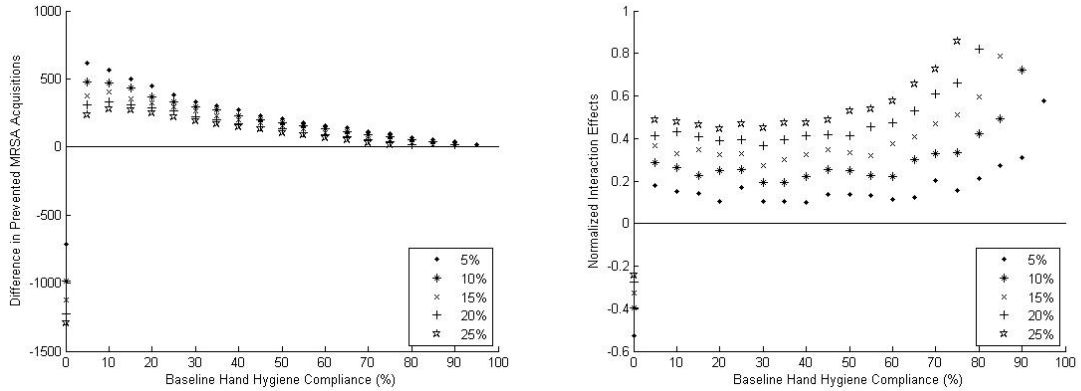


Figure 3.4: Summary of difference values between hand-hygiene compliance main effects and the main effect of increasing the nurse-to-patient ratio from 1:2 to 1:1 (left) and the corresponding normalized interaction effects (right). Non-significant results are shaded in gray ($\alpha = 0.05$).

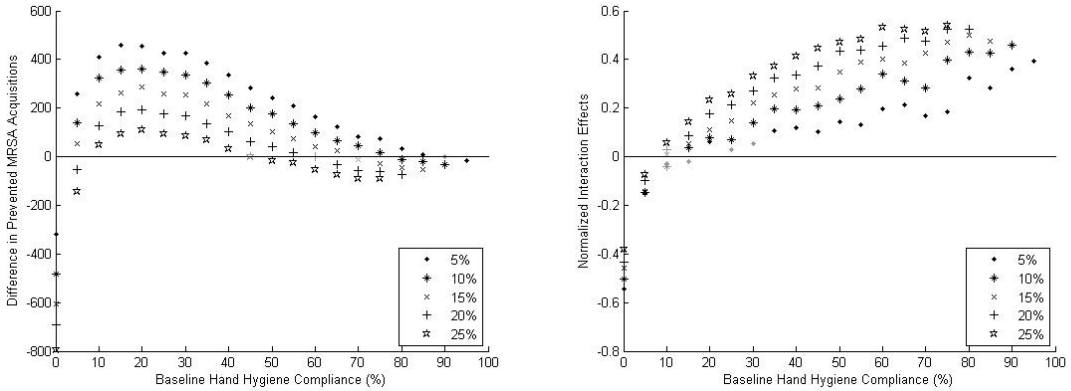


Figure 3.5: Summary of difference values between hand-hygiene compliance main effects and the main effect of increasing the nurse-to-patient ratio from 1:4 to 1:2 (left) and the corresponding normalized interaction effects (right). Non-significant results are shaded in gray ($\alpha = 0.05$).

As shown in Figure 3.4, increasing the nurse-to-patient ratio from 1:2 to 1:1 outperforms all other improvements in hand-hygiene compliance. These results are statistically significant, which suggests this implementation would almost certainly minimize transmission in an ICU setting. We do not provide figures that summarize improvements to a 1:1 ratio from ratios such as 1:3 or 1:4, as these changes are likely to demonstrate an even larger discrepancy between the effects of the two factors. In addition, improvements in hand-hygiene compliance from 0% to higher levels resulted in larger reductions in transmission than any associated change in nurse-to-patient ratio. However, it is unlikely that healthcare workers have a baseline compliance of 0%, so this aspect of the figures will not be discussed in additional detail.

In general, the figures show that increasing the nurse-to-patient ratio outperformed small improvements (i.e., those less than 10%) in hand-hygiene compliance, unless the baseline compliance level was extremely high. The benefit of increasing

the nurse-to-patient ratio appears to reach a maximum at lower hand-hygiene compliance levels, before tapering off and being overtaken by increases in hand-hygiene compliance from higher baseline levels. Larger increases in hand-hygiene compliance can prevent more MRSA acquisitions at lower baseline compliance levels when nurse-to-patient ratios are relatively low (i.e., 1:3 or 1:4).

The interaction effect plots in Figures 3.2 to 3.5 highlight some important trends when considering multiple infection control measures. These figures show that interaction effects increase as the baseline hand-hygiene compliance level becomes higher, which indicates that there is a diminishing return associated with combining increases in nurse-to-patient ratios with improvements in hand hygiene from high baseline levels. These trends support the conclusions from the main effect figures that hospitals that have been successful at improving hand-hygiene compliance can continue to reduce the incidence of MRSA transmission more effectively by further increasing the compliance of healthcare workers. At lower baseline compliance levels, interaction effects are much smaller, which indicates that improvements in compliance can be combined with increases in the nurse-to-patient ratio without losing a significant amount of the potential benefit. Interaction effects for increases to a 1:1 nurse-to-patient ratio are large for any improvement in compliance, and become very large for improvements from high baseline compliance levels. This trend suggests that if a 1:1 nurse-to-patient ratio is achieved, then hand-hygiene compliance becomes a less critical factor with respect to spreading MRSA among patients, assuming productive interactions (i.e., those that result in MRSA transmission) between healthcare workers are minimal.

3.4 Conclusions

We compared the relative effectiveness of improving the hand-hygiene compliance of nurses and increasing the nurse-to-patient ratio by applying a full 2^k factorial design to the output of an agent-based simulation. Our results showed that when baseline hand-hygiene compliance is low, it may be more effective to hire additional nurses than to rely on improvements in hand hygiene, unless those improvements are very large and are achieved very quickly. At high baseline levels of compliance, it may be more effective to continue focusing on improving compliance than to increase staffing levels.

Previous studies have not quantified the differences between the interventions in detail. This deficiency is addressed by the results in Figures 3.2 to 3.5, which compared the performance of the two factors across a wide range of performance levels. In addition, we have provided some indication as to how large an improvement in hand hygiene from a given baseline level is required to outperform a change in the nurse-to-patient ratio.

By applying factorial design methods iteratively over a wide range of parameter values, we demonstrated the effectiveness of this methodology for evaluating infection control measures. Main effects have a straightforward application, because they represent the effectiveness of various control measures in reducing MRSA acquisition rates. Interaction effects are also important, because they provide insight with respect to the efficiency of improving multiple control measures simultaneously. Some infection control strategies may seek to maximize effectiveness, whereas others

may require more efficiency, and these metrics help to differentiate between the two strategies. Ideally, with unlimited resources, hospitals would be able to institute multiple interventions aimed at reducing the incidence of hospital-acquired infections. However, hospitals and resident epidemiologists often face limited budgets and thus need to make choices among different interventions.

Our results did not include an analysis of the relative costs associated with each intervention, but this type of analysis could be performed in future studies. At first glance, it would seem less expensive to improve hand-hygiene compliance than to hire additional nurses. The average cost of an ICU nurse in the United States is approximately $62,733 [77]. However, improving hand hygiene and achieving a sustained increase in compliance has proven to be a difficult task, and there is little data to support how much improvement can be attained for a given investment. Many interventions have led to transient increases in hand-hygiene compliance, but few studies have been able to demonstrate sustained improvement, which brings into question whether it is possible to achieve a sustained increase in the hand-hygiene compliance of healthcare workers [12].

Our results reinforce the importance of basing decisions concerning infection control strategies on individual hospital circumstances, which may assist hospitals in reducing MRSA acquisition rates by taking into account their baseline hand-hygiene compliance rates and staffing ratios. There are situations when improving multiple factors is necessary, such as an outbreak, and other situations when improving one factor can reduce MRSA acquisition sufficiently well. Understanding the interaction effects between infection control measures can help determine the best approach. In

future work, we may investigate the relative effectiveness of a larger set of infection control measures, which would likely require a more robust experimental design methodology to account for non-linearities in multiple dimensions.

Chapter 4

Contribution of Inter-Facility Patient Movement to Overall MRSA Prevalence Levels

4.1 Overview

We investigate the effects of patient movement between hospitals and long-term care facilities (LTCFs) on the long-term prevalence of methicillin-resistant Staphylococcus aureus (MRSA). Typically, hospitals contain many interconnected units, and the length of stay for patients is usually short. Prevalence of MRSA varies widely between hospital units, but there are a number of infection control measures such as promoting hand hygiene for healthcare workers, active surveillance, and contact precautions that can be implemented to minimize transmission [33]. LTCFs house elder patients who are typically more susceptible to acquiring MRSA and are more likely to remain colonized due to both internal and external factors [35]. In addition, there is an objective to preserve the quality of life for residents in the facility. Thus, there are trade-offs when considering options for infection control [80].

Patients are frequently transferred between these two types of facilities. Therefore, it is important for infection control practitioners to recognize the impact of this movement on MRSA prevalence in both types of facilities [14]. Lesosky et al.

implemented a Monte Carlo simulation model of a healthcare facility network to explore the effects of patient transfers on MRSA prevalence levels [52]. However, our results focus on the effects on individual facilities, whereas the Lesosky model examined the effects of patient transfers at the system level. Smith et al. examined a similar problem using a mathematical model, but focused mostly on transmission dynamics between healthcare facilities and the community [79]. This work did not compare how a change in the transmission level of one facility affected the long-term prevalence of other healthcare facilities.

In this chapter, we address the following questions:

1. Can patient movement from a hospital affect the prevalence of MRSA in a LTCF?

2. Can patients from a LTCF increase the prevalence of MRSA in a hospital if they are sent to various units?

3. Can patients from a LTCF cause an increase in the prevalence level of MRSA in a particular hospital unit?

4.2 Methodology

We develop a hybrid model of a healthcare system using NetLogo (v4.1.1) [64], an open-source, agent-based modeling tool, that employs a mathematical model to simulate patient movement between hospitals and LTCFs. Transmission within each facility is modeled using a modified version of the susceptible-infected-recovered (SIR) equations (see Figure 4.1) [4]. We use these equations to simulate proportions

SIR

$$\frac{dS}{dt} = -\beta SI$$

$$\frac{dI}{dt} = \beta SI - \gamma I$$

$$\frac{dR}{dt} = \gamma I$$

Modified SIR

$$\frac{dU}{dt} = -\mu U + \mu u - \beta U(T+P) + \gamma_l P + \gamma_s T$$

$$\frac{dP}{dt} = -\mu P + ((1-u)\mu + \beta U(T+P))p - \gamma_l P$$

$$\frac{dT}{dt} = -\mu T + ((1-u)\mu + \beta U(T+P))r - \gamma_s T$$

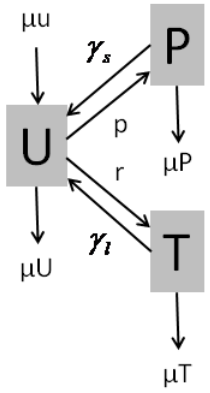


Figure 4.1: Susceptible-Infected-Recovered (SIR) and modified SIR Model Equations.

of uncolonized patients (U), persistently colonized patients (P), and transiently colonized patients (T) in each facility. The uncolonized patients in our model correspond to the susceptible patient state in the original model, whereas the persistently and transiently colonized patients are variations of the infected state. Infected patients are included within the respective colonized patient states, because there is limited data suggesting they are more likely to spread MRSA to others than colonized patients. There is no recovered state in this model, as patients who clear colonization become immediately susceptible to re-colonization.

The modified equations consist of admission and discharge (μ), transmission (β), and recovery (γ) rate terms. Positive terms represent an increase in the proportion of patients for the corresponding patient state, while negative terms represent a decrease. Patients can be admitted in any one of the three patient states. The pro-

portions of patients admitted into each state are represented by the terms u, p, and r, respectively. The transmission term (β) represents the rate at which uncolonized (i.e., susceptible) patients enter a colonized state. The terms p and r also represent the proportions of newly colonized patients who enter each respective state. Recovery rates for persistently and transiently colonized patients are represented by γ_l and γ_s, respectively, with $\gamma_l \ll \gamma_s$.

Each facility is modeled as an agent in a network of healthcare facilities, and has a unique state that consists of proportions of the three patient types. This agent-based formulation allows us to analyze various network configurations to determine the effect of patient movement on MRSA prevalence rates in each facility. Each scenario is configured as a network of facilities with directed links that represent patient movement from one facility to another (see Figure 4.2). The links have associated weights (m_{ij}) which, in conjunction with the facility size (N), determine the number of patients who move from one facility to another each time step. When patients are transferred from one facility to another, the number of patients in each state is representative of the relative proportions in the source facility. For example, if proportions of patients in the source facility are $U = 0.80$, $P = 0.05$, and $T = 0.15$, and 20 patients are transferred, then 16 patients will be uncolonized, one patient will be persistently colonized, and three patients will be transiently colonized.

The agent-based modeling framework enables the comparison of infection control strategies at one or more facilities that may be at risk of increased prevalence levels due to an influx of received colonized patients. One option for healthcare facilities is to start these patients on a decolonization regimen, which reduces the

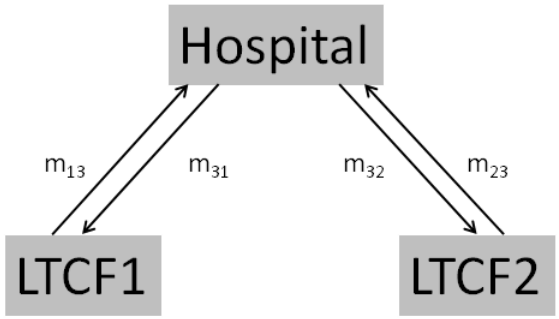


Figure 4.2: Example of an Inter-Facility Model.

number of colonized patients who will ultimately interact with susceptible patients. The regimen would likely consist of a mupirocin application to the nasal cavity, chlorhexidine bathing, or both [59, 84]. This control measure could be implemented in one of three ways. The first approach is to start all patients on the decolonization regimen immediately upon their arrival to the facility. These patients would not be isolated from the general population, and the benefits of decolonization would not be realized until the patients complete the process successfully, which takes two cycles of five-day treatments on average [13]. This approach, although simple to implement, would put a significant number of uncolonized patients through the decolonization process unnecessarily and could lead to increased resistance to the decolonization process.

The other two approaches screen arriving patients to determine their colonization status, so that only colonized patients undergo the decolonization process.

Culture-based screening can take up to three days for results, whereas polymerase chain reaction (PCR) testing can return results within a day. Culture-based screening is less expensive, but there may be an increased risk of secondary transmissions if the patient continues to interact with other patients or healthcare workers without control measures in place. A PCR test is more sensitive and allows for control measures to be implemented sooner, but it is more expensive [31].

For the first approach, all patients returning to the LTCF from the hospital begin a decolonization regimen immediately. This process returns them to the uncolonized state in 10 days. For the second approach, patients who arrive at the LTCF undergo a basic culture screening. If a positive test result is produced three days later, the patient starts a decolonization regimen and recovers in an additional 10 days. Uncolonized patients remain in the susceptible state. Finally, for the PCR case, all patients receive their test results in one day, and the colonized patients then start the decolonization process.

4.3 Results

Simulation experiments were conducted to address the three questions posed in the Overview. Parameters for the generic hospital, LTCF, and hospital unit are summarized in Table 4.1. These parameters apply to each agent facility in the model. These parameters were not based on specific healthcare facilities, but were chosen to represent relative sizes, admission and discharge rates, and transmission levels for typical hospitals and LTCFs in the United States [37, 61, 63]. Facilities are

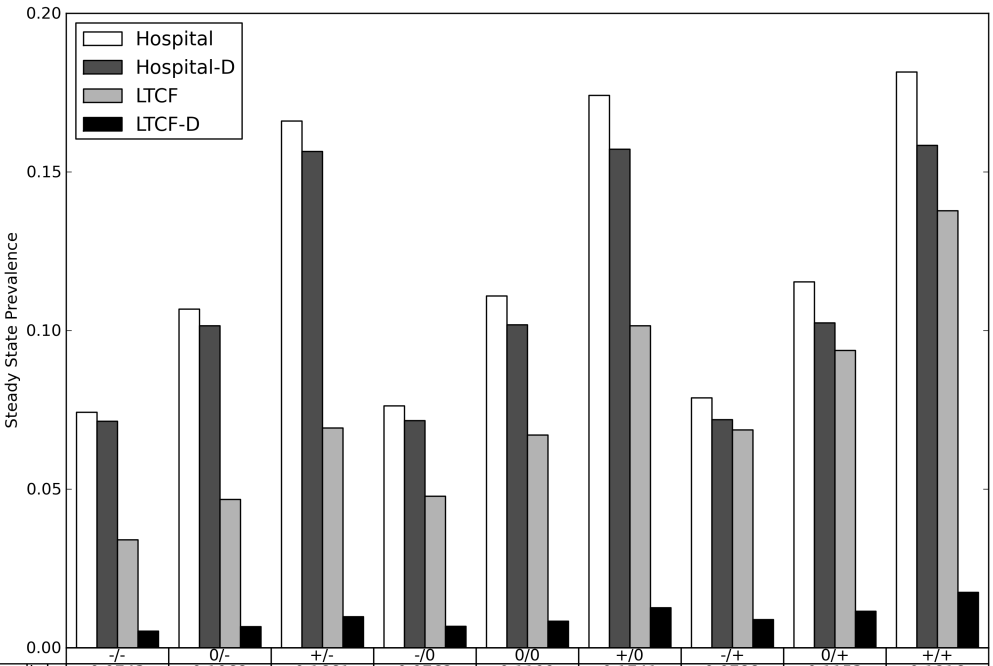
Table 4.1: Long-term care facility model parameters.

Parameter	Hospital	LTCF	Hospital Unit
N	300	100	20
μ	0.2	0.002	0.2 (high), 0.05 (low)
β	0.15 (low) 0.25 (medium) 0.35 (high)	0.05 (low) 0.075 (medium) 0.1 (high)	0.15 (low) 0.25 (medium) 0.35 (high)
u	0.9	0.9	0.9
p	0	0.6	0
r	1.0	0.4	1.0

initialized with an entirely susceptible population, and recovery rates for persistently and transiently colonized patients were held constant at 0.02 and 0.2, respectively. Prevalence is measured as the sum of proportions of transiently and persistently colonized patients, which implicitly includes infected patients. As a baseline, hospitals and LTCFs exchange nine patients at each time step, whereas hospital units and LTCFs exchange two patients, so that the size of each facility remains constant throughout the simulation period. Relaxing this assumption of equal exchange rates increased the variability of the transient prevalence (i.e., at each time step), however even large net patient flows had little effect on the steady-state prevalences in either facility type.

Can patient movement from a hospital affect the prevalence level of a LTCF?

The first set of experiments paired one hospital with one LTCF. All parameters were held constant except for the transmission rate parameter (β), which was varied for both facilities until all unique transmission level pairs were simulated. The results are summarized in Figure 4.3, which shows MRSA prevalence for both facility



	-/-	0/-	+/-	-/0	0/0	+/0	-/+	0/+	+/+
Hospital	0.0743	0.1068	0.1661	0.0762	0.1109	0.1741	0.0788	0.1153	0.1816
Hospital-D	0.0714	0.1015	0.1565	0.0716	0.1018	0.1572	0.0719	0.1025	0.1584
LTCF	0.0340	0.0468	0.0693	0.0478	0.0671	0.1016	0.0686	0.0938	0.1378
LTCF-D	0.0053	0.0066	0.0098	0.0067	0.0084	0.0126	0.0088	0.0115	0.0175

Figure 4.3: Comparison of steady-state prevalence levels for a hospital and a LTCF. Hospital-D and LTCF-D represent the scenario in which the LTCF has implemented a decolonization policy. The labels on the x-axis are expressed in pairs of hospital/LTCF transmission levels.

types for nine hospital-LTCF transmission level combinations. The transmission level of the hospital and LTCF specify each configuration. For example, the $+/0$ configuration means a hospital with a high transmission rate was exchanging patients with a LTCF that had a medium transmission rate. The $-/+$ configuration means a hospital with a low transmission rate was paired with a LTCF that had a high transmission rate. Transmission rate parameters differed for hospitals and LTCFs because each facility type had a different rate of patient turnover, which affected the steady state prevalence level.

Figure 4.3 shows that the steady-state prevalence of both facilities is strongly correlated to the transmission rate in the hospital. As a baseline, consider the first three cases ($-/-$, $0/-$, $+/-$) where a LTCF with a low transmission rate is receiving patients from a hospital with increasingly higher transmission rates. There is an increase in the steady-state prevalence of the LTCF from 3.4% to 4.7% to 6.9% as the transmission level of the hospital increases within the series. At higher LTCF prevalence levels, the increase in steady-state prevalence grows at a faster rate (e.g., transmission for the high prevalence LTCF increases from 6.9% to 9.4% to 13.8%). This trend suggests that LTCFs with higher transmission rates are more susceptible to outbreaks when receiving patients from high prevalence hospitals.

There are two key factors that explain how hospitals can cause increased prevalence levels in LTCFs. The first factor is the relative size difference between the two facility types. Hospitals are typically larger than LTCFs. Therefore, when patients are exchanged, the number of colonized patients sent from the hospital to the smaller LTCF represents a larger proportion than they did in the hospital, which creates an increased level of colonization pressure. The second key factor is the turnover rate. Patients in LTCFs are likely to reside in the facility for longer periods of time. Consequently, when a hospital sends patients to a LTCF, those patients are ultimately more susceptible to colonization because they spend more time in the facility. In addition, transmission continues to build momentum because the population is relatively stable.

LTCFs have some ability to protect themselves from colonized patients who are either newly admitted or returning from an acute-care hospital. We compared three

LTCF decolonization strategies for the same nine cases to determine their effect on steady-state prevalence of both facility types. The results of these experiments are given in Figure 4.3. The outcomes for all three decolonization strategies were approximately the same. Therefore, the results are summarized as a single data set. The similarity between these results is mostly attributable to the regularity with which patients move between facilities and start the decolonization process. After the first group of arriving patients successfully completes the decolonization process, a new group of patients will complete the process the next day, and so on, regardless of the screening test return time or the recovery period, which are both held constant for all patients. As a result, the long-term LTCF prevalences are approximately equal, because each scenario differs only in the initial time frame before patients start to become decolonized.

The key result from these experiments is that any decolonization strategy can significantly decrease the risk of increased prevalence levels for LTCFs receiving patients from hospitals. Under decolonization, the LTCF can identify patients who acquired MRSA in a hospital or returning patients who were already colonized. Hospital prevalence levels also showed marginal improvement under the LTCF decolonization program. Our model does not account for the cost of the various approaches in terms of the number of screening tests and decolonization supplies or the effect of each approach on a resident's quality of life. Decolonizing all patients has no cost with respect to screening, but requires more supplies for the regimen and may reduce the quality of care for patients who have adverse reactions to the process. The screening strategies reduce the number of patients who undergo decolonization,

which improves quality, but both have associated costs due to the testing procedures.

Can patients from a LTCF increase the prevalence of a hospital if they are sent to various units?

The answer to this question is derived from the results in Figure 4.3. It is clear from this figure that the steady-state prevalence of the hospital only changes significantly with its own transmission level. For example, comparing the first $(-/-)$, fourth $(-/0)$, and seventh $(-/+)$ cases shows a negligible change in the prevalence of the hospital as the transmission level of the LTCF increases. Even changes at higher transmission levels (cases 2, 5, and 8; cases 3, 6, and 9) lead to only small changes in steady-state prevalence. Therefore, changes in the transmission level of the LTCF have little to no effect on MRSA prevalence in the hospital when patients are sent to multiple units.

Facility size and patient turnover also explain why LTCFs have little effect on the prevalence level of hospitals as a whole. Hospitals are typically large in size, so a few additional colonized patients in the population will have little effect on the relative proportions of patient state. Colonization pressure will essentially remain unchanged. In addition, hospitals admit and discharge patients more frequently than LTCFs. Therefore, the hospital population is changing rapidly and the transmission process is continually interrupted.

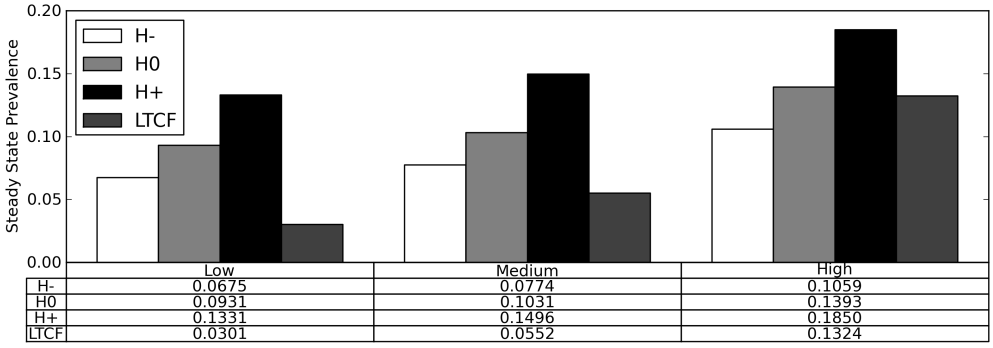
Can patients from a LTCF cause an increase in the prevalence level of a particular hospital unit?

In some cases, patients from a LTCF are sent to the same unit within a hospital on a consistent basis. This situation can occur in intensive care units or on geriatric floors, which usually receive patients who reside in LTCFs. For these cases, the population size is the size of the hospital unit, which is much smaller than the hospital and usually smaller than the LTCF. The turnover rates can vary in these types of units. Therefore, we conducted experiments using a high turnover rate and a low turnover rate (see Table 1). These experiments were conducted simultaneously, with a single LTCF exchanging patients with one low, one medium, and one high prevalence hospital unit. The results are summarized in Figure 4.4.

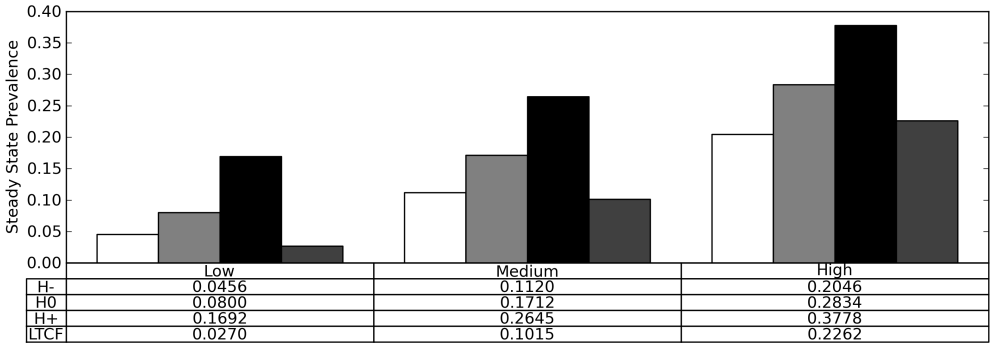
The data from the high turnover case in Figure 4.4 shows that a LTCF can have a noticeable effect on the prevalence within a specific hospital unit. In these types of units, the addition of one or two additional colonized patients may be sufficient to cause a significant outbreak. The low turnover case in Figure 4.4 presents a different outcome, as patients remain in the hospital unit for longer lengths of stay and, therefore, have an increased exposure to colonization. In all three cases of hospital transmission levels, there is a significant increase in the unit prevalence. In addition, there is a reciprocal effect that increases the LTCF prevalence to higher levels than in the high turnover case.

4.4 Conclusions

Our results suggest there are two primary factors that lead to sustained increases in MRSA prevalence when healthcare facilities are exchanging patients. The



	Low	Medium	High
H-	0.0675	0.0774	0.1059
H0	0.0931	0.1031	0.1393
H+	0.1331	0.1496	0.1850
LTCF	0.0301	0.0552	0.1324



	Low	Medium	High
H-	0.0456	0.1120	0.2046
H0	0.0800	0.1712	0.2834
H+	0.1692	0.2645	0.3778
LTCF	0.0270	0.1015	0.2262

Figure 4.4: Comparison of steady-state prevalence levels for three hospital units with low, medium, and high transmission levels that exchange patients with one LTCF with low, medium, and high transmission levels. High (top) and low (bottom) turnover rates were investigated.

patient departure rate, which incorporates the discharge and transfer of patients, plays the most significant role in whether or not a facility is susceptible to external influence. When patients have longer lengths of stay, they are visited more by healthcare workers who may have become transiently colonized. Thus, the risk to the facility is minimized when patients are discharged sooner. The facility size is also critical, because colonized patients represent larger proportions in smaller facilities and, therefore, contribute more to colonization pressure per capita. Consequently, small populations or facilities with low turnover rates are especially susceptible to increased prevalence levels. These facilities become more at risk when they receive patients from larger, high-prevalence facilities. These types of facilities can protect themselves by implementing control measures, such as bundling active surveillance with decolonization, which ultimately limit the interactions between susceptible and colonized individuals.

There are several limitations to our model and the simulation experiments we performed. The mathematical model is sensitive to the admission and discharge, transmission, and movement terms. The effect of the transmission parameter depends on the values of the other two parameters. When patients leave the facility at a high rate, transmission is less likely to occur than if the patient population was more stable. A more extensive study would explore wider ranges of these parameters to assess their relative effects on long-term prevalence levels. In addition, patient admissions and transfers occurred on a regular, periodic schedule in blocks of fixed sizes. In practice, these patient flows are less regular and probably impact the long-term effects of facility interactions and control measures.

Aside from the limitations, our model provides a unique perspective on exploring the effects of patient movement on long-term MRSA prevalence in hospitals and LTCFs. Combining agent-based and mathematical modeling techniques allowed us to adequately represent transmission dynamics and analyze the interconnectivity of multiple healthcare facilities that exchange patients. Future work may focus on developing an agent-based model for the transmission dynamics within each facility, which would not be limited by the restrictions of mathematical models and could begin to explore the effects of individual contact networks. In addition, we could explore actual networks of healthcare facilities that regularly exchange patients.

Chapter 5

Exploring the Effects of Network Structure and Healthcare Worker Behavior on the Transmission of Hospital-Acquired Infections

5.1 Overview

The problem of hospital-acquired infections has been well publicized, and infection control measures have been studied extensively [24]. However, research on this problem has not focused on the interactions among patients, nurses, and physicians. In general, we would like to have high ratios between the number of healthcare workers (HCWs) and patients in a specific unit. The underlying network of connections among patients, nurses, and physicians and the nature of those interactions are also important factors. This chapter, building on earlier work [8], investigates how network structure in a hospital unit can affect patient-to-patient transmission of infectious diseases.

With this network model, we hope to accomplish three key objectives. First, we model a network of patients in a hospital unit and the transmission of an infectious disease that passes among patients through the HCWs who care for them. The network structure plays an important role in transmission, and there may be observable trends related to the structure that are important to understand [47, 71]. Second, we investigate the effect of the behavior of HCWs, who may cover each

other's patients temporarily or transmit the disease among each other. The practice of sharing patients is common in hospitals, but it is not well understood how it can change the dynamics of transmission. Also, HCW-to-HCW transmission has not been studied extensively, because many diseases are not likely to spread among HCWs who have only casual contact with one another. Third, we move from a model with a stable population to one with patient turnover, and examine the effects on network structure, patient sharing, and HCW-to-HCW transmission. By focusing on these objectives, we hope to gain a better understanding of network structure, HCW behavior, and patient turnover, and then we can offer strategies to healthcare organizations that are likely to minimize transmission.

5.2 Methodology

Agent-based modeling and simulation (ABMS) is a methodology that focuses on the interactions among individuals and then aggregates that behavior into a system that can be analyzed [53]. Each individual agent can have different characteristics and behavior, which is an advantage of ABMS not afforded by equation-based modeling. We developed our patient network model using NetLogo 4.1.1, an open source agent-based modeling development platform (http://ccl.northwestern.edu/netlogo/), and conducted experiments to explore the effect of network structure and HCW behavior on transmission.

5.2.1 Baseline Conceptual Model

We begin with a model of a single hospital unit with a static patient population and two HCW types. We explicitly define only patients as agents, with a single state that indicates whether or not the patient is infected or colonized with some type of pathogen. There are two HCW types-nurses and physicians-and we model them implicitly through the transmission mechanism. Each patient has a primary nurse and a primary physician who provide care during his or her stay. We control the proportion of visits performed by each HCW type because, in practice, nurses are likely to visit patients more frequently. At each time step, a random number is drawn to determine whether nurses or doctors will visit patients. All HCWs of the selected type then visit a single patient in their respective cohorts at random.

Because HCWs care for multiple patients, there is an underlying network that connects patients who share nurses and a separate network for patients who share physicians. Figure 5.1 highlights the difference between a dense network-one that has few HCWs and many connections between patients-and a sparse network that has few connections. The patient network diagrams in Figure 5.1 show 20 patients in a hospital unit with two physicians whose patients are indicated by the black and gray colors of the agents, respectively. The dense case on the left of Figure 5.1 shows how many connections exist in the nurse network when there are only two nurses. In this case, there are 10 patients in each nurse cohort who are all connected. The sparse case shows that the number of connections between patients can be reduced significantly when there are 10 nurses. The same reduction could be realized with

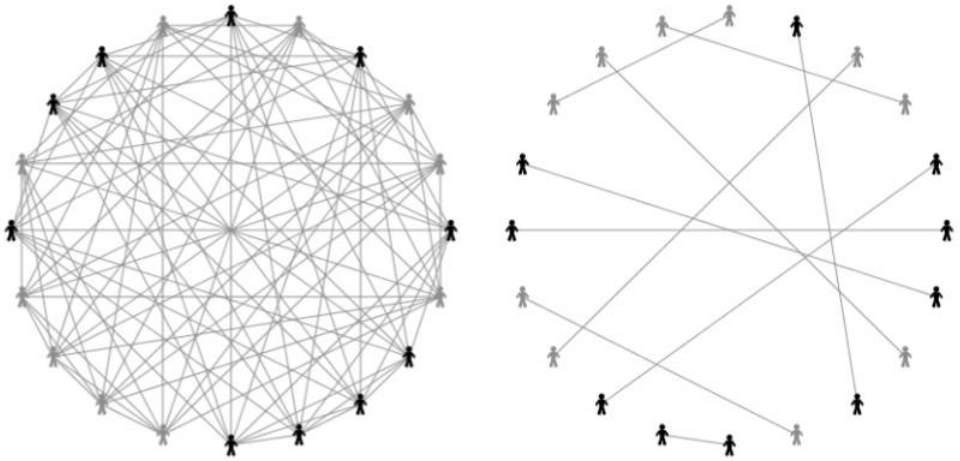


Figure 5.1: A dense patient network (left) and a sparse patient network (right). Patients who share a nurse are connected by a link whereas patients who share a physician are shaded with the same color. The dense network represents a hospital unit with two nurses and two physicians. Ten nurses and two physicians care for patients in the sparse network.

additional physicians, but hospitals are less likely to hire physicians for infection control purposes due to their high salaries.

Patients can become infected only if there is a source patient who shares a nurse or physician. Consequently, susceptibility implies that there is a non-infected patient who has at least one connection to an infected patient at some point in time. Patients who are connected by both a nurse and a physician have an increased likelihood of transmission if either HCW becomes infected. Each simulation is initialized with a single infected patient, and the simulation runs until all susceptible patients become infected. There is a single model parameter that defines the probability of an infected individual (i.e., a patient or a HCW) transmitting the pathogen to a susceptible individual (we call this parameter virulence). Virulence takes several factors into account, including the probability of transmission between susceptible and infected

agents, the probability that a HCW washes his or her hands, and the probability that the hand washing successfully removes the pathogen. This parameter can be adjusted to represent changes in any of these factors.

Patients can also be visited by secondary HCWs who may be covering the patient temporarily for the primary HCW. This type of sharing can occur in practice when HCWs are on a break or attending a meeting. Sharing also occurs more often in academic medical centers where HCWs are collaborating in the care of patients. Patient sharing modifies the nurse and physician patient networks and creates temporary paths for transmission to spread to the rest of the unit. We investigate several configurations of sharing to determine how transmission can be minimized under these circumstances. In our model, we specify nurse and physician sharing rates that control how often the secondary HCW visits the patient. This mechanism creates a dynamic patient network that contains connections that were not in the initial network. Our goal is to understand how these temporary connections affect the extent and rate of transmission in the unit.

We investigate the four configurations of patient sharing shown in Figure 5.2. Each configuration shows four cohorts of three patients. Within each cohort, all patients are connected because they share the same HCW. When HCWs share patients, temporary connections enter the network between patients in different cohorts, depending on the sharing configuration. In the first configuration there is no patient sharing. These cohort connections are bidirectional, because the secondary HCW can either introduce transmission into the shared cohort or acquire the disease from the shared cohort and bring it back to his or her primary patients. The second

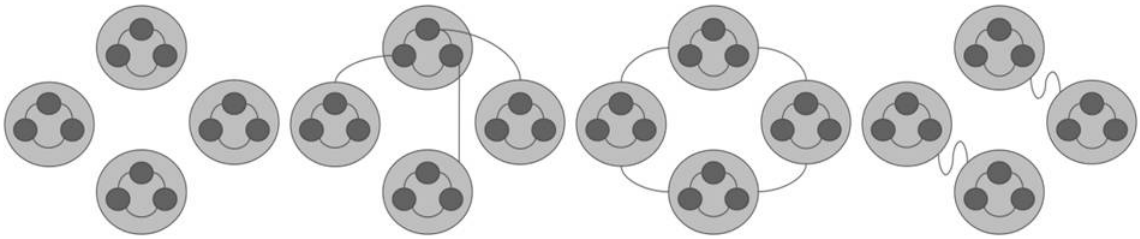


Figure 5.2: Patient sharing configurations (from left to right: none, random, revolving, and paired).

configuration is random sharing. For this case, any HCW other than the primary HCW is selected at random to visit the patient in need of care. This configuration creates a small-world effect [85] where transmission can occur between patients who did not initially share a connection. The third configuration is revolving sharing. This configuration creates a circular network in which each HCW covers all patients of one HCW and his or her patients are covered by a different HCW (e.g., HCW A covers for HCW B who covers for HCW C who covers for HCW A). In the fourth configuration, paired sharing matches two HCWs together who each serve as the secondary caregiver for the partner's primary patients. Paired sharing creates additional links in the network. However, this configuration tries to keep the network as disconnected as possible in order to limit transmission.

5.2.2 Initial Results

The first set of experiments focuses on the structure of the patient network. We restrict our experiments to a model of a 20-patient hospital intensive care unit (ICU). Each simulation is run until all susceptible patients have become infected.

Monte Carlo replications are conducted to account for the stochastic effects. The simulation time, in number of ticks (an arbitrary NetLogo unit of time), and the number of transmissions are recorded. In our model, ticks essentially equate to the amount of time between HCW visits to an individual patient, which depends on the number of patients, number of HCWs, number of visits each day, and the proportion of visits by nurses and physicians.

Our primary metric for these experiments is the ratio of the total number of ticks to the total number of transmissions, which we call the mean time to transmission. This metric is a better indicator of transmission than simply the number of ticks, because it adjusts for cases in which some patients are not connected to an infected patient. Obviously, it is in the best interest of the patient for hospitals to implement control measures that increase the mean time to transmission as much as possible. ABMS allows us to track the number of transmissions due to each HCW type, so that we can gain an understanding of the threat each poses to patients.

Given the number of patients and HCWs, we can compute the density of the nurse and physician networks as the ratio of connections between pairs of patients to the number of links in the complete network in which all patients are connected. For example, the sparse network in Figure 5.1 depicts an ICU with 20 patients, 10 nurses, and two physicians would have 10 nurse cohorts of two patients each and two physician cohorts with 10 patients each. The nurse network has 10 cohorts of two patients who are connected. The physician network has two cohorts of 10 patients who are all connected. The total number of connections in each physician cohort is $\binom{10}{2} = 45$ because each patient is connected to nine neighbors. The total

number of connections is $\binom{20}{2} = 190$. Thus, the nurse network has a density of $^{10}/_{190} = 0.0526$ and the physician network has a density of $^{90}/_{190} = 0.474$. The nurse network density in the dense network scenario with only two nurses, shown in Figure 5.1, is also 0.474. In general, density (d) for a unit with n total patients, k cohorts, and i_k patients in each cohort is given by

$$d = \frac{\sum_k \binom{i_k}{2}}{\binom{n}{2}} \tag{5.1}$$

The HCW network density varies according to the trend shown in Figure 5.3. This trend shows a diminishing marginal effect as the number of HCWs increases. This effect translates to a reduction in transmission, in both the number of patients susceptible to infection and the rate at which that transmission occurs. In addition, network density offers a normalized metric that can be used to compare hospital units with different sizes and configurations. These types of comparisons are not as easy to make using absolute parameters (e.g., HCW-to-patient ratios), and they become more difficult when comparing units that have different cohort sizes.

We performed experiments in which we varied the nurse and physician network densities to determine the effect on transmission. Assuming that nurses visit patients more often than physicians, we configured the model to have nurses perform 80% of the visits, while physicians performed 20% of visits. The effect of network density on our system is summarized in the form of contour plots that shade areas from light to dark as the value of the response increases. In Figure 5.4, we show the effect of network density on the mean time to transmission using a contour plot.

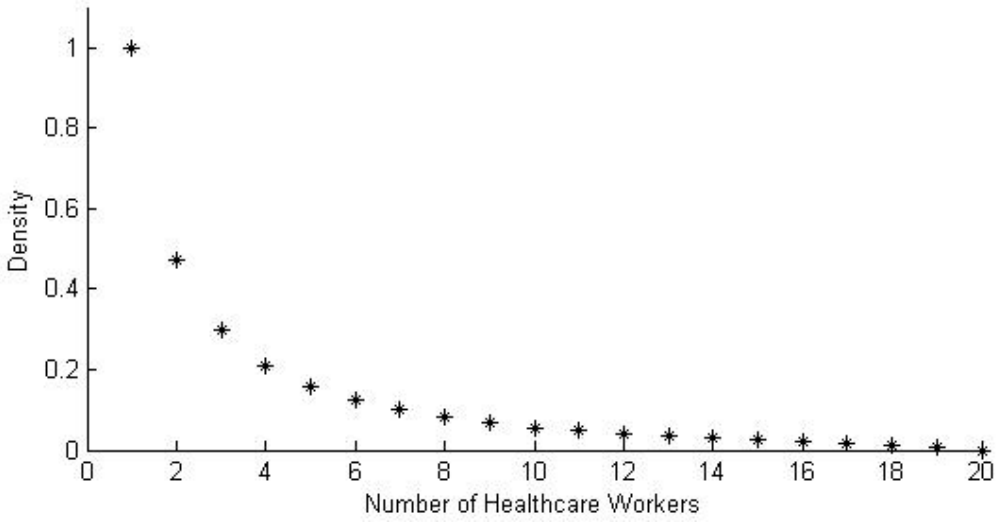


Figure 5.3: Density of the HCW network as a function of the number of HCWs in a 20-patient ICU.

For example, at point A, the nurse and physician network densities are both equal to 0.474, which corresponds to two nurses and two physicians in our 20-patient ICU. At this point, the system responded with a mean time to transmission of approximately 822 ticks, which is represented in the plot with a light gray color. At point B, we have a nurse network density of 0.126 and a physician network density of 0.158, which corresponds to a much sparser network. As a result, the mean time to transmission nearly doubles to approximately 1600 ticks, which is represented on the plot by a darker shade. Highly dense nurse networks–with density values of 0.3 and higher–allow transmission to reach the entire unit quicker. Even in moderately dense nurse networks–densities ranging from 0.1 to 0.3–we only observe a moderate increase in the mean time to transmission. In order to maximize the average time between transmissions, a nurse network density of 0.1 or less is required for our 20-patient ICU.

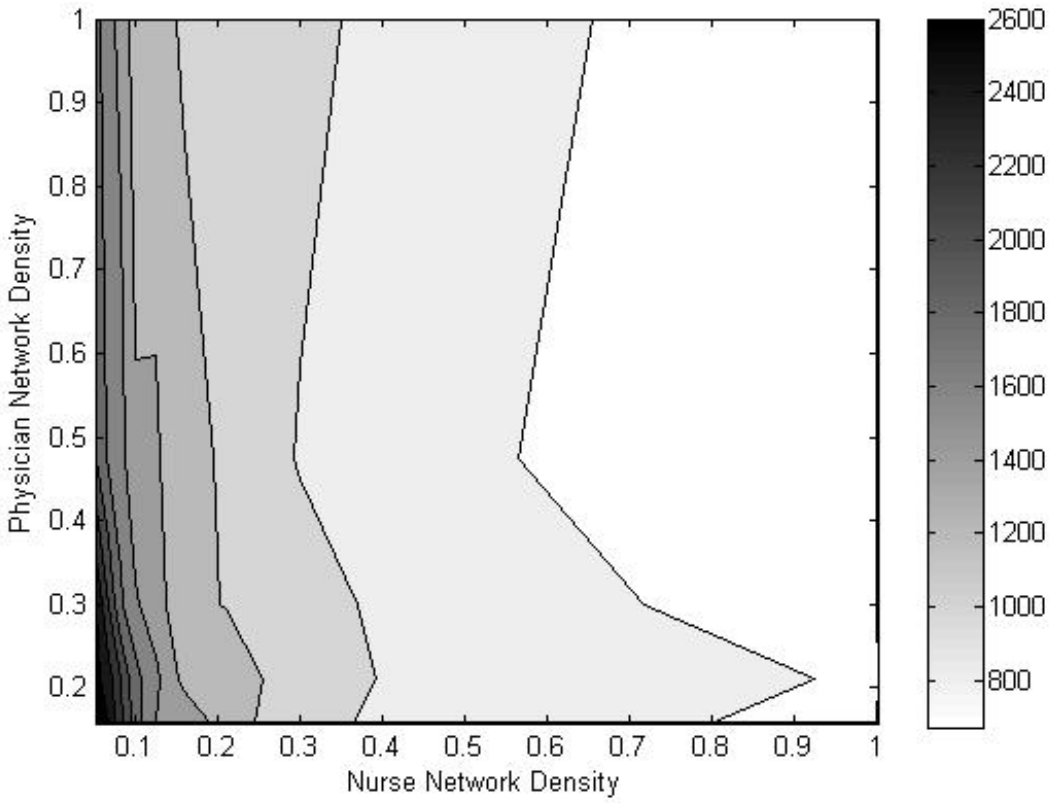


Figure 5.4: Mean time to transmission as a function of nurse and physician network densities.

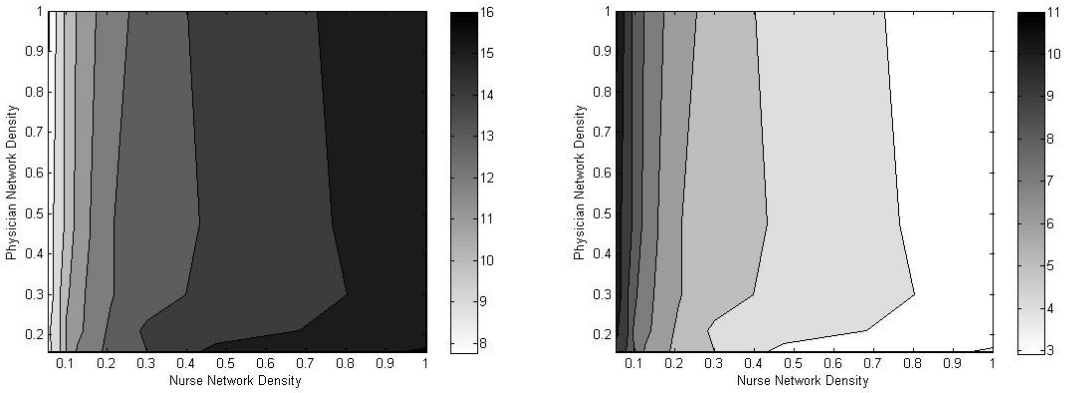


Figure 5.5: Transmissions due to nurses (left) and physicians (right) as a function of nurse and physician network densities

5.2.3 Cohort Alignment

In Figure 5.5, we show contours for the number of transmissions due to nurses and physicians. This figure shows that transmission is strongly related to nurse network density, but it is not affected significantly by physician network density. Despite the lower proportion of transmissions due to physicians, physicians still pose a significant threat in several ways. As shown in Figure 5.5, physician transmissions reach a maximum when the nurse network becomes very sparse. At low nurse network densities, physicians infect more patients because nurses do not infect patients as quickly. Physicians also pose a threat because each physician's network can overlap multiple nurse networks. Therefore, they can transmit the infection to multiple nurse cohorts, which is only possible when there is a HCW that visits many patients [81]. When the infection reaches a new cohort, nurses can spread the infection to the remaining patients. In effect, one physician transmission could lead to many additional transmissions by nurses. Due to the important roles that both nurses and physicians play in transmission, maximizing the mean time to transmission requires keeping both the nurse and physician network densities at their lowest possible levels.

In nearly all cases of network densities, every patient in the ICU eventually became infected due to the overlap of the nurse and physician networks. In an ICU with no physicians and a single index patient, we expect that only the patients who are cared for by the same nurse will become infected. Therefore, if we assign a specific number of nurses to a single physician who cares for all patients in the union

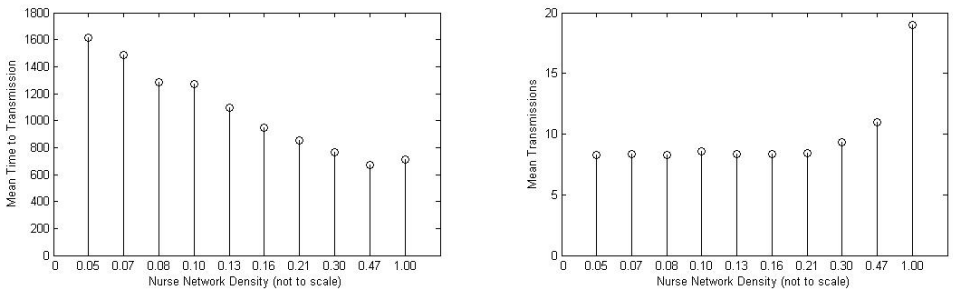


Figure 5.6: Mean time to transmission (left) and mean number of transmissions (right) as a function of nurse network density with cohort alignment.

of the nurse networks, then we are essentially segmenting the patient population in the unit. This control measure minimizes the potential of physicians-who typically are fewer in number and, therefore, see more patients-to spread an infection to patients in different nurse cohorts.

We conducted additional experiments using this alignment strategy. For simplicity, the results in Figure 5.6 are averaged over all physician network density values and are shown as only a function of nurse network density. With the exception of a fully connected nurse network (i.e., one nurse covering the entire unit), cohort alignment reduces the number of infected patients in the unit by approximately half on average, which equates to the number of patients under the care of a single physician. Depending on the nurse network density, this strategy reduces the number of transmissions from 19 to approximately nine. This reduction is likely to be larger in units with more than two physicians. As the mean number of transmissions stabilizes, we continue to see an improvement in the mean time to transmission as the nurse network density decreases.

5.2.4 Patient Sharing

The next experiments focus on the effect of patient sharing among HCWs in an ICU environment. These experiments use both HCW types, but only nurses share patients. In practice, nurses are more likely to cover each other's patients than physicians. Typically, there are a small number of physicians in each unit and there are limited options for sharing.

We simulated a 20-patient ICU with 10 nurses and two physicians using the four sharing configurations for nurses and sharing rates of 10%, 20%, and 30% of patient visits. These results are shown on the left side of Figure 5.7. It is clear from these results that no sharing is the ideal configuration, producing the longest mean times to infection and the fewest number of transmissions due to nurses, which is a good metric for transmission because nurses visit patients more often than physicians. Consequently, cases in which we see higher transmission numbers for physicians suggest that transmission is not occurring as quickly. Random sharing is the worst configuration, with the lowest mean times to transmission and the largest number of transmissions by nurses, especially at higher sharing rates. Random sharing allows transmission to occur between multiple cohorts that would otherwise be disconnected in terms of the HCWs who care for those patients. Transmission occurs very quickly in this case, because the disease can reach multiple cohorts easily and spread concurrently to other cohorts in the unit.

The revolving and paired sharing configurations perform much better than random sharing, although it is not entirely clear which of the two is better. Re-

volving sharing tries to maximize the time for infection to spread to the entire unit, whereas paired sharing tries to restrict transmission to the two cohorts that share nurses. However, physicians negate the intended effects of both strategies because they can spread the infection to other nurse cohorts as well. It appears that at low sharing rates (10% or less), revolving sharing is better than paired sharing, because transmission must pass through successive nurse cohorts to reach the entire unit. However, at higher patient sharing rates (20% or greater), the disease spreads through the network more quickly with revolving sharing. As a result, paired sharing becomes a better strategy, because the only path for infection to spread throughout the unit is through physicians.

The trends concerning patient sharing are more pronounced when the ICU is simulated without physicians. These results are presented on the right side of Figure 5.7, and they confirm that no sharing is the best configuration–only one patient who shares a nurse with the index patient becomes infected. The upward trend in the mean time to transmission for the case with no sharing is not related to the nurse sharing rate as it appears in the figure. Instead, this trend is attributable to the variability in the number of ticks for one patient to become infected. The results also confirm that random sharing is the worst configuration. It allows all patients in the ICU to eventually become infected at a significantly faster rate than the other three configurations. The revolving configuration produces higher mean times to transmission at low sharing rates, while paired sharing performs better at high sharing rates. Without physicians, paired sharing has an advantage. There is transmission to only two nurse cohorts, and there are high mean times to

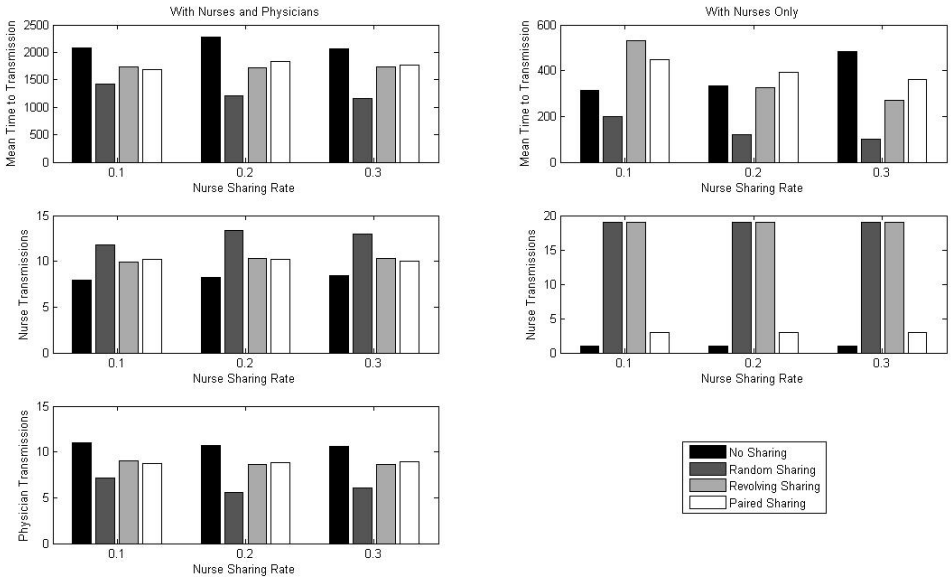


Figure 5.7: Mean time to transmission and mean number of transmissions as a function of sharing configuration and rate, with both nurses and physicians (left) and with nurses only (right).

transmission.

5.3 Healthcare Worker Transmission

Our experiments so far show that the risk of patient-to-patient transmission of an infectious disease-by way of HCWs who become transiently infected-can be reduced by minimizing the densities of the nurse and physician networks. However, achieving this goal requires more HCWs, who become significantly more likely to interact with each other. Some diseases are not likely to spread among HCWs because direct contact is needed (e.g., antibiotic-resistant bacteria), but other diseases may spread more easily (e.g., airborne diseases such as influenza).

5.3.1 Model Implementation

We expanded our patient network model to incorporate HCW-to-HCW transmission and investigated its effect relative to the effect of minimizing nurse and physician network densities. Each HCW is equally likely to interact with every other HCW, which corresponds to a complete network with equally weighted links. Modeling HCW interactions in this way represents the worst-case scenario for HCW-to-HCW transmission. We expect the effects of HCW-to-HCW transmission would be less significant for sparse interaction networks.

We define several parameters that are used to determine the number of HCW-to-HCW transmissions at each time step (see Table 5.1). The value of each parameter is calculated at each time step with the exception of the HCW-to-HCW transmission probability (p_t), which is the only user-defined parameter and is fixed at a constant value throughout the simulation. We define the concept of a relevant contact, which is a contact between one infected HCW and one uninfected HCW; this is the only type of contact that can potentially result in transmission. In contrast, contacts between two infected or two uninfected HCWs are not relevant.

At each time step, we allow every pair of HCWs to interact at most one time. We use the number of infected and uninfected HCWs to determine the number of HCW-to-HCW transmissions. Equations for determining the number of HCW-to-HCW transmissions are given in Table 5.2. In our model, we start by calculating the maximum number of relevant contacts (c_m) that would occur if every infected HCW interacted with every uninfected HCW. The probability of one relevant contact (p_c)

Table 5.1: Healthcare worker transmission parameters and variables.

Notation	Definition
n	Number of patients
m	Number of healthcare workers
m_u, m_c	Number of uninfected and infected healthcare workers
c_m	Maximum number of relevant contacts, given m_u and m_c
p_c	Relevant contact probability
p_t	Healthcare worker transmission probability
X	Random variable representing the number of relevant contacts
Y	Random variable representing the number of healthcare worker transmissions

is calculated as the ratio of the maximum number of relevant contacts to the total number of all contacts among HCWs. After c_m and p_c are calculated, we determine the number of relevant contacts (X) and the number of HCW-to-HCW transmissions (Y). X and Y are modeled as binomial random variables ($X \sim$ binomial(c_m, p_c), $Y \sim$ binomial(X, p_t)), and their probability density functions are given in Table 5.2. After the number of HCW-to-HCW transmissions (y) is determined at a particular time step, y uninfected HCWs are selected at random to become infected. During their next and potentially subsequent visits to patients, the infected HCWs will be at risk for transmitting the disease.

First, we demonstrate how our model of HCW-to-HCW transmission behaves for a dense network and a sparse network. Second, we explore how this extension affects patient-to-patient transmission in our 20-patient ICU setting. The dense network has four nurses ($d = 0.2105$) and two physicians ($d = 0.4737$). The sparse network has 10 nurses ($d = 0.0526$) and five physicians ($d = 0.1579$). Both networks

Table 5.2: Equations for determining the number of HCW-to-HCW transmissions and their impact.

Variable	**Equation**
Maximum Number of Relevant Contacts	$c_m = m_u \times m_c$
Relevant Contact Probability	$p_c = c_m \binom{m}{2}^{-1}$
Number of Relevant Contacts	$P(X = x) = \binom{c_m}{x} p_c^x (1 - p_c)^{c_m - x}; x = 0, 1, 2 \ldots c_m$
Number of HCW Transmissions	$P(Y = y) = \binom{x}{y} p_t^y (1 - p_t)^{x-y}; y = 0, 1, 2 \ldots x$
100(1-α)% Wilson Score Interval	$(p_{c,\alpha/2}, p_{c,1-\alpha/2}) = \frac{p_c + \frac{1}{2c_m} z^2_{1-\alpha/2} \pm z_{1-\alpha/2} \sqrt{\frac{p_c(1-p_c)}{c_m} + \frac{z^2_{1-\alpha/2}}{4c_m^2}}}{1 + \frac{1}{c_m} z^2_{1-\alpha/2}}$

are used for comparison purposes in the remaining experiments.

Given the total number of HCWs, we examine how the number of relevant contacts changes with different mixes of infected and uninfected HCWs. In Figure 5.8, we see that the relevant contact probability and the maximum number of relevant contacts are maximized when $m_c = {}^{m}/_{2}$ for m even and $m_c = {}^{(m-1)}/_{2} = {}^{(m+1)}/_{2}$ for m odd. We also see a significant increase in the potential for HCW-to-HCW transmission when there are more HCWs. In our example, there can be at most nine relevant contacts at each time step when there are six HCWs, but as many as 56 relevant contacts when there are 15 HCWs.

We can predict the probability of the number of relevant contacts (X) using our binomial probability model. The specific behavior changes with each time step, depending on values of c_m and p_c. We show the best and worst case scenarios in Figure 5.9. In this figure, we plot probability and cumulative density functions for X, based on minimum (non-zero) and maximum values for c_m and p_c. Clearly,

there is a much higher risk for HCW-to-HCW transmission when there are more HCWs. There is minimal risk with six HCWs, even for the maximum potential case. However, with 15 HCWs, the distribution is centered at approximately 30 relevant contacts per time step.

We compute an approximate confidence interval for X using the Wilson score interval [87]. The expression for computing the Wilson Score Interval (WSI) is given in Table 5.2, where c_m and p_c replace the standard binomial parameters n and p. The WSI can then be scaled by c_m to compute an interval for the expected value of the number of relevant contacts, E[X] (where E[X] = $c_m \times p_c$ for $X \sim$ binomial(c_m, p_c)). The intervals for the four cases in Figure 5.9 are given in Table 5.3 and show the range of values for E[X] that we expect for 95% of the simulation time steps. These intervals show the minimal risk for relevant contacts with few HCWs and the much larger risk for relevant contacts when there are many HCWs serving the same unit. The minimum potential for either case occurs when there is one or $m - 1$ colonized HCWs, which poses a similar risk regardless of the number of HCWs. The maximum potential for relevant contacts is much higher when there are many HCWs. For our sparse network with 15 HCWs, we expect approximately 30 relevant contacts (E[X] = 29.87) during each time step when there are an equal number of colonized and uncolonized HCWs. This risk is much higher than the five (E[X] = 5.4) relevant contacts we expect on average for the worst-case scenario in the dense network.

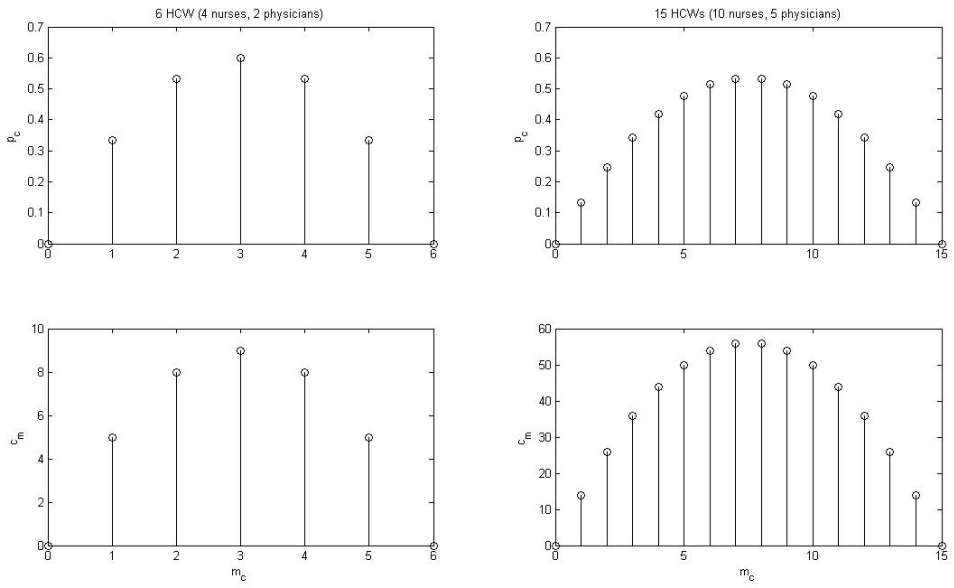


Figure 5.8: The relevant contact probability (p_c) and the maximum number of relevant contacts (c_m) are shown as functions of the number of infected HCWs (m_c) for a dense network (left) and a sparse network (right).

5.3.2 Simulation Results with HCW-to-HCW Transmission

In the remaining experiments, we explore how HCW-to-HCW transmission affects patient-to-patient transmission in the ICU. In Figure 5.10, we compare transmission dynamics with and without HCW-to-HCW transmission in our dense and sparse networks. Without HCW-to-HCW transmission, the dynamics follow the typical 'S-curve' pattern for disease-spread models, in which transmission is slow in the initial period, speeds up in the middle period, and slows again while the remaining patients become infected. This pattern holds for the dense and sparse cases with patients in the dense case becoming infected more quickly. Transmission for the cases with HCW-to-HCW transmission occurs at a much faster rate, with many patients becoming infected within a short period of time. These sharp increases

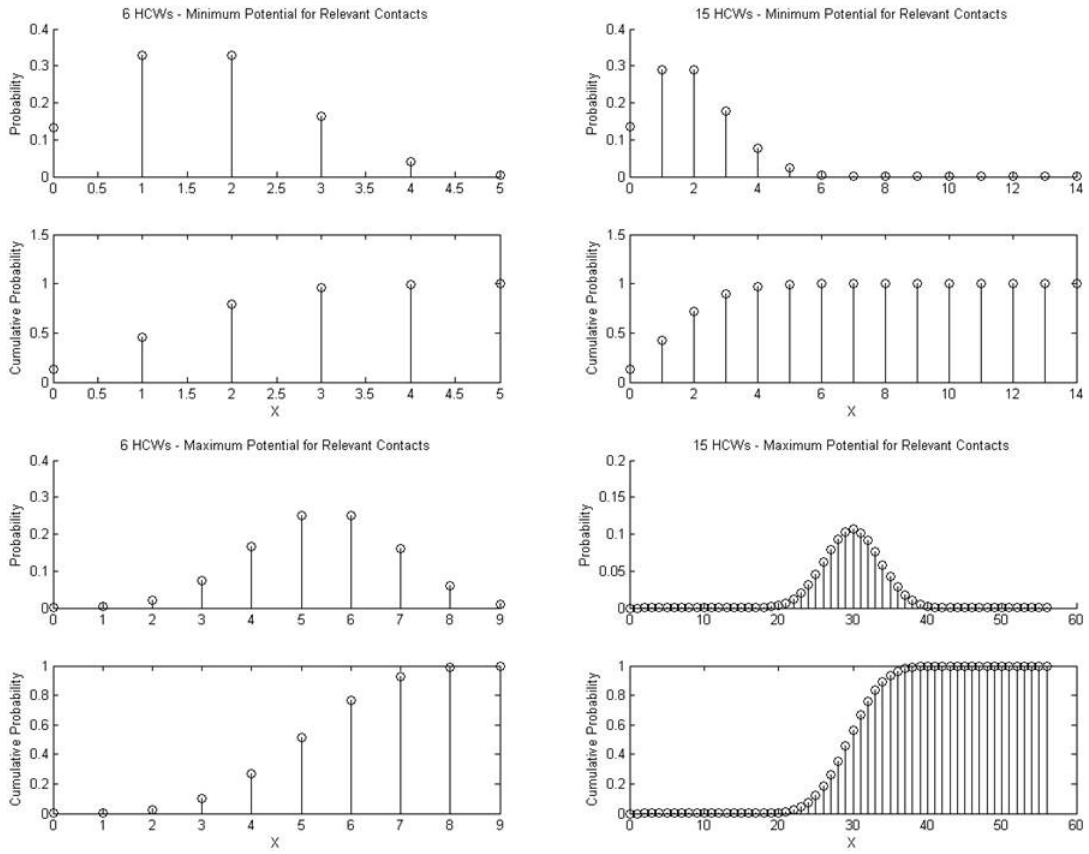


Figure 5.9: Minimum (top) and maximum (bottom) potential for the number of relevant contacts (X) for a dense network (left) and a sparse network (right), shown in the form of probability and cumulative density functions.

in transmission coincide with spikes in the number of HCW-to-HCW transmissions, indicating they are probably the cause for the outbreak. We also observe that transmission occurs more quickly in the sparse network, which is a result of the increased risk for relevant contacts with more HCWs. In both networks, transmission occurs much more quickly than in the cases without HCW-to-HCW transmission.

We introduced HCW-to-HCW transmission into our model and repeated our experiments with nurse and physician network densities. The results of these experiments are given in Figure 5.11 for three HCW-to-HCW transmission probabilities.

Table 5.3: 95% Wilson score intervals for the expected value of X (E[X]).

m (Number of HCWs)	Minimum Potential $c_m = 1$ or $c_m = m - 1$		Maximum Potential $c_m = m^2/4$ for m even $c_m = (m-1)^2/4$ for m odd	
	E[X]	95% WSI	E[X]	95% WSI
6	1.67	(-0.44, 2.76)	5.40	(1.51, 6.36)
15	1.87	(-0.90, 4.04)	29.87	(20.90, 35.06)

In addition to mean time to transmission, we show contour plots of the HCW-to-HCW transmission rates for each case. These plots show that the frequency for HCW-to-HCW transmission is maximized when both nurse and physician densities are small, which correspond to the cases with the most HCWs. These plots also show that the frequency of transmission increases with higher HCW transmission probabilities.

Based on these experiments, we observe that the best-case scenarios for all cases are much worse-in terms of mean time to transmission-than in the original experiments. At low HCW-to-HCW transmission probabilities, there appears to be an optimal combination of nurse and physician network densities where the mean time to transmission is maximized. This optimal configuration appears to occur in the region that combines low nurse network density with high physician network density. This result seems reasonable because, for nurses-who visit patients most often-network density still needs to be low to prevent transmissions, regardless of the increased risk for transmission between HCWs. Physicians contribute more to increasing the risk of HCW-to-HCW transmission than to reducing the risk of infecting patients directly, which explains the location of the optimal configuration. However,

as the HCW-to-HCW transmission probability increases, we begin to see a change in the overall response of the system. Two key trends emerge that demonstrate how HCW-to-HCW transmission dominates the benefits of sparse patient networks. In the first trend, the contours become symmetric about the axis of equal densities (i.e., nurse network density equals physician network density), which suggests that the nurse and physician densities no longer have a significant impact on transmission in the ICU. In the second trend, maximizing the mean time to transmission requires fewer HCWs, because the risk of HCW-to-HCW transmission is high. We do not suggest that hospitals should move towards reducing their staffing ratios to extremely low levels. Instead, for scenarios in which HCW-to-HCW transmission is a significant risk, hospitals need to take additional precautions such as providing masks or vaccinations for HCWs.

5.4 Patient Turnover

We extend the baseline model by incorporating patient admission and patient discharge (patient turnover). The baseline, static-population model provided a good indication of how network structure and HCW behavior affect transmission. However, in practice, patients are constantly moving in and out of various units. We accounted for this movement using two parameters: a patient turnover rate and an admission prevalence rate. The patient turnover rate was used at each tick to determine whether or not one patient, selected at random, would be replaced with a new patient. The admission prevalence rate was then used to determine whether

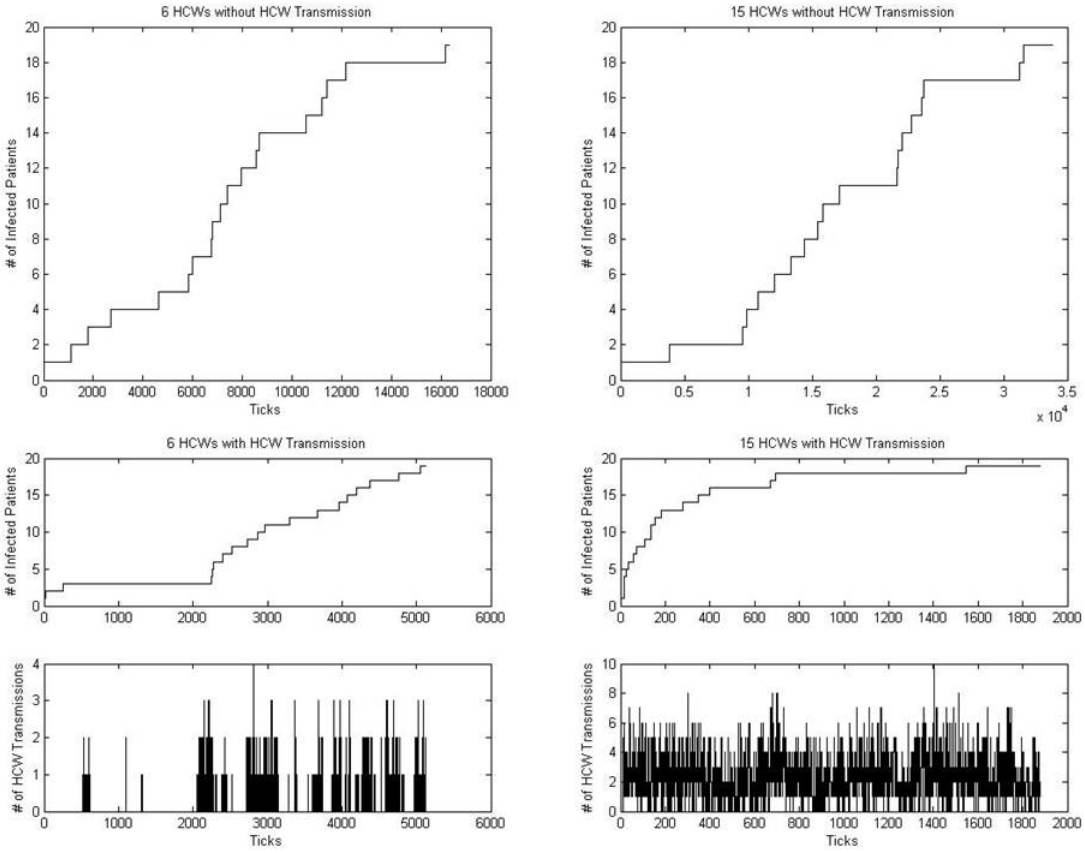


Figure 5.10: Transmission dynamics histories in a 20-patient intensive care unit for a dense network (left) and a sparse network (right) without (top) and with (bottom) HCW-to-HCW transmission. The HCW-to-HCW transmission probability (p_t) was held constant at 0.1. The number of HCW transmissions per time step is shown for the scenarios with HCW-to-HCW transmission.

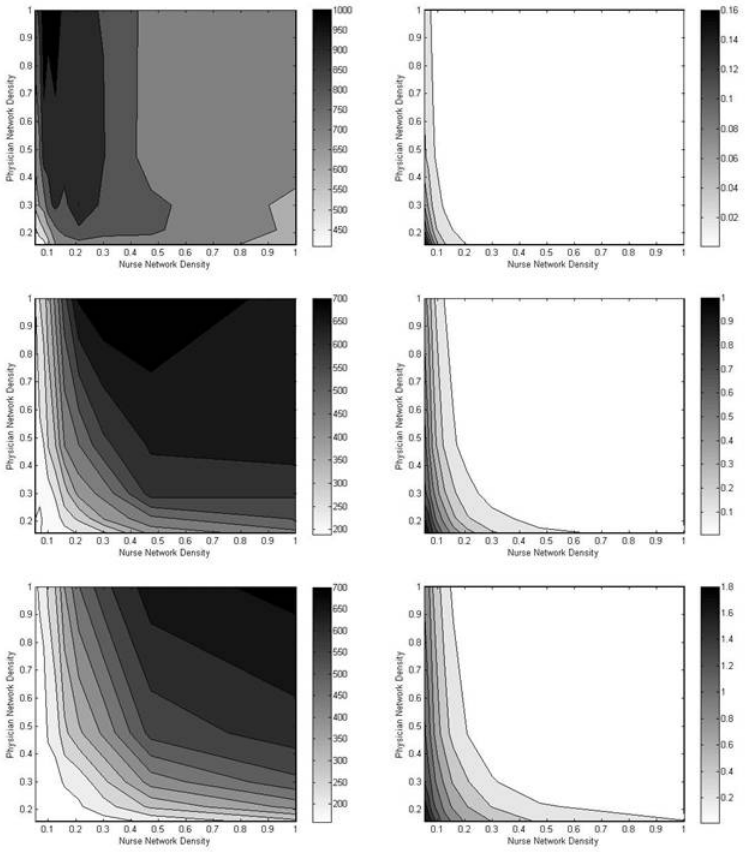


Figure 5.11: Contour plots for the mean time to transmission (left) and HCW transmission rate (right) as a function of nurse and physician network densities. HCW transmission rates are expressed as the number of HCW transmissions per tick. Results are shown for three values of the HCW transmission probability: $p_t = 0.01$ (top), $p_t = 0.05$ (middle), and $p_t = 0.1$ (bottom).

or not the new patient would already be infected.

5.4.1 Modified Transmission Dynamics

Introducing patient turnover changes the feasible outcomes for transmission dynamics. In the baseline model, the only possible outcomes occur when all patients become infected (we call this saturation) or when transmission stalls and the disease cannot reach other segments of the patient population. With patient turnover, transmission in the hospital can reach an endemic (steady) state in which a balance is established between infected patients entering and leaving the hospital. For our purposes, steady state is reached when a simulation runs 100,000 ticks or more without ending in saturation or extinction. Extinction can also occur. In this new outcome, all infected patients are discharged and transmission can no longer occur unless newly admitted patients are infected. In practice, extinction cases are more likely to look like endemicity, but we separate the two cases to highlight network configurations that are more effective at reducing transmission. Transmission dynamics outcomes are shown in Figure 5.12.

5.4.2 Simulation Results with Patient Turnover

We experiment first with patient turnover on the dense and sparse networks from Section 5.3, and then introduce patient sharing and HCW-to-HCW transmission to determine how transmission dynamics are affected. The patient turnover rate was set at values of 0 (no turnover), 0.01 (low turnover) and 0.1 (high turnover),

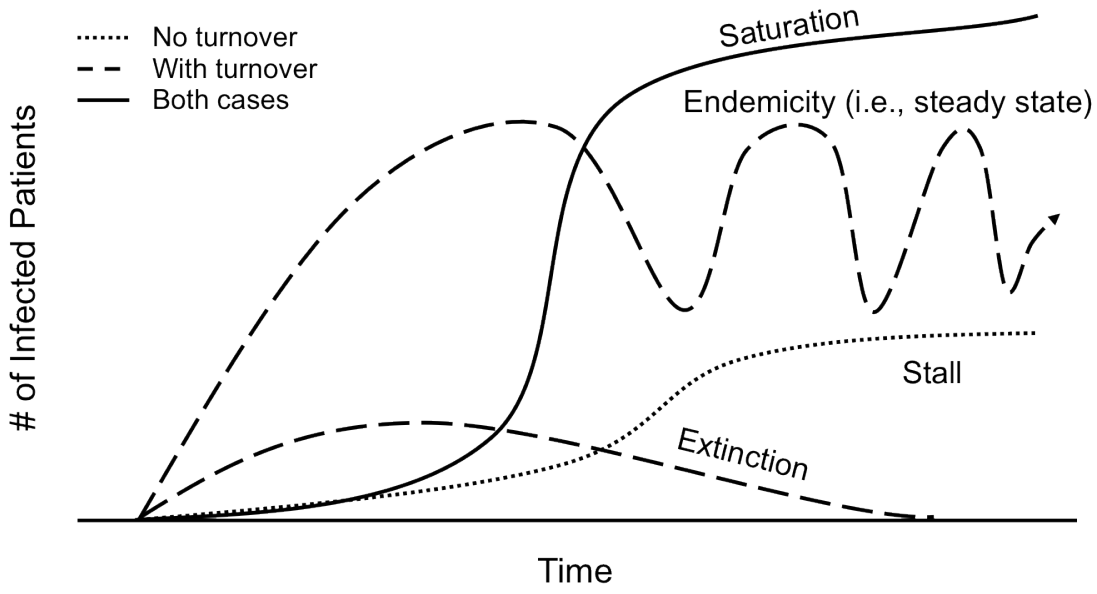


Figure 5.12: Feasible transmission dynamics with and without patient turnover.

while the admission prevalence for cases with turnover was set at 0.1 (low) and 0.5 (high). The results for the first experiment are given in Figure 5.13. The histograms show how the transmission dynamics change with increasing levels of turnover. For dense and sparse networks, the dynamics change from saturation to extinction as the patient turnover rate increases. Sparse networks lead to better outcomes in all cases (i.e., slower saturation times and faster extinction times). These trends support the general conclusion that shorter lengths of stay for patients can decrease the likelihood that they will acquire an infection at some point during their stay.

In the second experiment, we introduce patient sharing into the ICU used in Section 5.2.4 with turnover. The results are given in Figure 5.14, which shows an array of 12 stem-scatter plots for four sharing configurations and three patient turnover

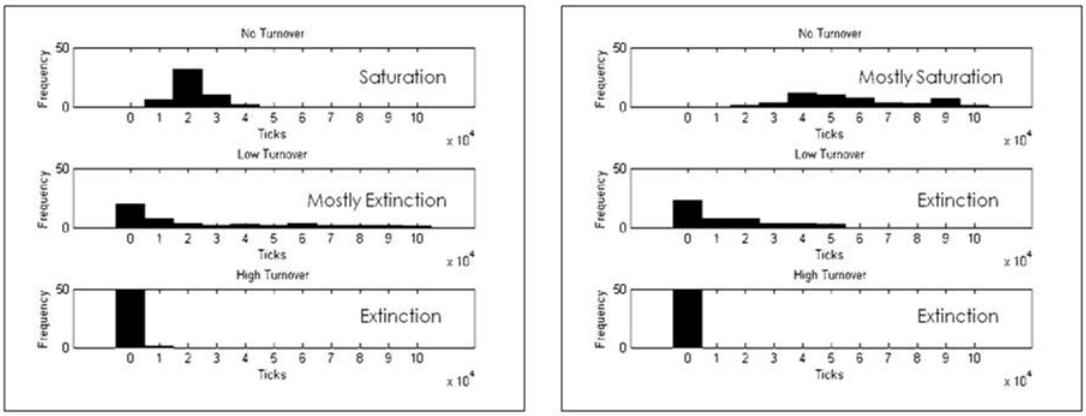


Figure 5.13: Transmission dynamics outcomes for a dense network (left) and a sparse network (right) with no turnover (top), low turnover (middle, rate = 0.01), and high turnover (bottom, rate = 0.1). Results are reported for 50 replications as histograms that show the frequency of simulations terminating at a specific number of ticks. The transmission dynamics outcome is noted for each plot (saturation, extinction).

levels. For these experiments, the results did not always lead to the same outcome as they did for the results in Figure 5.13 (i.e., saturation or extinction). Therefore, we show the number of ticks, plotted on the left-axis as stem plot values, and the number of infected patients, plotted on the right-axis as scatter plot values, to specify the outcome for each simulation replication. These results generally reflect the trends from the original experiments. For example, the four plots with no turnover (left column) show the saturation outcome in every case, with all 20 patients in the unit becoming infected within relatively short simulation times. However, with patient turnover introduced, the revolving and paired sharing configurations appear to perform as effectively as the case with no sharing. These three configurations result in a mix of endemic and extinction outcomes with low turnover, but lead to only extinction when the patient turnover rate increases. Random sharing results in an endemic state even for high turnover rates and is the worst configuration.

In the third experiment, we explore the effects of patient turnover in our dense and sparse network examples combined with HCW-to-HCW transmission, which was the dominant factor in the experiments with network density from Section 5.3. The results, given in Figure 5.15, show similar trends for the sparse network and the dense network. For the case with no turnover, both networks saturate for all simulation replications. The sparse network leads to faster rates due to the higher number of HCWs. With low turnover, there is a mix of saturation and endemic outcomes, which suggests that the effects of patient turnover and HCW-to-HCW transmission are approximately equal at this level. For the sparse network, there is a higher frequency of endemic outcomes, and even the saturation cases appear to

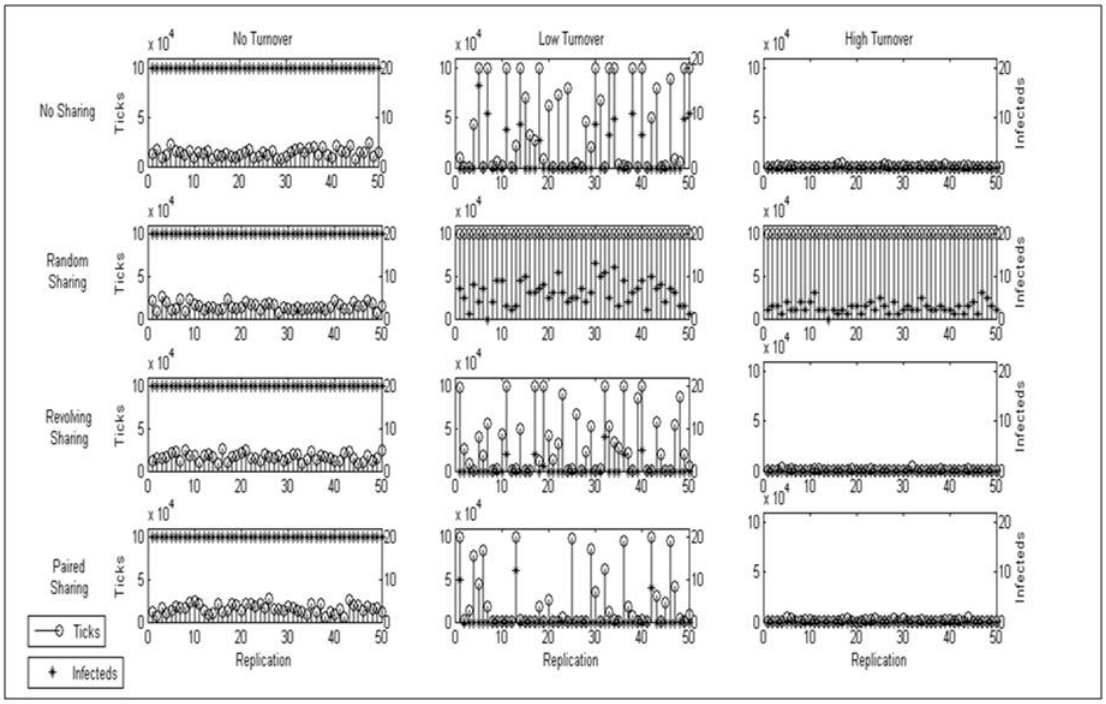


Figure 5.14: Summary of patient turnover effects on four configurations of patient sharing (none, random, revolving, and shared) for cases of no turnover (left), low turnover (middle, rate = 0.01), and high turnover (right, rate = 0.1). The patient sharing rate was held constant at 0.1. Each plot shows the number of ticks (stem plot values, left axis) and the number of infected patients (scatter plot values, right axis) at the end of each simulation replication, which indicate whether the transmission dynamics outcome was saturation (number of infected patients = 20), endemicity (number of ticks = 100,000), or extinction (number of infected patients = 0).

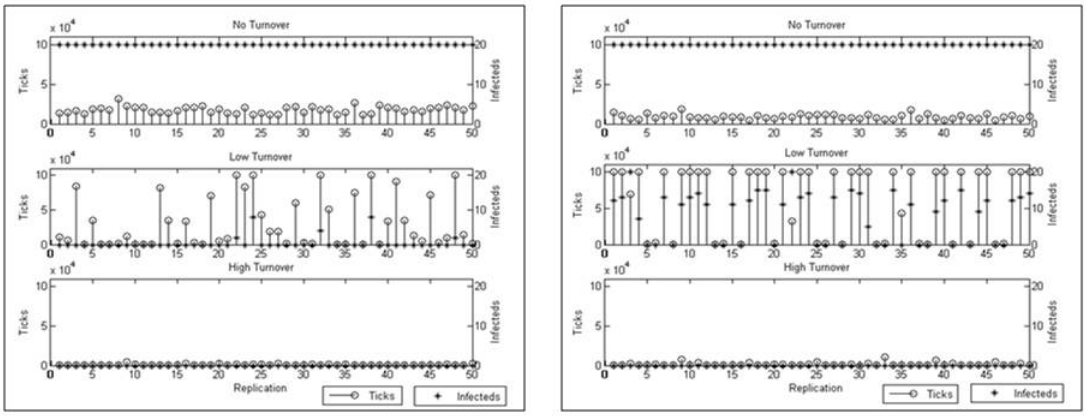


Figure 5.15: Summary of patient turnover effects on HCW-to-HCW transmission for a dense network (left) and a sparse network (right), for cases of no turnover (top), low turnover (middle, rate = 0.01), and high turnover (bottom, rate = 0.1). The HCW-to-HCW transmission probability (pt) was held constant at 0.1. Each plot shows the number of ticks (stem plot values, left axis) and the number of infected patients (scatter plot values, right axis) at the end of each simulation replication, which indicate whether the transmission dynamics outcome was saturation (number of infected patients = 20), endemicity (number of ticks = 100,000), or extinction (number of infected patients = 0).

occur quickly, as if the index patient was discharged before there was sufficient time for the disease to spread. Clearly, high turnover negates the effect of HCW-to-HCW transmission, leading to extinction very quickly.

For all experiments with patient turnover, we used an admission prevalence of 10%, that is, on average, 10% of newly admitted patients are infected. We repeated the patient turnover experiments with an admission prevalence of 50%. This level could represent an ongoing outbreak of a particular disease. Whereas higher rates of turnover lead to a higher probability of extinction, higher admission prevalence rates shift the outcome balance towards an endemic state or saturation, so that extinction

becomes very difficult. The balance between these two outcomes depends on several factors including network density, rates of turnover and admission prevalence, and the presence or absence of patient sharing and HCW-to-HCW transmission. The results of the experiments with patient turnover are given in Table 5.4. In most cases, an endemic state is the most likely outcome when newly admitted patients are more likely to be infected on admission. However, there are some scenarios for which there is a high probability of saturation. Units that combine low turnover with either dense networks–such as in the baseline model (42% of outcomes) or with patient sharing (from 30% to 62% of outcomes)–or HCW-to-HCW transmission (38% of cases for a dense network, 86% of cases for a sparse network) are especially susceptible to saturation. All units with high turnover are essentially protected from saturation outcomes.

5.5 Conclusions

Agent-based modeling provides a convenient framework for simulating the transmission of an infectious disease in a hospital. The results from this type of modeling effort can provide valuable insights to professionals in infection control. We used an agent-based model to simulate the effect of network density, HCW behavior, and patient turnover on transmission in an ICU. By examining the densities of the nurse and physician networks, we gained a new perspective on the effect of the implicit connections among patients who share a HCW. Our results showed that nurses and physicians pose significant threats to patients in terms of infection, but

Table 5.4: Results for network density (top), patient sharing (middle), and HCW-to-HCW transmission (bottom) with patient turnover and high admission prevalence (rate = 0.5).

Baseline Density Case	% of Outcomes (50 Replications)		
	Extinction	Endemicity	Saturation
Dense network with low turnover	8	50	42
Dense network with high turnover	8	90	2
Sparse network with low turnover	6	86	8
Sparse network with high turnover	0	98	2

Sharing Configuration	Turnover Level	% of Outcomes (50 Replications)		
		Extinction	Endemicity	Saturation
None	Low/High	0/12	38/86	62/2
Random	Low/High	0/0	50/100	50/0
Revolving	Low/High	6/2	60/98	34/0
Paired	Low/High	6/6	64/90	30/4

HCW-to-HCW Transmission Case	% of Outcomes (50 Replications)		
	Extinction	Endemicity	Saturation
Dense network with low turnover	10	52	38
Dense network with high turnover	6	88	6
Sparse network with low turnover	6	8	86
Sparse network with high turnover	8	86	6

they do so in different ways. Nurses can spread infection to patients quickly within their cohorts, but they are limited in their ability to spread infection to the entire unit. Physicians spread infection much more slowly than nurses. However, their infections can accelerate transmission in the unit because they can spread the disease to multiple nurse cohorts.

Our model enabled us to investigate the effects of HCW behavior. Patient sharing is often overlooked in epidemiological modeling and it tends to promote transmission. Revolving and paired sharing configurations offer better alternatives than random sharing, which is most likely how sharing occurs in practice. Physicians can negate the intended benefits of these configurations because their networks extend across multiple nurse networks, but structured sharing (i.e., revolving or paired sharing) between nurses can reduce the rate of transmission. Transmission between HCWs poses a significant threat as well, and can offset the benefits of high HCW-to-patient ratios. This type of transmission is probably relevant only for highly transmissible diseases, and would require additional precautions such as the use of masks and gloves. In future work, the effects of less uniform HCW interaction network configurations could be explored.

Finally, we compared how patient turnover affects transmission dynamics in a hospital unit. Our experiments showed that shorter lengths of stay could reduce the threat of hospital-acquired infections. However, random patient sharing, HCW-to-HCW transmission, and high admission prevalence rates all showed the capability to offset the benefit of patient turnover. Hospitals can alleviate these risks to some degree, by implementing structured sharing configurations, protecting HCWs from

infection by providing masks and vaccinations (when available), and isolating newly admitted, patients who are infected.

Chapter 6

Early Detection of Bioterrorism: A Combined Social and Spatial Network Analysis

6.1 Overview

Bioterrorism refers to the intentional release of viruses, bacteria, or other toxic biological agents, to cause illness or death in people, animals, or plants. Biological agents can be spread through the air, through water, or by food. The U.S. Centers for Disease Control and Prevention (CDC) list anthrax, botulism, plague, smallpox, tularemia, and viral hemorrhagic fever as Class A bioterror threats, which are currently known biological diseases that are likely to do the most damage [16].

Bioterrorism is considered a significant threat to the United States. According to Congress, the U.S. is more vulnerable to biological weapons than more traditional means of warfare. The Office of Emergency Preparedness estimates that 40 million Americans could die if a terrorist released smallpox into the American population. Anthrax could kill 10 million. New forms of deadly biological weapons could also kill millions [20].

From a public health perspective, early detection of a potential bioterrorism incident is vital, as it allows timely communication with epidemiologic investigators, health providers, laboratories, and other critical organizations that will attempt to

limit the damage. In addition, it improves the likelihood that an adequate supply of antimicrobial drugs, antitoxins, or vaccines will be available.

One of the main challenges in detecting bioterrorism incidents is that many candidate diseases present symptoms in a similar manner as the common cold or flu (e.g., anthrax, viral hemorrhagic fever). In all likelihood, health providers who are accustomed to dealing with everyday cases will not attribute patient symptoms to an unknown bioterror incident, but rather diagnose the condition as a common illness. The disease is then likely to continue spreading until the first unexpected deaths occur.

In this chapter, we propose two techniques for *quickly* differentiating between a bioterrorism incident and a seasonal flu epidemic according to their transmission patterns. Specifically, we construct a multilayered network that includes social and spatial components, and then compare the temporal diffusion of the disease throughout the population for each scenario. We assume that seasonal flu is only transmitted through human-to-human interactions, whereas transmission during a bioterrorism incident occurs exclusively because of a person's proximity to the location where the attack occurred. This assumption follows the common spread of currently known bioterror agents such as anthrax, hemorrhagic fever, and tularemia that are not highly transmissible between humans. We try to identify the best measures–those that provide an early detection time of a bioterror attack–that can be used as diagnostic metrics in practice, using results from simulating the disease transmission through the multilayered network.

6.2 Methodology

The simulation model has two major phases. The first phase involves the generation of the multilayered network. The second phase simulates an epidemic or bioterror scenario through the generated network. We discuss both phases in the sections that follow.

6.2.1 Multilayered Network Generation

We present a multilayered network of N_H humans and N_L locations. We use $H = \{H_1, H_2, \ldots, H_n\}$ to represent the set of human nodes with $P_{H_iH_j}$ as the probability of interaction (i.e., contact probability) between nodes H_i and H_j. The set $L = \{L_1, L_2, \ldots, L_m\}$ is used to represent all locations, with $d_{L_iL_j}$ designating the distance between locations L_i and L_j. Finally, each human node H_i is assigned a probability $P_{H_iL_j}$ of visiting each location L_j, which we refer to as the *human-location probability*. A schematic representation of the multilayered network is given in Figure 6.1.

There are several simplifying assumptions that we use in the generation of the multilayered networks. First, we model a closed population, so there are no individuals that enter or depart the network once it is formed. We use the Barabási-Albert (BA) algorithm for generating the human social network [1]. There are four steps in the BA algorithm.

1. Let m denote the number of nodes that each successive node connects to in the network.

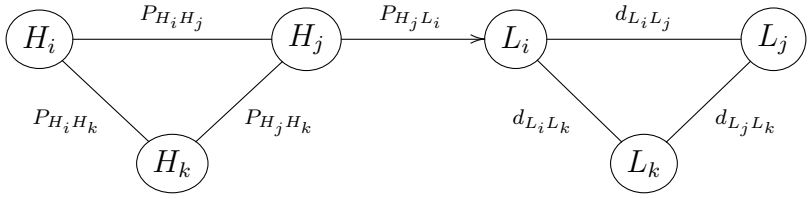


Figure 6.1: Schematic representation of the multilayered network. The human social network (depicted on the left) shows human nodes that interact with probabilities $P_{H_iH_j}$. The location network (depicted on the right) shows location nodes separated by distances $d_{L_iL_j}$. The human-location network has edges that connect human nodes to location nodes with probabilities $P_{H_jL_i}$.

2. Initialize the network with $m + 1$ nodes that are all connected to each other.

3. Add a new node to the network and choose m nodes at random to connect to the node being inserted, where the likelihood of each node being chosen is proportional to its current degree (i.e., $p_i = \frac{deg(i)}{\sum_j deg(j)}$ where $deg(i)$ is the degree of node i).

4. Repeat Step 3 until the network has N_H nodes.

In Figure 6.2, we show an unweighted human social network for $N_H = 10$ agents and different values of the BA algorithm density parameter (m). The BA algorithm generates a scale-free social network with a power-law degree distribution in which there are few human nodes with many contacts and many human nodes with only a few contacts. This algorithm has been used in other research as an appropriate model for human social networks [6].

The edges in the human social network are generated using the BA algorithm, but we need a model to determine how frequently each pair of humans interacts. We

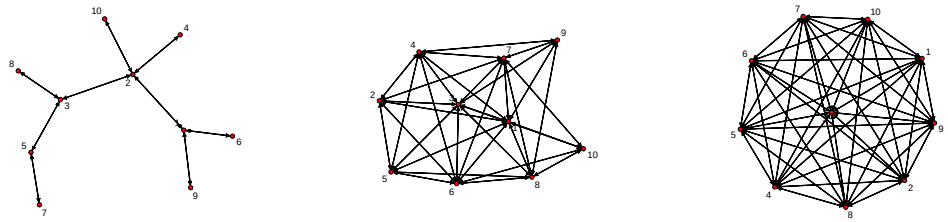


Figure 6.2: Barabási-Albert (BA) network examples, with the BA algorithm density parameter (m) ranging from 1 to 5 to 10 from left to right.

generate contact probabilities $P_{H_iH_j}$ between human nodes H_i and H_j for each edge using an exponential distribution with rate $\lambda = 10$, scaled between 0.1 and 1. The lower bound on the scaling ensures that the network remains fully connected. The exponential distribution generates a lot of edges with low contact probabilities (e.g., less than 0.2), which represent rare or occasional contact. The distribution also generates some edges with more frequent contact. These higher probability edges represent the frequency of contact we expect between family members, coworkers, and classmates. We set the initial contact probability to $P^{init}_{H_iH_j} \sim exp(10)$ and then scale the result according to

$$P_{H_iH_j} = \left[\frac{P^{init}_{H_iH_j}}{max\{P^{init}\}} \times \frac{10}{9} \right] + 0.1. \tag{6.1}$$

Once generated, the edges in the human social network and their contact probability weights are fixed (i.e., P_{H_iHj} is constant for all time t). The distribution of the human contact probability is given in Figure 6.3 and a sample human social

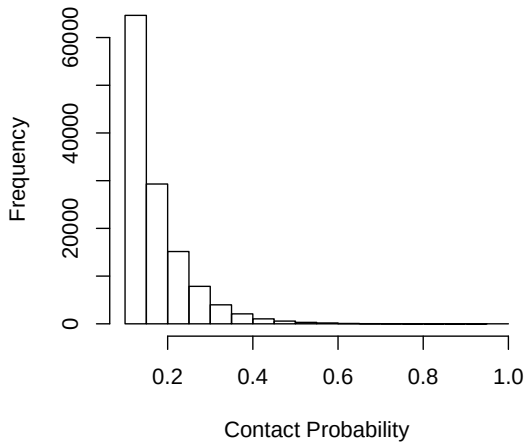


Figure 6.3: Contact probability distribution for a human social network with a BA algorithm density parameter equal to one. The frequency (i.e., likelihood) is shown for each contact probability bin plotted on the x-axis.

network instance with 40 human nodes is given in Figure 6.4.

The location nodes are arranged in a 2-dimensional Cartesian grid with width w and length l and a unit distance between adjacent nodes. This model was chosen to represent a geographic region where each location in the network contains the same size area. A sample location network with nine locations (i.e., $w = l = 3$) is shown in Figure 6.4.

The human-location agent probability is simulated from a multivariate mixed gamma distribution of the form

$$P_{H_i L_k}^{init} \sim \begin{cases} \Gamma(k^H = 10, \theta^H = 0.5; \Sigma) & , p^H = 0.1 \\ \Gamma(k^L = 1, k^L = 1; \Sigma) & , p^L = 0.9 \end{cases} \tag{6.2}$$

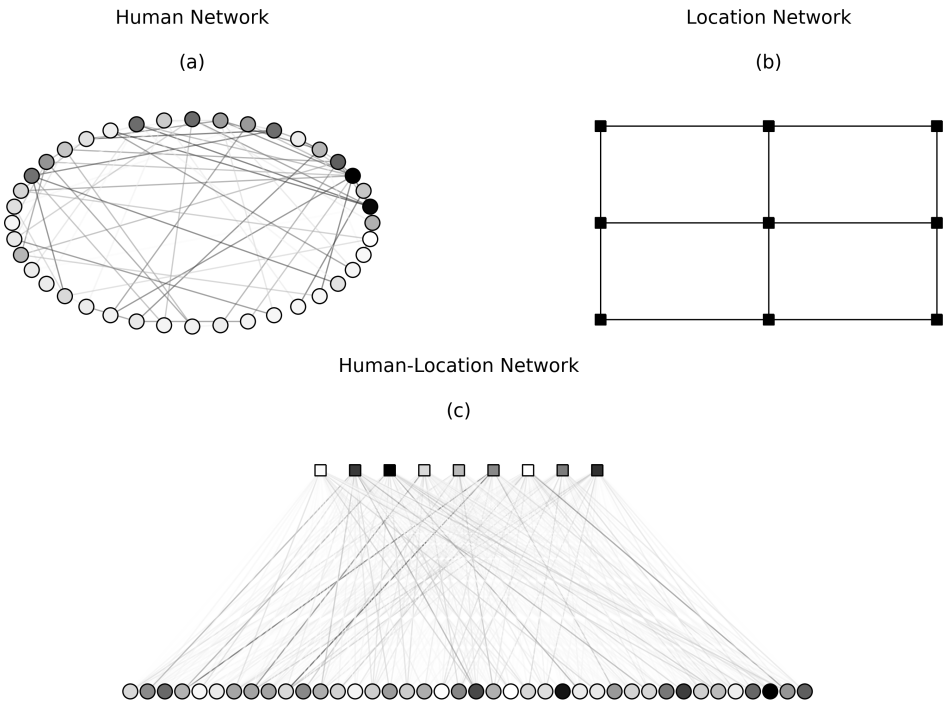


Figure 6.4: Sample multilayered network instance. In (a), we show a sample human social network instance with $N_H = 40$. The nodes are shaded according to the sum of weighted edges incident to each node, with darker nodes representing nodes that have more frequent contact with other nodes in the network. Edges are shaded according to their contact probability values, with darker edges representing higher contact probabilities. In (b), we show a sample location network instance for a 2-dimensional grid defined by $w = l = 3$. In (c), we show a sample human-location network instance as a bipartite graph. The location and human nodes are shaded (on separate scales) according the the sum of incident weighted edges and the edges are shaded according to their contact probability values.

with the covariance matrix Σ given by

$$\Sigma = \rho \times P_H^+ \tag{6.3}$$

where ρ is the correlation coefficient between social (i.e., human-human) and human-location contacts and P_H^+ is the nearest positive definite matrix to the contact probability matrix P_H. We calculate Σ using the algorithm developed by Higham [40]. The probability is then scaled in the range [0,1] to ensure that each human H_i visits a single location at any time t,

$$P_{H_i L_k} \leftarrow \frac{P_{H_i L_k}}{\sum_{L_l} P^{init}_{H_i L_l}} \tag{6.4}$$

The human-location agent probability distribution is given in Figure 6.5, which shows that there are many edges in the human-location network that represent rare or occasional travel to a particular location. In contrast to the social contact probability distribution, this distribution has a second peak that represents locations where people travel to relatively frequently, such as home, work, or school. A sample human-location network instance is given in Figure 6.4.

To illustrate the effect of the correlation coefficient ρ, in Figures 6.6 and 6.7, we plot realizations of the human-location network probabilities for different correlation coefficients ($\rho = 0$ and $\rho = 1$). We examine the location interaction probabilities of three human nodes, using the human contact probability matrix P_H in Equation 6.5. It is clear from 6.6 that when $\rho = 0$, we observe no correlation between the locations visited by any pair of nodes. However, when $\rho = 1$, there is a strong

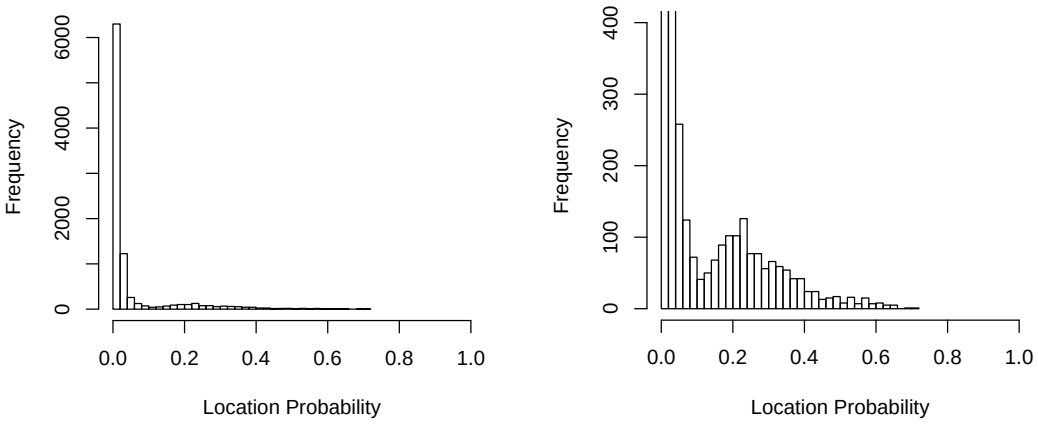


Figure 6.5: Distribution of the human-location contact probability for a multilayered network with a human-location density equal to 0.8. The entire distribution is shown on the left and a zoomed view of the higher location probabilities is shown on the right.

positive correlation between the locations visited by Agent 1 and Agent 2, which is represented by the large value in P_H (i.e., 0.9). There is still no correlation between Agent 1 and Agent 3 or between Agent 2 and Agent 3 because these pairs of human nodes have small correlation values in P_H.

$$P_H = \begin{pmatrix} 1.0 & 0.9 & 0.1 \\ 0.9 & 1.0 & 0.1 \\ 0.1 & 0.1 & 1.0 \end{pmatrix} \tag{6.5}$$

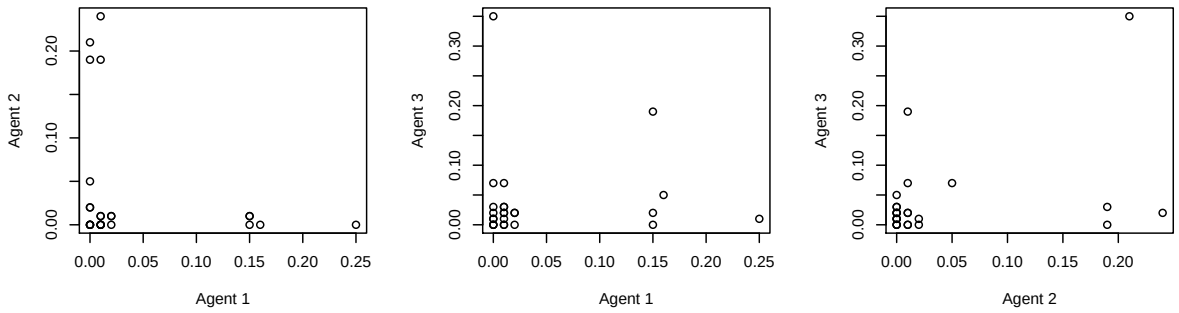


Figure 6.6: Example of human-location probability correlation with correlation coefficient $\rho = 0$. Each scatter plot shows the correlation between the human-location contact probabilities for a given pair of human nodes.

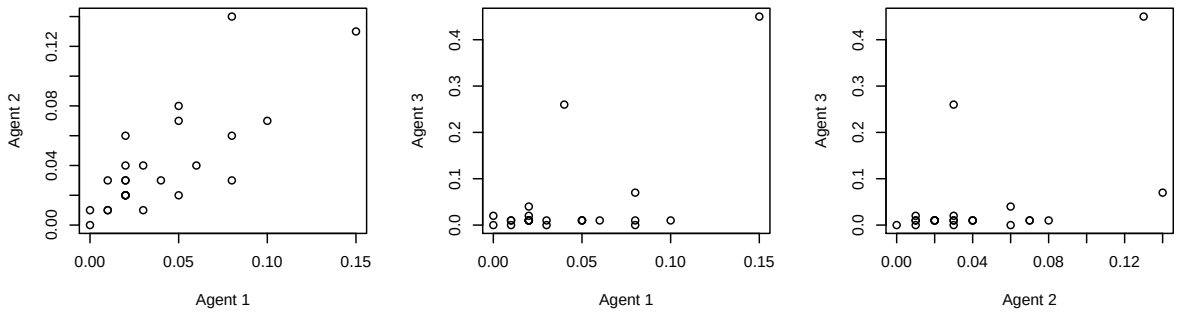


Figure 6.7: Example of human-location probability correlation with correlation coefficient $\rho = 1$. Each scatter plot shows the correlation between the human-location contact probabilities for a given pair of human nodes.

Table 6.1: Network simulation input parameters

Parameter	Description	Value
N_H	Number of human nodes	500
N_L	Number of location nodes	25
m	Barabási-Albert algorithm density parameter	1, 5, 10
w	Location grid width	5
l	Location grid length	5
d	Human-location network density parameter	0.5, 0.8, 1.0
ρ	Interaction probability correlation coefficient	0.0, 0.5, 1.0
p_t^H	Human transmission probability	0.1
p_t^L	Location transmission probability	0.01

6.2.2 Network Simulation Model

The network simulation model was implemented using Python[1], a dynamic object-oriented programming language, SimPy[2], a discrete-event simulation package, and NetworkX[3], a package that facilitates the creation and manipulation of networks. The simulation can be executed for an epidemic or bioterror scenario through a given network. Each disease scenario has different transmission dynamics, but simulations for both scenarios are run until the entire human population is infected. Table 6.1 summarizes the input parameters and the values used in the simulation.

In the epidemic scenario, we select $\lceil p_t^H \times \frac{N_H}{N_L} \rceil$ humans to be initially infected, where p_t^H is the transmission probability between human nodes. This formula for initially infected humans was chosen so that the initial susceptible population would be approximately the same size as the initial susceptible population in the bioter-

[1] www.python.org

[2] http://simpy.sourceforge.net/

[3] http://networkx.lanl.gov/

ror scenario, thereby creating a similar basis for comparison. The disease spreads randomly between human nodes, using the human contact probability and human transmission probability to determine whether transmission occurs. At each time t, every susceptible human node H_i has a probability of $1 - \sum_{H_k \in H} p_t^H \cdot P_{H_i H_k}$ of getting infected by one of its neighbors.

In the bioterror scenario, the attack occurs at a single location L_i and spreads to adjacent locations with probability p_t^L at each time step. Human nodes can only be infected at locations to which they travel. Human-to-human transmission does not occur in the bioterror scenario. Similar to the formulation used in the epidemic scenario, each human node H_i has a probability of $1 - \sum_{L_k \in L} p_t^H \cdot P_{H_i L_k}$ of getting infected at one of the locations.

Computational efficiency is a critical issue for simulating these scenarios, as the number of edges increases substantially for multilayered networks with many nodes. For both scenarios, we only iterate through the relevant edges in the network, that is, those edges that connect susceptible and infected nodes. The list of all relevant edges is updated at the end of each time step, and is shuffled prior to iterating so that the edges are not evaluated for transmission in the same order. In addition, we filter the multilayered network for each scenario so that we only iterate through the relevant sections of the network. Thus, we only iterate through the human social network in the epidemic scenario and through the location and human-location networks in the bioterror scenario.

When human nodes become infected, we track the time of infection and the source of the transmission, whether by a human node in the epidemic scenario or a

location node in the bioterror scenario. These data allow us to track the number of instantaneous infections and compute the cumulative number of infections at any point in time. By tracking the source of infection, we can construct an additional network with this information that will be used in the analysis of the simulation output.

6.3 Results

We performed simulation experiments on six parameterized networks using the parameters given in Table 6.1, which provides three human social network density cases for the epidemic scenario and three human-location density cases for the bioterror scenario. Ten instances of each parameterized network were generated. We performed 100 simulation replications of each scenario on each relevant network instance. The dynamics across all 100 simulation replications for one instance of each parameterized network are characterized in Figures **??** and **??**. Dynamics for the other network instances were similar except for some differences in the extreme (i.e., maximum and/or minimum) cases.

As shown in the dynamics plots in Figures **??** and **??**, there are some distinctive trends between the two scenarios. In both scenarios, it is clear that the infection spreads faster through networks that are more dense, although the trends are more significant in the epidemic scenario. The epidemic transmission dynamics follow the behavior defined by a mathematical model such as the susceptible-infected-recovered (SIR) model [48]. The cumulative number of transmissions in this case is

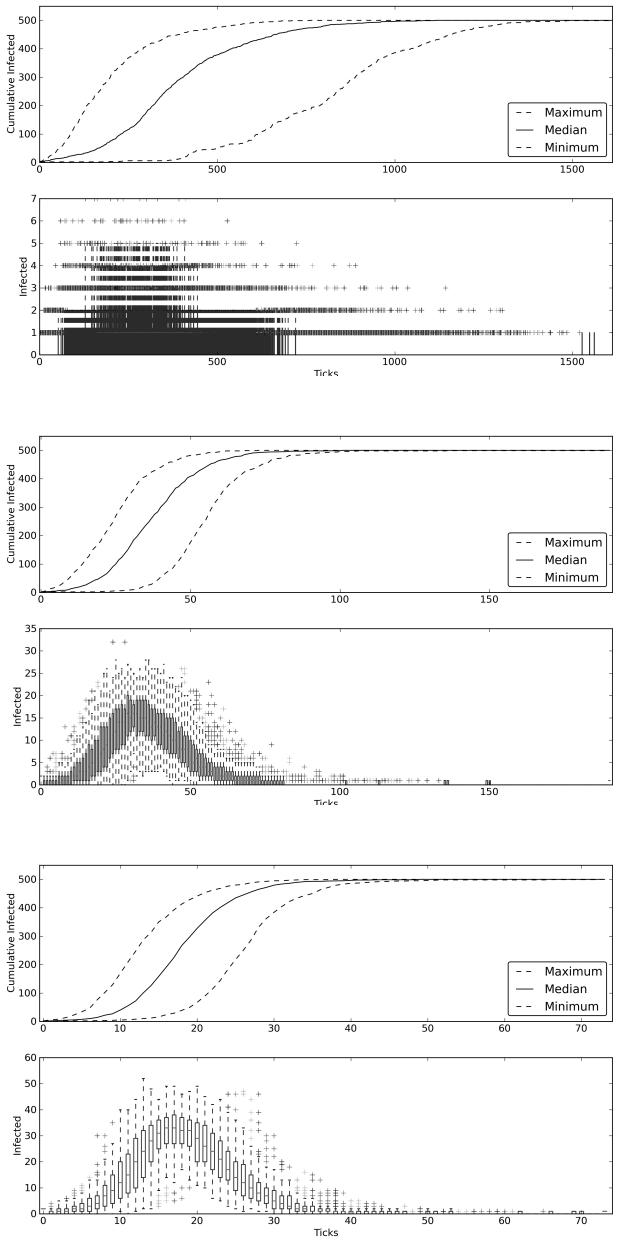


Figure 6.8: Transmission dynamics summaries for three parameterized epidemic network instances. Barabási-Albert algorithm density parameters (m) of 1, 5, and 10 were used to generate the human social networks. In the top plots, the median path (solid line) shows the typical transmission path for the epidemic on a given network instance, and the extreme cases (maximum and minimum, dashed lines) show the fastest and slowest dynamics observed in the batch of replications. The bottom plots show series of box-and-whisker plots that summarize the distribution of the number of newly infected humans at each time step, aggregated across all replications.

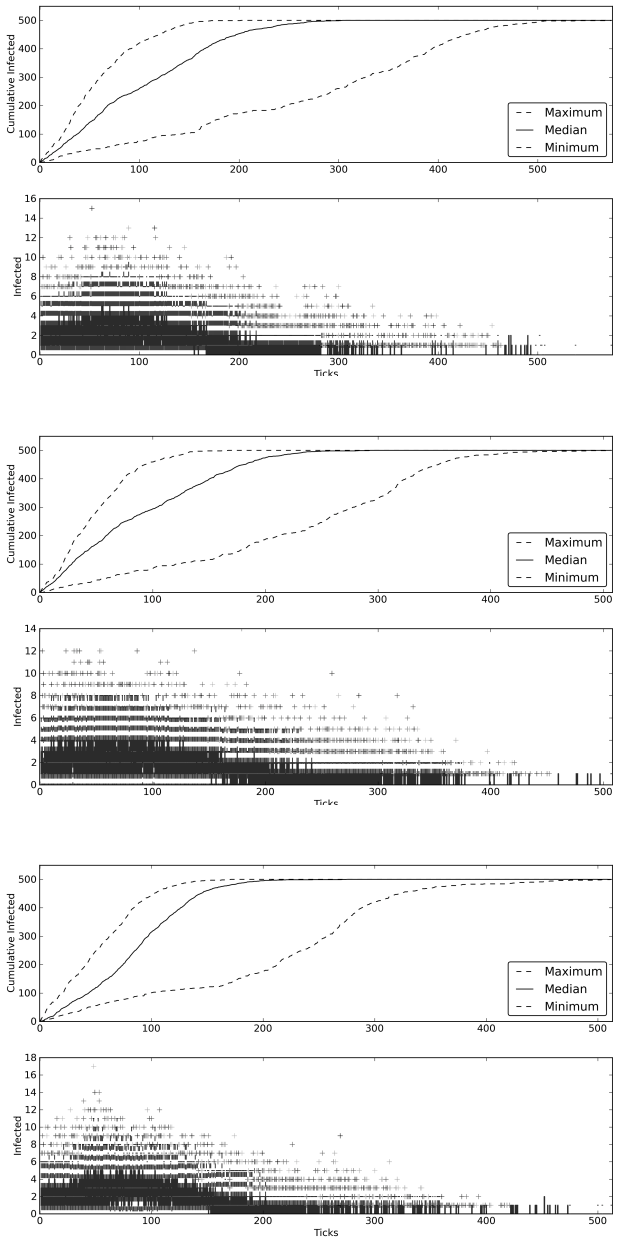


Figure 6.9: Transmission dynamics summaries for three parameterized bioterror network instances. Human-location network density parameters (d) of 0.5, 0.8, and 1 (top to bottom) were used to generate each instance. In the top plot, the median path (solid line) shows the typical transmission path for the outbreak on a given network instance, and the extreme cases (maximum and minimum, dashed lines) show the fastest and slowest dynamics observed in the batch of replications. The bottom plots show series of box-and-whisker plots that summarize the distribution of the number of newly infected humans at each time step, aggregated across all replications.

defined by an S-curve that begins with a slow rate of infection, transitions to a very fast infection rate during the middle period, then slows again while the remaining humans eventually become infected. These changes in the infection rate are reflected in the bottom plots in Figure **??**, in which the instantaneous number of infections are initially small, grow fairly large, and then become small again.

With the exception of the sparse case (i.e., BA density parameter equal to 1), we observe minimal variation with respect to the transmission curves in the epidemic scenario. The variation that does exist only relates to the time when the epidemic begins to accelerate throughout the population, not the peak transmission rate (i.e., maximum slope of the cumulative infection curve). This trend suggests that for a given population, there is a limit as to how quickly a disease can spread in an epidemic scenario. Surveillance efforts can then focus on monitoring the number of reported cases and trigger a response if a threshold infection rate is surpassed, because the infection rate will continue to increase in an uncontrolled population until the susceptible pool is sufficiently depleted.

The transmission dynamics in the bioterror scenario are quite different. Transmission in this scenario begins immediately, and continues to occur at a nearly constant rate until the susceptible population becomes very small. Detecting an outbreak in this scenario is more difficult because the transmission rate may never cross the type of threshold that would be appropriate for an epidemic. There is much more variation in the transmission dynamics than in the epidemic case. This variation suggests that control measures would have to be robust enough to prevent transmission that occurs very quickly or spreads quite gradually over a prolonged

period of time.

Differentiating the transmission dynamics in these two scenario is fairly easy by analyzing Figures **??** and **??**. In practice, these data would not be available. We would observe only one instance of a transmission dynamics curve as the outbreak was occurring. We examine two approaches for addressing this issue under different levels of certainty regarding the structure of the multilayered network.

6.3.1 Detection Under Social Network Certainty

We begin our analysis with the assumption that we know the underlying human social network. This assumption, although unrealistic for large populations, may have some applications for small, isolated settings such as in schools or within business or community organizations. In addition, there are techniques for estimating the structure of larger social networks using a combination of contact tracing and survey data [6]. We can explore how the disease spreads throughout these smaller or estimated networks, which may provide some insight as to how to detect a bioterror attack in larger populations.

We construct a secondary network for each simulation instance under the assumption that the disease spreads through the human population in an epidemic manner (i.e., by direct human-to-human transmission). Each secondary network is initialized as an empty network, and nodes (including initially infected nodes) are added to the network as they become infected during the simulation. This structure creates a spanning tree over all nodes. Each node added to the secondary network

must connect to a node that has already been added, and this node is chosen so that the distance between the two nodes in the original network is minimized. For example, if two nodes are directly connected in the original network, they are one degree removed from each other. Each node that sits on the shortest path between two nodes creates an additional degree removed. We denote the inserted node and its selection as a most likely pair.

With this construction, we calculate the shortest paths between all pairs of human nodes in the network and use these distances to generate the secondary network. We set the distance of each edge in the secondary network equal to the square of the number of degrees removed between each most likely pair. This model imitates a gravitational model in which the force between two masses decreases with the square of the distance between them. The secondary network is being rewarded with a low cost (in the form of small distances) for adding nodes in a way that is consistent with an epidemic outbreak. However, if a bioterror attack is being modeled by the secondary network, edges with longer distances will be added because successively infected nodes will not necessarily be close to each other in the original network. A sample secondary network is shown in Figure 6.10.

This construction assumes that we know the structure of the human social network in terms of the human nodes and interaction edges, but not necessarily the interaction probabilities. We are not likely to know all of this information during an actual outbreak, but this unweighted version of the human social network is easier to approximate. We can investigate how the structure of these networks evolves over time, which may provide insight as to how we can determine whether a bioterror

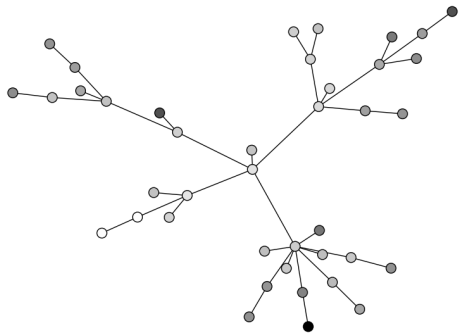


Figure 6.10: Sample secondary network instance for a multilayered network with 40 human nodes. Nodes are shaded according to their respective infection times, with darker nodes becoming infected at later times.

attack has occurred.

We explore how the structure of the secondary network evolves over time. We compute a very simple metric, the total secondary network length, to try to differentiate between the two scenarios. This metric is calculated by summing the edge distances as nodes are added to the secondary network. It requires minimal computational effort. It describes the state of the entire network, in contrast to closeness centrality or other node-specific metrics.

We begin our analysis with a baseline network using a BA algorithm density parameter equal to 5, a human-location density of 1.0, and a correlation coefficient equal to zero. We explore the dynamics of the secondary network when the human social network density changes. Figure 6.11 shows a comparison for each scenario of how the total secondary network length evolves for 100 simulation replications on a

single network instance with BA algorithm density parameters of 1, 5, and 10.

As we can see from Figure 6.11, total secondary network length is an excellent metric for differentiating between the two scenarios. The total secondary network length distinguishes between the two scenarios with high reliability when less than 5% of the population is infected. Other metrics may be as useful, but they would likely require significantly more computation. There is a clear separation between the two scenarios for all BA algorithm density values. The separation is most significant for the sparse human social network (i.e., BA algorithm density equal to 1), but the separation becomes smaller as the human social network becomes more dense. However, even in the most dense case, the total secondary network length grows much faster in the early stages of the outbreak for the bioterror scenario. The secondary network was designed to make this trend apparent, and the results confirm the intentions.

Early in the simulation of a bioterror attack, the secondary network is small (i.e., few humans have been infected) and each additional human that becomes infected will likely incur a penalty when they are inserted into the secondary network. As a result, longer edges are added to the secondary network when only a small proportion of the population has become infected. As the simulation progresses, the secondary network will grow larger, and most of the remaining susceptible human nodes will have an immediate neighbor from the original network to connect to when they become infected. This trend becomes apparent once the slope of the total secondary network length becomes equal to one. This behavior occurs for all of the epidemic simulations because, by definition, each newly infected node was

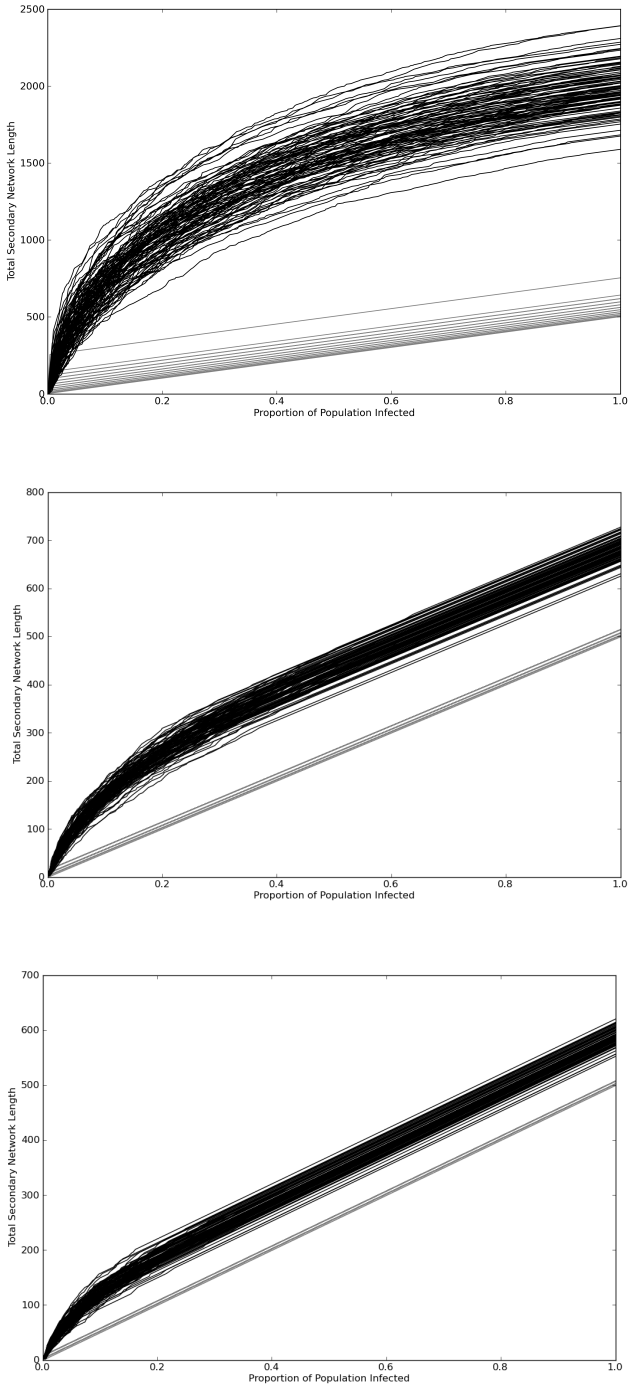


Figure 6.11: Total secondary network length comparison for different human social networks. Barabási-Albert algorithm density parameters of 1, 5, and 10 (top to bottom) were used to generate the human social networks from which the secondary networks were constructed. Epidemic results for 100 simulation replications are plotted in gray for a single network instance. Bioterror results for 100 simulation replications are plotted in black.

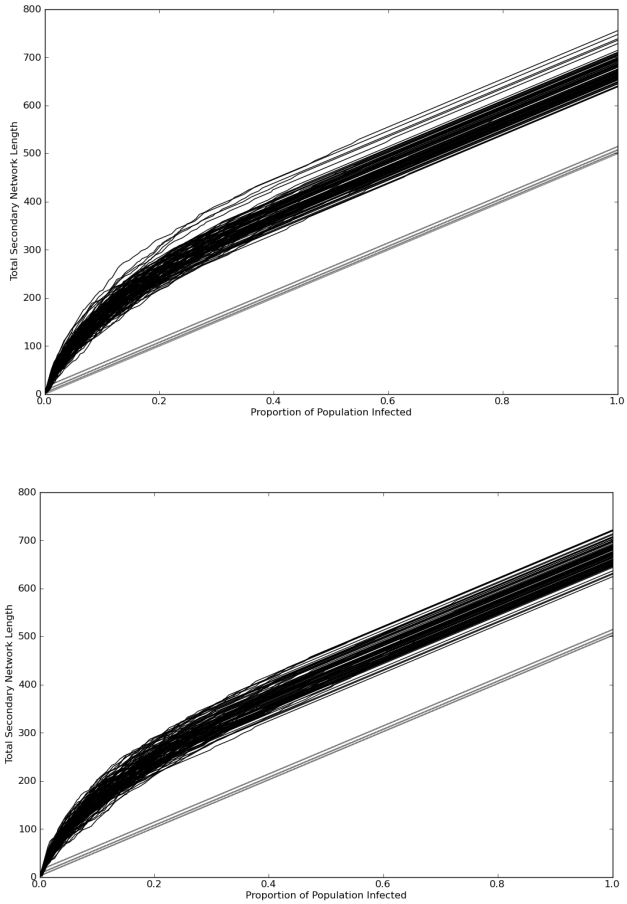


Figure 6.12: Total secondary network length comparison for different correlation coefficients. Correlation coefficients of $\rho = 0.5$ and 1.0 were used to generate the human-location networks. Epidemic results for 100 simulation replications are plotted in gray for a single network instance. Bioterror results for 100 simulation replications are plotted in black.

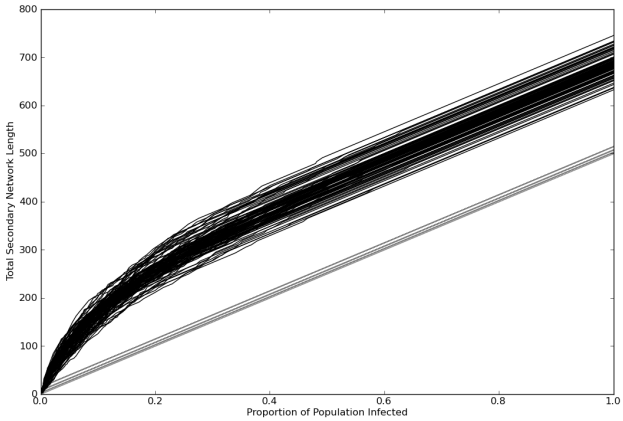


Figure 6.13: Total secondary network length comparison for a human-location network density of 0.5. Epidemic results for 100 simulation replications are plotted in gray for a single network instance. Bioterror results for 100 simulation replications are plotted in black.

infected by another node that was already added to the secondary network. The only penalized edges (i.e., those with distance greater than one) that are added to the secondary network in the epidemic scenario are between initially infected human nodes (i.e., index nodes). When an epidemic spreads through a larger population, there would likely be many more than two index nodes (which is what we used for our simulations). However, secondary network edges would be longer as well because distances between non-neighboring nodes would be longer. In this case, the total secondary network length curves for each scenario would likely shift up and the separation would be at least as large.

We explored the effect of the correlation coefficient and the human-location network density on our ability to differentiate between the two scenarios. Figure 6.12 shows the comparison for correlation coefficient values $\rho = 0.5$ and $\rho = 1.0$. As apparent from the figures, the correlation coefficient has little effect on the results

compared to the baseline case. Figure 6.13 shows the results for when the human-location density is decreased to 0.5, which, from Figure **??**, increases the amount of time for the entire population to become infected from a bioterror attack. However, changing this parameter has little effect on the effectiveness of our methods compared to the baseline case.

6.3.2 Detection Under Social Network Uncertainty

In Section 6.3.1, we assumed that the human social network is known. In practice, this information is not available to healthcare providers. Whereas the size of the social network and its density can be estimated using scale-free approximations, the contacts between people and the locations they visit are difficult to determine for the entire population. In this section, we develop a methodology to distinguish between the disease source that relies only on readily available information. This information includes network size (i.e., human population size and number of locations) and the cumulative number of infected humans, which could be computed by monitoring arrivals to physician's offices and emergency departments.

The cumulative infection curves output from the simulation model (see Figures **??** and **??**) provide information that is not available to healthcare providers and depends on the simulation model. We preprocess the curves to remove this information. Although the outbreak starts at time $t = 0$ in the simulation, healthcare providers can only potentially detect an outbreak upon the arrival of the first infected human to a physician's office or emergency department. We trim the process

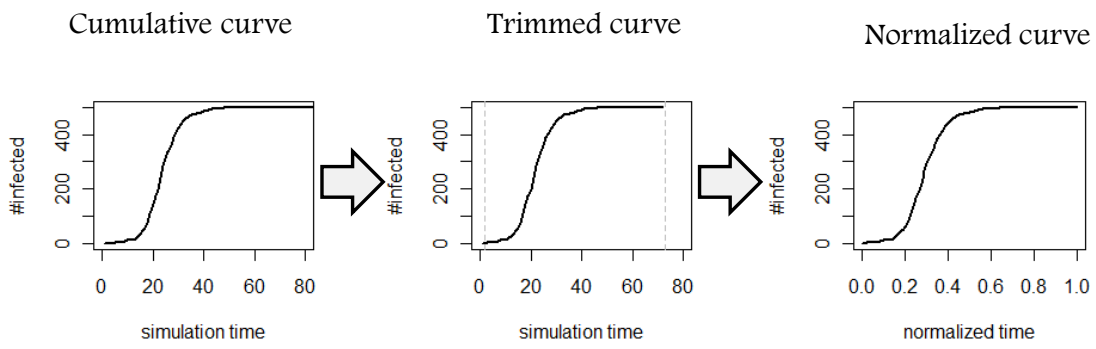


Figure 6.14: Schematic illustration of infection curve preprocessing.

to start at the first infection and end when the entire population has been infected. Second, the actual simulation time (counted as ticks) is a function of the infection probability. Transforming the simulation time to actual time (e.g., hours or days) requires a priori knowledge of this probability. Different infection probabilities produce curves with similar shapes, but on a different scale. We re-scale the simulation time to the interval [0,1], which denotes the fractional time between the onset of the disease and the time when the entire population is infected (herein denoted *time fractions*). An illustration of the preprocess step is given in Figure 6.14.

Now, we can compare the preprocessed infection curves of epidemic outbreaks to those of bioterror outbreaks and examine the differences in the transmission dynamics. As discussed in the beginning of Section 6.3, there are some clear distinctive trends between the two scenarios. In Figure 6.14, we show that the same trends are preserved even after preprocessing the curves. We quantify these differences and show that these differences are statistically significant between the two scenarios.

We use Functional Data Analysis (FDA) to distinguish between the epidemic

and bioterror scenarios. FDA is a set of flexible techniques to capture the dynamics in curves, surfaces, or images [72]. One of the most commonly used FDA methods is an adaptation of classical Principal Component Analysis (PCA) known as Functional PCA (fPCA). fPCA demonstrates the way in which a set of functional data varies from its mean. It quantifies the difference from the mean of each individual functional datum [50]. We apply fPCA to the observed, preprocessed, cumulative infection curves to capture differences between the two disease scenarios.

fPCA is similar in nature to classical PCA. However, rather than operating on data vectors, it operates on functional objects, which in our context corresponds to cumulative infection curves. Classical PCA is a mathematical procedure that projects a set of data vectors $X = \{X_1, \ldots, X_n\}$ ($X_i = \{x_{i1}, \ldots, x_{ip}\}$) onto a new space Y ($\{Y_1, \ldots, Y_n\}$ and $Y_i = \{y_{i1}, \ldots, y_{ip}\}$) that maximizes the variance along each component in the new space. At the same time, it renders the individual components of the new space orthogonal to each another. The result is a new coordinate system such that the greatest variance lies on the first coordinate (i.e., the *first principal component*), the second greatest variance lies on the second coordinate (i.e., *second principal component*), and so on. Formally, the new space Y is computed as follows:

$$Y^T = (X^T - \bar{X})W \tag{6.6}$$

where W is the matrix of eigenvectors of XX^T. Common practice is to choose only those eigenvectors that correspond to the largest eigenvalues, that is, those that

capture most of the information (variation) in the observed data.

The functional version of PCA is similar in nature, except that it operates on a set of continuous curves rather than discrete vectors. Consequently, summations are replaced by integrals, and eigenvectors by eigenfunctions

$$Y^T = \int_t (X(t)^T - \bar{X})W(t) \tag{6.7}$$

.

The continuous curves are approximated by a fine discrete grid, such that each time point on the curve becomes a component in the observed data vector. In other words, data vectors $X = \{X_{t_1}, ..., X_{t_n}\}$, where t_i is the curve level time i (in our context, *time fraction* at time i), are used as a proxy to $X(t)$. Then, classical PCA is carried out on the discrete curve [43, 73].

We apply the discrete version of fPCA to the preprocessed cumulative infection curves. We find that the first three eigenvectors capture nearly 99% of the variation in the data, and they have a meaningful interpretation in our context. In Figure 6.15, we plot these components. Each figure depicts the *loads* of the infection curves on each principal component at each time fraction, that is, how much of its information is captured by each component. The first principal component captures the *speed* or the *rate* at which the population gets infected, relative to the average infection curve. It exhibits a high load in the middle time fractions and a low load at the low and high time fractions. In terms of the simulation, we observe small loads in the first component at the beginning of the outbreak, during which few individuals

get infected. Then the outbreak accelerates and humans get infected at a high rate, which is represented by the peak load at a time fraction of approximately 0.35. Finally, as the number of susceptible humans decreases, the rate of infection slows and the load of the first component decreases back to zero.

The second and third principal components capture the *velocity* and *acceleration* of the infection curves, respectively. Velocity is the slope of the curve (i.e., first derivative), which becomes largest halfway between the initial infections and when the infection rate is maximum (in Figure 6.15, the largest load is at fractional time 0.2). Acceleration is the rate of change of the velocity component (i.e., second derivative).

To analyze the effectiveness of fPCA in differentiating between disease scenarios, we consider a subset of the networks defined in Table 6.1, beginning with five network instances defined by the BA algorithm and human-location density parameters being equal to one. We use the remainder of the networks to evaluate the performance of our methods and to examine its robustness to changes in the population (e.g., human social network density). Our analysis in this section consists of three steps:

1. Perform a visual comparison of PCA results from the two scenarios.

2. Perform formal analysis to quantify the properties (i.e., speed, velocity, and acceleration) of each scenario.

3. Perform evaluation and robustness tests.

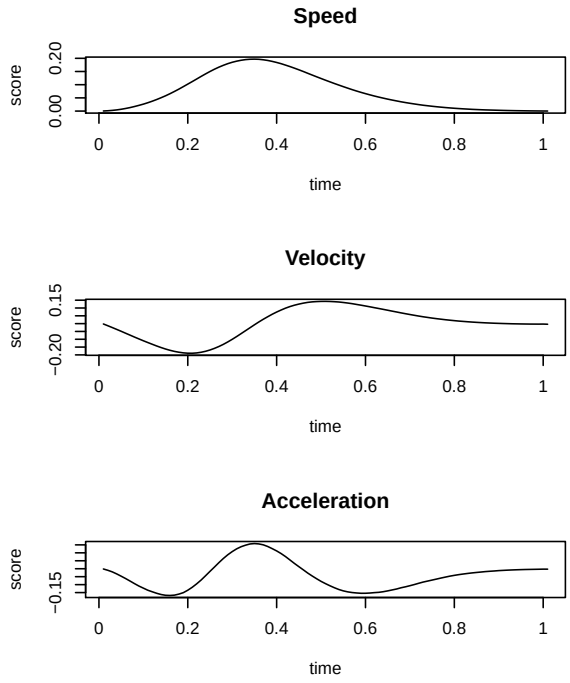


Figure 6.15: Infection curve loads for the first three principal components. The first principal component is representative of the speed (i.e., rate) of the infection. The second principal component is representative of the velocity of the infection. The third principal component is representative of the acceleration of the infection.

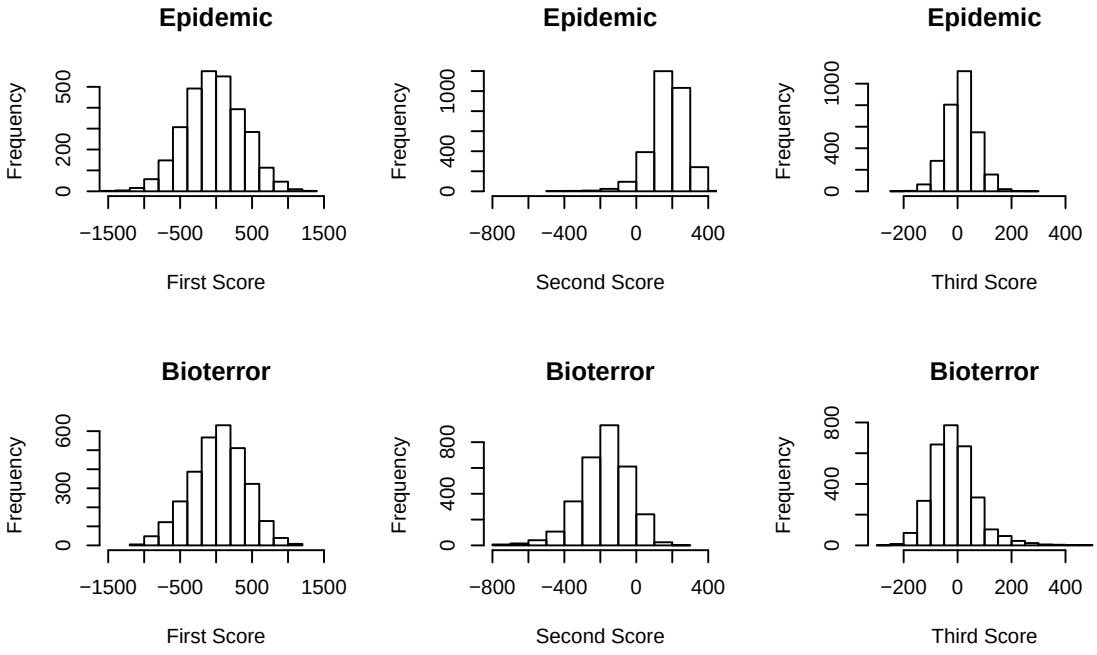


Figure 6.16: Principal component score distributions for the epidemic and bioterror scenarios.

To visually compare the properties of the curves, we examine the distribution of the first three principal component *scores* of each curve, that is, their projection on the new coordinate system. In Figure 6.16, we plot this distribution for each disease scenario. Although the scenarios do not differ by speed, they differ significantly by velocity and acceleration. The epidemic scenario generates cumulative infection curves with a higher average velocity when compared to the bioterror case. In contrast, the bioterror curve has higher velocity and acceleration *variability*.

We use ordinary logistic regression analysis to quantify the effect of using PCA to differentiate between the two disease scenarios. The predictive model is a function of the first three principal components, as follows:

$$\hat{Scenario} = \frac{1}{1 + exp^{-(\beta_0 + \beta_1 PCS_1 + \beta_2 PCS_2 + \beta_3 PCS_3 + \varepsilon)}} \tag{6.8}$$

The estimated model is given in Table 6.2. As expected, the coefficient of the first principal component score (β_1) is close to zero because both scenarios have approximately equal loads in this component. However, the coefficients of the second (β_2) and third principal component scores (β_3) are an order of magnitude larger, which implies that these components are more important in differentiating between the two scenarios.

The performance of the model on the training set is outlined in the confusion matrix in Table 6.3, which summarizes the number of correctly and incorrectly predicted outcomes for each disease scenario. The accuracy of the model is greater than 95%. The error rate distributes evenly between negative and positive rates. To evaluate the model, we examine its performance on a validation set of five networks with the same density parameters as those in the training set. For each replication of a network in the training set, we compute the first three principal components scores according to the loadings in Figure 6.15 and then predict the scenario using the model in Equation 6.8. We summarize the performance in Table 6.3. The overall accuracy is greater than 95%, and the positive and negative error rates are statistically equal.

In practice, the multilayered network parameters can be estimated, but are essentially unknown. Therefore, it is important to see how our model performs when the estimates are inaccurate. We examine the robustness of the model to changes

Table 6.2: Estimated parameters of the logistic model

Parameter	**Estimate**	**Standard Deviation**	***p* value**
β_0	0.7309185	0.1147268	1.88e-10
β_1	0.0020193	0.0001709	$<$ 2e-16
β_2	-0.0272249	0.0008508	$<$ 2e-16
β_3	-0.0279236	0.0014694	$<$ 2e-16

Table 6.3: Performance of the logistic model on training and validation data sets

	Training		**Validation**	
	Response		*Response*	
Prediction	Epidemic	Bioterror	Epidemic	Bioterror
Epidemic	956	40	473	24
Bioterror	44	960	27	476

in the network parameters. Specifically, we decrease the density of the human social network and the human-location network. The performance of the model is evaluated on these new networks *without* recomputing the principal component loadings or the logistic model parameters. Table 6.4 summarizes the performance.

In Table 6.4, we observe that the model does not perform well for sparse human social networks. When the BA density parameter decreases from 10 to 5, we observe only a slight increase in the false negative rate (i.e., identifying epidemic outbreaks as bioterror attacks). However, the false negative rate increases significantly when the BA algorithm density parameter decreases to 1. In this case, the model predicts that nearly 78% of the outbreaks are bioterror attacks. These predictions occur because infectious agents in the epidemic scenario can only infect a few contacts at most, which results in a gradual infection curve that is more representative of bioterror transmission dynamics (refer to the similar transmission dynamics between the top

Table 6.4: Robustness of the logistic model to changes in human social and human-location network densities

	$m = 1$		$m = 5$	
	Response		*Response*	
Prediction	Epidemic	Bioterror	Epidemic	Bioterror
Epidemic	82	6	188	7
Bioterror	118	194	12	193

	$d = 0.5$		$d = 0.8$	
	Response		*Response*	
Prediction	Epidemic	Bioterror	Epidemic	Bioterror
Epidemic	190	13	190	6
Bioterror	10	187	10	194

plot in Figure **??** and all plots in Figure **??**). The model is much less sensitive to changes in the density of the human-location network. We observe a slight increase in the false positive rate (i.e., identifying a bioterror attack as an epidemic) when the density decreases from 1 to 0.5.

We have demonstrated how FDA can capture the dynamics in the cumulative infection curves and distinguish between the disease scenarios reliably. However, early detection of the outbreak sources is the critical factor in preventing diffusion of the disease to the entire population. We pose the question of how early we can detect the infection source, or more specifically: What proportion of the population will get infected before we can accurately differentiate between the two scenarios? What is the level of confidence for this differentiation?

To answer these questions, we examine the cumulative infection curves up to the time when $N_H \times f$ of the population becomes infected, for different fractions

$f \in (0, 1]$. An example of the cumulative infection curve comparison for $f = 0.1$ is given in Figure 6.17, which shows that the infection curves for the two scenarios are very distinguishable. We find that when $f \geq 0.07$, that is 7% of the population is infected, the model predicts at least 95% of the disease scenarios accurately. When $f = 0.05$, detection performance only decreases to 93%. Finally, when only 2% of the population is infected (i.e., $f = 0.02$), the model accuracy is still 83%. Overall, this method can be employed when very small proportions of the population have become infected and provides extremely reliable predictions. In addition, this method requires no information about the structure of the multilayered network.

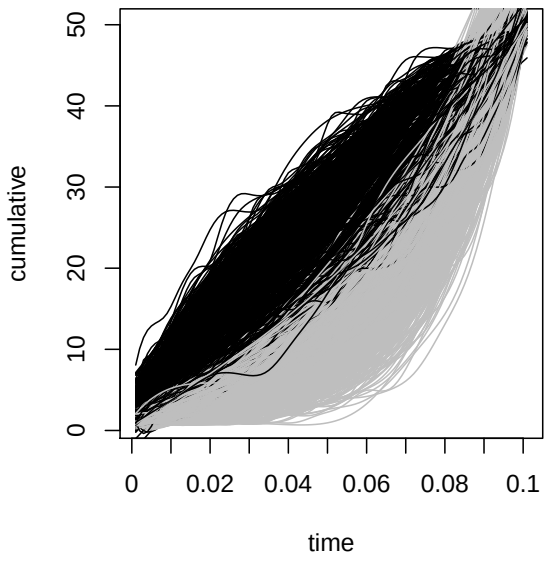


Figure 6.17: Comparison of cumulative infection curves after 10% of the population is infected. Epidemic simulations are plotted in gray and bioterror simulations are plotted in black.

6.4 Conclusions

Early detection of a bioterror attack can facilitate the isolation of the source location, which will prevent large numbers of people from becoming infected, reduce the demands on the healthcare system to provide care, and minimize the impact on schools and businesses. If a bioterror attack is incorrectly identified as an epidemic, control measures aimed at preventing human-to-human transmission will not necessarily protect people from getting infected at a compromised location. These misdirected efforts will drain resources that will be needed to address the attack once it is identified. We have provided two techniques that may assist healthcare providers in distinguishing bioterror attacks from epidemics. In addition, these methods could be used to identify other types of infections that spread from a single location, such as food poisoning or an infected water source.

When information is available about the structure of the human social network, we can construct secondary networks and monitor how the structure evolves with time. Our results showed that we can differentiate between the disease scenarios under these circumstances when very small proportions of the population are infected (i.e., less than 5%). Secondary networks for bioterror attacks will have longer total lengths because newly infected humans may not have any significant interaction with others that have already been infected. In epidemic scenarios, newly infected humans will always have contact with at least one other human that has already been infected and, therefore, secondary network lengths will be smaller. For cases when less information is known about the social network (e.g., larger popula-

tions), we may be able to estimate the structure of the network, but by doing so we lose accuracy in differentiating between the two scenarios. Future work may explore how this uncertainty affects the detection capability of these methods.

In most cases, we will have little or no information about the human social network, and we have to rely on the information available about the number of people infected each day. Functional data analysis can provide detailed information about the particular behavior of a given transmission curve. By using historical data, we can generate a series of benchmark epidemic infection curves and then monitor the principal components of an ongoing outbreak to identify bioterror attacks.

Both techniques demonstrated extremely high accuracy in differentiating between the two scenarios for several parameterized multilayered networks. The social network method was most effective for sparse human social networks, which would be easier to characterize. On the other hand, functional data analysis was effective for sufficiently dense human social networks. These types of networks are representative of urban populations that are most likely to be targeted by a bioterror attack. In addition, both methods were able to identify bioterror attacks when only a small proportion of the population became infected ($f < 10\%$), which is a critical capability for implementing effective infection control measures. These techniques would be fairly easy to implement in practice, depending on the available information, and could be incorporated into syndromic surveillance systems to monitor infections in real time [39]. These measures would ensure that appropriate infection control measures are taken as soon as possible.

Chapter 7

Conclusions

Building on the foundation established by systems dynamics and discrete event simulation, agent-based modeling has provided new insight to epidemiological problems by modeling individuals and the interactions between them. This perspective has facilitated analysis at both the individual and system levels, which is not typically possible using other methods. The greatest value of agent-based models is that they can be used as virtual environments to evaluate policy alternatives, some of which would be infeasible or unethical to experiment with in practice. This capability can help healthcare organizations make better decisions, which can potentially lead to a higher quality of care for patients and extensive cost savings.

Overall, agent-based models are exceptionally well-suited for testing potential infection control measures and can provide some indication of success before implementing any particular strategy. Process-oriented models have a more established record of contribution, and they have produced relatively widely accepted results concerning the effectiveness of common infection control measures. The foundation has been established for network models of transmission, but there are still many areas to explore, particularly in determining which network measures are the most effective in predicting an outbreak and differentiating between different types of transmission. Both static and dynamic network measures may be useful, but the

ease of calculation is a critical issue that will determine whether these techniques are used in practice. However, the benefit of modeling the explicit connections between individuals has been demonstrated in the unique results from these types of models.

There are several key results that agent-based models have reinforced and several more that are unique to this methodology. The hand-hygiene compliance of HCWs is critical to preventing outbreaks of infectious diseases, but in most cases, handwashing is not sufficient to control transmission. Additional measures such as active surveillance, patient isolation, or higher staffing ratios become necessary. To the extent possible, minimizing patient lengths of stay and HCW visits to patients can also reduce the risk of secondary infections, as susceptible patients are exposed less to potential infections and infected patients have fewer opportunities to infect HCWs.

Agent-based models have also provided insight into the relative threats posed by HCWs that care for patients in a hospital. Nurses and other HCWs that visit patients frequently are at a high risk of becoming at least transiently colonized or infected, and can quickly spread disease to other patients in their care. Physicians and other HCWs that visit many patients can potentially create multiple pockets of infection that could lead to an entire unit becoming infected. Maintaining high staff-to-patient ratios and assigning patients to HCWs in a structured way can offset these dangers. Future studies may be able to suggest methods for identifying individuals at a high risk of infection because of their location in a network, and appropriate measures could be taken to prevent that person from becoming infected before isolation or quarantine measures become necessary.

As with the other modeling methodologies, the challenge in developing agent-based models is first generating realistic dynamics for a particular scenario, and then analyzing those patterns to identify the best strategies for intervention. Determining the appropriate level of detail for an agent-based model can be a difficult challenge. Extremely detailed models require many parameters, and it is difficult to determine accurate values for these parameters from empirical data and the research literature. Parameters are often set at reasonable values and sensitivity analyses are performed to address any shortcomings associated with poor selection, but these analyses can become very time consuming if model run times are large. Models with few parameters mostly avoid this problem, but are sometimes challenged due to their lack of detail and complexity. One solution to this issue is to use hybrid models, which incorporate systems dynamics and agent-based methods. In Chapter 4, we modeled a network of healthcare facilities whose states were affected by both their interaction with other facilities and an internal differential equations model. These models take advantage of the strengths of each methodology, most importantly the lower computational demands of systems dynamics and the heterogeneity afforded by agent-based modeling. For all levels of complexity, agent-based models can incorporate and generate large quantities of data. As a consequence, statistical rigor, efficient data analysis techniques, and visualization are all critical to producing insightful results and communicating those findings to healthcare professionals.

The other major challenge with agent-based models is related to validation. The level of validation required for a particular agent-based model is primarily driven by its ultimate purpose. Models that are intended to produce accurate quantitative

(i.e., predictive) results may require extensive validation, whereas more qualitative (i.e., illustrative) models have less stringent requirements. Validation is difficult for some applications because they lack empirical data as a baseline for comparison. Therefore, it is not always possible to know what simulation outcomes should look like.

When empirical data are available, there are several approaches for using them to validate a model [29]. The best of these approaches primarily involves incorporating empirical data, historical perspective, and subject matter expertise in the selection of the best possible set of model assumptions, parameter settings, and initial conditions. By combining these data sources, agent-based models are more likely to use appropriate values for input parameters and generate reasonable system responses. However, validation is more difficult when the underlying dynamics are not as well-understood or there is no historical reference, as is generally the case for epidemic and bioterror scenarios. For these cases, an agent-based model has more value in demonstrating relative trends or generating alternative outcomes for comparison, and less value as a predictive model. For any validation approach, it is always necessary to perform sensitivity analyses to identify any parameters that cause non-proportional changes in the system.

Moving forward, new and existing agent-based models should continue to build on the most relevant achievements of other models, including those from other fields that have taken advantage of the methodology. Making these models widely available is a practice that can help to accelerate progress. This evolutionary process is another advantage of agent-based modeling, as system dynamics and discrete event

simulations are more difficult to augment with additional complexity. Each individual model may have limitations, but as the research grows, many of those limitations can eventually be eliminated, thereby increasing the acceptance of these methods by healthcare organizations.

Bibliography

[1] Albert, R., A.L. Barabási. 2002. Statistical mechanics of complex networks. *Rev Mod Phys* 74(1):4797.

[2] Angus, D.C., W.T. Linde-Zwirble, C.A. Sirio, A.J. Rotondi, L. Chelluri, R.C. Newbold III, J.R. Lave, and M.R. Pinsky. 1996. The effect of managed care on ICU length of stay: implications for medicare. *J Amer Med Assoc* 276:1075-1082.

[3] Antia, R., M. Lipsitch. 1997. Mathematical models of parasite responses to host immune defences. *Parasitology* 115 Suppl:S155-67.

[4] Austin, D.J., R.M. Anderson. 1999. Studies of antibiotic resistance within the patient, hospitals and the community using simple mathematical models. *Phil Trans Roy Soc Lond B* 354(1384):721-738.

[5] Austin, D.J., M.J. Bonten, R.A. Weinstein, S. Slaughter, and R.M. Anderson. 1999. Vancomycin-resistant enterococci in intensive-care hospital settings: transmission dynamics, persistence, and the impact of infection control programs. *Proc Natl Acad Sci U S A* 96(12):6908-6913.

[6] Bansal, S., B.T. Grenfell, L.A. Meyers. 2007. When individual behaviour matters: homogeneous and network models in epidemiology. *J Roy Soc Interface* 4(16):879-891.

[7] Barnes, S., B. Golden, E. Wasil. 2010. MRSA Transmission Reduction Using Agent-Based Modeling and Simulation. *INFORMS J Comput* 22(4):635-646.

[8] Barnes, S., B. Golden, E. Wasil. 2010. A dynamic patient network model of hospital-acquired infections. In: B. Johansson, S. Jain, J. Montoya-Torres, J. Hugan, E. Yücesan, eds. *Proc 2010 INFORMS Winter Simul Conf* 2249-2260.

[9] Beggs, C.B., C.J. Noakes, S.J. Shepherd, K.G. Kerr, P.A. Sleigh, K. Banfield. 2006. The influence of nurse cohorting on hand hygiene effectiveness. *Am J Infect Control* 34:621-626.

[10] Beggs, C.B., S.J. Shepherd, K.G. Kerr. 2008. Increasing the frequency of hand washing by healthcare workers does not lead to commensurate reductions in staphylococcal infection in a hospital ward. *BMC Infect Dis* 11:1-11.

[11] Bootsma, M.C.J., O. Diekmann, M.J.M. Bonten. 2006. Controlling methicillin-resistant Staphylococcus aureus: quantifying the effects of interventions and rapid diagnostic testing. *Proc Natl Acad Sci U S A* 103(14):5620-5625.

[12] Boyce, J.M., D. Pittet, Healthcare Infection Control Practices Advisory Committee, Society for Healthcare Epidemiology of America, Association for Professionals in Infection Control, Infectious Diseases Society of America, Hand Hygiene Task Force. 2002. Guideline for Hand Hygiene in Health-Care Settings: recommendations of the Healthcare Infection Control Practices Advisory Committee and the HICPAC/SHEA/APIC/IDSA Hand Hygiene Task Force. *Infect Control Hosp Epidemiol* 23:S3-40.

[13] Buehlmann, M., R. Frei, L. Fenner, M. Dangel, U. Fluckiger, A.F. Widmer. 2008. Highly effective regimen for decolonization of methicillin-resistant Staphylococcus aureus carriers. *Infect Control Hosp Epidemiol* 29(6):510-516.

[14] Caffrey, C. 2010 Potentially preventable emergency department visits by nursing home residents: United States, 2004. *NCHS Data Brief* Apr(33):1-8.

[15] Carley, K.M., D. Fridsma, E. Casman, A. Yahja, N. Altman, L-C. Chen, B. Kaminsky, D. Nave. 2006. BioWar: Scalable Agent-based Model of Bioattacks. *IEEE T Syst Man Cy A* 36(2):252-265.

[16] U.S. Centers for Disease Control and Prevention. 2011. Bioterrorism Agents/Diseases. Available at http://www.bt.cdc.gov/agent/agentlist-category.asp [accessed November 28, 2011].

[17] Christley, R.M. 2005. Infection in Social Networks: Using Network Analysis to Identify High-Risk Individuals. *Am J Epidemiol* 162(10):1024-1031.

[18] Cooper, B.S., G.F. Medley, G.M. Scott. 1999. Preliminary analysis of the transmission dynamics of nosocomial infections: stochastic and management effects. *J Hosp Infect* 43(2):131-147.

[19] Cooper, B.S., G.F. Medley, S.P. Stone, C.C. Kibbler, B.D. Cookson, J.A. Roberts, G. Duckworth, R. Lai, S. Ebrahim. 2004 Methicillin-resistant Staphylococcus aureus in hospitals and the community: stealth dynamics and control catastrophes. *Proc Natl Acad Sci U S A* 101(27):10223-10228.

[20] Congressional Record. July 14, 2000: 10685-10686.

[21] Cosgrove, S.E., Y. Qi, K.S. Kaye, S. Harbarth, A.W. Karchmer, Y. Carmeli. 2005. The impact of methicillin resistance in Staphylococcus aureus bacteremia

on patient outcomes: mortality, length of stay, and hospital charges. *Infect Control Hosp Epidemiol* 26:166-174.

[22] Cromer, A.L., S.C. Latham, K.G. Bryant, S. Hutsell, L. Gansauer, H.A. Bendyk, R. Steed, M.C. Carney. 2008. Monitoring and feedback of hand hygiene compliance and the impact on facility-acquired methicillin-resistant Staphylococcus aureus. *Am J Infect Control* 36:672-677.

[23] Cummings, D., D.S. Burke, J.M. Epstein, R.M. Singa, S. Chakravarty. 2004. Toward a Containment Strategy for Smallpox Bioterror: An Individual-Based Computational Approach. *Brookings Inst Pr* 1-55.

[24] Curtis, L.T. 2008. Prevention of hospital-acquired infections: a review of non-pharmacological interventions. *J Hosp Infect* 69(3):204-219.

[25] D'Agata, E.M.C., P. Magal, D. Olivier, S. Ruan, G.F. Webb. 2007. Modeling antibiotic resistance in hospitals: the impact of minimizing treatment duration. *J Theor Biol* 249(3):487-499.

[26] Duggan, J.M., S. Hensley, S. Khuder, T.J. Papadimos, L. Jacobs. 2008. Inverse correlation between level of professional education and rate of handwashing compliance in a teaching hospital. *Infect Control Hosp Epidemiol* 29:534-538.

[27] Duke-Sylvester, S.M., E.N. Perencevich, J.P. Furuno, L.A. Real, H. Gaff. 2008. Advancing epidemiological science through computational modeling: a review with novel examples. *Annales Botanici Fennici* 45:385-401.

[28] Eubank, S. 2005. Network based models of infectious disease spread. *Jpn J Infect Dis* 58(6):S9-13.

[29] Fagiolo, G., A. Moneta, P. Windrum. 2007. A Critical Guide to Empirical Validation of Agent-Based Models in Economics: Methodologies, Procedures, and Open Problems. *Comput Econ* 30(3):195-226.

[30] Fone, D., S. Hollinghurst, M. Temple, A. Round, N. Lester, A. Weightman, K. Roberts, E. Coyle, G. Bevan, S. Palmer. 2003. Systematic review of the use and value of computer simulation modelling in population health and health care delivery. *J Public Health* 25(4):325-335.

[31] French, G.L. 2009. Methods for screening for methicillin-resistant Staphylococcus aureus carriage. *Clini Microbiol Infect* 15 Suppl(7):10-16.

[32] Fridkin, S.K., S.M. Pear, T.H. Williamson, J.N. Galgiani, W.R. Jarvis. 1996. The role of under-staffing in central venous catheter-associated bloodstream infections. *Infect Control Hosp Epidemiol* 17:150-158.

[33] Furuno, J.P., J.N. Hebden, H.C. Standiford, E.N. Perencevich, R.R. Miller, A.C. Moore, S.M. Strauss, A.D. Harris. 2008. Prevalence of methicillin-resistant Staphylococcus aureus and Acinetobacterbaumannii in a long-term acute care facility. *Am J Infect Control* 36(7):468-471.

[34] Girou, E., S. Loyeau, P. Legrand, F. Oppein, C. Brun-Buisson. 2002. Efficacy of handrubbing with alcohol based solution versus standard handwashing with antiseptic soap: randomised clinical trial. *BMJ* 325:362.

[35] Griffin, F.A.. 5 Million Lives Campaign. 2007. Reducing methicillin-resistant Staphylococcus aureus (MRSA) infections. *Jt Comm J Qual Patient Saf* 33(12):726-731.

[36] Grundmann, H., S. Hori, B. Winter, A. Tami, D.J. Austin. 2002. Risk factors for the transmission of methicillin-resistant Staphylococcus aureus in an adult intensive care unit: fitting a model to the data. *J Infect Dis* 185:481-488.

[37] Hall, M.J., C.J. DeFrances, S.N. Williams, A. Golosinskiy, A. Schwartzman. 2010. National Hospital Discharge Survey: 2007 summary. *Natl Health Statist Reports* 2010(29):1-20,24.

[38] Harris, A.D., J.C. McGregor, E.N. Perencevich, J.P. Furuno, J. Zhu, D. Peterson, J. Finklestein. 2006. The Use and Interpretation of Quasi-Experimental Studies in Medical Informatics. *J Amer Med Informatics Assoc* 13(1):16-23.

[39] Henning, K.J. 2004. Overview of Syndromic Surveillance: What is Syndromic Surveillance? *Morbidity and Mortality Weekly Report* 53(Suppl):5-11.

[40] Higham, N.J. 1988. Computing a nearest symmetric positive semidefinite matrix. *Linear Algebra Appl* 103:103-118.

[41] Honda, H., M.J. Krauss, C.M. Coopersmith, M.H. Kollef, A.M. Richmond, V.J. Fraser, D.K. Warren. 2010. Staphylococcus aureus nasal colonization and subsequent infection in intensive care unit patients: does methicillin resistance matter? *Infect Control Hosp Epidemiol* 31:584-591.

[42] Hotchkiss, J.R., D.G. Strike, D.A. Simonson, A.F. Broccard, P.S. Crooke. 2005. An agent-based and spatially explicit model of pathogen dissemination in the intensive care unit. *Crit Care Med* 33(1):168-176.

[43] Jank, W., I. Yahav. 2010. E-Loyalty Networks in Online Auctions. *Ann Appl Stat* 4(1):151-178.

[44] Jun, J.B., S.H. Jacobson, J.R. Swisher. 1999. Application of discrete-event simulation in health care clinics: A survey. *J Oper Res Soc* 50(2):109-123.

[45] Iwashyna, T.J., A.A. Kramer, J.M Kahn. 2009. Intensive care unit occupancy and patient outcomes. *Crit Care Med* 37:1545-1557.

[46] Kanagarajah, A., P. Lindsay, A. Miller, D. Parker. 2006. An Exploration into the Uses of Agent Based Modeling to Improve Quality of Health Care. In: A. Minai, D. Braha, Y. Bar-Yam, eds. *Proc 6th International Conf Complex Systems* 1-10.

[47] Keeling, M. 2005. The implications of network structure for epidemic dynamics. *Theoret Population Biol* 67:1-8.

[48] Kermack, W.O., A.G. McKendrick. 1927. A Contribution to the Mathematical Theory of Epidemics. *Proc Roy Soc Lond A* 115:700-721.

[49] Kirkland, K.B., J.M. Weinstein. 1999. Adverse effects of contact isolation. *Lancet* 354:1177-1178.

[50] Kneip, A., K.J. Utikal. 2001. Inference for Density Families Using Functional Principal Component Analysis. *J Amer Stat Assoc* 96(454):519-542.

[51] Koelling, P., M.J. Schwandt. 2005. Health Systems: A Dynamics SystemBenefits From System Dynamics. In: Kuhl ME, Steiger NM, Armstrong FB, Joines JA, eds. *Proc 2005 INFORMS Winter Simul Conf* 1321-1327.

[52] Lesosky, M., A. McGeer, A. Simor, K. Gren, D.E. Low, J. Raboud. 2011. Effect of Patterns of Transferring Patients among Healthcare Institutions on Rates of Nosocomial Methicillin-Resistant Staphylococcus aureus Transmission: A Monte Carlo Simulation. *Infect Control Hosp Epidemiol* 32(2):136-147.

[53] Macal, C.M., M.J. North. 2007. Agent-based modeling and simulation: Desktop ABMS. In: S.G. Henderson, B. Biller, M.-H. Hsieh, J. Shortle, J.D. Tew, R.R. Barton, eds. *Proc 2007 INFORMS Winter Simul Conf* 95-106.

[54] McBryde, E.S., A.N. Pettitt, D.L.S. McElwain. 2007. A stochastic mathematical model of methicillin resistant Staphylococcus aureus transmission in an intensive care unit: predicting the impact of interventions. *J Theor Biol* 245(3):470-481.

[55] McCaughey, B. 14 Aug 2008. Hospital Infections: Unacceptable and Preventable. *Wall Street Journal* A11.

[56] McKenzie, F.E., E.M. Samba. 2004. The role of mathematical modeling in evidence-based malaria control. *Am J Trop Med Hyg* 71:94-96.

[57] Meng, Y., R. Davies, K. Hardy, P. Hawkey. 2010. An application of agent-based simulation to the management of hospital-acquired infection. *J Simul* 4(1):60-67.

[58] Meyers, L.A., B. Pourbohloul, M.E.J. Newman, D.M. Skowronski, R.C. Brunham. 2005. Network theory and SARS: predicting outbreak diversity. *J Theor Biol* 232(1):71-81.

[59] Milstone, A.M., C.L. Passaretti, T.M Perl. 2008. Chlorhexidine: expanding the armamentarium for infection control and prevention. *Clin Infect Dis* 46(2):274-281.

[60] Moland, E.S., N.D. Hanson, J.A. Black, A. Hossain, W. Song, K.S. Thomson. 2006. Prevalence of newer beta-lactamases in gram-negative clinical isolates collected in the United States from 2001 to 2002. *J Clin Microbiol* 44:3318-3324.

[61] Moss A. 2004. National Nursing Home Survey. 2006:1-23. Available at http://www.cdc.gov/nchs/fastats/nursingh.htm [accessed July 5, 2011].

[62] Muto, C.A., J.A. Jernigan, B.E. Ostrowsky, H.M Richet, W.R. Jarvis, J.M. Boyce, B.M. Farr. 2003. SHEA guideline for preventing nosocomial transmission of multidrug-resistant strains of Staphylococcus aureus and enterococcus. *Infect Control Hosp Epidemiol* 24:362-386.

[63] National Center for Health Statistics. 2010. Health, United States. 2010:1-563. Available at http://www.cdc.gov/nchs/fastats/hospital.htm [accessed July 5, 2011].

[64] NetLogo Agent-Based Modeling Software. Available at http://ccl.northwestern.edu/netlogo/ [accessed March 12, 2009].

[65] Ojajarvi, J. 1980. Effectiveness of hand washing and disinfection methods in removing transient bacteria after patient nursing. *J Hyg* 85:193-203.

[66] Ong, B., M. Chen, V. Lee, J. Tay. 2008. An individual-based model of influenza in nosocomial environments. *Comput Sci, Int Conf Comput Sci 2008*, Part I, Lecture Notes Comput Sci 5101:590599.

[67] Pittet, D., S. Hugonnet, S. Harbarth, P. Mourouga, V. Sauvan, S. Touveneau, T.V. Perneger. 2000. Effectiveness of a hospital-wide programme to improve compliance with hand hygiene. Infection Control Programme. *Lancet* 356:1307-1312.

[68] Python Programming Language. Available at www.python.org [accessed August 22, 2009].

[69] Raboud, J., R. Saskin, A. Simor, M. Loeb, K. Green, D.E. Low, A. Mcgeer. 2005. Modeling Transmission of Methicillin Resistant Staphylococcus Aureus Among Patients Admitted to a Hospital. *Infect Control Hosp Epidemiol* 26(7):607-615.

[70] Raboud, J., R. Saskin, K. Wong, C. Moore, G. Parucha, J. Bennett, K. Green, D. Low, M. Loeb, A. Simore, A. McGreer. 2004. Patterns of handwashing behavior and visits to patients on a general medical ward of healthcare workers. *Infect Control Hosp Epidemiol* 25:198-202.

[71] Rahmandad, H., J. Sterman. 2008. Heterogeneity and Network Structure in the Dynamics of Diffusion: Comparing Agent-Based and Differential Equation Models. *Manag Sci* 54(5):998-1014.

[72] Ramsay, J.O., B.W. Silverman. 1997. *Functional Data Analysis.* New York: Springer.

[73] Ramsay, J.O., B.W. Silverman. 2005. *Functional Data Analysis, Second Edition.* New York: Springer.

[74] Committee to Reduce Infection Deaths. Available at www.hospitalinfection.org [accessed August 22, 2009].

[75] Ridenour, G., R. Lampen, J. Federspiel, S. Kritchevsky, E. Wong, M. Climo. 2007. Selective use of intranasal mupirocin and chlorhexidine bathing and the incidence of methicillin-resistant Staphylococcus aureus colonization and infection among intensive care unit patients. *Infect Control Hosp Epidemiol* 28:1155-1161.

[76] Robotham, J.V., D.R. Jenkins, G.F. Medley. 2007. Screening strategies in surveillance and control of methicillin-resistant Staphylococcus aureus (MRSA). *Epidemiol Infect* 135(2):328-342.

[77] Salary.com. Career Advancement Tools and Resources. Available at http://swz.salary.com/salarywizard/layouthtmls/swzl_compresult_national_HC07000274.html [accessed July 13, 2010].

[78] Sebille, V., S. Chevret, A.J. Valleron. 1997. Modeling the spread of resistant nosocomial pathogens in an intensive- care unit. *Infect Control Hosp Epidemiol* 18:84-92.

[79] Smith, D.L., J. Dushoff, E.N. Perencevich, A.D. Harris, S.A. Levin. 2004. Persistent colonization and the spread of antibiotic resistance in nosocomial pathogens: resistance is a regional problem. *Proc Natl Acad Sci U S A* 101(10):3709-3714.

[80] Smith, P.W., G. Bennett, S. Bradley, P. Drinka, E. Lautenbach, J. Marx, L. Mody, L. Nicolle, K. Stevenson. 2008. SHEA/APIC guideline: infection prevention and control in the long-term care facility, July 2008. *Infect Control Hosp Epidemiol* 29(9):785-814.

[81] Temime, L., L. Opatowski, Y. Pannet, C. Brun-Buisson, P.-Y. Bolle, D. Guillemot. 2009. Peripatetic health-care workers as potential superspreaders. *Proc Natl Acad Sci U S A* 106(43):18420-18425.

[82] Temime, L., L. Kardas-Sloma, L. Opatowski, C. Brun-Buisson, P.-Y. Bolle, D. Guillemot. 2010. NosoSim: an agent-based model of nosocomial pathogens circulation in hospitals. *Proc Comput Sci* 1(1):2245-2252.

[83] Warren, D.K., R.M. Guth, C.M. Coopersmith, L.R. Merz, J.E. Zack, V.J. Fraser. 2006. Epidemiology of methicillin-resistant Staphylococcus aureus colonization in a surgical intensive care unit. *Infect Control Hosp Epidemiol* 27:1032-1040.

[84] Watanakunakorn, C., C. Axelson, B. Bota, C. Stahl. 1995. Mupirocin ointment with and without chlorhexidine baths in the eradication of Staphylococcus aureus nasal carriage in nursing home residents. *Am J Infect Control* 23(5):306-309.

[85] Watts, D.J., S.H. Strogatz. 1998. Collective dynamics of small-world networks. *Nat* 393(6684):440-442.

[86] Weber, S.G., S.S. Huang, S. Oriola, W.C. Huskins, G.A. Noskin, K. Harriman, R.N. Olmsted, M. Bonten, T. Lundstrom, M.W. Climo, M. Roghmann, C.L. Murphy, T.B. Karchmer. 2007. Legislative mandates for use of active surveillance cultures to screen for methicillin-resistant Staphylococcus aureus and vancomycin-resistant enterococci: position statement from the Joint SHEA and APIC Task Force. *Infect Control Hosp Epidemiol* 28:249-260.

[87] Wilson, E.B. 1927. Probable inference, the law of succession, and statistical inference. *J Amer Statist Assoc* 22 209-212.

[88] Wilson, P.R., G. Bellingan, M. Singer. 2005. Isolation of patients with MRSA infection. *Lancet* 365(9467):1304-1305.