MATHEMATICAL MODELING OF *CLOSTRIDIUM DIFFICILE* TRANSMISSION IN HEALTHCARE SETTINGS

Eric T. Lofgren, MSPH

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the UNC Gillings School of Global Public Health (Epidemiology).

Chapel Hill
2013

Approved by:

David J. Weber, MD, MPH

Stephen R. Cole, MPH, PhD

Nina H. Fefferman, PhD

William A. Rutala, PhD

Annelies Van Rie, MD, PhD

Abstract

ERIC T. LOFGREN: Mathematical Modeling of *Clostridium difficile* Transmission in Healthcare Settings
(Under the direction of Dr. David J. Weber)

Clostridium difficile is a frequent source of healthcare-associated infection, especially among patients on antibiotics or proton pump inhibitors (PPIs). The rate of *C. difficile* infection (CDI) has been steadily rising since 2000 and now represents a major burden on the healthcare system in terms of both morbidity and mortality. However, despite its public health importance, there are few mathematical models of *C. difficile* which might be used to evaluate our current evidence base or new control measures.

Three different data sources were analyzed to provide parameters for a mathematical model: a cohort of incident CDI cases in the Duke Infection Control Outreach Network (DICON), a hospital-level surveillance time series, also from DICON, and inpatient records from UNC Healthcare, all from 7/1/2009 to 12/31/2010. Using estimates from these data, as well as from the literature, a pair of compartmental transmission models, one deterministic and the other stochastic,

were created to evaluate the potential effect of the use of fecal transplantation as a treatment to prevent CDI.

The analysis of the cohort of incident cases suggested that ICU patients experience a greater burden of mortality while infected with *C. difficile* and have longer lengths of stay and times until death, suggesting this population as one of special interest. Two interventions were simulated using the stochastic model: the use of fecal transplantation to treat CDI and prevent recurrent cases and the use of fecal transplantation after treatment with antibiotics or PPIs to prevent the development of CDI. Simulation results showed that treating patients with CDI was effective in preventing recurrence but not in reducing the overall number of incident cases of CDI. Transplantation after treatment with antibiotics or PPIs had no effect on preventing recurrence and a statistically significant reduction in incident cases that did not reach clinical significance.

These results suggest that routine fecal transplantation for patients with CDI may be an effective treatment to prevent recurrence. Mathematical models such as the one described in this dissertation are powerful tools to evaluate potential interventions, suggest new directions for study, and understand the dynamics of infection on a population level.

Acknowledgements

I would like to gratefully acknowledge the guidance, support and encouragement of my doctoral advisor, David Weber, and the members of my committee during my time at UNC, as well as Nina Fefferman for her continued mentorship and collaboration, and for logging many miles between North Carolina, Washington, D.C. and New Jersey while on my committee.

My gratitude extends to Deverick Anderson and Rebekah Moehring at DICON for their enthusiasm and willingness to provide both data and clinical expertise, to Jane Smith for references and advice about intensive care nursing and to my friends at UNC for many hours of mutual support, complaining and insightful discussions in the student room.

This dissertation would not have been possible without funding from the Merck Investigator Initiated Studies Program, the North Carolina Translational and Clinical Sciences Institute (UL1RR025747), the PiCloud Academic Research Program, and NIH Training Grant 2T32A1070114 through the Department of Epidemiology.

Finally, I would like to thank the makers of Dr. Pepper and those members of the StackOverflow and CrossValidated communities for their invaluable assistance in many a late-night coding session.

To my parents, for teaching me to ask questions, and to Jessica for seeing this through with patience, love and understanding.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

CDI	Clostridium difficile infection
CDC	Centers for Disease Control
DICON	Duke Infection Control Outreach Network
FMT	Fecal Microbiota Transplantation
HAI	Healthcare associated infection
HCP	Healthcare personnel
HR	Hazard ratio
ICU	Intensive care unit
MCMC	Markov Chain Monte Carlo
MRSA	Methicillin-resistant *Staphylococcus aureus*
OR	Odds ratio
PCR	Polymerase chain reaction
PPI	Proton pump inhibitor
RT	Relative time
UNC	University of North Carolina
VRE	Vancomycin-resistant enterococci
95% CI	95% Confidence Interval

Chapter 1: Background and Specific Aims

Clostridium difficile is a frequent source of healthcare-associated infection (HAI), especially among patients on prolonged antibiotic treatment regimens or other conditions involving the disruption of normal gut flora. *C. difficile* is the most commonly recognized etiological agent for healthcare-associated diarrhea, and consequences of infection range from uncomplicated diarrhea to colitis and death. *C. difficile* infection (CDI) is also a problem of growing importance. The appearance and spread of a relatively rare strain identified as group BI by restriction endonuclease analysis, ribotype 027 by polymerase chain reaction, and North American pulse-field gel electrophoresis type 1 – often abbreviated as BI/NAP1/027 in Canada and soon after the U.S. and Europe, beginning in 2000 has been associated with a marked increase in CDI cases in these areas. The BI/NAP1/027 strain is characterized by a high mortality rate, which may be the result of increased virulence, increased antibiotic resistance, or both[1-3].

Rates of CDI-related hospitalizations and fatalities have been steadily rising (Figure 1-1). In a recent report[4], CDI eclipsed methicillin-resistant *Staphylococcus aureus* (MRSA) as the leading source of HAIs within the Duke Infection Control Outreach Network (DICON) group of hospitals. While a later study from the same

group of hospitals (Moehring et al, unpublished) found no evidence of a continued increase in the 2009 to 2010 period, infection rates remain elevated.

Beyond the burden of morbidity and mortality, CDI represents a significant drain on the healthcare resources of the United States and abroad. In 2009, there were an estimated 336,565 cases in the United States based on discharge data from the Nationwide Inpatient Sample (NIS) from the Healthcare Cost and Utilization Project (HCUP), Agency for Healthcare Research and Quality. Recent estimates place the cost of a single *C. difficile* infection at $2,000 per case for mild and uncomplicated cases to upwards of $90,000 in the most severe cases[5,6], with an estimated total burden on the U.S. healthcare system of over $500,000,000.

Many problems in hospital infection control are difficult to study empirically, for both practical and methodological reasons. As the purpose of a hospital is to treat, rather than study, patients, interventions to halt the spread of an infection are not done in a stepwise fashion, trying each potential intervention in turn. Instead, they are often deployed as a "bundle" of interventions, and once the spread of infection has been eliminated or lessened, we are left only with knowing that some component or components of that bundle were successful.

Compounding this practical problem are two serious violations of normal statistical assumptions, and as such conventional observational methods. First, patients cannot be considered statistically independent from one another. They are correlated by a number of factors, including the staff they are treated by, the ward and even hospital they are in. Second, the exposure status of one patient is not

independent of the disease status of another – indeed an infected patient acts as the source of exposure for a currently uninfected patient (dependent happenings).

These problems may be addressed with the use of mathematical models, which model the theoretical process by which infection is translated from one patient to another. These serve as virtual, quantitative environments within which controlled, repeatable experiments can be conducted. However, these models are not without their own assumptions. This dissertation seeks to compose a rigorous mathematical model of *C. difficile* transmission within a healthcare setting and provide a systematic evaluation of some of those assumptions within the context of hospital infection control.

Aim 1: The elucidation of the parameter estimates governing models of C. difficile transmission.

Aim 1a: Estimation of incidence, time until death and time until discharge from cohort and surveillance time-series data obtained from the Duke Infection Control Outreach Network (DICON).

Aim 1b: Estimation of non-CDI specific parameters, such as the overall time until discharge, proportion of admissions with active CDI and exposure to CDI-risk factors such as proton pump inhibitors (PPIs) and fluoroquinolones from administrative data obtained from the Carolina Data Warehouse for Health.

Rationale: In order to comprehensively compare different model types, a set of validly estimated parameters that will be used throughout the analysis must first be obtained. This aim provides as many parameter estimates as possible using modern epidemiological methods.

Aim 2: Comparison of parameter estimation methods and sources of between- and within-model sources of uncertainty.

Aim 2a. Comparison of the estimates obtained from fitting a deterministic compartmental model to an incidence time series to an approach directly linking an incidence estimate from a regression model to the corresponding parameter within the mathematical model.

Aim 2b. Comparing the results of a deterministic compartmental model with the parameters obtained in Aims 1 and 2a to an identically parameterized stochastic compartmental model.

Rationale: Traditionally, the estimation of transmission parameters is done from within a mathematical model itself, fitting a parameter or group of parameters using a method such as least-squares or maximum likelihood. However, it is possible, at least in principle, to "back in" to a model parameter directly from incidence estimates. This aim seeks to compare these two techniques, and evaluate the possibility of leveraging multiple incidence estimates from a number of sites, studies, meta-analyses etc. to provide a more robust parameter estimate.

This aim also seeks to explore issues around parameter uncertainty. In order for a deterministic model fit to a time series to be an appropriate representation of a stochastic process, the deterministic results themselves must be a good approximation of the most common outcomes of the stochastic process. By fitting an identically parameterized stochastic model, we may evaluate how well this assumption holds.

Aim 3: Application of dynamic transmission models to hospital infection control practice, assessing the impact of colonized/active CDI cases admitted from the community, hand washing compliance and other variables of interest to infection control on the modeled transmission of *C. difficile*.

Rationale: While Aim 1 and 2 seek to establish a robust model of *C. difficile* transmission with carefully considered assumptions, the purpose of Aim 3 is to apply this model to public health practice. Several major questions in infection control, such as to what extent approaching 100% hand hygiene compliance might effect transmission within the hospital, what role asymptomatic carriers of *C. difficile* play in transmission, etc. can be addressed using this model, providing possible mechanistic explanations for findings from observational studies or suggesting new directions for inquiry.

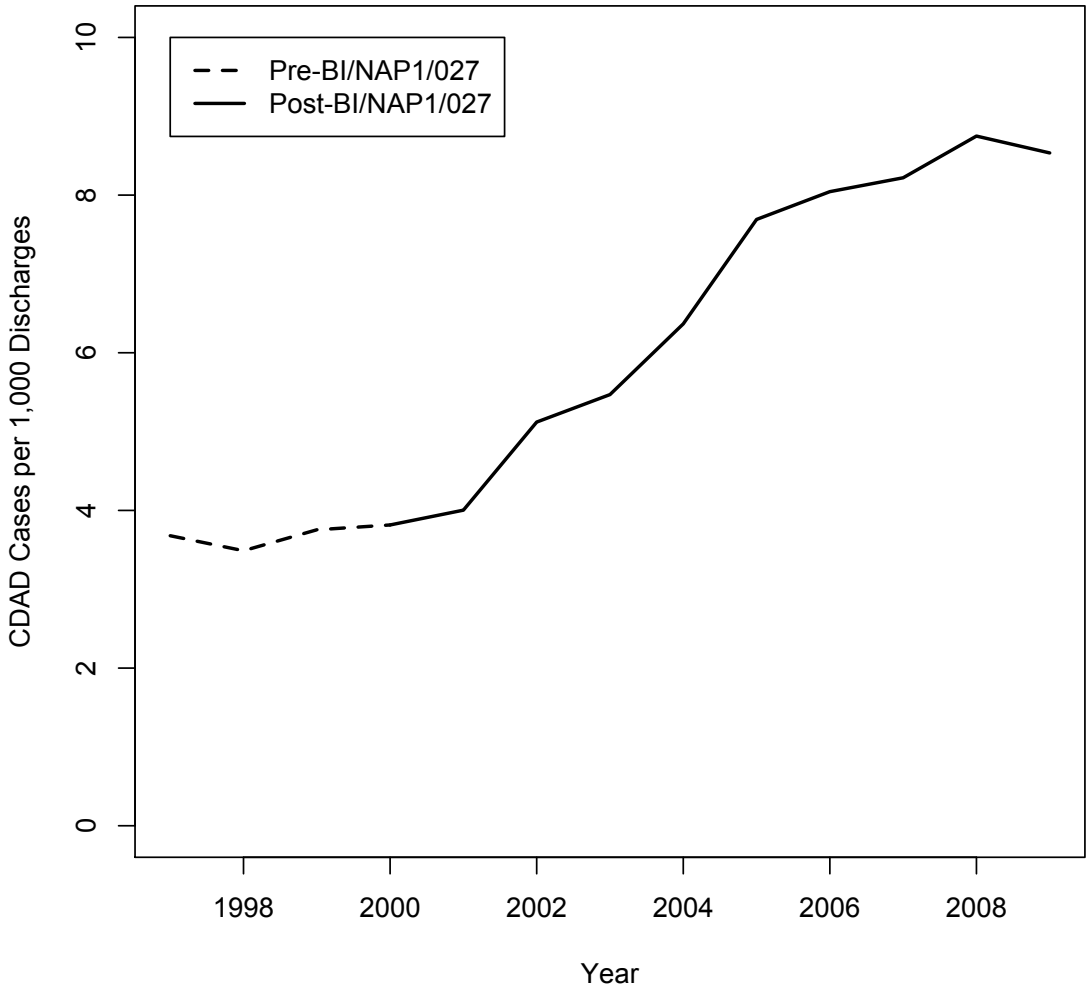

Figure 1-1. Discharge rate for C. difficile-associated disease per 1,000 hospital discharges, 1997-2009. Data from weighted national estimates from HCUP Nationwide Inpatient Sample (NIS).

Chapter 2: The Use of Mathematical Models to Study Healthcare-associated Infection

Hospitals are inherently difficult settings in which to conduct observational research, which complicates the study of healthcare-associated infections. Because patients are treated by the same set of doctors, nurses and other healthcare personnel (HCP) and share the same environment, the assumption of independence between patients that underlies many traditional statistical methods is questionable. Healthcare-associated infections are also a clear case of the problem of "dependent happenings" – when the exposure status of one individual is dependent on the disease status of those around them. Patients with HAIs shed infectious material into the environment, contaminate HCP hands, and transmit infections directly as they come into contact with other patients. Finally, as research is of secondary concern when compared to patient care, much of the observational research that is done is based on examining a "bundle" [7] or collection of interventions all tried simultaneously to arrest an epidemic within a hospital. While useful, these studies can only demonstrate that one or more components of the bundle were successful, not highlight the role of any one intervention.

Mathematical models are ideal for addressing these challenges. Designed to capture the dynamics of a system – like a hospital – as a whole, they are unburdened by assumptions of independence between patients by explicitly modeling how patients interact with one another. In doing so, they extend the empirical evidence we do have from observations based on individuals to ask research questions about the system as a whole. Mathematical models also provide a means to examine a system under the effects of multiple interventions, such as a hospital using a prevention bundle, and evaluate the effect of each intervention in turn[8], or in combination, in a repeatable, quantitative environment.

Model Structure and Composition

Mathematical models, at their core, are an attempt to quantitatively describe the way in which a system under study – such as the transmission of disease within a hospital – works. For any one research question, there are potentially infinite ways to describe a system with varying levels of complexity, different sets of assumptions, and exploring different facets of the underlying processes that drive what we observe in the real world. A few commonly used types of models are discussed below.

Deterministic Compartmental Models

Deterministic compartmental models are by far the most commonly used models in mathematical epidemiology today. Patients are grouped into a series of health states known as *compartments* with the rates that govern the transition

between these compartments known as *parameters*, which are often denoted by a Greek letter. For example, the simplest model used in mathematical epidemiology is the so-called "SI" model, where patients are grouped into one of two compartments – either susceptible (S) or infectious (I). Movement between these two compartments is governed by the parameter β ("beta"), the product of the rate of contact between individuals and the probability that that contact will result in successful transmission from a infected individual to a susceptible one. The collection of compartments and the parameters associated them are expressed as a system of ordinary differential equations, one for each compartment as they change over time. For example, for the SI model described above:

$$\begin{aligned} \frac{dS}{dt} &= -\beta SI \\ \frac{dI}{dt} &= \beta SI \end{aligned} \qquad \textbf{(Eq. 1)}$$

These types of models have been extensively used in the epidemiological literature since the work of Kermack and McKendrick[9], though they were more recently popularized by Anderson and May[10,11].

The use of these models in the study of HAIs is somewhat more recent, reflecting the emergence of pathogens such as MRSA and *C. difficile* as major threats to public health (e.g. [12-14]). Most are single-ward adaptations of the Ross-Macdonald model, a model originally developed for malaria control research[15]. These models assume that disease acquisition is the result of indirect interaction with HCP

"vectors" carrying infection from an infected patient to a susceptible patient (Figure 2-1).

In their most basic form, these models represent both the HCP and patient populations as two compartments – one for members of each group who are not yet colonized or contaminated (depending on which pathogen is being modeled) and one for those who are. Transmission occurs when a contaminated/colonized member of one group comes into contact with an uncontaminated/colonized member of the other group. The reasons for using this particular form are myriad. The analogy between HCPs and vectors is intuitive and easy to understand, the model – and extensions of it – have proven remarkable flexible in describing a variety of disease systems, and the equations that make up the Ross-Macdonald model are extensively studied and analytically tractable, allowing for results based purely on the mathematical properties of the system, rather than on any particular combination of parameter values.

These models do, however, have several major assumptions and limitations inherent to them that may threaten their validity. First, they assume uniform, random contact between compartments that interact with one another – for example, that all HCPs see all patients. While it is possible to segment the population into more and more refined compartments to address this assumption (e.g. splitting HCP compartments into compartments for nurses, residents, medical students, technicians, etc.), this is extremely cumbersome mathematically after a small number of such divisions, and the assumption remains true *within* a given

interaction between two compartments. Extending the example used previously, even if medical students only see a particular group of patients, the contact between them is uniform and random. Second, these models assume strictly patient-to-HCP-to-patient transmission, without a role for indirect transmission through the environment, which may play a substantial role in the transmission of many healthcare-associated infections[16]. Third, these models are memory-less – they do not follow the course of individual patients moving through the model, but rather model the behavior of the population within each compartment as a group. This means that while you can track outcomes such as how many infections occur, how many HCPs have transient hand contamination, etc. you cannot say *which* individuals have these outcomes. Beyond this, these types of models allow for non-integer numbers of individuals in each compartment; it is perfectly possible to have 0.20 infected patients, regardless of the biological plausibility of such a scenario. Finally, these models are purely deterministic – there is no mechanism to account for probability and randomness within the system. While in large populations these deterministic results might be considered the "average" outcome of the underlying stochastic reality of the real world, for small populations this assumption is problematic, making deterministic models frequently inappropriate for use in settings like the modeling of infection transmission within a single ward or hospital.

Stochastic Compartmental Models

Adopting the same compartmental framework, but allowing for the effects of randomness is a logical and relatively straightforward extension of the deterministic

compartmental model. There are a number of ways to accomplish this. Perhaps the most accessable is the use of an "event-driven approach". Instead of modeling the movement of the population between compartments at very small time increments, as the deterministic model does, this approach instead uses the same baseline rates as the deterministic model to probabilistically generate a list of events and when they occur, adjusting the numbers in each compartment accordingly. For example, this approach might determine that at time t = 5, a susceptible patient is infected. Thus, when the simulation reaches t = 5, the population of the susceptible compartment is reduced by one, and the infected compartment increased by one. This approach has the two-fold advantage of incorporating randomness in the system and forcing compartment populations to be integer-valued, better reflecting the nature of reality[17]. One well-known algorithm for implementing such a stochastic model is Gillespie's Direct Method[18]. This method works by calculating the time until the next event based on the cumulative rates of all possible events in the model, converting each of these rates into probabilities, and then randomly selecting one of these possibilities using a random number generator. This process is then repeated through time until the simulation comes to an end.

The use of such a stochastic approach has several key features. In small populations, random fluctuations may play an important role – for example, while a deterministic model may predict 0.20 infected patients, a stochastic model must express the infected patient population as either one or zero, and if that number is zero, infection transmission becomes impossible. This process is known as

stochastic extinction, and it plays a major role in the dynamics of infection within small populations. However, purely mathematical analysis of the system becomes somewhat more difficult. As a single simulation of the system represents only one potential probabilistic outcome out of many, it is necessary to run many simulations of the same model. This increases the amount of computation time necessary to obtain results and necessitates statistical analysis to analyze results rather than being able to rely purely on mathematical results. Additionally, while stochastic compartmental models address some of the assumptions behind their deterministic cousins, they share the assumption of random mixing between compartments, and transmission from person-to-person (or patient-to-HCP-to-patient in the case of healthcare-associated infections).

Network and Agent-based Models

Addressing the assumptions shared by both forms of compartmental model requires the adoption of considerably more computationally sophisticated models. As these models are not the focus of this dissertation, they will be discussed only briefly.

Network models explicitly model individuals and the interactions between them, lifting the assumption of random mixing. Individuals within the model have specific health states very similar to those in compartment-based models – susceptible, infected, etc. Disease transmission takes place not through the random encounters between infected and susceptible individuals, but across the links between them. These models have a number of strengths – they are especially useful

for examining research questions about how contact between individuals effects disease transmission. For example, the effect of individuals preferentially forming links based one who that person knows[19] or the impact of dividing HCP into teams that see only specific patients or other cohorting strategies[20]. Unlike compartmental models, they are not "memory-less" – they can, and indeed must, model individual patients as unique entities. This allows for the potential to include patient covariates such as age, race or gender to modify their disease risk, and modeling individuals allows for the analysis of the population in a way that is directly analogous to a real-world population, allowing them to be analyzed and interpreted as virtual cohorts [21]. This individual-level modeling also allows for the introduction of sophisticated forms of stochasticity, such as alterations in mixing patterns, and parameters arising from complex probability distributions that elude compartmental models.

These strengths do however have drawbacks. The mathematics behind the dynamics of infection across a network are complex, usually necessitating simulation-based implementations that, compared to compartmental models, are more difficult to program and more computationally intensive. Beyond the complexity of their implementation, these models also have somewhat more burdensome data requirements. In addition to parameters detailing the natural history of the disease, researchers must also specify the contact patterns between individuals. Empirically obtained network data is relatively rare and difficult to collect[22], and while artificially generated networks that follow certain empirically-derived distributions (e.g.[23]) may be used in place of direct data, this only replaces

the random mixing assumption of compartmental models with a more refined assumption about how individuals interact, rather than avoiding the assumption completely. Finally, these models generally are meant to model person-to-person transmission, though it is possible to model environmental transmission by the links between individuals representing shared contact with the same environment, rather than direct contact. However, this type of abstraction does not model the environment itself, only the shared use of it.

Perhaps the most flexible type of model is an agent-based model, where individuals are modeled within an environment, with their behaviors and interactions governed by a set of rules. These rules can be extremely complex, based on the agent's environment, current health status, the status of those around them, etc. (e.g. [24]), or quite simple. For example, rules allowing random mixing which allows agent-based models to act as an individual-based implementation of a compartmental model system. Agent-based models share many of the same strengths with network-models, arising from the representation of individuals within the model as discrete entities, while also allowing for the direct modeling of individuals within their environment, more complex behavioral patterns, etc. This flexibility comes at the cost of even higher computational complexity than network models and greater requirements for data from which to derive parameter values to describe an ever more complex system.

Meta-Population Models

Meta-population models are extensions of the any of the models described above. A meta-population model breaks up a larger population, such as an entire healthcare system (e.g.[25,26]) into individual hospitals and long term care facilities, or a single hospital into its constituent wards, modeling the movement of individuals both within the smaller sub-models and between them. These types of models are best thought of as a collection of smaller models, with the corresponding assumptions, strengths and weaknesses that accompany them.

Modeling *C. difficile* Transmission

The current literature on the modeling of *C. difficile* is relatively sparse, even when compared to other HAIs such as MRSA, which has thus far dominated the mathematical modeling literature as concerns hospital infections (e.g. [13,25-30] among others. A selection of the models that do exist is discussed below and summarized in Table 2-1.

Starr and colleagues argue in two separate papers[31,32] that a mathematical modeling approach is needed to understand the dynamics of *C. difficile*, citing the lack of independence between a given patient's level of exposure and the number of infected patients present in a hospital. The earlier of the two papers[31] outlines a model of *C. difficile* transmission that is entirely patient centric, possessing five compartments to describe a patient's health state: "Resistant, uncolonized" for patients who are uncolonized with the organism and who possess some resistance

to colonization – for example patients with healthy gut flora, "Resistant, colonized" for patients who, despite having resistant traits have become colonized with *C. difficile*, "Susceptible, uncolonized" for patients who, while not yet colonized are at elevated risk for such an event, "Susceptible, colonized" for those higher risk patients who have become colonized, and "Toxin-positive diarrhea" for those who have developed clinically evident disease. This early paper does not however actually attempt to model the disease process itself – rather, it suggests using a mathematical modeling and population-level approach more as a conceptual basis from which to consider infection control measures for *C. difficile* and recognizes that the effect of any intervention will depend on the context in which it is attempted.

The later paper[32] implements this model using a Markov Chain Monte Carlo-based implementation in an attempt to capture both the stochasticity inherent to small-population outbreaks and to draw inferences about so-called "hidden states" within the model, notably patient health states that are not normally observed, such as pre-clinical colonization status. The paper is notably lacking in mathematical detail, though the authors present an important result: environmental contamination alone is capable of driving small (<4 patients) sporadic outbreaks, but the larger outbreaks seen in many hospital settings also require some person-to-person transmission process. It should be noted that the "environment" in this model appears to be modeled as a constant, fixed "infective pressure" on patients, and does not change based on the number of infected patients in the ward. Given that environmental contamination arises from patients shedding *C. difficile* into the

environment, this assumption seems problematic. Additionally, there is no account for incoming prevalent infected or colonized cases from the community, which may play an important role in seeding wards that would otherwise be free of infection[33].

A more recent model by Starr *et al*.[34] uses a similar approach with considerable more methodological detail examines a number of potential infection control strategies. Building off a 390 patient, two-ward hospital data set, the authors find that interventions that influence patient susceptibility to *C. difficile* (i.e. antibiotic stewardship) are considerably more effective at reducing infection compared to environmental decontamination strategy or strategies that interrupt transmission. Again however, several caveats to this finding are needed. The environment is, again, modeled as a constant background colonization pressure, rather than a dynamic source of transmission that varies with the number of sick patients within the ward. Finally, despite having different patient compartments for patients not on antibiotics and on antibiotics, the MCMC algorithm estimated an essentially identical posterior transition rate from immune to susceptible to *C. difficile* (0.012 (95% Credible Interval: 0.00081, 0.01670) for patients not on antibiotics and 0.013 (95% CI: 0.0078, 0.020) for patients on antibiotics), and from there models these patients as identical. These findings are contrary to the generally accepted clinical evidence that antibiotic exposure puts one at increased risk for developing *C. difficile*[35].

Grima *et al.*[36] examine the use of non-antibiotic treatments for *C. difficile*, such as fecal transplant or currently-unsuccessful use of tolevamer as a substitute

for therapy with vancomycin or other antibiotics. However, the focus of their model is not *C. difficile*, but rather on the impact a reduction in the number of antibiotics used to treat the disease would have on the prevalence of other antibiotic resistant organisms, such as vancomycin-resistant enterococci (VRE). The authors find that a decrease in antibiotics for the treatment of *C. difficile* results in a lower rate of subsequent colonization with VRE. The model however has some serious shortcomings. The model is deterministic, which makes it difficult to distinguish if the differences between model scenarios would manifest in a noticeably different clinical experience. Beyond this, the model assumes that neither VRE nor *C. difficile* can be acquired in patients not on antibiotics, thus positing that the use of non-antibiotic treatments for *C. difficile* completely interrupt transmission in all circumstances – an extremely strong assumption. Finally, because the model is focused on VRE, clinically relevant outcomes for CDI patients, such as recurrence, are ignored.

Finally, a model of *C. difficile* transmission in a hospital by Lanzas *et al*[37] represents a model closest in form to the compartmental models detailed in later chapters. The model treats patients as being in one of five states – resistant (not at risk of *C. difficile* colonization), susceptible (susceptible to colonization having been treated with antibiotics), asymptomatically colonized with protection (colonized, but not progressing to CDI), asymptomatically colonized without protection (colonized and progressing to CDI) and diseased. The actual transmission is abstracted away, making the modeling of HCP-oriented interventions difficult to

implement. Building from parameter estimates in the literature and a hospital-level data set, and using both deterministic and stochastic implementations of the model, the authors explore which parameters are influential in determining the basic reproductive number (R_0) for the model – a measure of how many secondary infections will arise from a single primary infection and other measures of influence parameter values have on model outcomes. They found, in contrast to Starr *et al*.[34], that parameters governing transmission had a much greater impact on overall rates of colonization and disease compared to those governing susceptibility. The authors suggest this may have been because the widespread use of antibiotics in the model eliminated the progression from resistant to susceptible as a meaningful rate-limiting step in the transmission process.

Further Directions

The mathematical modeling landscape for *C. difficile* is relatively undeveloped, despite the disease's considerable burden on the healthcare system. There have been few if any attempts to actively model clinical interventions. Rather, current models have predominantly sought to describe the infection process mathematically and quantify which parameters most influence the model – a useful process to identify targets for intervention, but one step removed from actual clinical impact. Despite growing evidence for the role surface contamination plays in the transmission of *C. difficile*[16,38,39], little modeling has been done focusing on the role of the environment. Those models that do consider the role of environmental contamination abstract it to be a constant background colonization risk –

understandable given the difficult compartmental models have in directly modeling environmental contamination. This abstraction does however ignore the fact that that, just as in person-to-person transmission, environmentally mediated transmission is a dynamic process. A patient's exposure to environmental sources of transmission is a function of the number of other infected patients proximate to the patient in both space and time, rather than a constant background exposure. These types of models will be necessary in the future to both fully model the role of the environment in the transmission of *C. difficile* and to model interventions on the environment itself, such as disinfection and cleaning.

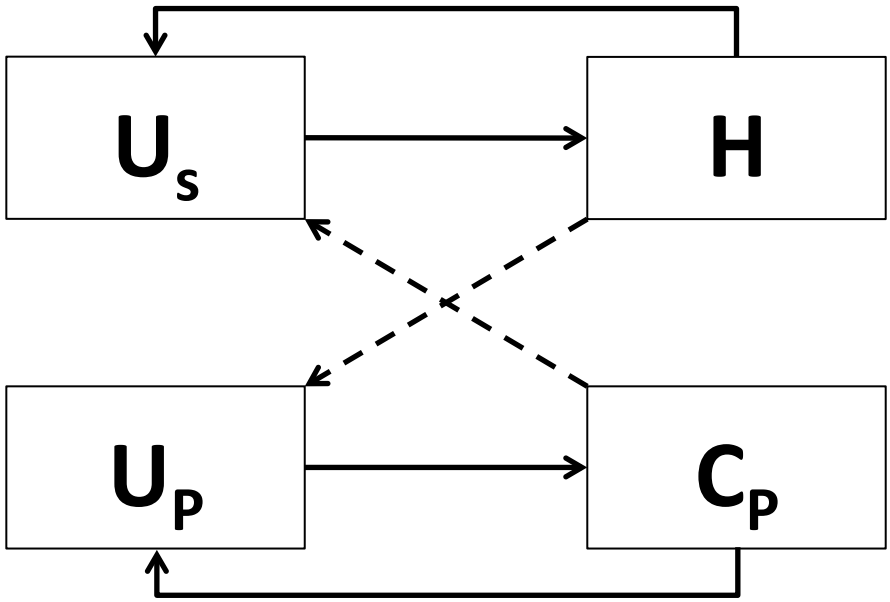

Figure 2-1. Flow diagram for Ross-Macdonald-style healthcare-associated infection model. Healthcare personnel are denoted as either uncolonized/contaminated (Us) or colonized/contaminated (H). Similarly, patients are shown as either uncolonized/contaminated (Up) or colonized/contaminated (Cp). Solid arrows indicate available paths to move between compartments, while dashed arrows indicate pathways of disease transmission.

Table 2-1. Key Mathematical Modeling Papers Focusing on Healthcare-Associated *Clostridium difficile* Transmission.

Author (Year)	Journal	Model Type	Key Findings
Starr *et al.* (1997)	Lancet	Compartmental	None (Conceptual framework).
Starr *et al*. (2001)	Clinical Microbiology and Infection	Compartmental (Stochastic -MCMC)	Different outbreak signatures are obtained for different types of transmission (environmental, person-to-person, etc.).
Starr *et al.* (2008)	Journal of Hospital Infection	Compartmental (Stochastic -MCMC)	Patient susceptibility is more important for infection than transmission rates.
Lanzas *et al.* (2011)	Infection Control and Hospital Epidemiology	Compartmental (Deterministic and Stochastic – Gillespie)	Transmission rates dominate infection process; patient susceptibility parameters play a relatively minor role.
Grima *et al.* (2012)	Computational and Mathematical Methods in Medicine	Compartmental (Deterministic)	Use of non-antibiotic treatments for *C. difficile* reduces rates of VRE.

Chapter 3: Data Sources

The data for this study comes from three sources, each covering a different aspect of *C. difficile* transmission. The data from these sources, as well as estimates from the literature, inform the parameter estimates used in the mathematical model described in later chapters.

Individual level *C. difficile* cohort

A cohort of 609 patients over 18 years of age was drawn from patients in 28 hospitals within the DICON network with incident, hospital-onset, healthcare facility-associated CDI, as defined by the CDC surveillance criteria[40] between 7/1/2009 and 12/31/2010 (Figure 3-1).

This data set included patient admission, discharge, laboratory-based diagnosis times, outcomes including death and discharge, and patient demographics and whether or not the case arose while the patient was in an intensive care unit (ICU). These data are used to quantify and describe the outcomes of individual patients once they develop CDI, such as their average length of stay, all-cause case fatality rates, etc.

Hospital level *C. difficile* surveillance

The patient-level CDI cohort is supplemented by a weekly surveillance time-series of 31 hospitals within the DICON network, consisting of the overall number of hospital-onset, healthcare facility-associated cases over the same time period of 7/1/2009 to 12/31/2010. In addition to the overall burden in terms of number of cases, this data set includes patient-day denominators for the hospital overall and within the ICU specifically, as well as whether or not the hospital was using polymerase chain reaction (PCR) or non-molecular diagnostic tests to diagnose cases. In total, this surveillance time series consisted of 1805 CDI cases over 4,038,447 patient days, 344,471 of which were within the ICU. This data set's primary utility is in describing the overall incidence of CDI within a hospital. As some parameters in the model are unknown, and the model largely based on disparate information sources, it is useful to verify that the burden of disease as seen in the model corresponds, at least roughly, to the burden of disease experienced by actual hospitals.

Non-CDI Patient data

The datasets above describe, at varying levels of aggregation, the experience of CDI patients within the DICON hospitals. However, a mathematical model of *C. difficile* transmission requires information about the disposition of presently healthy patients within the hospital – how they are prescribed drugs that put them at risk for CDI, admissions and discharge rates, etc. A data set of billing records for all inpatients within the UNC Healthcare System was obtained for patients admitted

between 7/1/2009 and 12/31/2010. This data set contained the records of 42,093 patients, 452 of which were in the ICU, and included demographic information, admission and discharge times, complaints present on admission, and flags for prescriptions that might place a patient at risk for the development of CDI (Table 3-1).

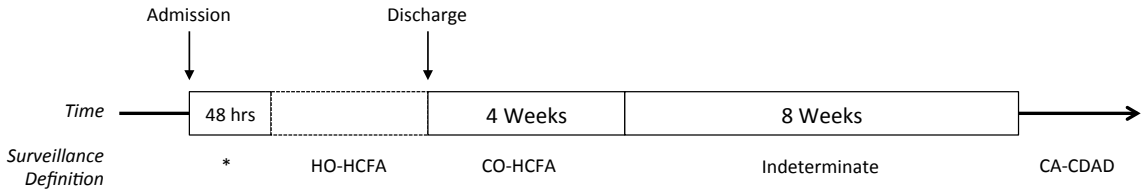

Figure 3-1. CDC/NHSN timeline-based *C. difficile* surveillance definitions. *C. difficile* infections are divided into one of four categories: Healthcare Facility Onset-Healthcare Facility Associated (HO-HCFA), Community Onset-Healthcare Facility Associated (CO-HCFA), Indeterminate or Community Associated (CA-CDAD) cases based on the time of onset of disease. Patients developing CDI within 48 hours of admission are considered CO-HCFA if they had a previous discharge from a healthcare facility within 4 weeks, CA-CDAD otherwise.

Table 3-1. Drug Prescriptions for UNC Healthcare System Inpatients Identified as High-Risk for Development of *Clostridium difficile* Infection

Brand Name	Generic Drug Name	Type
Ciloxan	Ciprofloxacin	Fluoroquinolone
Cipro	Ciprofloxacin	Fluoroquinolone
Floxin	Ofloxacin	Fluoroquinolone
Levaquin	Levofloxacin	Fluoroquinolone
Nexium	Esomeprazole Magnesium	Proton Pump Inhibitor
Noroxin	Norfloxacin	Fluoroquinolone
Ofloxacin	Ofloxacin	Fluoroquinolone
Omeprazole	Omeprazole	Proton Pump Inhibitor
Prevacid	Lansoprazole	Proton Pump Inhibitor
Prilosec	Omeprazole	Proton Pump Inhibitor
Protonix	Pantoprazole Sodium	Proton Pump Inhibitor
Vigamox	Moxifloxacin	Fluoroquinolone

Chapter 4: Estimating All-Cause Mortality and Length of Stay in Incident, Healthcare Facility-associated *Clostridium difficile* Cases Using Parametric Mixture Models

Introduction

Clostridium difficile infection (CDI) is a rapidly increasing cause of healthcare-associated infections (HAI). Discharge data from the Healthcare Cost and Utilization Project Nationwide Inpatient Sample demonstrated that approximately 336,000 cases of CDI occur annually in the US[41]. This number of cases would cost approximately $500 million per year[5,6]. In contrast to other HAI, CDI incidence has actually increased despite prevention efforts in the US, Canada and Europe[4].

The design and analysis of interventions to control *C. difficile* is complicated by the setting in which infection takes place. Hospitalized patients are often non-independent, single intervention studies are rare and difficult to conduct, and infected patients act as a source of exposure for other patients, in addition to having their own outcomes. In this kind of environment, mathematical and cost-effectiveness models are widely used, and hospital policy set on scarce data. There is a need for unbiased epidemiological estimates of patient outcomes, including length

of stay, all-cause mortality, and other estimates that quantify the experience of a patient suffering from CDI.

Quantifying these outcomes presents a three-fold problem. First, infection events cannot be considered independent, necessitating analytic techniques that account for clustering within a hospital. Second, to facilitate the use of these estimates in mathematical models, cost-effectiveness research, and other applications, rates or hazards must be directly estimated. Finally, patients may experience several mutually exclusive outcomes (such as death or discharge from the hospital). To address this final problem competing risks approaches must be employed. Conventional competing risk analysis (i.e., a cause-specific survival model) estimates the time to one outcome, while treating the other outcomes as censored[42]. These estimates address a particular question; namely, in the case of death versus discharge from a hospital, they estimate the time until death if no one were ever discharged or the time until discharge if no one ever died while in the hospital. While in some settings this approach might be acceptable or even desirable, in the case of CDI we wish to estimate the time until death given the observed levels of discharge, and the time until discharge given the observed levels of mortality.

In this study we describe an application of parametric mixture survival models to estimate two survival outcomes and address the problems enumerated above. We model the relative survival times of death from any cause and discharge between patients in the intensive care unit (ICU) and those in the general hospital

population from a multi-hospital cohort of CDI patients. In addition to the relative survival times between those patient groups, we also estimate the proportion that die while in the hospital for ICU patients and non-ICU patients and the odds ratio for those proportions.

Materials and Methods

Study Population

We used a cohort of 609 adult (>18 years of age) incident cases of CDI admitted between 7/1/2009 and 12/31/2010 obtained from infection control surveillance data from 28 hospitals within the Duke Infection Control Outreach Network, a group of hospitals that shares infection control expertise and data in the southeastern United States[43]. The maximum number of cases from a single hospital was 74, the minimum 1, and the median number of cases per hospital was 13. All cases were hospital onset, healthcare facility-associated, as defined by the CDC's surveillance guidelines[40]. Specifically, cases must have arisen more than 48 hours after admission.

Survival Times and Outcomes

The study had two competing, mutually exclusive outcomes of interest: death from any cause and hospital discharge within 180 days. The origin of time at risk was defined as the date of a positive test for *C. difficile*. The event time was given as the date of discharge from the hospital or date of death. The single patient with an event time greater than 180 days was censored at 180 days, and the 12 patients

with an unknown event time were considered interval censored from 12 hours after diagnosis to 180 days after diagnosis. Patients with identical diagnosis and discharge dates were assumed to spend 12 hours in the hospital.

Exposure Definition and Covariate Selection

Patient ICU status was determined at the time of diagnosis of CDI. Patients were either in the hospital's ICU at the time of their infection and thus exposed or were not, and are unexposed, regardless of whether or not their treatment subsequently ended or involved the ICU.

Inverse probability weights were used to control for confounding by patient characteristics measured at hospital admission[44]. Using such weights, rather than regression adjustment, allows estimated curves to represent the marginal survival, rather than survival conditional on covariates[45]. Variables considered for inclusion in the model were patient age, whether or not they were on dialysis, if the patient had been hospitalized within 12 weeks prior to admission, if that prior admission had been in the same institution as the current admission, if the patient had been previously diagnosed with *C. difficile*, the patient's gender and race, source of admission (where the patient was prior to admission), which medical specialty was primarily responsible for the patient (i.e. medicine, surgery, obstetrics/gynecology, etc.), if the patient had been discharged from any hospital within the past year, and whether the CDI case was a new episode, a recurrent episode, or a continuation per CDC definitions.

Potential confounders were included in the weighting model if they were marginally associated ($p<0.20$) with either death or discharge using a Weibull or log-normal parametric survival model, respectively. A quadratic term for the sole continuous variable (age at admission) and bivariate interactions that marginally improved model fit as evaluated by a likelihood ratio rest ($p<0.20$) were included. Multiple imputation was used to handle missing covariate values, which resulted in 119 cases with at least one missing variable. Thirty imputations were based on a multivariate normal model including all variables used in the substantive analysis, including outcomes. Imputations were combined using Rubin's canonical variance estimator[46].

Parametric Mixture Model

We modeled time to death and time to discharge as a mixed survival function, $S_D(t)$ and $S_N(t)$ respectively, where $S_D(t) + S_N(t) = 1$ at $t = \infty$, indicating all patients having experienced one of the two outcomes of interest. These two functions, as well as the proportion of patients who died (π) and who were discharged ($1-\pi$) give a probability that an event time T taking place at $T < t$ of $P(T < t) = \pi[1-S_D(t)] + (1-\pi)[1-S_N(t)]$. Details on the theory and implementation of this type of model have been previously published[47]. Briefly, these functions can be estimated using maximum likelihood methods, with the likelihood of a given individual *i* expressed as follows (Equation 2):

$$L_i = [\pi_i f_D(t_i)]^{\delta_i(1-\zeta_i)} \times [\pi_i (S_D(t_{i1}) - S_D(t_{i2}))]^{\delta_i \zeta_i} \times [(1-\pi_i) f_N(t_i)]^{\theta_i(1-\eta_i)} \times$$
$$[(1-\pi_i)(S_N(t_{i1}) - S_n(t_{i2}))]^{\theta_i \eta_i} \times [\pi_i S_D(t_i) + (1-\pi_i) S_N(t_i)]^{1-\delta_i-\theta_i}$$

Where $f_D(t)$ and $f_N(t)$ are the probability density functions for death and discharge, δ and θ are indicators for death = 1 and discharge = 1, and ζ and η are indicators for interval censored times for death and discharge. For interval censored observations, t_{i1} and t_{i2} indicate the two times bracketing the censored interval, where $t_{i1} < t_{i2}$ and in this study specifically, $t_{i1} = 0.5$ and $t_{i2} = 180$ days. Weighting is incorporated by multiplying the natural log of L_i by individual i's weight.

The survival functions used in the mixture model may be any parametric functions. Previous studies have used exponential[48], log-normal[49], and generalized gamma survival functions[50], differing functions for each outcome[47], and non-parametric extensions of the Kaplan-Meier method[51], among others. In this study, a Weibull function for death and a log-normal function for discharge were used. This choice mirrored the best fitting parametric models used in the single outcome models discussed below, and in the confounder selection process.

Robust standard errors with clustering by hospital were calculated to account for non-independence between patients in the same hospital. From this model, five main estimates are obtained: The ratios of the mean survival times for death (RT_D) and discharge (RT_N) between the ICU cases and non-ICU cases, the proportions who died in hospital for the ICU cases and non-ICU cases (π_1 and π_0,

respectively) and the odds ratio of the mixing proportions (OR_{π}), which provides a relative measure of mortality between the ICU groups.

For comparison purposes, the cohort was also analyzed using conventional competing risks analysis, with each outcome modeled independently and patients experiencing the other event being considered censored at their event time. A parametric Weibull survival model was used to model time until death, and a log-normal survival model used to model discharge. As with the mixture model, robust standard errors were used to account for non-independence arising from clustering by hospital. All analysis was done using SAS 9.2 (SAS Institute, Cary, NC).

Results

The characteristics of the ICU and non-ICU patient populations are summarized in Table 4-1. There were 160 (26.2%) ICU patients and 449 (73.7%) non-ICU patients. In the ICU population, 42 patients (26.3%) died, while in the non-ICU population 43 patients (9.6%) died. The remaining patients were discharged from the hospital. Figure 4-1 provides a graphical depiction of the distribution of exposures, outcomes and survival times in a 20% random sample of the cohort.

Several factors were at least moderately associated with one of the two outcomes, including the patient's age, gender and race, along with the source of admission, whether or not they were a surgical patient, if the patient was on dialysis, and whether or not this was a new case of CDI (in contrast to a continuing or recurrent case) (Table 4-2). Interactions between patient's race and gender, age,

whether or not this was a new CDI case, and dialysis status, between patient's gender and both surgical and dialysis status and between admission source and patient age and dialysis status were also found to result in moderately superior model fit for the outcome specific models.

Parametric Survival Models

Using a conventional competing risks approach, the relative time to death for the ICU versus non-ICU populations, RT_D was 0.65 (95% CI: 0.36, 1.17), suggesting that ICU patients died marginally more swiftly than their counterparts in the general hospital population. Concurrently, the relative time until discharge, RT_N was 2.30 (1.66, 3.18), reflecting longer lengths of stay within the exposed population (Figure 4-2).

The mixing proportion in the ICU population (π_1) was 0.28 while the mixing proportion in the non-ICU population (π_0) was 0.10. The odds ratio of the mixing proportions (OR_π) was 3.38 (95% CI: 1.84, 6.19), capturing the substantially higher burden of mortality in ICU patients compared to those in the general hospital population. Comparing the mean event times between ICU and non-ICU patients, RT_D was 1.97 (95% CI: 0.96, 4.01) and RT_N was 1.88 (95% CI: 1.40, 2.51) (Figure 4-3).

The robust standard errors typically resulted in a slight inflation of a parameter's uncertainty and performed similarly to standard errors obtained using a nonparametric bootstrap method (not shown). Compared to the multiply imputed

data used in the primary analysis, estimates using complete cases were less precise and resulted in considerably different effect estimates (Table 4-3).

These estimates, in contrast to those from the conventional models described above, indicate that despite the higher severity of illness that might reasonably be assumed in patients admitted to the ICU, they experience longer times to both death and discharge than patients in the general hospital population. The differences in estimates between the two models are summarized in Table 4-4.

Discussion

The purpose of this study was to examine the outcomes experienced by patients with CDI as a mixture of two simultaneously occurring survival processes, rather than as two disjoint events. We believe that this approach captures the actual disposition of patients within the hospital in a more realistic fashion. The use of a weighted, parametric mixture model allows for the estimation and prediction of survival times, produces marginal effect estimates and covariate adjusted survival curves and is free from the proportional hazards assumption. This assumption is however exchanged for the necessity of correctly specifying the underlying distribution of event times as well as proportional survival times.

This method also illustrates the potential for incorrect estimation in conventional survival analysis when both outcomes are of interest for informing prevention efforts. The conventional competing risks model found a reduced time until death for ICU patients. In essence, this method conflates the proportion of

patients who died with the time with which it took them to die. In our study, the mixture approach separates these two processes into two separately estimated parameters. In doing so, our study suggests that while ICU patients may experience a greater burden of mortality, their survival times appear longer. The conventionally estimated RT_D of 0.65 (95% CI: 0.36, 1.17) is not only on the other side of the null from the RT_D estimated in the mixture model, 1.97 (0.96, 4.01), but does not include the mixture model's point estimate within its 95% confidence interval. The origin of this difference is in the survival functions estimated by the two methods. The conventional method forces the survival functions for both outcomes to equal zero at $t = \infty$, whereas the mixture approach only forces the functions to equal the mixing proportion of their respective outcomes at $t = \infty$, a less stringent requirement. By treating patients who did not experience the outcome as censored, the conventional approach also pushes the estimated survival functions out toward the tail of their distribution. This difference in the shape of the survival curves manifests itself as drastically different survival probabilities. For example, the probability of death at 90 days is 0.275 for ICU and 0.014 for non-ICU patients in the mixture model, and 0.765 for ICU and 0.622 for non-ICU patients using a conventional approach. These differences in the estimated survival functions are the source of the disparate estimates of RT_D and RT_N. The potential downstream effects of such a difference in the estimated survival times on administrative decisions, mathematical or cost-effectiveness models, etc. are significant.

Focusing on the results of the parametric mixture model, this study demonstrated that ICU patients infected with *C. difficile* experience both longer times until death and longer overall lengths of stay post-infection, as well as experiencing a burden of mortality three times that of their non-ICU peers. The study estimates suggest that ICUs demand additional resources and attention from an invention prevention perspective, despite *C. difficile* being a hospital-wide problem, unlike HAIs associated with a particular device or procedure. The ICU has a proportionately large volume of adverse outcomes, and ICU patients longer length of stay may contribute more to the contamination of the hospital environment and have implications for impacting in-hospital transmission of CDI. Patients have been shown to shed *C. difficile* into the environment continuously after infection, even after their symptoms have subsided[52]. Because of their longer time within the hospital, ICU patients have increased opportunities to shed *C. difficile* spores into the environment. Whether this higher individual-level potential for shedding is outweighed by the considerably larger number of spore-shedding patients within the general hospital population who are hospitalized for shorter time periods warrants further examination.

This study has several limitations. Though the surveillance data used has information on whether or not a given patient died within the hospital, it cannot necessarily be assumed that these deaths were attributable to *C. difficile*, either solely or as part of a constellation of ailments. Rather than an estimation of the impact of CDI on patient mortality, this study is instead an estimation of the

patient's potential impact on the hospital's environment during their infection, until it is interrupted favorably by a discharge from the hospital or unfavorably by the patient's death. While less patient-centric than many studies, these types of estimates are crucial for the study and prevention of hospital acquired infections, where patients are not only important in the prevention of their own adverse outcomes, but represent sources of infection risk to future patients.

Additionally, as patients who have been discharged from the hospital may experience a recurrent infection and die, be treated outside the DICON network, or be readmitted into a DICON hospital with a different exposure status than their original infection, some outcome misclassification may have occurred. These limitations are however inherent to the difficult task of conducting observational studies within a hospital setting and occur regardless of what analytical methods are used.

Our approach, which allows for the separate estimation of the timing of an event and the frequency with which it occurs, provides a more nuanced view of the outcomes experienced by CDI patients. As interest in healthcare associated infection prevention increases, so too does the need for more sophisticated analytic techniques to reflect the complexity surrounding patients, providers, and the environment of a healthcare facility.

Table 4-1. Baseline Characteristics of 609 Incident *Clostridium difficile* Infection Cases Within the DICON Hospital Network, Southeastern USA, 2009-2010.

Variable	ICU			Non-ICU		
	N	%	Mean (SD)	N	%	Mean (SD)
Age (Years)			66.99 (14.33)			70.04 (15.20)
Dialysis						
Yes	21	13.1		37	8.2	
No	135	84.4		407	90.7	
Missing	4	2.5		5	1.1	
Gender						
Female	66	41.3		218	48.6	
Male	83	51.9		187	41.7	
Missing	11	6.9		44	9.8	
Admission Source						
Home	99	61.9		290	64.6	
Nursing Home	27	16.9		99	22.1	
Hospital	20	12.5		10	2.2	
Other	13	8.1		45	10.0	
Missing	1	0.6		5	1.1	
New CDI Episode						
Yes	156	97.5		440	98.0	
Nod	4	2.5		9	2.0	
Missing	0	0		0	0	
Race						
White	83	51.9		241	53.7	
Black	58	36.3		108	24.1	
Other	3	1.9		11	2.5	
Missing	16	10.0		89	19.8	
Surgical Patient						
Yes	8	5.0		36	8.0	
No	139	86.9		361	80.4	
Missing	13	8.1		52	11.6	

Abbreviations: SD, standard deviation; ICU, intensive care unit. DICON, Duke Infection Control Outreach Network; CDI, *C. difficile* infection..

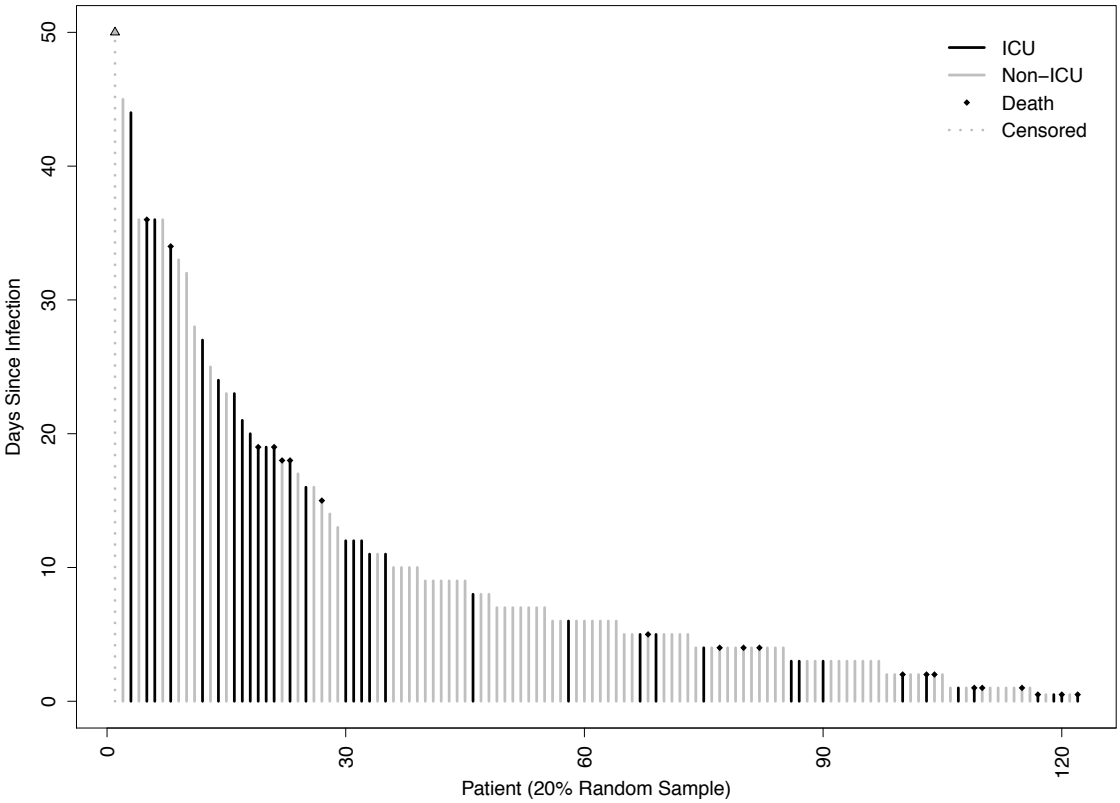

Figure 4-1. Diagram of survival times and outcomes for a 20% random sample of a cohort of 609 incident *Clostridium difficile* Infection cases within the DICON hospital network, Southeastern USA, 2009-2010. Dark lines indicate ICU cases, while lighter grey lines indicate non-ICU cases. Dotted lines cases with unknown outcome times, treated as censored from t = 0.5 to t = 180 days. Lines terminating in diamonds indicate patients that died.

Table 4-2. Association Between Patient-level Covariates and Time Until Death or Discharge in a Cohort of 609 Incident *Clostridium difficile* Infection Cases Within the DICON Hospital Network, Southeastern USA, 2009-2010.

	Death			Discharge		
Variable	RT_D	95% CI	*P value**	RT_N	95% CI	*P value**
Age (Years)	0.98	0.96, 0.99	0.005	1.00	0.99, 1.01	0.89
Dialysis			0.98			0.16
Yes	1.01	0.54, 1.89		0.78	0.56, 1.10	
No	*ref*			*ref*		
Gender						
Female	0.98	0.62, 1.55		0.79	0.64, 0.96	
Male	*ref*			*ref*		
Admission Source			0.93			0.026
Home	0.91	0.41, 2.04		0.99	0.70, 1.39	
Nursing Home	0.91	0.37, 2.45		0.97	0.66, 1.42	
Hospital	1.08	0.36. 3.29		2.00	1.16, 3.45	
Other	*ref*			*ref*		
New CDI Episode			0.66			0.072
Yes	1.30	0.41, 4.14		0.77	0.53, 1.12	
Nod	*ref*			*ref*		
Race			0.11			0.89
White	0.65	0.16, 2.64		0.85	0.44, 1.64	
Black	1.04	0.24, 4.46		0.87	0.44, 1.70	
Other	*ref*			*ref*		
Surgical Patient			0.47			0.17
Yes	1.52	0.49, 4.67		0.77	0.53, 1.12	
No	*ref*			*ref*		

Abbreviations: RT_D, relative difference in mean time until death; RT_N, relative difference in mean time until discharge; CI, confidence interval; DICON, Duke Infection Control Outreach Network; CDI, *C. difficile* infection.

*Chi-square test of the overall effect of the variable on time to death or discharge

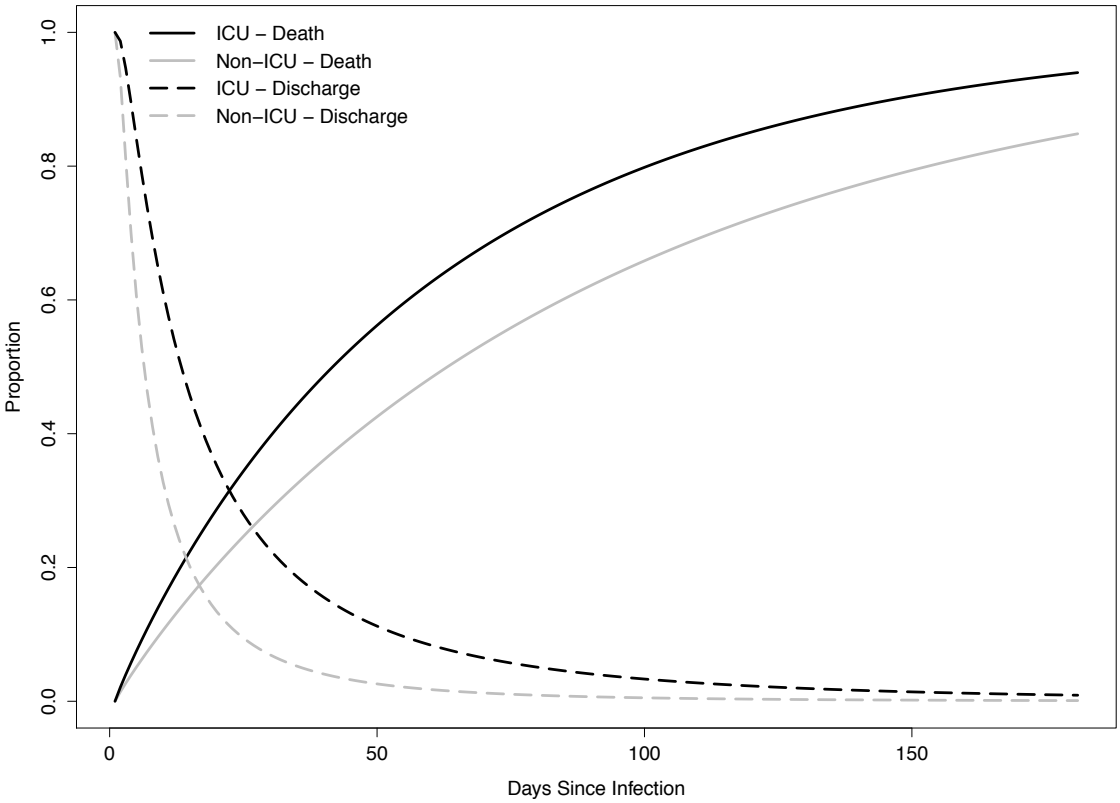

Figure 4-2. Cause-specific parametric survival curves for time until death (A) and time until discharge (B) by ICU-exposure status in a cohort of 609 incident *Clostridium difficile* Infection cases within the DICON hospital network, Southeastern USA, 2009-2010.

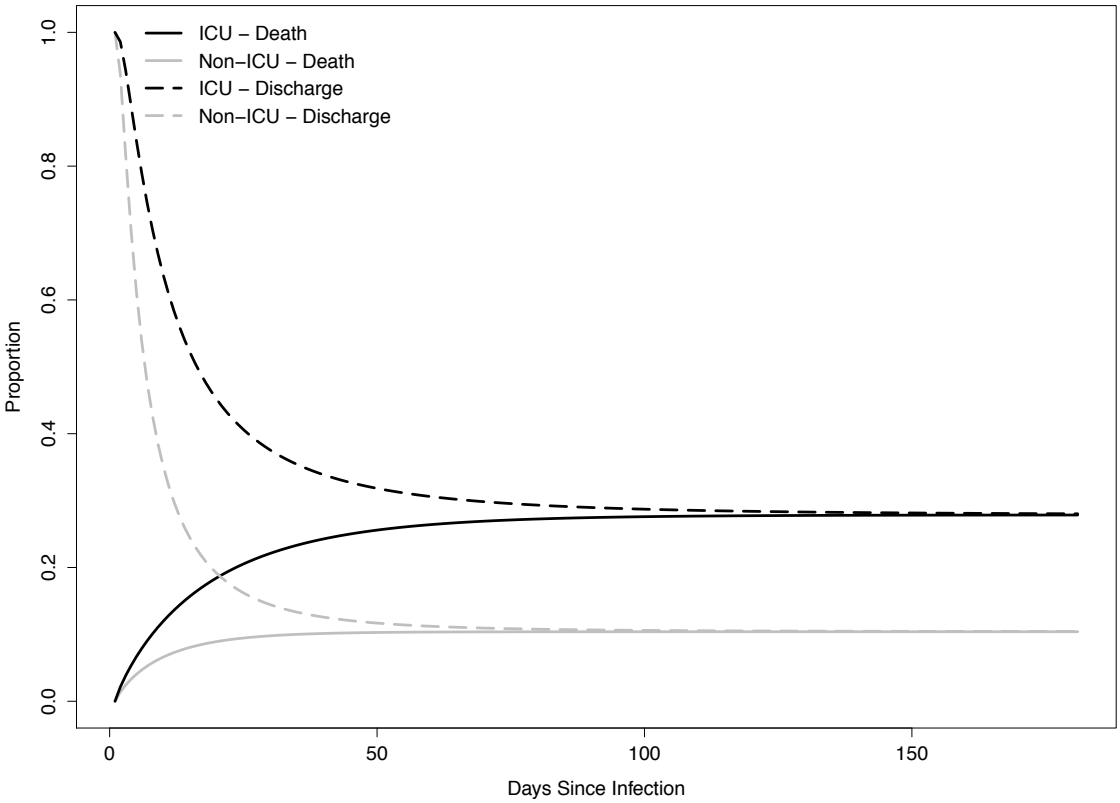

Figure 4-3. Times to death and discharge estimated using parametric mixture models in a cohort of 609 incident *Clostridium difficile* Infection cases within the DICON hospital network, Southeastern USA, 2009-2010. Black lines denote ICU patients, while grey lines denote non-ICU patients. Solid lines are 1 minus the survival function for death, and dashed lines are the survival function for discharge.

Table 4-3. Difference in Estimates from Multiple-Imputation versus Complete Cases Analysis for Time Until Death, Discharge, and Mixing Odds Ratio Comparing ICU patients to non-ICU Patients From a Cohort of 609 Incident *Clostridum difficile* Infection Cases Within the DICON Hospital Network, Southeastern USA, 2009-2010.

Model	RT_D	95% CI	RT_N	95% CI	π_1	π_0	OR_π	95% CI
Complete Case*	1.27	0.59, 2.74	2.03	1.36, 3.05	0.25	0.10	2.94	1.18, 7.33
Multiple Imputation*	1.97	0.96, 4.01	1.88	1.40, 2.51	0.28	0.10	3.38	1.84, 6.19

Abbreviations: RT_D, relative difference in mean time until death; RT_N, relative difference in mean time until discharge; R_π, odds ratio of mixing proportions in the ICU and non-ICU patient population; CI, confidence interval; DICON, Duke Infection Control Outreach Network; CDI, *C. difficile* infection.
*Adjusted for patient's age, gender and race, location prior to admission, whether or not patient was a surgical patient or on dialysis, and if this was a new CDI episode.

Table 4-4. Estimates Obtained for Time Until Death, Discharge, and Mixing Odds Ratio Comparing ICU patients to non-ICU Patients From a Cohort of 609 Incident *Clostridum difficile* Infection Cases Within the DICON Hospital Network, Southeastern USA, 2009-2010.

Model	RT_D	95% CI	RT_N	95% CI	π_1	π_0	OR_π	95% CI
Cause-Specific (Crude)	0.72	0.39, 1.35	2.45	1.86, 3.25	-	-	-	-
Cause-Specific (Adjusted)*	0.65	0.36, 1.17	2.30	1.66, 3.18	-	-	-	-
Mixture Model (Crude)	2.24	1.25, 4.02	2.01	1.50, 2.69	0.26	0.10	3.36	1.85, 6.11
Mixture Model (Adjusted)*	1.97	0.96, 4.01	1.88	1.40, 2.51	0.28	0.10	3.38	1.84, 6.19

Abbreviations: RT_D, relative difference in mean time until death; RT_N, relative difference in mean time until discharge; R_π, odds ratio of mixing proportions in the ICU and non-ICU patient population; CI, confidence interval; DICON, Duke Infection Control Outreach Network; CDI, *C. difficile* infection.
*Adjusted for patient's age, gender and race, location prior to admission, whether or not patient was a surgical patient or on dialysis, and if this was a new CDI episode.

Chapter 5: A Mathematical Model to Evaluate the Routine Use of Fecal Transplantation to Prevent Incident and Recurrent *Clostridium difficile* Infection

Introduction

Clostridium difficile is a frequent source of healthcare-associated infection (HAI), especially among patients receiving treatment regimens involving antibiotics[35] or proton pump inhibitors (PPIs)[53,54], or with other conditions involving the disruption of normal gut flora. The rate of *C. difficile* infection (CDI) in the United States has been steadily rising since 2000, causing an estimated 336,565 cases in the United States in 2009[41]. In some healthcare facilities, CDI has eclipsed methicillin-resistant *Staphylococcus aureus* as the leading source of HAIs[4]. Of special concern is the development of recurrent CDI, which may be a complicated, long-term condition typified by repeated bouts of severe diarrhea that involves treatment with antibiotics such as oral vancomycin, fidaxomicin or metronidazole[35].

Because altering the indigenous flora of the intestinal tract causes CDI, there has been an interest in recolonizing the intestinal tract with introduced donor bacteria obtained from either healthy donor stool[55,56] or synthesized as a pure culture[57]. This procedure, referred to as fecal microbiota transplantation (FMT), restores the bacterial ecology that typically keeps *C. difficile* in check. Both

uncontrolled case reports[56,57] and a small clinical trial[55] have shown encouraging results; however FMT is still largely reserved for specialized intervention in difficult or refractory cases. Further, the implications of routine intestinal recolonization as a standard course of treatment for the prevention of recurrent or incident CDI has not been widely explored.

Mathematical models are ideal for studying such hypothetical scenarios. They can provide a repeatable, quantitative environment with which to evaluate evidence, guide policy creation, discover critical thresholds upon which the success of interventions may depend, and suggest new directions for observational studies and clinical trials. These strengths are difficult or impossible to duplicate with empirical research within a hospital. Critically, one patient's outcome influences another's exposure, which violates traditional statistical assumptions of independence. Finally, mathematical models are capable of scaling up the independent, individual level observations that emerge from clinical research up to the population level. In this way, we may study how these individuals interact with one another and influence the transmission process without a risk to patient safety.

In order to evaluate the impact of routine intestinal flora recolonization in patients with CDI, we developed a mathematical model describing the transmission of *C. difficile* within an intensive care unit (ICU), with the capability to test the impact of FMT for prevention of recurrent *C. difficile* or initial infection due to in-hospital transmission.

Methods

Data Sources

Hospital data was obtained from three separate sources, all consisting of patient records between 7/1/2009 and 12/31/2010. The first dataset was a cohort of 609 adult patients with incident CDI was extracted from prospectively collected HAI surveillance data from 28 community hospitals in the Duke Infection Control Outreach Network (DICON). This data set included admission, discharge and diagnosis times, and outcomes including death and discharge, and patient demographics. The second data set included weekly surveillance time-series from 31 DICON-affiliated hospitals within the DICON network, consisting of the overall number of hospital-onset, healthcare facility associated CDI cases as defined by CDC surveillance criteria[40], patient-day denominator data for the hospital as a whole hospital patient-day denominator data, ICU patient-days, and whether the hospital was using a non-molecular diagnostic test or a diagnostic test based on polymerase chain reaction (PCR). In total, these series consist of 1805 cases and 344,471 ICU patient-days. Finally, a third data set included hospital billing records for 452 inpatients discharged from the ICU within the UNC Healthcare System, consisting of discharge times, orders for drugs that place patients are risk for CDI such as PPIs or fluoroquinolones, coded diagnoses present on admission and demographics.

Transmission Model

The transmission of *C. difficile* through an intensive care unit was modeled as a series of compartments representing patient health and treatment states (Figure 5-1). Healthcare personnel (HCP) were modeled as being either Uncontaminated (U_S) or Contaminated (H), representing hands or gloves contaminated by vegetative *C. difficile* or spores. Patients could be in one of six compartments. Compartment U_P represented uncolonized patients who were not on high-risk medications for CDI, and U_A represented uncolonized patients who were on high-risk medications. Similarly, C_P and C_A represented low- and high-risk patients who were previously exposed to the organism. Compartment D represented patients who have developed CDI. Finally, some of the scenarios we considered had an additional compartment, C_T, which represented patients under prophylactic treatment using FMT to prevent an initial infection.

The interactions and transitions between these compartments are governed by a series of eight differential equations (Eq 3):

$$\frac{dU_S}{dt} = \iota H - \rho\sigma_P C_P \frac{U_S}{N} - \rho\sigma_D D \frac{U_S}{N} - \rho\sigma_A C_A \frac{U_S}{N} - \rho\sigma_P C_T \frac{U_S}{N}$$
$$\frac{dH}{dt} = \rho\sigma_P C_P \frac{U_S}{N} + \rho\sigma_D D \frac{U_S}{N} + \rho\sigma_A C_A \frac{U_S}{N} + \rho\sigma_P C_T \frac{U_S}{N} - \iota H$$

$$\frac{dU_P}{dt} = -\rho\psi U_P \frac{H}{N} - \theta_P U_P + \nu_{U_P}(\theta_P(U_P + C_P) + \theta_A(U_A + C_A + C_T) + \zeta D + \gamma D)$$
$$\frac{dU_A}{dt} = -\rho\psi U_A \frac{H}{N} - \theta_A U_A + \nu_{U_A}(\theta_P(U_P + C_P) + \theta_A(U_A + C_A + C_T) + \zeta D + \gamma D)$$
$$\frac{dC_P}{dt} = \rho\psi U_P \frac{H}{N} - \kappa C_P - \theta_P C_P + \nu_{C_P}(\theta_P(U_P + C_P) + \theta_A(U_A + C_A + C_T) + \zeta D + \gamma D)$$
$$\frac{dC_A}{dt} = \rho\psi U_A \frac{H}{N} - \phi C_A - \kappa\tau C_A - \theta_A C_A + \nu_{C_A}(\theta_P(U_P + C_P) + \theta_A(U_A + C_A + C_T) + \zeta D + \gamma D)$$
$$\frac{dC_T}{dt} = \phi C_A - \kappa C_T - \theta_A C_T$$
$$\frac{dD}{dt} = \kappa(C_P + C_T) + \kappa\tau C_A - \zeta D + \nu_D(\theta_P(U_P + C_P) + \theta_A(U_A + C_A + C_T) + \zeta D + \gamma D)$$
$$- \gamma\chi(1-\omega)\eta D - \gamma\chi\omega(1-\eta)D - \gamma(1-\chi)(1-\omega)D - \gamma(1-\chi)\omega D$$
$$M = U_A + U_p + C_A + C_p + C_T + D$$
$$N = U_s + H + U_A + U_p + C_A + C_p + C_T + D$$

The definitions and values of the parameters in this model are detailed in a later section.

Patients were admitted into U_P, U_A, C_P, C_A or D. Colonized patients (C_P and C_A) and patients with CDI (D) shed infectious material that may contaminate hands of HCP, and uncontaminated patients (U_P and U_A) are subsequently colonized when cared for by HCP with contaminated hands. HCP could decontaminate their hands by washing them after contact with either the patient or the environment immediately surrounding them.

As there is evidence of a surface contamination component to *C. difficile* transmission[16], contacts between patients and HCP were modeled as direct care tasks, which could involve contact with the patient's environment as well as physical interaction between the patient and HCP. Once colonized, patients could progress to CDI (C_X to D). All patients were eventually discharged from the hospital. Three possible outcomes were tracked for patients with CDI – death, a healthy discharge, and a discharge that resulted in the development of recurrent CDI.

The model made several simplifying assumptions. First, all HCP were assumed to interact with all patients within the ICU, and patients were assumed not to interact with each other. It is not known whether disruption of intestinal flora places patients at greater risk for developing CDI once colonized or at greater risk of colonization and thus subsequent infection[58,59]. Therefore, colonization once exposed to *C. difficile* and development of infection after colonization is treated as a single process within the model. Additionally, we assumed that patients who were placed on antibiotics or PPIs were prescribed those medications immediately on arrival into the ICU. Additionally, we assumed that the medication-induced disruption of the normal gut flora was immediate and lasted beyond the discontinuation of treatment. This effectively meant that once a patient became high risk, they remained so unless an active intervention (such as FMT) was made to recolonize their intestinal tract.

Parameterization and Model Calibration

The transmission model was parameterized using a combination of estimates from the literature and the data sets discussed above. Specific values, and the sources they are derived from, are described in Table 5-1 below. A more detailed discussion of parameterization may be found in Appendix A.

The underlying hazard of developing CDI for low-risk patients was estimated by fitting a deterministic version of the mathematical model above to the DICON surveillance time series. Because the case counts were not ICU-specific, the proportion of all reported cases in the time series data that arise within the ICU was assumed to be equal to the proportion of cases within the DICON cohort data that arose in the ICU. Hospitals that did not report individual level data within the cohort were assigned the proportion of the hospital with the closest number of total patient-days.

Based on prior research indicating a 56% increase in the number of reported cases within these hospitals that switched from non-molecular to PCR diagnostic tests[60], weeks where non-molecular tests were in use had their case numbers inflated by 1.56. These adjusted time series were then used to calculate an expected weekly rate for a twelve-bed ICU (the size of ICU used in this model) and a corresponding estimate of cumulative incidence over the one and a half years the surveillance data was collected. The cumulative incidence curves for each hospital were then averaged and weighted by the total number of ICU patient-days per hospital. This weighted average cumulative incidence curve was used as an estimate

of a "typical" level of infection for the modeled ICU. The cumulative incidence produced by the mathematical model was then fit to this weighted curve using least-squares regression to obtain an estimate of the hazard of developing CDI in colonized, low-risk patients.

Simulations

The mathematical model described above was applied to a single twelve-bed ICU consisting of single patient rooms with four registered nurses and a single intensivist, based on average size and staffing information and best-practice guidelines for ICUs[72-74]. Admissions were fixed to be equal to discharges to maintain a steady patient population. Several different potential treatment regimens were considered (Table 5-2). First, we created a baseline scenario, modeling no routine use of fecal transplantation. Second, we modeled a series of scenarios depicting the systematic use of FMT after CDI in order to prevent recurrent cases, treating 20, 40, 60, 80 and 100 percent of cases. Third, we modeled a series of scenarios examining the use of FMT prophylactically to prevent incident infections, treating contaminated high-risk (C_A) patients immediately after the conclusion of their treatment regimen, moving them to a new, low-risk category (C_T). These scenarios considered treating 20, 40, 60, 80 and 100 percent of all high-risk patients or just those patients on fluoroquinolone antibiotics. Finally, both the second and third treatment strategies used in combination, treating with FMT patients both post-CDI and post-high risk medication.

Deterministic models do not fully capture the transmission dynamics of small populations, such as the experience of a single, 12-bed ICU. Therefore, each of the treatment scenarios described above were modeled using 1,000 stochastic simulations of the equation system by means of Gillespie's Direct Method[18]. This approach converts the deterministic rates of the differential equation model into probabilities, and then repeatedly simulates the movement of individuals within the system over time based on these probabilities. The effect of this is two-fold. First, individuals within the models are treated as discrete units – no fractions of patients exist in compartments. Second, because individuals are treated as discrete units and the model becomes probabilistic, variations due to random chance may arise. While in large population models the differences between these two approaches are small, in small populations this variability plays an important role in understanding the disease dynamics in the real world.

As a consequence of simulating this system as a random Monte Carlo process, it is possible for the patient population to "die out" when enough discharged patients are not replaced to reduce the patient population to zero, or for the patient population size to grow larger than 12 patients. These outcomes, while unrealistic, are important for understanding the level of possible variability inherent in the model system. The simulations were run over a one-year timespan.

Two primary outcomes were tracked in all scenarios: the number of incident infections and the number of infections that develop into recurrent cases. Note that in many simulations we expected the number of recurrent cases to be higher than

the number of incident infections. The model handled both incident infections that arise in the ICU and prevalent infections on admission, both of which could develop into recurrence. The results of stochastic models are frequently non-normally distributed, differences between treatment groups were analyzed with non-parametric Kruskal-Wallis tests. Simulations were written in Python, and all statistical analysis was performed in R.

Results

The median, 25th and 75th percentile number of recurrent and incident cases for all modeled scenarios are reported in Table 5-2. The baseline, no-intervention model produced results similar to the known epidemiology of CDI. Infection rates were low, but periodic outbreaks of *C. difficile* occurred, as did periods of no infection. Despite high levels of patient colonization and sustained transmission of *C. difficile* within the ICU, hand contamination of HCP was rare and short lived. An example showing the development of a typical simulation over time is shown in Figure 5-2.

The impact of FMT in different proportions of post-CDI patients to prevent recurrence is shown in Figure 5-3. The treatment resulted in a statistically significant ($p<0.001$) difference in the number of recurrences among the different treatment groups. The median number of recurring cases ranged from 2 (Interquartile Range (IQR)=6) for no treatment to a median of 0 (IRQ=1) when 100% of patients were treated. Treatment did not result in a significant difference in the number of incident infections, regardless of what proportion was treated

(p=0.35). The median number of incident cases was 0 for all scenarios (IQR=1), save the 80% treatment level which had an IQR of 2.

The results of treating patients prophylactically after discontinuation of antibiotic therapy or PPIs had similar results. Figure 5-4 shows the results of the latter scenario for the different proportions of treatment. Neither approach resulted in a statistically significant difference in recurrence, regardless of the proportion treated (p=0.47 and 0.97 respectively); all scenarios had a median number of recurrent cases of 2.

The difference in incident infections was statistically significant when the treatment group was limited to patients receiving only antibiotics over all levels of treatment (p=0.004) but not antibiotics, PPIs or both (p=0.09). In both treatment scenarios however, this difference did not result in tangibly different model outcomes from a clinical perspective. In scenarios treating only patients on antibiotics, all treatment levels had a median of 0 incident cases (IQR=1). Simulations with 0% treatment did have a higher maximum number of incident cases (N=16) than models treating 20% to 100% of cases (N =10, 13, 11, 10, 14 respectively). Similar patterns were seen for simulations treating patients on both antibiotics and PPIs (not shown).

Combining both prophylactic treatment and post-infection treatment protocols resulted in a statistically significant difference in recurrent cases over the proportion of patients treated (p<0.001). The median number of recurring cases ranged from 2 (IQR=6) for no treatment to a median of 0 (IQR=1) when all patients

were treated. This strategy also resulted in a statistically significant difference in incident infections (p=0.007), though as with the purely prophylactic scenarios, this difference did not manifest in a change in median incidence, as all treatment levels had a median number of cases of 0 (IQR=1). However, the no treatment scenario had a higher maximum number of incident cases (N=16) compared to treatment levels of 20%-100% (N=11, 7, 9, 11, 9 respectively). These results are shown in Figure 5-5.

Discussion

Our unique study using mathematical modeling found that the widespread use of FMT resulted in a marked reduction in discharges that would go on to develop into recurrent cases. Importantly, this reduction was seen in all modeled scenarios ranging from relatively low levels of treatment (20% of patients) to very high levels of treatment (100% of patients) with no apparent threshold effect. This widespread evidence of a positive effect suggests that these results should be robust not only to varying levels of treatment, but also to lower levels of efficacy, as the two are mathematically equivalent.

Unsurprising, as post-CDI treatment to prevent recurrent cases is an entirely post-infection process, the modeled intervention had very little impact on incident, hospital-acquired infections. Some secondary effects may be seen if fecal transplantation becomes a regularly used treatment, by way of a reduced number of recurrent cases resulting in fewer admissions with prevalent recurring *C. difficile*. Capturing this phenomenon would require modeling not only a single ward but an entire local healthcare system, which is beyond the scope of this study.

The evidence for a positive effect of using FMT after high-risk medications to prevent incident infection is less apparent. While treating patients on antibiotics or both antibiotics and proton-pump inhibitors resulted in statistically significant or nearly significant results at an alpha equals 0.05 level, these results appear to have little tangible clinical impact on the number of incident infections and no impact on the occurrence of recurrent infections. There is also very little evidence for a synergistic effect between the two treatment strategies. Scenarios that explored the use of post-CDI and post-high risk medication FMT simultaneously were very similar to that of post-CDI FMT alone. Taken as a whole, these results indicate that routine use of fecal transplantation represents a promising tool to prevent complicated, recurring episodes of *C. difficile*, but techniques to recolonize the intestinal tract alone will be insufficient to contain the spread of *C. difficile* within a healthcare system.

This study is not without limitations. Many of the states within the model, such as whether or not a patient has come into contact with *C. difficile*, are not regularly observed within hospitals, and thus some of the outcomes of the model cannot be directly verified. As with all mathematical models, the results of the study are dependent on the assumptions about the natural history of *C. difficile* infection and the values of the parameters used. The purpose of this study, however, is not to provide precise predictions of future levels of infection, but rather to examine the overall impact of fecal transplantation as a routine treatment option when dealing

with *C. difficile*. Within this scope, the model structure and parameter estimates represent the state-of-the-art knowledge of in-hospital *C. difficile* transmission.

This study represents, to the best of our knowledge, the first use of a mathematical model to quantify the potential effects of fecal transplantation for the treatment and prevention of CDI. These types of models represent a useful method for evaluating the potential impact of new treatment approaches in areas of limited clinical and empirical evidence. Our results suggest that routine intestinal recolonization is a powerful tool for the prevention of recurrent infection. When combined with other infection control measures such as improved surface disinfection and antibiotic stewardship, fecal transplant has great potential to produce a substantial reduction in the burden of *C. difficile*, especially in reducing highly morbid and difficult to manage recurrent infections.

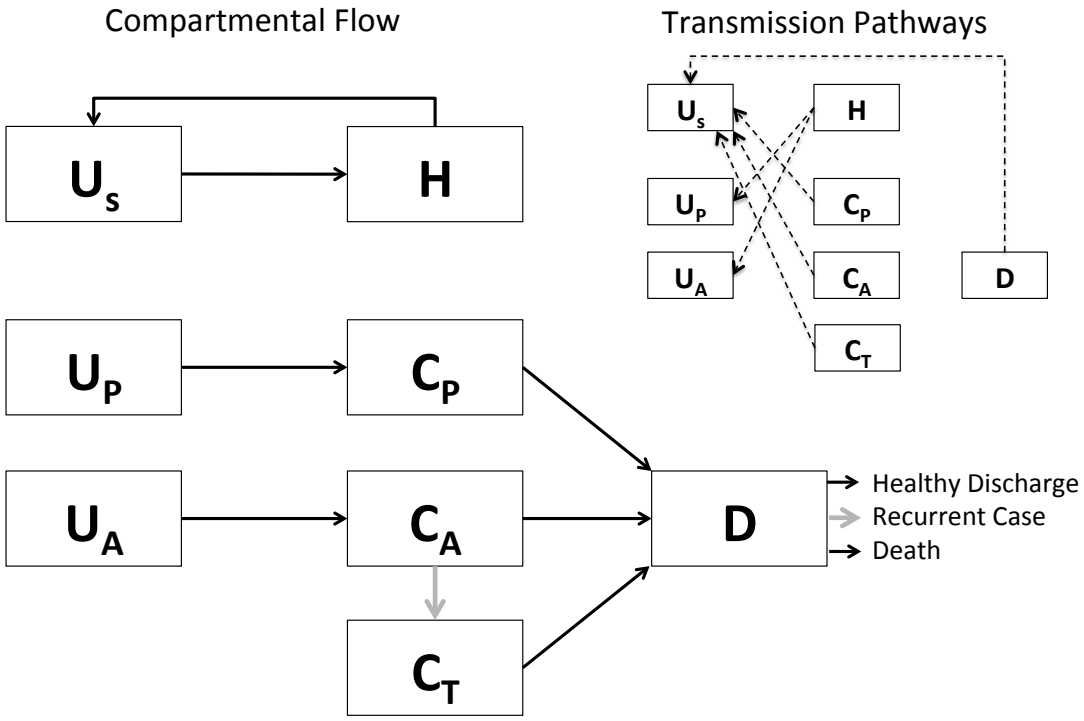

Figure 5-1. Schematic representation of the compartmental flow of a mathematical model of the use of fecal transplantation to prevent incident and recurrent *C. difficile*. Inset indicates the potential routes of bacterial contamination between patients and healthcare workers, while grey arrows indicate the movements within the model influenced by the simulated intervention. HCPs are classed as uncontaminated (U_S) or contaminated (H), and patients into low-risk, uncolonized (U_p), low-risk, colonized (C_P), high-risk, uncolonized (U_A), patients with CDI (D) and patients undergoing FMT (C_T).

Table 5-1. Parameters for a Mathematical Model of the Use of Fecal Transplantation to Prevent *Clostridium difficile* Infection and Recurrence

Symbol	Description	Value (Units)	Source
ι	Handwashing Rate	9.365 hand washes or glove changes per hour	[61-63]
ρ	Contact rate between patients and HCP	4.244 direct care tasks per patient per hour	[64]
σ_i	Probability a healthcare provider's hands are contaminated by contact with a patient of type *i*	Low Risk: 0.35, High Risk: 0.35 Active Infections: 0.50	[33,52,59,65,66]
Ψ	Probability of transmission from a contaminated HCP hands to a uncontaminated patient's skin	0.90*	[33]
θ_i	Discharge rate for an uninfected patient of type *i*	High Risk: 1/ 12.006 days Low Risk: 1/ 3.318 days	UNC Healthcare Billing Data
ζ	Hourly probability of death for a patient with an active *C. difficile* infection	0.000625	DICON Cohort Data
Υ	Hourly probability of discharge for a patient with an active *C. difficile* infection	0.00188	DICON Cohort Data
ν_i	Proportion of admitted patients who are of patient type *i*	C_P: 0.00447, C_A: 0.0155, U_P: 0.209, U_A: 0.727, D: 0.044	UNC Healthcare Billing Data, [35,37]
κ	Hazard of developing CDI in low-risk, contaminated patients	0.000268	DICON Surveillance Data
τ	Relative Risk of developing CDI due to high-risk treatment	3.37	[53,59,67]
Φ	Probability of receiving post-treatment fecal transplant to prevent incident infection or recurrence	Antibiotics-only: 0.0011 Antibiotics & PPIs: 0.00169	UNC Healthcare Billing Data, [68-71]
χ	Percent of eligible patients receiving fecal transplant	0 – 100 (varies by scenario)	
ω	Probability of a discharged patient developing recurrence	0.30	[54]
η	Probability of fecal transplant in moving patient to low-risk category	0.938	[55]

Abbreviations: CDI, *C.* difficile infection, HCP, Healthcare Personnel; PPI, Proton Pump Inhibitor; DICON, Duke Infection Control Outreach Network, UNC, University of North Carolina

*Assumed to be highly efficient based on general agreement between skin sampling and hand culture methods, indicating a minimal loss of contamination between contamination from a patient skin site to deposition on another surface.

Table 5-2. Patient Outcomes from a Mathematical Model of the Use of Fecal Transplantation to Prevent *Clostridium difficile* Infection and Recurrence in a Simulated 12-bed ICU Over One Year.

Scenario	Percent Treated	Median Recurrence (25th & 75th Percentiles)	p-value*	Median Incidence (25th & 75th Percentiles)	p-value*
Baseline	0%	2 (0,6)		0 (0,1)	
Post-Infection	20%	1 (0,4.25)	< 0.001	0 (0,1)	0.35
	40%	1 (0,4)		0 (0,1)	
	60%	1(0,3)		0 (0,1)	
	80%	0 (0,2)		0 (0,2)	
	100%	0 (0,1)		0 (0,1)	
Prophylactic (Antibiotics)	20%	2 (0,5)	0.47	0 (0,1)	0.004
	40%	2 (0,6)		0 (0,1)	
	60%	2 (0,5)		0 (0,1)	
	80%	2 (0,6)		0 (0,1)	
	100%	1 (0,5)		0 (0,1)	
Prophylactic (Antibiotics and PPIs)	20%	2 (0,6)	0.97	0 (0,1)	0.09
	40%	2 (0,6)		0 (0,1)	
	60%	2 (0,6)		0 (0,1)	
	80%	2 (0,7)		0 (0,1)	
	100%	2 (0,6)		0 (0,1)	
Combined	20%	1 (0,6)	< 0.001	0 (0,1)	0.007
	40%	1 (0,5)		0 (0,1)	
	60%	1 (0,3)		0 (0,1)	
	80%	0 (0,2)		0 (0,1)	
	100%	0 (0,1)		0 (0,1)	

Abbreviations: PPI, Proton Pump Inhibitor
*Kruskal-Wallis one-way analysis of variance test

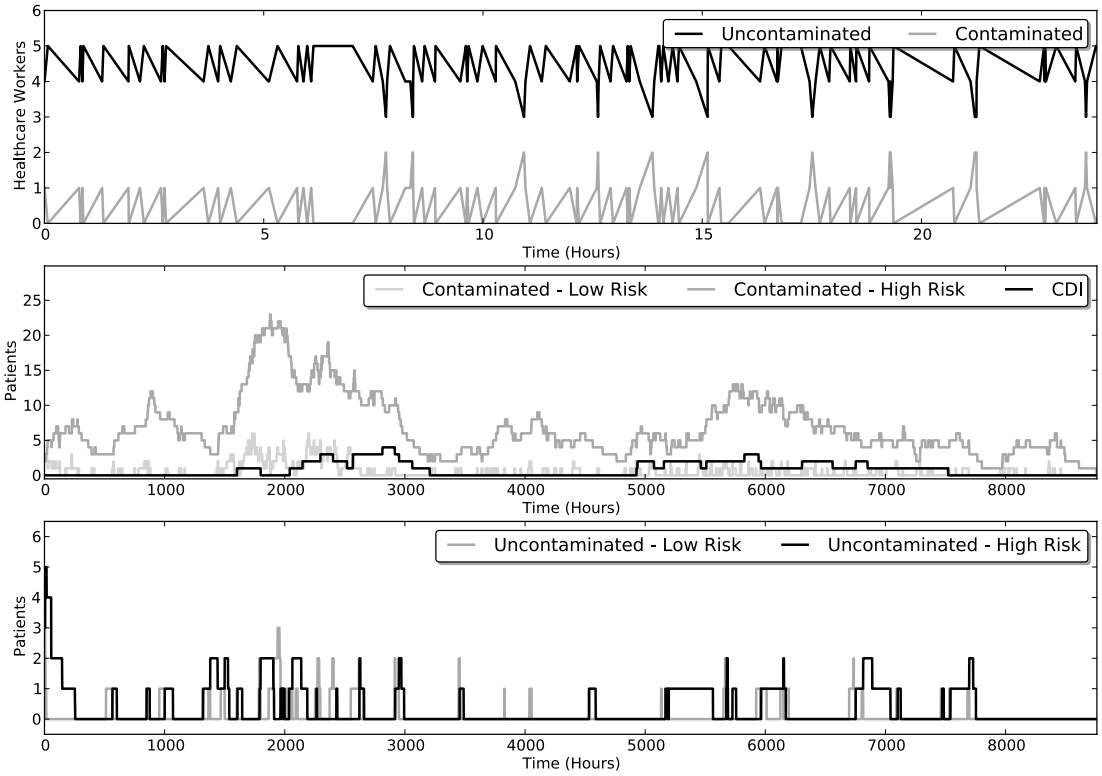

Figure 5-2. A single stochastic realization of a mathematical model of the use of fecal transplantation to prevent incident and recurrent *C. difficile*. The top panel shows the level of hand contamination in healthcare workers over a 24 hour period, while the bottom two panels depict the number of patients and their current health state over a one-year period.

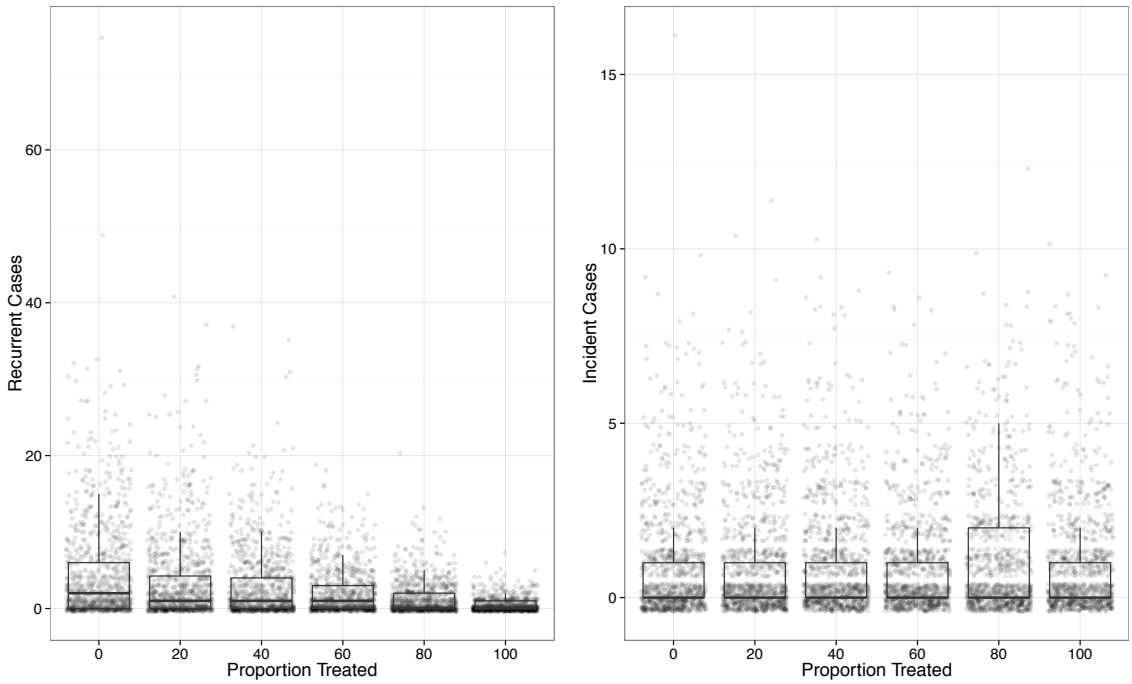

Figure 5-3. Simulated recurrent and incident cases of *C. difficile* under six levels of post-CDI FMT to prevent the development of recurrence. All simulation outcomes are show, with the results summarized with box-and-whisker plots depicting the median, 25th and 75th percentiles, and 1.5 times the interquartile range.

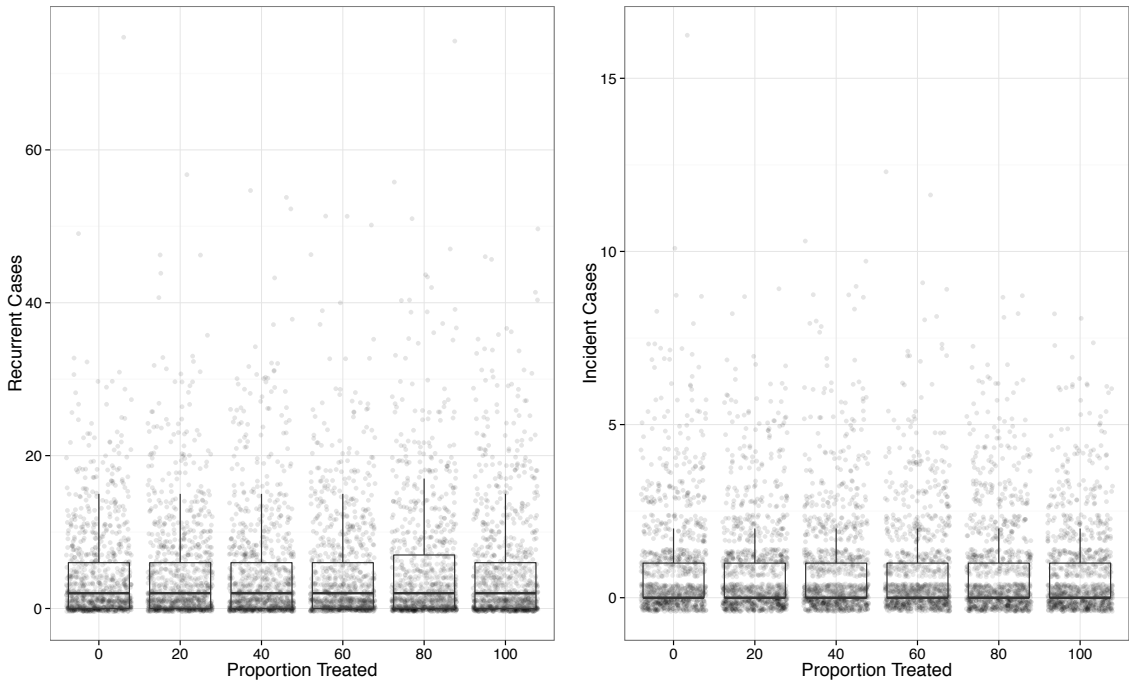

Figure 5-4. Simulated recurrent and incident cases of *C. difficile* under six levels of post-high risk medication FMT to prevent the development of infection and recurrence among patients on antibiotics or proton pump inhibitors. All simulation outcomes are show, with the results summarized with box-and-whisker plots depicting the median, 25th and 75th percentiles, and 1.5 times the interquartile range.

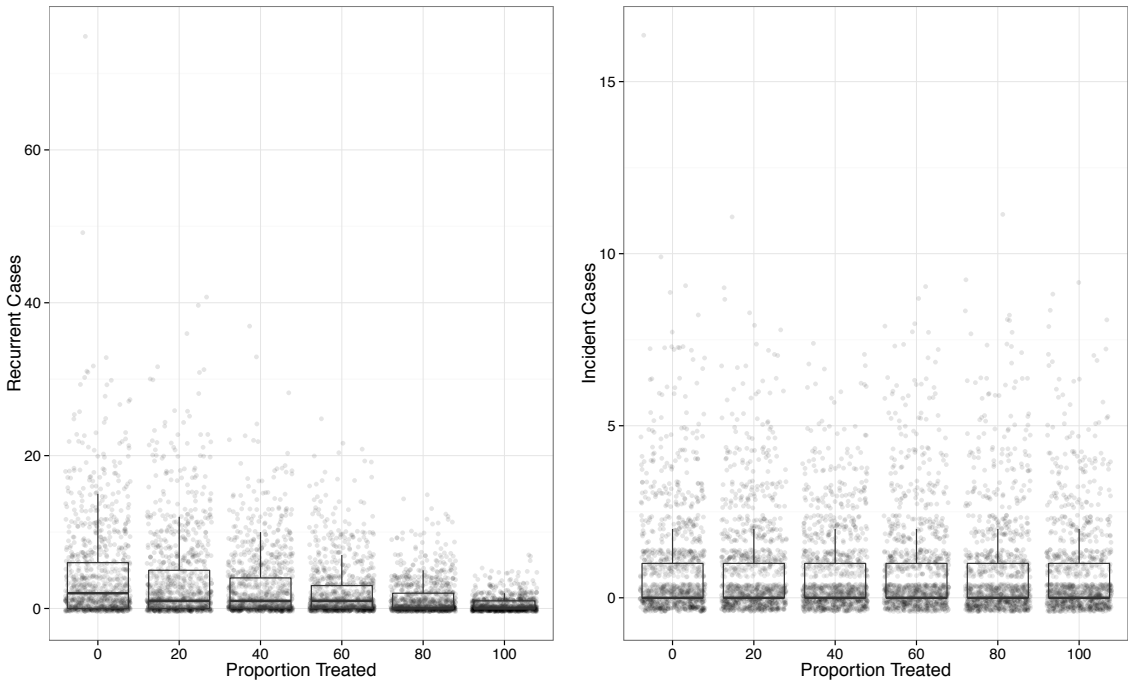

Figure 5-5. Simulated recurrent and incident cases of *C. difficile* under six levels of combined post-high risk medication FMT to prevent the development of infection and recurrence among patients on antibiotics or proton pump inhibitors and post-CDI FMT to prevent the development of recurrence. All simulation outcomes are show, with the results summarized with box-and-whisker plots depicting the median, 25th and 75th percentiles, and 1.5 times the interquartile range.

Chapter 6: Conclusions and Future Directions

This dissertation has explored the use of mathematical models to study the transmission of *C. difficile* in healthcare settings by combining mathematical and computational methods with modern observational epidemiology. The two should be thought of as mutually reinforcing disciplines, interacting with one another in a continual cycle of research, rather than viewing one as a lesser substitute when the other is untenable for a specific question. Mathematical models help organize and formalize the present state of knowledge about a disease system, discovering areas where there is currently an absence of empirical data. Empirical studies, in turn, provide unbiased parameters for mathematical models that can then be used to extend the results of those same studies into new and unexplored areas.

To accomplish this, parameters for a mathematical model were estimated using parametric mixture models from a cohort of CDI cases from the DICON hospital network (Aim 1a) and parametric accelerated failure time models from inpatient billing data from UNC Healthcare (Aim 1b), as well as estimates from the literature and fitting a deterministic compartmental model to DICON surveillance data (Aim 2a). From these parameters, identically formulated deterministic and stochastic compartmental models (Aim 2b) were used to examine the potential

impact of FMT as a routine treatment for the prevention of either incident or recurrent CDI (Aim 3).

Key Findings

Parametric Mixture Models for CDI Outcomes

The parametric mixture models used in Chapter 4 represent a method for handling mutually exclusive competing risks that disentangles the time until an event occurs and the frequency with which it occurs. Using a cohort of 609 incident, healthcare-associated CDI cases, we estimate that the all-cause case mortality rate for ICU patients was 28% for cases in the ICU, compared to 10% for cases outside the ICU (OR = 3.38 (95% CI: 1.84, 6.19)). Compared to patients not in the ICU, CDI patients in the ICU experienced longer times until death (RT_D = 1.97 (95% CI: 0.96, 4.01)) and longer times until discharge (RT_N = 1.88 (95% CI: 1.40, 2.51). This is in contrast to the conventional cause-specific approach to modeling competing risks, which overestimated the difference in time until discharge (RT_N = 2.30 (95% CI: 1.66, 3.18)) and not only underestimated the difference in time until death, but estimated it on the other side of the null, suggesting ICU patients had a shorter – though not statistically significant – time until death (RT_D = 0.65 (95% CI: 0.36, 1.17)).

These results have importance both from an infection control and epidemiological methods standpoint. Patients within the ICU experience longer times until both death and discharge and experience a higher burden of mortality.

As a result, despite their relatively small numbers, they represent a population of concern for the prevention of CDI, despite CDI being often thought of as a hospital-wide problem. Beyond the higher frequency of negative outcomes, patients who remain in the hospital longer have a greater opportunity to shed infectious material into the environment and to contaminate the hands of HCP that care for them. From a methodological perspective, these findings suggest that in some circumstances cause-specific models, which conflate the time until an event occurs and frequency with which it occurs may lead to errant estimates.

Mathematical Models of Routine Fecal Transplantation

Using the parameters obtained from the DICON cohort and surveillance data, as well as UNC Healthcare billing data and estimates from the literature, a compartmental model of *C. difficile* transmission within an ICU was developed. This model explored the impact of using fecal transplantation to restore the intestinal flora of patients who had undergone flora-disrupting treatments involving antibiotics or proton pump inhibitors. Two different potential interventions were simulated: the regular use of FMT to treat CDI cases to prevent recurrence, and the use of FMT as a post-treatment prophylactic to prevent the development of CDI for patients at the end of course of antibiotics or PPIs. Over a number of modeled scenarios, the model consistently found that post-CDI fecal transplantation has a positive, statistically significant impact on the number of recurrences but no impact on incident cases. The evidence for the use of FMT as a prophylactic measure to prevent the development of CDI is somewhat less definitive. Prophylactic treatment

did not result in a decrease in recurrence, and while it did result in a borderline significant decrease in incident cases, this difference did not manifest in a decrease that would meaningfully impact clinical care.

Strengths and Limitations

This dissertation represents a fusion of observational epidemiology and mathematical modeling, and is the first study we are aware of which uses these two techniques to examine the use of routine fecal transplantation for the prevention of incident or recurrent CDI. It provides a "best of both worlds" approach, using modern observational methods to provide robust, unbiased parameter estimates for the mathematical model and using the mathematical model in turn to extend the results of the limited observational evidence available for the clinical impact of fecal transplantation and to suggest new directions for empirical research.

As with all studies, this dissertation has limitations. While extensive effort went in to insuring the parameter estimates obtained from observation data were unbiased, there is the potential for residual confounding and unknown misclassification is ever present. Additionally, because of the nature of *C. difficile* as a healthcare-associated infection for which hospitals are not reimbursed by insurance or CMS, there is no single source of data that follows the whole cohort of hospital inpatients from admission to discharge. As such, a working picture of an intensive care unit must be assembled from disparate sources and then be assumed to be capable of being meaningfully melded together into a single coherent model. The similarity of the two sources of data, both major healthcare systems in the

Southeastern USA over the same time period is intended to minimize the effect of this merging of data, but it is possible that the parameters arising from one data source reflect a different local "reality" of *C. difficile* transmission compared to those arising from the other. Finally, the mathematical model used in this dissertation is subject to the assumptions outlined in Chapter 2 and is only valid insofar as those assumptions are not violated. These limitations are however, as much a call for further research as they are caveats about the present findings. This dissertation reflects the most current knowledge of *C. difficile*, derived from an extensive review of both the mathematical modeling and clinical literatures, as well as the analysis of several datasets using modern epidemiological methods. The modeling results within it are not meant to be exact predictions of case counts, but rather a first attempt at quantifying the impact of adopting a novel treatment as the standard of care for *C. difficile*, and within this more limited intention, are as robust as possible.

Future Research

Several avenues of continued research present themselves based on the results of this dissertation. Many of the limitations of the current mathematical model lie in the assumptions arising from its compartmental model formulation. As much as this is a challenge to the present model, it is an opportunity for future research, as the findings of this model can evaluated using more sophisticated models that reflect a more realistic process of mixing between patients and HCPs as well as environmental contamination.

The purpose of such an evaluation is two-fold. First, it allows for a progressively more robust examination of a clinically relevant problem, providing better and more realistic predictions of the impact of fecal transplantation as a routine therapeutic option. Second, the sequential examination of the same research question, using largely the same parameters but with progressively more sophisticated model forms represents a relatively poorly explored area of research: sensitivity analysis not of a particular parameter or parameters, but of model structure as a whole.

Mathematical models are descriptions of the dynamics of a system, a representation of how we believe an infection process to work. Even with perfect parameter values, the choices a researcher makes in how to represent this system have a major impact on the outcomes of the model. Despite this, much of the current focus in the sensitivity analysis of mathematical models is focused on the variation of particular parameters while keeping the more fundamental structure of the model constant. Extending sensitivity analysis to questions of model structure would benefit the field not only by evaluating the robustness of certain results to changes in model type, but also in an increased understanding of what types of questions require more sophisticated model types such as network or agent-based models and what types are adequately addressed with more approachable implementations.

Finally, mathematical models are flexible tools, and even for a system as well studied as *C. difficile* there are a wealth of questions that lend themselves well to

modeling and can be examined with small modifications to the same basic model structure. Some examples of these questions might be the potential impact of new surface decontamination procedures, the role of staffing levels and ICU design (closed vs. open ICUs), and novel therapies for the treatment or prevention of *C. difficile*, among others. As many of these questions involve the dynamic interaction of patients, HCP and the environment that elude statistical techniques, mathematical models will play a critical role in eventually understanding their respective impact on the spread of a pathogen of serious and growing public health concern.

Appendix: Model Parameterization

What follows is a more detailed look at the parameter values used in the mathematical model described in Chapter 5. Such a description would distract from the overall message of that chapter, but may be of interest to the reader.

Discharge Rates

Rates of death and discharge for CDI cases were obtained using estimates from Chapter 4. The length of stay for non-CDI patients was determined using parametric survival models using a technique similar to that discussed in Chapter 4, but for a single outcome using the subsample of ICU patients from the UNC Healthcare inpatient billing data set. By default, the rates in a Gillespie Direct Method-based stochastic model result in exponentially distributed waiting times, with a single rate parameter $\boldsymbol{\lambda}$ equal to the reciprocal of the average time until death, discharge, etc. An alternate formulation however, using n compartments for each transition between health states allows the waiting times to take the form of a gamma distribution with the shape parameter $\kappa = n$ and a scale parameter θ = the average time until death, discharge, etc. [75]. Because many of best fitting parametric survival models were not exponential, but rather log-normally distributed, it may be necessary to use the more flexible gamma distribution to provide simulated waiting times comparable to those obtained from data. However, as a gamma distribution is not a survival distribution generally supported in available software, a gamma distribution cannot be fit directly.

Instead, a gamma distribution approximating a log normal distribution with parameters μ and σ was obtained by minimizing the Kullback-Leibler divergence, an information theoretic measure of the difference between two probability distributions [76]. Each log-normal survival curve obtained from the accelerated failure time models was approximated by a gamma distribution meant to minimize K-L divergence, with κ constrained to be a positive, non-zero integer (necessary for implementation as a series of compartments). Weibull distributed survival times were refit as exponentially distributed survival times for the purposes of model parameterization. Comparisons of the empirically estimated density functions and survival curves to their approximated counterparts are shown in Figures A1 – A4. All approximated gamma distributions had κ =1, which reduces to an exponential distribution, confirming the default assumptions of the Gillespie implementation as justified.

The estimated length of stay for high risk, non-CDI patients was 12.006 days, while for low-risk patients it was 3.318 days, which translate to hourly rates of 0.00347 and 0.0126 respectively. The rate of discharge for patients with CDI is a combination of the probability of being safely discharged from the hospital and the time it takes to be discharged. Based on the estimates in Chapter 4, it is estimated that patients have a 72.15% chance of being discharged, with a mean time until discharge of 15.768 days. Similarly, patients with CDI have a 27.85% chance of dying while in the hospital, with a mean time until death of 18.66 days. Combining

the two, we obtain an hourly probability of death for CDI patients of 0.000625 and an hourly probability of discharge of 0.00188.

Contact Rate

The contact rate between patients and HCP was estimated from studies using the 'Work Observation Method by Activity Timing' (WOMBAT) method. These studies can estimate the number of patient care tasks a HCP engages in per hour that involve direct contact with the patient or their immediate environment. The estimates used in the model described in Chapter 5 are obtained from a study of the time use patterns of 100 Canadian critical care HCP [64]. Based on additional data provided by Ballermann and colleagues, this study estimates that nurses perform 11.92 direct care tasks per hour and doctors perform 3.253 direct care tasks per hour. Combined with the staffing levels described in the model and assuming random mixing and a uniform care load between patients, this results in 4.244 direct care tasks per patient per hour. As these tasks include all tasks related to direct patient care, including those that do not involve touching the patient, they represent an attempt to capture HCP interaction with the potentially contaminated physical environment as well, within the limitations of a compartmental model.

Handwashing Rate

The rate of hand washing is a composite of the number of times a hand washing opportunity presents itself to a HCP (after all direct care tasks involving the patient or their environment), a rate of compliance, and a rate at which washing

ones hands successful rids them of *C. difficile* contamination. The first two are obtained from the per HCP contact rate discussed above and a study of hand hygiene compliance within Duke Hospital[61].

Typically in studies of the efficacy of a cleaning product against microbial contamination (such as hand washing agents), the results are reported in terms of a $\log_{10}$ reduction in overall colony forming units (CFU) between the pre-wash and post-wash surfaces. Based on estimates of the efficacy of the mechanical agitation from a water-only wash (0.76 $\log_{10}$ reduction) (Edmonds et al., In Press), along with one for soap that had already been previously adjusted for the effect of water to obtain the reduction due solely to the introduction of soap (0.90 $\log_{10}$ reduction)[62] we obtain a combined 1.66 $\log_{10}$ reduction in bacterial load. For models directly modeling the surface, this reduction estimate can be used in this form. However, for the compartmental models discussed in Chapter 5, this estimate must be converted to a percent-efficacy.

Based on an initial estimated bacterial load of 3.20 $\log_{10}$ CFU on worker hands, the corresponding 1.66 $\log_{10}$ reduction results in ~1.54 $\log_{10}$ CFU remaining, a 97.8% reduction in the overall number of CFUs. This conversion makes the assumption that bacterial load is independent of the probability of infection – that is that an HCP's hands are either clean or not, and that the corresponding reduction in bacterial load translates well to a probability that a HCP's hands are below the contamination level necessary for efficient transmission. As there is little evidence regarding the infectious dose of *C. difficile*, the accuracy of this assumption is

difficult to verify. However, the generally accepted efficacy of soap and water washes as the primary means of hand decontamination in the case of *C. difficile* lends some credence to the high efficacy estimate produced with this method, and the influence of this assumption on the model's results is straightforward to examine using sensitivity analysis.

Probability of contamination of staff hands by colonized patient of type *i*

There is ample evidence that *C. difficile* can be found contaminating the hands of HCP after contact with a colonized or infected patients or the surrounding environment. A number of studies performed by Donskey and colleagues have directly examined the rate of acquisition of *C. difficile* spores on HCP hands after contact with contaminated skin sites[33,52,65,66]. While relatively small studies, they consistently report a ~50% contamination rate on gloved hands after contact with patients with active CDI. The estimation of the contamination rate for asymptomatic carrier patients is somewhat more difficult. In a study comparing asymptomatic carriers to patients with active CDI, Riggs et al.[33] report a skin contamination rate by swabbing sample sites of 61% vs. 78%. Based on the finding from the same paper that 57% of the patients with contaminated skin sites passed that contamination onto gloved hands, we can estimate a 35% chance of hand contamination on gloved hands from touching the skin of an asymptomatic carrier. No data could be found examining the contamination rates for asymptomatic carriers with high risk factors for developing CDI such as exposure to antibiotics or PPIs, and as such both types of

patient, once colonized, are assumed to be equally capable of shedding spores onto their skin and the surrounding environment.

Probability of contamination of staff hands by colonized patient of type *i*

In contrast to the probability that a patient shedding *C. difficile* spores contaminates a HCP's hands, the probability that contact with a contaminated HCP results in the colonization of a previously colonized patient is less well studied. Generally speaking however, it appears from studies estimating the bacterial burden arising from touching a shedding patient's skin by pressing a gloved hand first to the patient and then to a culture plate[33] that gloved hands are a very efficient medium for passing *C. difficile* from one receptive surface to the next, and thus we assume at least transient skin colonization to be very likely when a contaminated HCP comes in contact with an uncolonized patient.

Duration of High Risk Treatment

The duration of treatment for patients on fluoroquinolones was assumed to be 7 days, in line with treatment guidelines for a number of different potential diseases, including skin and soft tissue infections, sexually transmitted infections and community-acquired pneumonias[68-70,77].

The duration of treatment for the use of PPIs for gastric acid suppression is a somewhat more challenging task – treatment for various diseases using PPIs range from a single dose to multiple months or years of intermittent use. The duration of treatment with PPIs was assumed to be 4 weeks – the minimum length of treatment

for a number of diseases including gastroesophageal reflux disease, the healing of duodenal ulcers and NSAID-related gastric ulcers, although twice as long as the maximum recommended course for the eradication of *H. pylori* in conjunction with a course of antibiotics[71,78-80]. Note that this time is considerably higher than the average length of stay for patients, hence the assumption within the model that, when combined with the time it would take for intestinal flora to recover, that once a patient's flora have been disrupted, they remain disrupted for the duration of their stay. However, it is still important to estimate this parameter, as it is used in several of the fecal transplantation scenarios – patients are treated prophylactically immediately *after* the conclusion of their course of antibiotics or PPIs.

To reflect the mix of patients taking one or both drugs, the overall duration of treatment is the weighted average of the two drug-specific treatment durations. Patients being treated with both types of drugs were assumed to be in the high-risk category for the longer of the two treatments (i.e. patients on both a fluoroquinolone and a PPI are assumed to be in the high risk category for 4 weeks). Within the ICU patient population in the UNC Healthcare inpatient, 14.38% of patients had an order for a fluoroquinolone drug alone, 29.20% for a PPI alone and 34.07% for the two in combination. Using the treatment durations above, we thus obtain a weighted average of high-risk treatment duration of 24.66 days.

Proportion of Admitted Patients Who Are Of Patient Type *i*

The proportion of patients admitted into the D compartment, representing patients with CDI was determined to be 0.044, based on the proportion of patients

in the ICU who had *C. difficile* colitis entered as complaint present on admission in the UNC Healthcare inpatient billing data.

Based on the inpatient billing data and the assumption that patients are prescribed antibiotics or PPIs immediately upon admission to the ICU, the estimated proportion of high-risk patients is 0.7765. Based on estimates that ~ 2% of the community carries *C. difficile* as part of their commensal gut bacteria, this results in the following admission proportions for each patient type: $C_P = 0.00447$, $C_A = 0.01553$, $U_P = 0.209196$, $U_A = 0.726804$.

Relative Risk of CDI for Patients on Fluoroquinolone and PPIs

While many models assume that patients not on antibiotics or PPIs are incapable of developing CDI, the model in Chapter 5 assumes that they are instead at a significantly lower risk of developing CDI. Based on a number of meta-analyses[53,59,67], we obtain relative risks (RR) of 1.74 for PPIs alone and 3.40 for fluoroquinolone alone. Assuming the two are additive on the log-scale, this yields a relative risk for patients on both drugs of 5.92. Using the same proportions as used to determine average treatment duration, the log weighted average RR for a patient in the high risk category is 1.215. Exponentiating this weighted average RR results in an RR for high-risk patients of 3.37.

Hazard of Developing CDI in Low-Risk Patients

The model presented in Chapter 5 had a single unknown parameter for which no values could be found in the literature – the hazard of developing CDI in

patients who are not in the typical risk categories, such as those on antibiotics or PPIs. In order to obtain a value of this parameter that results in a level of infection comparable to that of the experience of a "real life" hospital, the model was fit to the DICON surveillance time series. As described in Chapter 5, rather than fitting the parameter (κ) to each hospital series and then attempting to pool those estimates, κ was fit to a single time series of the weighted average cumulative incidence of CDI for the entire data set. The best fitting value of κ was found by minimizing the sum of squared errors between the time series and the predicted cumulative incidence produced by a deterministic implementation of the model using an adaptive mesh refinement to find a best fitting value. The fit between the data and the best fitting value of κ is shown in Figure A-5.

Probability of a Discharged Patient Developing Recurrence

Estimates of the proportion of patients who go on to develop recurrence vary widely in the literature. One meta-analysis suggested that recurrence could occur in 10% to 40% of cases[54], citing previous studies on the topic. The model is parameterized with a value on the higher end of this range (0.30). Because it is applied to all scenarios, the specific value of this parameter, as long as it is in a reasonable range, should not effect the results of the modeled scenarios in relation to each other. However, if recurrence is considerably more rare, it may be more difficult to detect a significant difference between scenarios, as model realizations with zero recurrent cases will become more frequent. However it is settings where recurrence is more common that it is a pressing clinical concern and where interest

in preventative interventions is more active, hence the selection of a parameter on the higher end of the range.

Probability of Fecal Transplant in Moving Patient to Low-Risk Category

The efficacy of fecal transplant was estimated based on the percentage of patients enrolled in a recent clinical trial who were cured without relapse after one or more infusion of donor feces[55]. We assume a similar efficacy in our model, where 93.8% of patients are cured without relapse after having their intestinal flora restored.

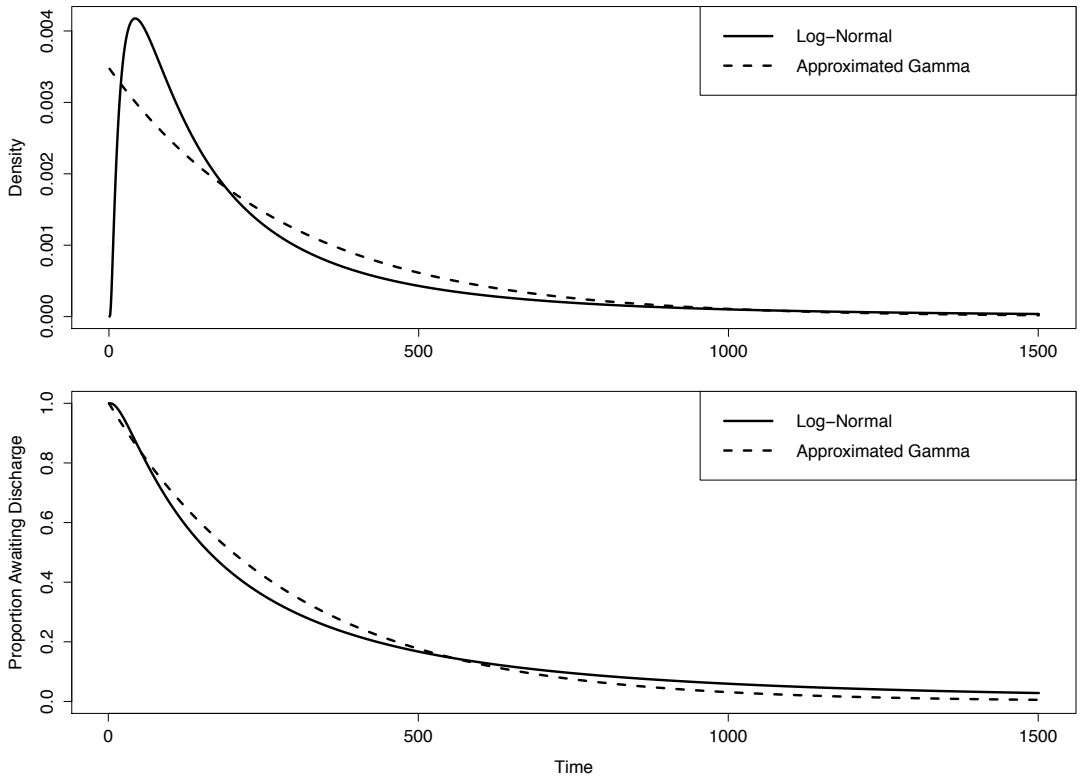

Figure A-1. Comparison of density functions and survival curves for an empirically estimated log-normal length of stay for high-risk patients without CDI and a corresponding approximate gamma distribution.

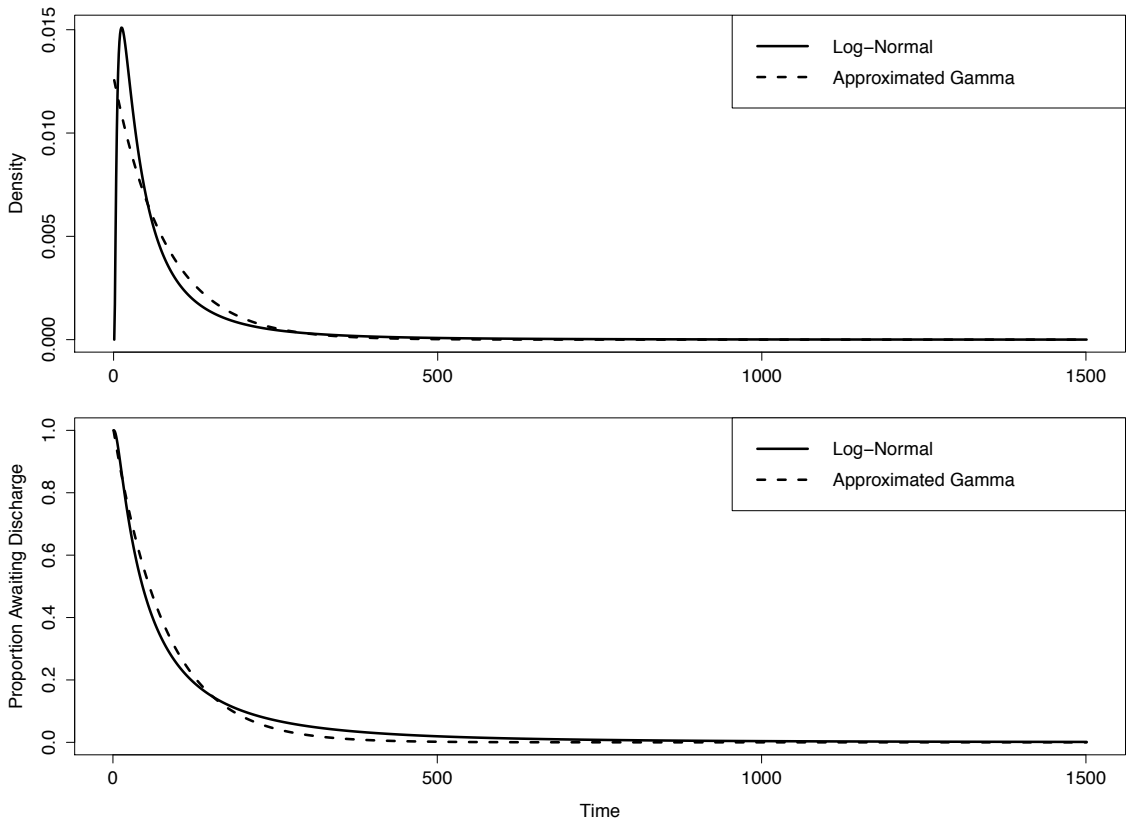

Figure A-2. Comparison of density functions and survival curves for an empirically estimated log-normal length of stay for low-risk patients without CDI and a corresponding approximate gamma distribution.

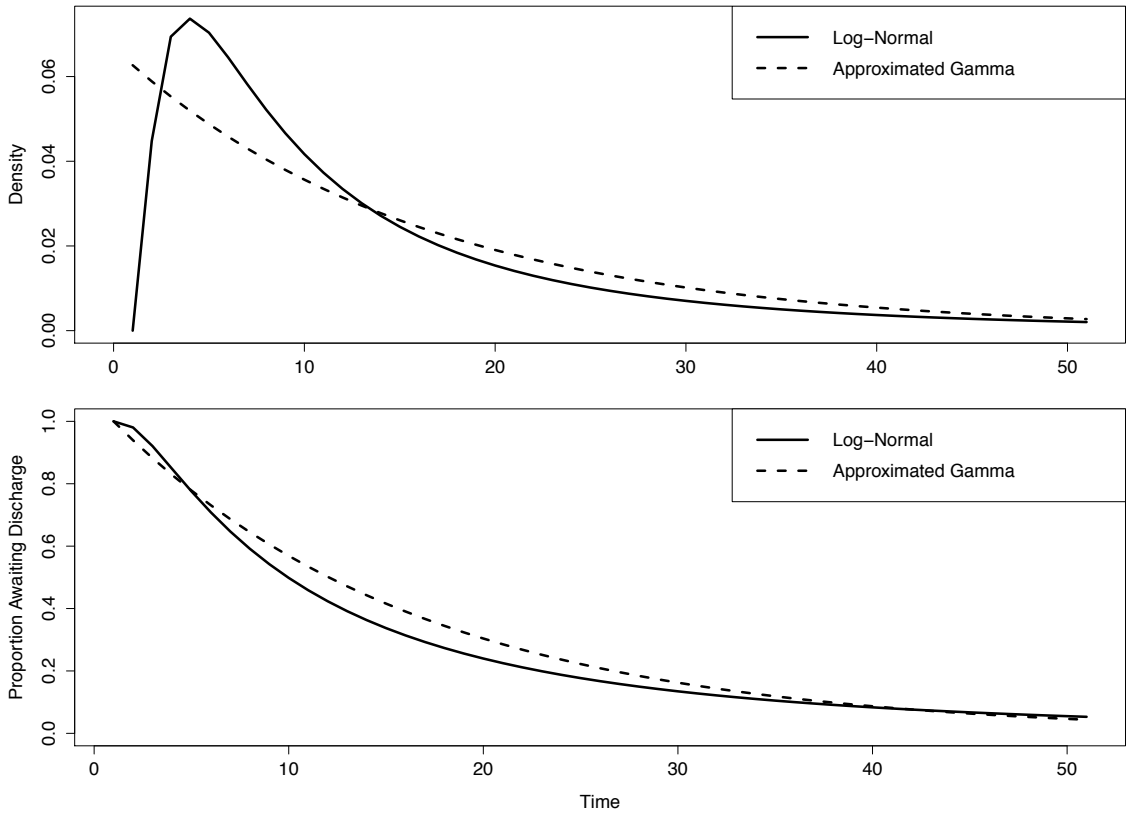

Figure A-3. Comparison of density functions and survival curves for an empirically estimated log-normal time until discharge for patients with CDI and a corresponding approximate gamma distribution.

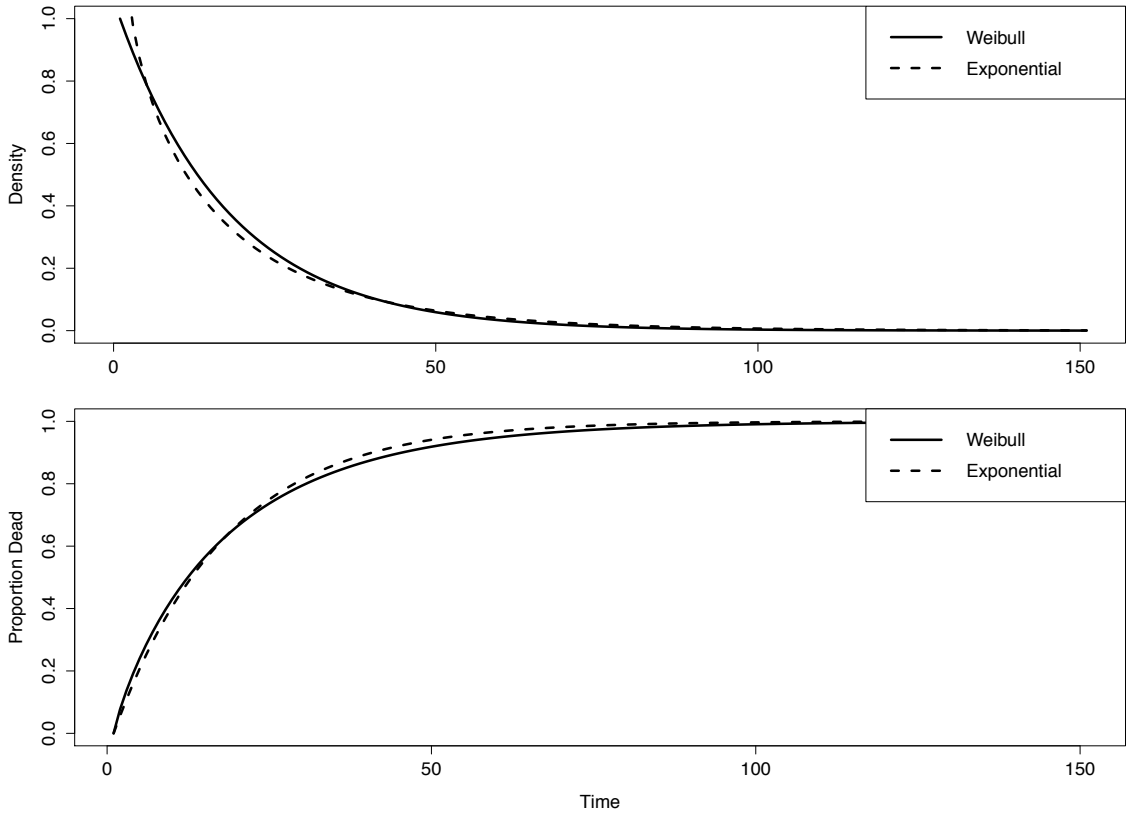

Figure A-4. Comparison of density functions and survival curves for an empirically estimated Weibull time until death for patients with CDI and a corresponding exponential fit.

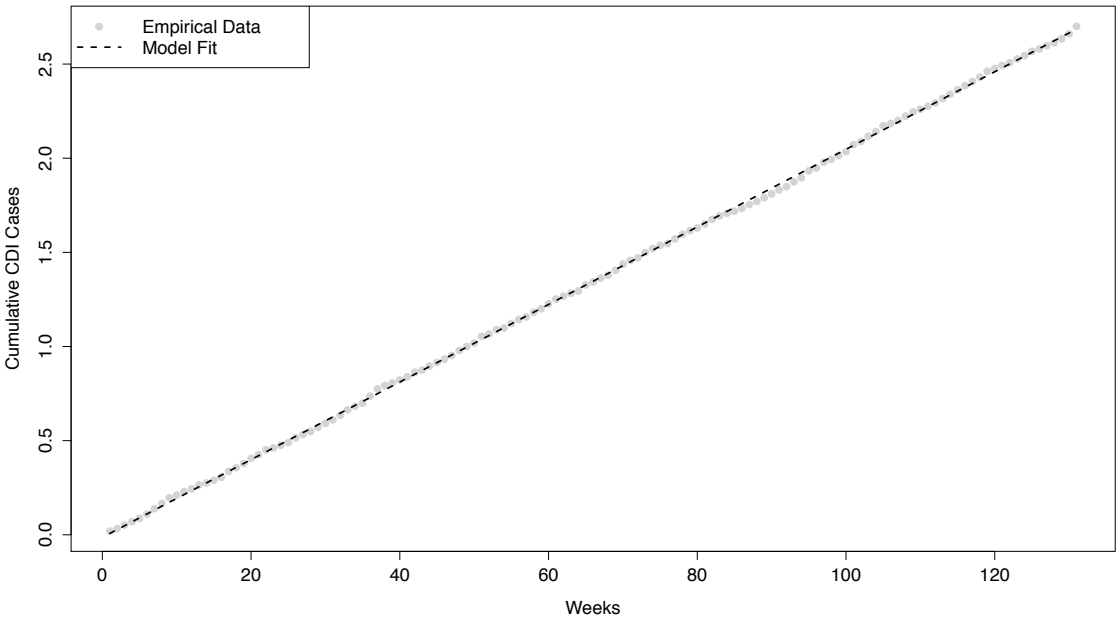

Figure A-5. Comparison of weekly weighted cumulative CDI incidence time series from 31 hospitals in the Duke Infection Control Outreach Network, and the predicted cumulative incidence obtained by fitting a deterministic model of *C. difficile* transmission to this data

References

1. McDonald LC, Killgore GE, Thompson A, et al. An epidemic, toxin gene-variant strain of Clostridium difficile. *N. Engl. J. Med.* 2005;353(23):2433–2441.

2. Bartlett JG, Perl TM. The New Clostridium difficile— What Does It Mean? *N. Engl. J. Med.* 2005;353(23):2503–2505.

3. Loo VG, Poirier L, Miller MA, et al. A predominantly clonal multi-institutional outbreak of Clostridium difficile-associated diarrhea with high morbidity and mortality. *N. Engl. J. Med.* 2005;353(23):2442–2449.

4. Miller BA, Chen LF, Sexton DJ, Anderson DJ. Comparison of the burdens of hospital-onset, healthcare facility-associated Clostridium difficile Infection and of healthcare-associated infection due to methicillin-resistant Staphylococcus aureus in community hospitals. *Infect. Control Hosp. Epidemiol*. 2011;32(4):387–390.

5. Ghantoji SS, Sail K, Lairson DR, DuPont HL, Garey KW. Economic healthcare costs of Clostridium difficile infection: a systematic review. *J. Hosp. Infect.* 2010;74(4):309–318.

6. McGlone SM, Bailey RR, Zimmer SM, et al. The economic burden of Clostridium difficile. *Clin. Microbiol. and Infect*. 2011; 18(3):282-289

7. Jain R, Kralovic SM, Evans ME. Veterans Affairs initiative to prevent methicillin-resistant Staphylococcus aureus infections. *N. Engl. J. Med.* 2011; 264(15): 1419-1430

8. Gurieva T, Bootsma M. Successful Veterans Affairs initiative to prevent methicillin-resistant Staphylococcus aureus infections revisited. *Clin. Infect. Dis.* 2012; 54(11): 1618-1620

9. Kermack W, McKendrick A. A Contribution to the Mathematical Theory of Epidemics. *Proc. R. Soc. London A*. 1927;115(772):700–721.

10. Anderson RM, May RM. Population biology of infectious diseases: Part I. *Nature*. 1979; 280: 361-367.

11. May RM, Anderson RM. Population biology of infectious diseases: Part II. *Nature*. 1979; 280: 445-461.

12. Austin DJ, Anderson RM. Studies of antibiotic resistance within the patient, hospitals and the community using simple mathematical models. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* 1999;354(1384):721–738.

13. Lipsitch M, Bergstrom CT, Levin BR. The epidemiology of antibiotic resistance in hospitals: paradoxes and prescriptions. *PNAS*. 2000;97(4):1938–1943.

14. Sébille V, Chevret S, Valleron AJ. Modeling the spread of resistant nosocomial pathogens in an intensive-care unit. *Infect. Control Hosp. Epidemiol*. 1997;18(2):84–92.

15. Smith DL, Battle KE, Hay SI, Barker CM, Scott TW, McKenzie FE. Ross, macdonald, and a theory for the dynamics and control of mosquito-transmitted pathogens. *PLoS Pathog.* 2012; 8(4)

16. Weber D, Rutala W, Miller M, Huslage K. Role of hospital surfaces in the transmission of emerging health care-associated pathogens: Norovirus, Clostridium difficile, and Acinetobacter species. *Am. J. Infect. Control.* 2010; 38(5) S25-S33

17. Bartlett MS. Deterministic and Stochastic Models for Recurrent Epidemics. In: Neyman J, ed. *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*. Berkeley and Los Angeles: University of California Press; 1956:81–108.

18. Gillespie DT. Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.* 1977;81(25):2340–2361.

19. Fefferman N, Ng K. How disease models in static networks can fail to approximate disease in dynamic networks. *Phys. Rev. E*. 2007;76(3):31919.

20. Ueno T, Masuda N. Controlling nosocomial infection based on structure of hospital social networks. *J. Theor. Biol.* 2008;254(3):655–666.

21. Christley R, Pinchbeck G. Infection in social networks: using network analysis to identify high-risk individuals. *Am. J. Epidemiol.* 2005; 162(10): 1024-1031

22. Morris M. Overview of Network Survey Designs. In: Morris M, ed. *Network Epidemiology*. Oxford University Press; 2004:8–21.

23. Curtis DE, Kanade G, Pemmaraju SV. Analysis of hospital health-care worker contact networks. *5th UK Social Networks Conference*. July 3-5, London.

24. Lee BY, Brown ST, Cooley P, et al. Simulating school closure strategies to mitigate an influenza epidemic. *J Public Health Manag. Pract*. 2010;16(3):252–261.

25. Lee BY, McGlone SM, Wong KF, et al. Modeling the spread of methicillin-resistant Staphylococcus aureus (MRSA) outbreaks throughout the hospitals in Orange County, California. *Infect. Control Hosp. Epidemiol*. 2011;32(6):562–572.

26. Smith DL, Dushoff J, Perencevich EN, Harris AD, Levin SA. Persistent colonization and the spread of antibiotic resistance in nosocomial pathogens: resistance is a regional problem. *PNAS*. 2004;101(10):3709–3714.

27. Raboud J, Saskin R, Simor A, et al. Modeling Transmission of Methicillin-Resistant Staphylococcus aureusAmong Patients Admitted to a Hospital. *Infect. Control Hosp. Epidemiol*. 2005;26(7):607–615.

28. McBryde ES, Pettitt AN, McElwain DLS. A stochastic mathematical model of methicillin resistant Staphylococcus aureus transmission in an intensive care unit: predicting the impact of interventions. *J. Theor. Biol.* 2007;245(3):470–481.

29. D'Agata EMC, Magal P, Olivier D, Ruan S, Webb GF. Modeling antibiotic resistance in hospitals: the impact of minimizing treatment duration. *J. Theor. Biol.* 2007;249(3):487–499.

30. Perencevich EN. Deconstructing the Veterans Affairs MRSA prevention bundle. *Clin. Infect. Dis*. 2012;54(11):1621–1623.

31. Starr JM, Rogers TR, Impallomeni M. Hospital-acquired Clostridium difficile diarrhoea and herd immunity. *Lancet*. 1997;349(9049):426–428.

32. Starr JM, Campbell A. Mathematical modeling of Clostridium difficile infection. *Clin. Microbiol. Infect.* 2001;7(8):432–437.

33. Riggs MM, Sethi AK, Zabarsky TF, Eckstein EC, Jump RLP, Donskey CJ. Asymptomatic Carriers Are a Potential Source for Transmission of Epidemic and Nonepidemic Clostridium difficile Strains among Long-Term Care Facility Residents. *Clin. Infect. Dis.* 2007;45(8):992–998.

34. Starr JM, Campbell A, Renshaw E, Poxton IR, Gibson GJ. Spatio-temporal stochastic modelling of Clostridium difficile. *J. Hosp. Infect.* 2009;71(1):49–56.

35. Cohen SH, Gerding DN, Johnson S, et al. Clinical practice guidelines for Clostridium difficile infection in adults: 2010 update by the society for healthcare epidemiology of America (SHEA) and the infectious diseases society of America (IDSA). *Infect. Control Hosp. Epidemiol*. 2010;31(5):431–455.

36. Grima DT, Webb GF, D'Agata EMC. Mathematical model of the impact of a nonantibiotic treatment for Clostridium difficile on the endemic prevalence of vancomycin-resistant enterococci in a hospital setting. *Comput. Math Methods Med*. 2012.

37. Lanzas C, Dubberke ER, Lu Z, Reske KA, Gröhn YT. Epidemiological model for Clostridium difficile transmission in healthcare settings. *Infect. Control Hosp. Epidemiol*. 2011;32(6):553–561.

38. Shaughnessy MK, Micielli RL, DePestel DD, et al. Evaluation of hospital room assignment and acquisition of Clostridium difficile infection. *Infect. Control Hosp. Epidemiol*. 2011;32(3):201–206.

39. Hacek DM, Ogle AM, Fisher A, Robicsek A, Peterson LR. Significant impact of terminal room cleaning with bleach on reducing nosocomial Clostridium difficile. *Am J Infect Control*. 2010;38(5):350–353.

40. McDonald LC, Coignard B, Dubberke E, et al. Recommendations for surveillance of Clostridium difficile-associated disease. *Infect. Control Hosp. Epidemiol*. 2007;28(2):140–145.

41. *HCUPnet.* Agency for Healthcare Research and Quality; 2006. Available at: http://hcupnet.ahrq.gov/. Accessed September 17, 2012.

42. Lau B, Cole SR. Competing risk regression models for epidemiologic data. *Am. J. Epidemiol*. 2009; 170(2): 244-256.

43. Anderson DJ, Miller BA, Chen LF, et al. The network approach for prevention of healthcare-associated infections: long-term effect of participation in the Duke Infection Control Outreach Network. *Infect. Control Hosp. Epidemiol*. 2011;32(4):315–322.

44. Cole SR, Hernán MA. Constructing inverse probability weights for marginal structural models. *Am. J. Epidemiol*.2008;168(6):656–664.

45. Cole SR, Hernán MA. Adjusted survival curves with inverse probability weights. *Comput. Methods Programs Biomed*. 2004;75(1):45–49.

46. Little RJA, Rubin DB. The Analysis of Social Science Data with Missing Values. *Sociol Method Res*. 1989;18(2-3):292–326.

47. Lau B, Cole SR, Gange SJ. Parametric mixture models to evaluate and summarize hazard ratios in the presence of competing risks with time-dependent hazards and delayed entry. *Stat. Med*. 2011;30(6):654–665.

48. Cox DR. The analysis of exponentially distributed life-times with two types of failure. *J. R. Stat. Ser. B Stat. Methodol.* 1959; 21(2):411-421

49. Cole SR, Li R, Anastos K, et al. Accounting for leadtime in cohort studies: evaluating when to initiate HIV therapies. *Stat. Med.* 2004;23(21):3351–3363.

50. Checkley W, Brower RG, Muñoz A, NIH Acute Respiratory Distress Syndrome Network Investigators. Inference for mutually exclusive competing events through a mixture of generalized gamma distributions. *Epidemiology*. 2010;21(4):557–565.

51. Ghani AC, Donnelly CA, Cox DR, et al. Methods for estimating the case fatality ratio for a novel, emerging infectious disease. *Am. J. Epidemiol*. 2005;162(5):479–486.

52. Sethi AK, Al-Nassir WN, Nerandzic MM, Bobulsky GS, Donskey CJ. Persistence of skin contamination and environmental shedding of Clostridium difficile during and after treatment of C. difficile infection. *Infect. Control Hosp. Epidemiol*. 2010;31(1):21–27.

53. Kwok CS, Arthur AK, Anibueze CI, Singh S, Cavallazzi R, Loke YK. Risk of Clostridium difficile infection with acid suppressing drugs and antibiotics: meta-analysis. *Am. J. Gastroenterol.* 2012;107(7):1011–1019.

54. Garey KW, Sethi S, Yadav Y, DuPont HL. Meta-analysis to assess risk factors for recurrent Clostridium difficile infection. *J. of Hosp. Infect*. 2008; 70(4): 298-304.

55. van Nood E, Vrieze A, Nieuwdorp M, et al. Duodenal infusion of donor feces for recurrent Clostridium difficile. *N. Engl. J. Med.* 2013;368(5):407–415.

56. You DM, Franzos MA, Holman RP. Successful treatment of fulminant Clostridium difficile infection with fecal bacteriotherapy. *Ann. Intern. Med.* 2008;148(8):632–633.

57. Petrof EO, Gloor GB, Vanner SJ, Weese SJ. Stool substitute transplant therapy for the eradication of Clostridium difficile infection:"RePOOPulating"the gut. *Microbiome*. 2013; 1(3).

58. Rupnik M, Wilcox MH, Gerding DN. Clostridium difficile infection: new developments in epidemiology and pathogenesis. *Nature Reviews Microbiology*. 2009;7(7):526–536.

59. Owens RC, Donskey CJ, Gaynes RP, Loo VG, Muto CA. Antimicrobial-associated risk factors for Clostridium difficile infection. *Clin. Infect. Dis*. 2008;46(S1):S19–31.

60. Moehring RW, Lofgren ET, Obure J, Anderson DJ. Impact of Switch to Polymerase Chain Reaction (PCR) in Rate of C. difficile Infection (CDI): Multicenter Before-After Study using a Semi-Parametric Interrupted Time Series Model. In: Abstracts of IDWeek 2012; October 2-6, 2012;San Diego, CA. Abstract 36149.

61. Chen LF, Carriker C, Staheli R, et al. Observing and Improving Hand Hygiene Compliance: Implementation and Refinement of an Electronic-Assisted Direct-Observer Hand Hygiene Audit Program. *Infect. Control Hosp. Epidemiol*. 2012;34(2): 207-210.

62. Oughton MT, Loo VG, Dendukuri N, Fenn S, Libman MD. Hand hygiene with soap and water is superior to alcohol rub and antiseptic wipes for removal of Clostridium difficile. *Infect. Control Hosp. Epidemiol*. 2009;30(10):939–944.

63. Jabbar U, Leischner J, Kasper D, et al. Effectiveness of alcohol-based hand rubs for removal of Clostridium difficile spores from hands. *Infect Control Hosp Epidemiol*. 2010;31(6):565–570.

64. Ballermann MA, Shaw NT, Mayes DC, Gibney RTN, Westbrook JI. Validation of the Work Observation Method By Activity Timing (WOMBAT) method of conducting time-motion observations in critical care settings: an observational study. *BMC Med Inform. Decis. Mak*. 2011;11:32.

65. Guerrero DM, Nerandzic MM, Jury LA, Jinno S, Chang S, Donskey CJ. Acquisition of spores on gloved hands after contact with the skin of patients with Clostridium difficile infection and with environmental surfaces in their rooms. *Am J Infect Control*. 2012;40(6):556–558.

66. Bobulsky GS, Al-Nassir WN, Riggs MM, Sethi AK, Donskey CJ. Clostridium difficile Skin Contamination in Patients with C. difficile-Associated Disease. *Clin. Infect. Dis.* 2008;46(3):447–450.

67. Pépin J, Saheb N, Coulombe MA. Emergence of fluoroquinolones as the predominant risk factor for Clostridium difficile–associated diarrhea: a cohort study during an epidemic in Quebec. *Clin. Infect. Dis.* 2005; 41(9): 1254-1260.

68. Mandell LA, Wunderink RG, Anzueto A, et al. Infectious Diseases Society of America/American Thoracic Society Consensus Guidelines on the Management of Community-Acquired Pneumonia in Adults. *Clin. Infect. Dis.*2007;44(S2):S27–S72.

69. Warren JW, Abrutyn E, Hebel JR, Johnson JR, Schaeffer AJ, Stamm WE. Guidelines for antimicrobial treatment of uncomplicated acute bacterial cystitis and acute pyelonephritis in women. Infectious Diseases Society of America (IDSA). *Clin. Infect. Dis.*1999;29(4):745–758.

70. Giordano P, Weber K, Gesin G, Kubert J. Skin and skin structure infections: treatment with newer generation fluoroquinolones. *Ther. Clin. Risk Manag*. 2007;3(2):309–317.

71. Vanderhoff BT, Tahboub RM. Proton pump inhibitors: an update. *Am Fam Physician*. 2002;66(2):273–280.

72. Kane RL, Shamliyan TA, Mueller C, Duval S, Wilt TJ. The Association of Registered Nurse Staffing Levels and Patient Outcomes. *Med. Care*. 2007;45(12):1195–1204.

73. Ward NS, Read R, Afessa B, Kahn JM. Perceived effects of attending physician workload in academic medical intensive care units: a national survey of training program directors. *Crit. Care Med.* 2012;40(2):400–405.

74. Thompson DR, Hamilton DK, Cadenhead CD, et al. Guidelines for intensive care unit design*. *Crit. Care Med.* 2012;40(5):1586–1600.

75. Lloyd AL. Sensitivity of Model-Based Epidemiological Parameter Estimation to Model Assumptions. *Mathematical and Statistical Estimation Approaches in Epidemiology*. 2009.

76. Bernardo JM. Approximations in statistics from a decision-theoretical viewpoint. In: Viertl R, ed. *Probability and Bayesian Statistics*.1987:53–60.

77. Centers for Disease Control and Prevention. Sexually Transmitted Diseases Treatment Guidelines, 2010. *MMWR Morbid Mortal Wkly Rep*. 2010;59(RR-12):1–116.

78. Pharmaceuticals A. PRILOSEC Prescribing Information. 2012:1–12.

79. Pharmaceuticals T. PREVACID Prescribing Information. 1012:1–28.

80. Eisai Inc. *ACIPHEX Prescribing Information*. 2012:1–15.