Mathematical studies of the dynamics of antibiotic resistance

Mathematical studies of the dynamics of antibiotic resistance

Wiskundige studies over de dynamiek van antibioticaresistentie

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht
op gezag van de Rector Magnificus, Prof. dr. W. H. Gispen,
ingevolge het besluit van het College voor Promoties in het openbaar
te verdedigen op maandag 23 mei 2005 des middags te 12.45 uur

door

Martinus Christoffel Jozef Bootsma

geboren op 18 december 1977, te Almelo

promotores: Prof. Dr. O. Diekmann
Faculteit der Wiskunde en Informatica
Universiteit Utrecht

Prof. Dr. M. J. M. Bonten
Faculteit der Geneeskunde
Universiteit Utrecht

Dit proefschrift werd mede mogelijk gemaakt door financiële steun van
ZonMW en de Universiteit van Utrecht

2000 Mathematics Subject Classification: 92C60, 92D30

ISBN 90-393-2619-3

Contents

Chapter 1

General introduction

Infectious diseases, also called communicable diseases, have always tormented humans. The bubonic plague in the fourteenth century killed more than one third of the people of Europe and the Spanish invasion in Mexico led to a decrease of the local population of 95% in 75 years, mostly due to the introduction of infectious diseases. However, infectious diseases are certainly not something from the past alone. Recent examples of human infectious diseases, all with different epidemiology, are the Acquired Immune Deficiency Syndrome (AIDS), infection with multi-drug resistant microorganisms like Methicillin-resistant *Staphylococcus aureus* (MRSA), the Severe Acute Respiratory Syndrome (SARS) and variant Creutzfeld-Jakob disease caused by prions and linked to BSE in cows.

Although we focus on human infectious diseases in this thesis, infectious diseases are not specific for humans, also animals and plants suffer from them. Recent examples of these are the Phocine Distemper Virus, the virus that killed a substantial part of the seal population in north-western Europe in 1988 and 2002, foot and mouth disease, classical swine fever, bovine spongiform encephalopathy (BSE) also called the mad cow disease, avian influenza and a fungus (*Ophiostoma ulmi*) that causes the elm tree disease.

Antibiotic resistance

In this thesis we focus on colonization of humans with bacteria resistant for antimicrobial agents. Colonization is defined as carriership of the pathogen which can lead to infection, defined as an inflamating

response to an infecting agent or its products. The large-scale introduction of penicillin during and after WWII and the development of many new antimicrobial agents gave the impression that bacterial infections were no longer a serious threat for humans, as these infections could now be cured. However, ever since the application of antimicrobial agents, bacterial strains resistant to these agents are seen with increasing frequency and treatment of bacterial infections with antimicrobial agents has no longer a guaranteed success.

The bacteria perspective

For most types of resistance, bacteria have to carry additional genes that encode for the mechanisms that lead to resistance [Andersson and Levin, 1999]. These additional genetic elements have to be duplicated upon cell division and therefore cell division of resistant bacteria should be slower and requires more resources compared with cell division of non-resistant bacteria of the same strain. This means that, in the absence of the antimicrobial agent, non-resistant bacterial strains should have a selective advantage compared to the resistant strains. However, this so called cost of resistance may be overcome by compensating mutations [Andersson and Levin, 1999]. Moreover, plasmids and transposons (small mobile genetics elements that can be transferred between bacteria (even of different strains)) which carry resistance genes may be an evolutionary cheap way for bacteria to eliminate the cost of resistance [Bergstrom et al., 2000]. When a patient carries other types of bacteria, for instance non-pathogenic commensal strains, and these commensal strains have mobile genetic elements like plasmids and transposons with resistance genes against the given antimicrobial agents, there is a serious risk that the pathogenic strain acquires the resistance genes from the harmless bacteria. Therefore, in the human perspective, resistance in non-pathogenic bacteria is also a serious problem.

The human perspective: how to reduce it?

The most obvious way to reduce the frequency of carriage of resistant bacteria is a prudent use of antimicrobial agents [Stewart et al., 1998], [Austin et al., 1999a]. This will reduce the selective advantage of resistant strains and may hence lead to a lower frequency of resistant strains compared to the frequency of susceptible strains. However, also other

strategies may lead to a lower level of resistance. For instance, multi-drug therapy might be more effective in preventing resistance than single drug therapy despite the fact that the total amount of antimicrobial agents prescribed per patient is higher. The idea behind this mechanism is that in case bacteria suffer from one type of antimicrobial agent, acquiring a genetic element that leads to resistance of the bacteria to that antimicrobial agent is sufficient to survive. (The acquisition may take place by mutations or by acquisition of genetic elements via transposons or plasmids from other strains.) When bacteria are exposed to more antimicrobial agents, a single acquisition is not sufficient for survival and when the probability for bacteria of acquiring resistance against all used antimicrobial agents is extremely small, this will prevent the development of resistance strains.

Another strategy to prevent resistance in hospitals is the so-called cycling of antimicrobial agents, the sequential use of different antibiotics in a hospital. Cycling is based on the idea that once the frequency of resistance for an antimicrobial agent rises above a certain level, the hospital switches to another antimicrobial agent for which no resistance is present and which has a similar spectrum, i.e., it is active against the same type of bacteria. When resistance against this agent has increased, the frequency of bacteria resistant for the original antimicrobial agent should have decreased (for instance because patients who carried these resistant bacteria are discharged) and re-introduction of the original agent should be successful. However, the effect of this measure is controversial [Bergstrom et al., 2004; Bonhoeffer et al., 1997].

Another approach to prevent infections in immunocompromised patients is selective digestive decontamination (SDD) [Jonge et al., 2003]. Immunocompromised patients receive antimicrobial agents that eradicate potentially pathogenic gram-negative bacteria in the gut but leave the anaerobic flora intact, thereby preventing bacteria from colonizing the gut (this concept is called colonization resistance [Vollaard, 1991]). With a high level of hygiene in the (intensive care) unit and, ideally, no potentially pathogenic bacteria in the gut, there are no pathogenic bacteria left to cause infections. However, when resistance against the antimicrobial agent is already present, as might be the case in hospitals, these resistant strains will have a selective advantage and SDD may lead to an increase in the number of infections with resistant bacteria.

The available data and how to extract useful information from it?

To study the effect of interventions, like SDD or cycling, is not easy. There are many different antimicrobial mechanisms and many different types of bacteria that all interact differently. But even if we focus on one specific type of bacteria there are many problems. One of the problems in the epidemiology of resistant bacteria is that there is not just a single clone of each bacterium but they exist in all kind of varieties. Not only are there differences in antimicrobial resistance, e.g., vancomycin-resistant *Enterococcus faecium* (VRE) [Hayden, 2000] [Austin et al., 1999b] vs. vancomycin-susceptible *Enterococcus faecium* (VSE) and Methicillin-resistant *Staphylococcus aureus* (MRSA) [Austin and Anderson, 1999] vs Methicillin-susceptible *Staphylococcus aureus* (MSSA), but strains can differ in many other characteristics, e.g., in adherence factors, either or not toxin-producing, virulence factors [Willems et al., 2001], the level of resistance against an antimicrobial agent and so on. Large databases with information on different clonal groups of these bacteria are available and genotyping is required to analyze the epidemiology of these different clonal groups as standard culturing on plates (with antibiotics to test the susceptibility) is not sufficient to determine whether strains collected from different patients are identical. However, genotyping is an expensive method and other methods may be useful as a complementary tool (see Section 5). Another basic problem in epidemiology, but also for clinicians, is that a negative microbiological culture does not necessarily mean that the microorganisms of interest are not present because a culture may also be false negative [D'Agata et al., 2002a] either because the microorganism of interest was not present at the body-site at which the swap was performed, but was present at other sites, or because the test itself is imperfect. Moreover, a positive culture does also not imply that the patient is persistently colonized with the bacteria, as colonization can be transient.

The possible role of mathematical models

What has mathematics to do with the previous discussion? At first sight maybe not much, but if one faces an outbreak or epidemic with an infectious disease, one would like to know the answer to questions like: "What will be the final size of the epidemic?" or "What will be the effect of intervention measures?" One can think of intervention measures

like the removal of all elm trees in a specific area to stop the spread of the elm tree disease, the preventive removal of cattle in case of foot and mouth disease, to quarantine infectious individuals in case of SARS [Wallinga and Teunis, 2004] or the screening on admission of patients with a higher risk of colonization with MRSA [Verhoef et al., 1999]. To answer these questions, mathematical models of the spread of infectious diseases can be useful and these models can also contribute to answering questions whether the interventions are cost-effective.

Before starting to model an infectious disease, one has to have a question in mind to which an answer is sought. However, even if the question is well-defined, there are many aspects involved. One has to have a basic understanding of the disease, but also the question how detailed the model should be is important. A very detailed description that stays close to reality often has the disadvantage that the model has many parameters. If the value of many of these parameters is not known very well, the predictions of the model may be inaccurate if one of the estimates of the parameters is way off. Yet, such a complicated model may serve as a tool to determine which parameters play an important role and should be determined more accurately. Often a very simplistic model is chosen which is a caricature of reality, but which, hopefully, gives the correct relation between the parameters of interest. In modeling, the interaction between models and experiments is crucial. Ideally, a model is a tool for constructing a hypothesis which can be tested experimentally. The experiments may lead to the observation that the model is not completely satisfactory and the revised model may, again, lead to a new hypothesis which can be tested.

If one has a basic idea which elements are important and should be incorporated in the model, there is still an important decision to make. What is the influence of chance in the process , i.e., should we use deterministic models or stochastic models?

With the word deterministic we mean that when we know the initial situation, we can predict, with certainty, subsequent behaviour. In reality this is not the case and probability is important, e.g., an individual has a certain probability to acquire the influenza virus during an epidemic but whether or not this individual acquires the virus cannot be predicted beforehand with certainty. To acquire influenza one has to be in the vicinity of someone who has influenza. Apart from heterogeneity between individuals (some have many contacts with people, some only a few, some are (partly) immune for influenza, others not), chance

processes play a role, e.g., whether or not there was an infectious individual in the bus you took? When there are many infectious individuals and many susceptibles, you still are not able to predict with certainty whether or not a given individual will acquire influenza. However, if you are interested in the total number of infectious individuals at a given time, the individual chances to acquire influenza will average out and it is possible to predict relatively accurately how many individuals are infectious at a certain time. (Compare this with throwing a coin. If you throw a fair coin 10000 times, you can be quite sure that the number of heads that will appear will be close to 5000.) When either the number of susceptibles or the number of infectious individuals is small, chance will become important. (If one throws a fair coin 2 times, it is not unlikely that you will only see heads.) In general, any model for infectious diseases is stochastic in nature at the individual level. However, if the numbers of individuals in each of the categories is large, stochastic fluctuations will be small and a deterministic model is likely to give good results.

For small hospital units (intensive care units), a deterministic model is unlikely to be satisfactory. Suppose one knows that on average 3.2 of the 8 patients are colonized. Such a number does not mean much without knowledge about the statistical fluctuations. For instance, are there often zero colonized patients in the unit while all patients are colonized at other days, or is the number of colonized patients fairly constant. This is crucial information for designing intervention policies. What does it mean that many patients in a unit are colonized? Is it just a 'normal' statistical fluctuation or is it an extreme value and is it likely that something is wrong with the hygiene in the unit? There is in fact a general problem with intervention studies: If one notices a high prevalence of a certain pathogen, intervention measures are taken. But what can be concluded from a decrease in the prevalence? Is the decrease the result of the intervention measure, or was there, by chance, a high prevalence and was it likely that the prevalence would have decreased also in the absence of the intervention measures? Therefore an epidemiological study without knowledge of the typical fluctuations, i.e., without data from a baseline period, can be useless.

Examples of the kind of conclusions one can deduce from mathematical models

We will now discuss two well-known mathematical models as they both play a role in this thesis.

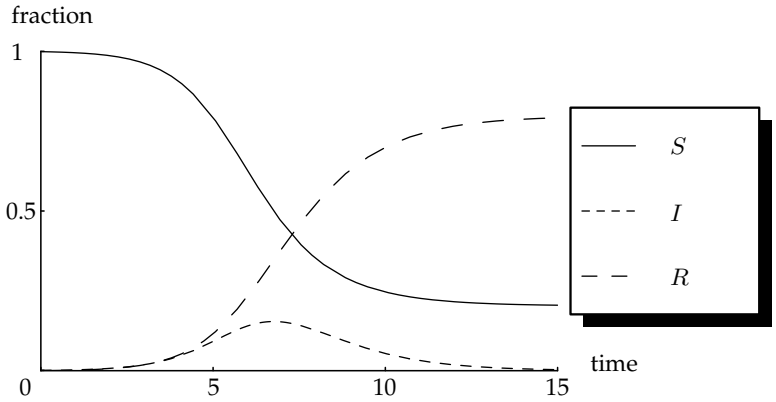
The first model is the simple case of the Kermack-McKendrick model [Kermack and McKendrick, 1927] which is a reasonable model for an epidemic that lasts for a short period of time, i.e., birth and death of individuals is of minor importance during the course of the epidemic. An example to which the Kermack-McKendrick model can be applied is measles. A typical measles infection of a child has the following pattern. After contact with the saliva of a person who has measles, the measles virus can enter the body via the respiratory system. After infection, the virus multiplies inside the body. After 8 till 14 days the person becomes ill and, more importantly for the spread of the disease, the person becomes infectious. After about two weeks the immune system beats the virus and the individual becomes resistant against the measles virus. To model the spread of the disease, we look at a population of N individuals. In the Kermack-McKendrick model, the population is divided into three categories. Individuals who have never been in contact with the virus (labeled with an 'S' of susceptibles), individuals who are infectious (labeled with an 'I') and the group of individuals who have had the disease but are not infectious anymore, either because they have become resistant or because they died. This group is labeled with an 'R' for 'Removed'. Because of the names of the groups, the simple version of the Kermack-McKendrick model is also called the SIR-model. (A more realistic model for the spread of measles should incorporate the latent period (the time from infection to when the individual is infectious to others). This requires the introduction of an extra category of individuals and leads to the so-called SEIR-models where the 'E' stands for exposed. However, within a closed population the final size is not influenced by a latent period.)

The SIR-model assumes that the number of contacts per unit of time between susceptibles and infectives is proportional to the number of infectives and the number of susceptibles in the population. The constant of proportionality contains information about how many contacts individuals have per day and which fraction of the contacts actually leads to infection. This constant is denoted with the symbol β/N where division by N ensures that the number of contacts per unit of time of an

individual does not depend on N. Another parameter in the model determines how long individuals remain infectious. The probability for an infectious individual to lose the contagiousness is denoted by α. Note that this debatable assumption implies that the infectious period is exponentially distributed; in fact the parameter α is chosen as $(\text{average period of infectiousness})^{-1}$. We ignore birth and death in the population, which is a reasonable assumption because the duration of a measles-epidemic is short (a few months). Therefore the number of newborns and the number of deaths during the epidemic is small compared to the size of the population. With these assumptions, we can write down a system of three differential equations that governs, for large N, the dynamics of the disease when we know the size of each of the three categories at a given time.

$$\begin{aligned} \frac{dS}{dt} &= -\beta \frac{IS}{N} \\ \frac{dI}{dt} &= \beta \frac{IS}{N} - \alpha I \\ \frac{dR}{dt} &= \alpha I \end{aligned} \tag{1.1}$$

Although the derivation of this model is not complicated, the system of differential equations is already too complicated to solve analytically. However, a numerical solution can be determined without any problem. For $\alpha = 1$ and $\beta = 2$ we obtain as solution



where we assumed that initially almost all individuals are susceptible.

Although an explicit solution for the SIR-model is not available, other characteristics of the epidemic can be calculated exactly. For instance, one can determine whether or not the pathogen is 'contagious' enough to start an epidemic. To answer this question, we introduce the basic reproduction number R_0 [Diekmann et al., 1990]. Suppose a population has never been in contact with the pathogen of interest

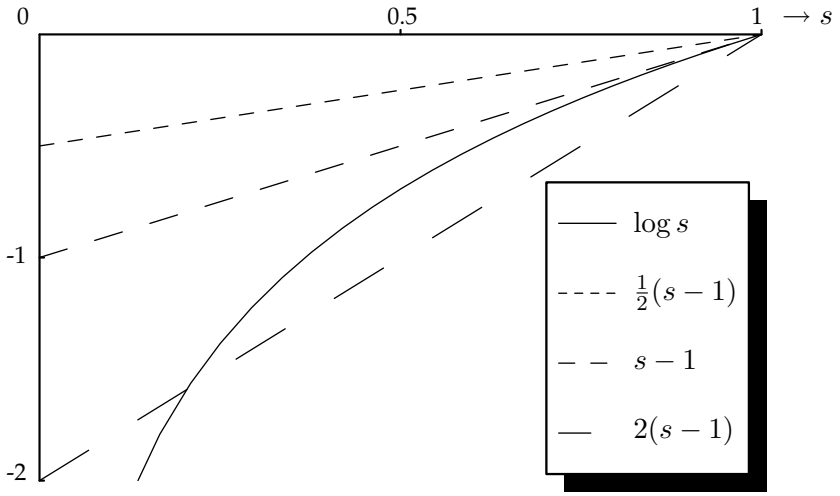
and an infectious individual enters the population. R_0 is then defined as the expected number of individuals that are infected by the index case. If we ignore the depletion of susceptibles in the population, each of the infected individuals will infect on average again R_0 susceptibles. Therefore, the disease will spread in the population when $R_0 > 1$ while the pathogen becomes extinct if $R_0 < 1$. For the SIR-model, the calculation of R_0 is straightforward. Each infectious individual has a probability α per day to lose the infectiousness. Therefore, given that an individual becomes infectious, the expected length of the infectious period is $\frac{1}{\alpha}$. Per day, an infectious individual infects $\beta S/N$ individuals, but the fraction of susceptibles in the population is 1 directly after the introduction of the pathogen. Therefore an infectious individual infects on average $R_0 = \frac{\beta}{\alpha}$ susceptibles. When we know R_0, we can determine which fraction of the population is still susceptible after an epidemic. Instead of looking at absolute numbers of individuals, we look at fractions and we denote with s, i and r the fraction of the individuals that belong to the respective categories. One can verify that at the following quantity does not change over time, it is conserved:

$$\frac{\alpha}{\beta} \ln s - s - i \tag{1.2}$$

Before the epidemic (denote this time with $-\infty$) the fraction of susceptibles equals 1 and there are no infectious individuals. Similarly, when the epidemic is over (denote this time with $+\infty$), there are no infectious individuals. The equation $\frac{\alpha}{\beta} \ln s(-\infty) - s(-\infty) - i(-\infty) = \frac{\alpha}{\beta} \ln s(\infty) - s(\infty) - i(\infty)$ reduces to:

$$\ln s(\infty) = R_0 \left(s(\infty) - 1\right) \tag{1.3}$$

With this equation, we can determine the fraction of the population that is still susceptible after the epidemic. In the figure below we plotted the left hand side of equation (1.3) as a function of s and we plotted the right hand side for $R_0 = 0.5, 1$ and 2.



From this graph we note that there are two distinct situations, i.e., $R_0 < 1$ and $R_0 > 1$. If $R_0 < 1$ we see that the only solution is $s = 1$. This means that everybody in the population is still susceptible. This is logical as the pathogen was not contagious enough to start an epidemic. If $R_0 > 1$ equation (1.3) has two solutions. One solution is $s = 1$ when there has been no epidemic, for instance because no infectious individual has entered the population. The other solution shows that even when an epidemic has occurred, a fraction of the population has escaped infection. These individuals are still susceptible after the epidemic. However, the larger the R_0-value in this model, the smaller the fraction of susceptibles after the epidemic.

Although this model is relatively simplistic (the article of Kermack-McKendrick also deals with far more general models), this model sketches how an infectious disease can spread in the population.

This model, though, is not suitable for all questions about infectious-disease-dynamics one would like to study. The duration of an epidemic like HIV is so long that the population dynamics (natural birth and death) cannot be ignored. Other diseases have a long latency period in which the patient is not at all or hardly infectious, e.g., tuberculosis for which the latency period can be lifelong. Such a latency period can have important consequences for the spread of a disease. Moreover, in case of a latency period, the latency period need not be exponentially distributed. On the contrary, it has often a more or less fixed duration. This questions the validity of the standard type of differential equations which implicitly assume exponential distributions for the infectious pe-

riod and for the latency period. Often the general conclusions which can be drawn from these simplistic models are quite robust, though quantitative answers based on these models are not very accurate. Another problem is that patients do not become completely immune but only to a specific strain (influenza). Other infections, like Herpes simplex, are recurrent. After recovery, the virus hides inside the body and may cause symptoms again when the immune system is weakened. Other diseases are caused by pathogens that have a lifecycle in different hosts, e.g., for malaria. Both humans and mosquitos can carry the malaria parasite. There is no direct transmission of the parasite between humans or between mosquitos, but a parasite free mosquito can acquire infection on biting a human who carries the parasite in the blood and uncolonized humans can acquire infection when they are bitten by an infected mosquito. The Ross-Macdonald model, see e.g., Chapter 14 of [Anderson and May, 1991], takes this into account.

Recently [Austin et al., 1999b] the Ross-Macdonald model was applied for the spread of vancomycin-resistant enterococci in an intensive care unit. In this application the health care workers (HCW) serve as vectors (like the mosquitos) who can transfer the pathogen from one patient to another. In a similar way as for the Kermack-McKendrick-model, we can construct a system of differential equations for the application of the Ross-Macdonald model to an intensive care unit. This system describes the simplest variant of the model in which re-infection is not important.

$$\begin{aligned}
\tfrac{dS_1}{dt} &= -\beta_1 \tfrac{I_2}{I_2+S_2} S_1 + \alpha_1 I_1 \\
\tfrac{dI_1}{dt} &= \beta_1 \tfrac{I_2}{I_2+S_2} S_1 - \alpha_1 I_1 \\
\tfrac{dS_2}{dt} &= -\beta_2 \tfrac{I_1}{I_1+S_1} S_2 + \alpha_2 I_2 \\
\tfrac{dI_2}{dt} &= \beta_2 \tfrac{I_1}{I_1+S_1} S_2 - \alpha_2 I_2
\end{aligned} \tag{1.4}$$

(Note that this deterministic model for HCW and patients relies on the assumption that both the number of HCW and the number of patients are fairly large.) If we assume that the number of patients is constant (N_1) and the number of HCW is constant (N_2), the system reduces to a system of 2 differential equation because $S_1+I_1 = N_1$ and $S_2+I_2 = N_2$.

$$\begin{aligned}
\tfrac{dI_1}{dt} &= \beta_1 (N_1 - I_1) \tfrac{I_2}{N_2} - \alpha_1 I_1 \\
\tfrac{dI_2}{dt} &= \beta_2 (N_2 - I_2) \tfrac{I_1}{N_1} - \alpha_2 I_2
\end{aligned} \tag{1.5}$$

When there is good hand hygiene, HCW's will carry the bacterium for a short period of time (till the next disinfection procedure) while the pa-

tients are colonized for a much longer period (typically during the rest of their stay). This scale difference allows us to use a so-called quasi-steady state approximation. Id est, in the differential equation for the colonized health care workers, one can assume that the number of colonized patients is constant and that the number of colonized HCW will have its equilibrium value. The equilibrium value for the number of colonized HCW, given the number of colonized patients, can be calculated:

$$I_2^{qss} = \frac{\beta_2 N_2 I_1}{\alpha_2 N_1 + \beta_2 I_1} \tag{1.6}$$

If we use this approximation in the differential equation for the number of colonized patients we obtain:

$$\frac{dI_1}{dt} = \frac{\beta_1 \beta_2 N_2}{\alpha_2 N_1 + \beta_2 I_1}(N_1 - I_1)I_1 - \alpha I_1 \tag{1.7}$$

If $\frac{\beta_1 \beta_2 N_2}{\alpha_2 N_1 + \beta_2 I_1}$ is not fluctuating much, e.g., if $\alpha_2 \gg \beta_2 N_1$, we have a model in which the health care workers are not explicitly taken into account anymore, in the sense that the dynamics of the model would be the same if patients would transfer the bacterium directly to each other instead of via HCW's. Based on longitudinal data of patient colonization it is impossible to determine whether HCW were involved in bacterial transmission. This observation is used in Chapter 5 where we only try to distinguish between infection routes for which the corresponding acquisition rates have fundamentally different mathematical expressions.

For small units, the deterministic approach is no longer correct and a Markov chain approach can solve the problem of a finite number of patients. However, a direct translation of the differential equation (1.7) into a Markov model approach is not appropriate. Due to stochastic fluctuation, at some moment no colonized patients will be left in the unit and according to (1.7) all patients will remain uncolonized from that moment on. Therefore, we incorporate that a fraction of the admitted patients is colonized. Let $p_i(t)$ be the probability of having i colonized patients in the unit at time t. The parameter a denotes the rate at which an uncolonized patient is discharged and replaced by a patient who is colonized on admission. The parameter c denotes the rate at which a colonized patient loses colonization or is replaced by an uncolonized patient see [Pelupessy et al., 2002]. The parameter β is the

transmission parameter. The master equation will be:

$$\begin{aligned}
\tfrac{d}{dt}p_0(t) &= -aNp_0(t) + cp_1(t) \\
\tfrac{d}{dt}p_i(t) &= \Big(a(N-i+1) + \tfrac{\beta}{N}(i-1)(N-i+1)\Big)\, p_{i-1}(t) \\
&\quad - \Big(a(N-i) + \tfrac{\beta}{N}i(N-i) + ci\Big)\, p_i(t) + c(i+1)p_{i+1}(t) \\
\tfrac{d}{dt}p_N(t) &= \Big(a(N-1) + \tfrac{\beta}{N}(N-1)\Big)\, p_{N-1}(t) - cNp_N(t).
\end{aligned} \tag{1.8}$$

This thesis

In Part I we discuss models for the spread of nosocomial antibiotic-resistant bacteria with admission of colonized individuals from the extramural population. In Chapter 2 we will discuss an analytical model and in Chapter 3 we will focus on colonization with Methicillin-resistant *Staphylococcus aureus* (MRSA). We use both an analytical and a simulation model. Both models suggest that isolation of identified carriers of MRSA in hospitals combined with either screening on admission of high-risk patient or the screening of contact patients in case of the identification of an unexpected MRSA carrier in the hospital, may be sufficient to prevent high levels of MRSA in the hospitals. However, the so-called Dutch search and destroy policy in which both interventions are applied ensures that the current low prevalence level of MRSA in the Netherlands is far less sensitive to changes in the parameter values.

In Part II we use real hospital data to draw conclusions for specific pathogens/diseases. In Chapter 4 we use a simple observation to disentangle the phenomena that patients who acquire an infection are likely to stay longer in a unit and that patients who stay longer in a unit are more likely to acquire an infection. In Chapter 5 we use likelihood methods in a Markov chain approach to distinguish between different infection routes on the basis of the fluctuations in the prevalence. This method is applied to data for colonization with two different pathogens. This method is also used to determine optimal culture frequencies.

Open problems, future work, outlook

There are still many questions to be answered. What is the typical site of colonization or is it important to take colonization at different sites

into account? This requires within-host dynamics, as some compartment of the body can acquire colonization from another compartment. This is important when bacteria are harmless in their normal compartment (e.g., in the gut) but can cause infections when other compartments are reached (e.g., in case of ventilator-associated pneumonia, the lungs). Also for the modeling of SDD the compartmental structure may be essential. Another point requiring more attention is the influence of the quantity of potentially pathogenic bacteria on the infectivity of a patient. A related question is how the initial dose of the pathogen influences the likelihood of acquiring colonization or infection.

Another important point is the length of the infectious period. Do colonized patients become infectious directly after acquisition of colonization, or does a latency period exists? Is colonization persistent till discharge or can colonization be transient and how is the infectious period influenced by the use of antimicrobial agents in case of infection? Does it eradicate colonization or can colonization persist, although at a lower level than before treatment?

Another point is a more detailed distinction between infection routes. Is contamination of the environment relevant? What is the role of visitors? Can the pathogen persist in the air for a long period of time or is airborn transmission not important at all? Although mathematics alone is certainly not able to answer these questions, it may help to better understand the dynamics of these acquisition routes and collaboration between physicians, epidemiologists, statisticians and mathematicians can be fruitful.

Part I

Modeling

Chapter 2

Hospitals as the driving force for antibiotic resistance

2.1 Introduction

Antibiotic resistance has emerged as an important health care problem in the last decades. Difficulties in patient treatment are becoming more and more apparent, both for infections within hospital settings as in the community at large. In several studies the dynamics of antibiotic-resistant microorganisms were analyzed theoretically. Some studies specifically addressed the effects of antibiotic consumption [Bonhoeffer et al., 1997; Austin et al., 1997; Austin et al., 1999a], whereas other analyses were restricted to dynamics within hospital settings [Lipsitch et al., 2000; Austin et al., 1999b]. Only recently has a start been made with the theoretical investigation of the interaction between within-hospital dynamics and community dynamics [Cooper et al., 2004]. In the rest of this article we focus on species for which the spontaneous development of resistance is insignificant as compared to transmission. For vancomycin-resistant *Enterococcus faecium* (VRE) or Methicillin-resistant *Staphyloccus aureus* (MRSA) this assumption holds, but for many other antibiotic-resistant microorganisms, e.g., for *Mycobacterium tuberculosis*, spontaneous development of resistance, due to patients who do not complete their medication regime, cannot be neglected.

The characteristics of the hospital setting and the community set-

ting differ markedly. The hospital setting only covers a small proportion of the total population, with rapid turnover of patients, that are exposed to high levels of antibiotics and multiple potential sources of resistant bacteria. As hospitalized patients are usually ill, they have a higher risk to develop infections with antibiotic-resistant bacteria. In contrast, within the community, antibiotic use per individual is much lower and for many microorganisms individual-to-individual transmission will hardly ever occur, e.g, for MRSA. Nevertheless, individuals carrying antibiotic-resistant bacteria will be discharged from the hospital and some will, at some later time, be re-admitted to a hospital, potentially re-introducing antibiotic-resistant microorganisms.

Infection control measures, such as identification of carriers of antibiotic-resistant bacteria and preventing spread of these microorganisms to other patients, are used to limit the rise of antibiotic-resistance. Although an infection with an antibiotic-resistant microorganism is the most relevant entity, the epidemiological dynamics of antibiotic-resistant microorganisms are primarily determined by colonized patients. The adjective 'colonized' indicates that an individual is a carrier of the microorganism, without necessarily suffering from infection. Infections only represent the tip of the iceberg, as only a fraction of colonized patients will develop an apparent infection. Moreover, among colonized patients one can distinguish between those who can spread the microorganism to their surroundings (most importantly to other patients), and those who cannot and so are not infectious. Progression from the just colonized state to the spreader state can be facilitated by antibiotic use, creating a strong selection for the resistant microorganisms, leading to overgrowth and potential for transmission. Unfortunately, there are no diagnostic tools to distinguish between the colonized and the spreader status, and, therefore, the distinction is mainly theoretical.

The aim of the present study is to evaluate how the community reservoir and the hospital reservoir of antibiotic-resistant microorganisms influence each other, and what can be expected from different infection prevention strategies. Specific questions that have been addressed are: What fraction of the population will eventually carry antibiotic-resistant microorganisms? How long will it take to reach a stable situation? What are the consequences of prevalence of antibiotic-resistant microorganisms in the community for the endemic prevalence of antibiotic-resistant microorganisms in the hospital? For this purpose we have used a deterministic model in which we compartmentalize

the population into six subpopulations and distinguish individuals on the basis of their colonization status. We also introduce some heterogeneity of the population and analyze the effects of spatial structure of hospitals and their associated communities.

The present paper is a prelude to data analysis, see chapter 3. It tries to develop tools for risk assessment, or, more precisely, for the assessment of the efficacy of various potential control measures to prevent the spread of antibiotic resistant pathogens at long time scales and large spatial scales. Admittedly this is a hesitant start, not a final result. We hope it stimulates others to improve the situation. We face a complex problem with far reaching implications of major importance and frightening prospects, so we think that a modest attempt at improving our arsenal of tools is worth the effort.

2.2 Basic Model

To investigate how an antibiotic resistant microorganism can spread in the population we consider the following model (see Table 2.6 for a list of symbols). We subdivide the population into six groups. First we make a distinction between individuals who are hospitalized, we label them with a T for (treated), and individuals in the community (non-hospitalized), labeled with an U (for untreated). We also distinguish individuals on the basis of their colonization status. We restrict ourselves to three classes. Different names are used for these categories in medical and mathematical literature. In this paper we will adopt the medical names and below we will mention the mathematical names between square brackets. Uncolonized individuals [susceptibles] (labeled with the subscript 0), people who are colonized [infected] with the microorganism of interest, but who are not able to spread the microorganism, not even when they are in the hospital (labeled with the subscript $\frac{1}{2}$) and people who harbour the microorganism in great quantities and so are spreaders [infectious] (labeled with the subscript 1) provided they are in hospital. Slightly abusing notation, we shall use the symbol $T_{\frac{1}{2}}$ both to denote a colonized individual in the hospital and, in differential equations to be formulated below, to denote the number of such individuals. Mutatis mutandis the same applies to similar symbols. We make the following assumptions (also see Figure 2.1):

- After introduction of the resistant strain into the population,

spontaneous development of resistance is insignificant compared to transmission.

- Newborn individuals do not carry the strain and enter the population in the community at rate λ.

- Individuals can only acquire colonization when they are a patient in the hospital, see e.g., [Salgado et al., 2003], by contacts, direct or indirect, with a contagious patient (T_1), leading to transition from T_0 to $T_{\frac{1}{2}}$. The probability per unit of time to acquire colonization is proportional to the number of T_1 individuals, with as constant of proportionality the transmission rate parameter β.

- Only hospitalized colonized individuals can become spreaders. We denote the rate of this transition by α. (The idea is that this transition is caused by antibiotic treatment, which improves the competitive success of the resistant strain by (partially) eliminating sensitive strains.)

- Transition from the state with great quantities of antibiotic resistant microorganisms to the colonized state and from the colonized state to the uncolonized state only occurs in the community (for instance due to a cost of resistance for antibiotic resistant microorganisms such that they are outcompeted by non-resistant microorganisms). We denote the rate at which the high-load is lost by χ (whose value will be high) and the rate of getting completely rid of the strain by ω. (We assume that type 1 individuals cannot turn into type 0 without first becoming type $1/2$.) We neglect decolonization in the hospital settings as resistant strains often have a selective advantage in hospitals but also because individuals stay in the hospital only a short period. If the decolonization rate is not extremely large, decolonization is most likely to occur in the community.

- Individuals are discharged from hospital at rate σ and are hospitalized at rate ν.

- The death rate, μ, is independent of the state of the individuals.

When the population is large, so that a deterministic model applies, these assumptions lead to the following system of differential equations:

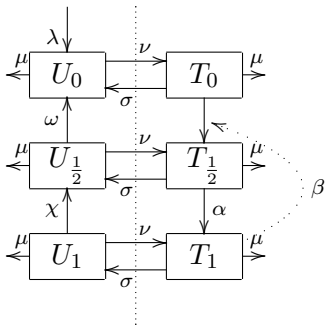


Figure 2.1: Basic model

$$\begin{array}{rcl}
\frac{d}{dt}U_0 = & \lambda - (\mu+\nu)\,U_0 + \omega U_{\frac{1}{2}} + \sigma T_0 \\
\frac{d}{dt}U_{\frac{1}{2}} = & -(\mu+\omega+\nu)\,U_{\frac{1}{2}} + \chi U_1 + \sigma T_{\frac{1}{2}} \\
\frac{d}{dt}U_1 = & -(\mu+\chi+\nu)\,U_1 + \sigma T_1 \\
\frac{d}{dt}T_0 = & -(\sigma+\mu)\,T_0 + \nu U_0 - \beta T_0 T_1 \\
\frac{d}{dt}T_{\frac{1}{2}} = & -(\sigma+\mu+\alpha)\,T_{\frac{1}{2}} + \nu U_{\frac{1}{2}} + \beta T_0 T_1 \\
\frac{d}{dt}T_1 = & -(\sigma+\mu)\,T_1 + \nu U_1 + \alpha T_{\frac{1}{2}}
\end{array} \tag{2.1}$$

According to these equations, the total population size (N) stabilizes at $\frac{\lambda}{\mu}$. The total number of individuals in the community U (= $U_0+U_{\frac{1}{2}}+U_1$) and the total number of patients in the hospital T (= $T_0+T_{\frac{1}{2}}+T_1$) satisfy:

$$\frac{d}{dt}U = \lambda - (\mu+\nu)U + \sigma T \tag{2.2}$$

$$\frac{d}{dt}T = \nu U - (\sigma+\mu)T \tag{2.3}$$

from which we derive the equilibrium:

$$U^* = \frac{\sigma+\mu}{\sigma+\mu+\nu}\frac{\lambda}{\mu} \tag{2.4}$$

$$T^* = \frac{\nu}{\sigma+\mu+\nu}\frac{\lambda}{\mu} \tag{2.5}$$

We shall henceforth assume that N, U and T have their equilibrium values from the beginning.

Instead of working with the size of the subpopulations in terms of absolute numbers, we now rescale them to fractions of the population

in and outside the hospital, i.e. $\widetilde{U}_i = \frac{U_i}{U^*}$ and $\widetilde{T}_i = \frac{T_i}{T^*}$ with $i \in \{0, \frac{1}{2}, 1\}$. For convenience we also rescale the time such that patients stay on average one unit of time in the hospital, i.e. $\tau = \sigma t$. We obtain:

$$\begin{array}{l} \frac{d}{d\tau}\widetilde{U}_{\frac{1}{2}} = -\frac{\mu+\omega+\nu}{\sigma}\widetilde{U}_{\frac{1}{2}} + \frac{\chi}{\sigma}\widetilde{U}_1 + \frac{\nu}{\sigma+\mu}\widetilde{T}_{\frac{1}{2}} \\ \frac{d}{d\tau}\widetilde{U}_1 = -\frac{\chi+\mu+\nu}{\sigma}\widetilde{U}_1 + \frac{\nu}{\sigma+\mu}\widetilde{T}_1 \\ \frac{d}{d\tau}\widetilde{T}_{\frac{1}{2}} = -\frac{\sigma+\mu+\alpha}{\sigma}\widetilde{T}_{\frac{1}{2}} + \frac{\sigma+\mu}{\sigma}\widetilde{U}_{\frac{1}{2}} + \frac{\beta\nu}{\sigma(\sigma+\mu+\nu)}(1-\widetilde{T}_{\frac{1}{2}}-\widetilde{T}_1)\widetilde{T}_1 \\ \frac{d}{d\tau}\widetilde{T}_1 = -\frac{\sigma+\mu}{\sigma}\widetilde{T}_1 + \frac{\sigma+\mu}{\sigma}\widetilde{U}_1 + \frac{\alpha}{\sigma}\widetilde{T}_{\frac{1}{2}} \end{array} \tag{2.6}$$

2.2.1 Determination of R_0

R_0 is defined as the expected number of individuals who acquire colonization from a spreader introduced in a 'virgin' hospital (and with the rest of the world being 'virgin' as well), and become spreaders themselves. To calculate R_0, we do the following: Suppose a spreader enters the hospital, while all other individuals are uncolonized. R is defined as the expected number of colonized patients due to contact with this spreader. P is defined as the probability that a hospitalized colonized patient will ever become a spreader. A hospitalized spreader remains hospitalized for an average period of $1/(\mu+\sigma)$. On average, such a patient spreads the antibiotic resistant microorganism to βT_0 patients per unit of time. Therefore, per admission a spreader will spread the antibiotic resistant microorganism to $\beta T_0/(\mu+\sigma)$ patients. With probability $\frac{\sigma}{\sigma+\mu}$, the spreader will not die during hospitalization. When this is the case, there are two ways in which this individual can become a spreader again (see Figure 2.1). It may return to the hospital without losing its high-load status, i.e. label 1, this occurs with probability $\frac{\nu}{\mu+\chi+\nu}$, or it may first turn into a low-load individual (label $\frac{1}{2}$) (probability $\frac{\chi}{\mu+\chi+\nu}$), return to the hospital (probability $\frac{\nu}{\mu+\omega+\nu}$) and regain the spreader status (label 1) due to antibiotic treatment (probability P).

Mathematically this is expressed by the formula:

$$R = \frac{\beta T_0}{\mu+\sigma} + \frac{\sigma}{\sigma+\mu}\left\{\frac{\nu R}{\mu+\chi+\nu} + \frac{\chi}{\mu+\chi+\nu}\frac{\nu}{\mu+\omega+\nu}PR\right\} \Rightarrow$$
$$R = \frac{\beta T_0\,(\mu+\chi+\nu)\,(\mu+\omega+\nu)}{(\mu+\sigma)\,(\mu+\chi+\nu)\,(\mu+\omega+\nu) - \sigma\nu\,(\mu+\omega+\nu) - \sigma\nu\chi P}$$

P can be determined from the observation that a hospitalized colonized patient can become a spreader during its current hospitalization (probability $\frac{\alpha}{\alpha+\mu+\sigma}$) or the patient can be discharged without acquiring the

spreader status (probability $\frac{\sigma}{\alpha+\mu+\sigma}$) but acquire the spreader status during a subsequent hospitalization. This results in the following equation for the probability P:

$$P = \frac{\alpha}{\alpha+\mu+\sigma} + \frac{\sigma}{\alpha+\mu+\sigma}\frac{\nu}{\mu+\omega+\nu}P \Rightarrow$$
$$P = \frac{\alpha(\mu+\omega+\nu)}{(\alpha+\mu+\sigma)(\mu+\omega+\nu)-\sigma\nu}$$

By definition of R and P we have that $R_0 = RP$. In a colonization free population we have that $T_0 = T^*$. Therefore:

$$R_0 = \frac{\alpha\beta\lambda\nu(\omega+\nu+\mu)(\mu+\chi+\nu)}{\mu(\sigma+\mu+\nu)\{[(\alpha+\mu+\sigma)(\omega+\nu+\mu)-\sigma\nu][(\mu+\sigma)(\mu+\chi+\nu)-\sigma\nu]-\sigma\nu\chi\alpha\}} \tag{2.7}$$

If $R_0 > 1$ an epidemic will occur with certainty because the model is deterministic, while if $R_0 < 1$ no large outbreak will occur. For later convenience we define

$$R_A = \lim_{\omega,\chi\to\infty} R_0(\omega,\chi) \tag{2.8}$$

as the basic reproduction ratio in the 'decoupled' situation that there are no carriers outside the hospital and accordingly no colonized patients are admitted to the hospital. This is the reproduction ratio per admission to the hospital.

2.2.2 Determination of the steady state

To determine the behaviour of the solutions of the system of differential equations (2.6), we first would like to know the steady states of (2.6). If we put all right hand sides of (2.6) equal to 0, we find two solutions:

- the trivial solution in which the infective agent is absent:

$$\widetilde{T}_0 = 1 \qquad \widetilde{T}_0 = 1 \qquad \widetilde{T}_{\frac{1}{2}} = \widetilde{U}_{\frac{1}{2}} = \widetilde{T}_1 = \widetilde{U}_1 = 0 \tag{2.9}$$

- the non-trivial steady state:

$$\begin{aligned}
\widetilde{T}_1 &= +\frac{\alpha\beta\nu(\chi+\mu+\nu)(\omega+\mu+\nu)}{\beta\nu((\alpha+\mu+\sigma)(\mu+\nu+\chi)-\nu\sigma)} + \\
&\frac{-(\sigma+\mu+\alpha)(\mu+\nu+\sigma)((\sigma+\mu)(\mu+\chi+\nu)-\nu\sigma)(\omega+\mu+\nu)+\nu\sigma(\mu+\nu+\sigma)(\alpha\chi-\nu\sigma+\nu(\sigma+\mu)(\chi+\mu+\nu))}{\beta\nu((\alpha+\mu+\sigma)(\mu+\nu+\chi)-\nu\sigma)} \\
\widetilde{T}_{\frac{1}{2}} &= \frac{(\mu+\sigma)(\mu+\nu+\chi)-\nu\sigma}{\alpha(\mu+\nu+\chi)}\widetilde{T}_1 \\
\widetilde{U}_1 &= \frac{\nu\sigma}{(\mu+\sigma)(\mu+\nu+\chi)}\widetilde{T}_1 \\
\widetilde{U}_{\frac{1}{2}} &= \frac{\nu\sigma(\alpha\chi-\nu\sigma+(\sigma+\mu)(\mu+\nu+\chi))}{\alpha(\mu+\sigma)(\mu+\nu+\omega)(\mu+\nu+\chi)}\widetilde{T}_1
\end{aligned} \tag{2.10}$$

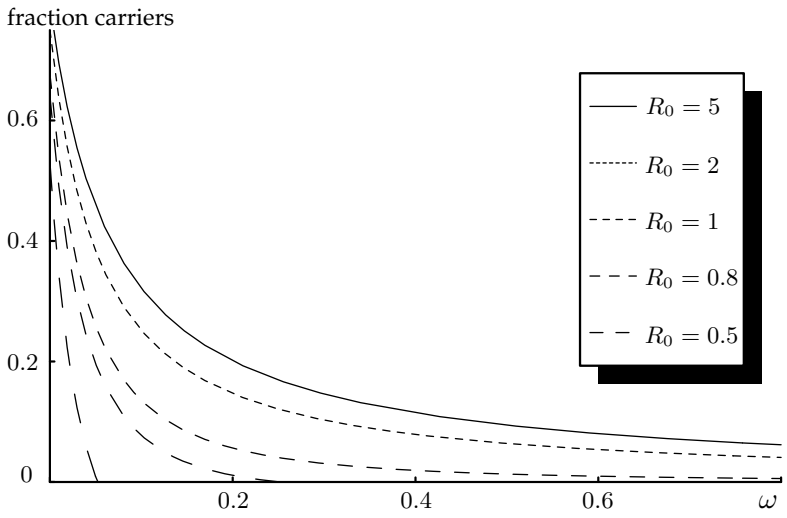


Figure 2.2: The equilibrium fraction of carriers in the populations function of ω for $\chi \to \infty$, $\alpha = 26, \nu = \frac{1}{15}, \mu = \frac{1}{70}, \sigma = 26$ and five different values of β such that, up from below, $R_A = 0.5, 0.8, 1, 2, 5$. If $R_A \geq 1$, endemicity exists for all values of the decolonization rate ω. For $R_A < 1$, re-admission of carriers is essential for the pathogen to persist and the decolonization rate should not be too high for endemicity.

which is only relevant when the size of all subgroups is non-negative. This is the case if and only if $R_0 \geq 1$. Using Routh-Hurwitz criteria, we find that for $R_0 < 1$ the only equilibrium, the trivial one, is stable, while for $R_0 > 1$ the trivial equilibrium is unstable but the non-trivial equilibrium is stable[1]. See Figure 2.2 for a graphical and quantitative illustration of (2.10)

2.2.3 Time scales

It is not possible to solve the differential equations (2.6) analytically. Therefore we try to approximate the solution by exploiting differences in the typical time scale of the various processes. Due to the rescaling of the time in deriving (2.6), the time scale for the length of stay in the hospital (and therefore also for the length of the infectious period), is $\mathcal{O}(1)$. As individuals generally spend only a short period of their

[1] A Mathematica notebook with the proof is available from the authors

life in a hospital, we assume that the time scale for the life time and the periods between hospitalizations are of order $\mathcal{O}(\frac{1}{\epsilon})$ with ϵ a small parameter. As a consequence, the fraction of the individuals in the hospital $\frac{T^*}{T^*+U^*} = \mathcal{O}(\epsilon)$. As $R_0 \sim \beta T^* = \beta(U^*+T^*)\frac{T^*}{U^*+T^*}$ we have to assume that $\beta(U^*+T^*) = \mathcal{O}(\frac{1}{\epsilon})$. (Otherwise R_0 would be $\mathcal{O}(\epsilon) < 1$ which would mean that the antibiotic resistant microorganism cannot persist. See section 2.3.3 for a more complicated situation in which this is not valid.) We also assume that the typical duration of colonization outside the hospital is long compared with the duration of a stay in a hospital. These assumptions can also be written as:

$$\omega = \epsilon\hat{\omega} \quad \mu = \epsilon\hat{\mu} \quad \nu = \epsilon\hat{\nu} \quad \beta = \frac{\hat{\beta}}{\epsilon} \tag{2.11}$$

where symbols with a hat all have the same order of magnitude.

This leads to the following system of differential equations:

$$\begin{aligned}
\tfrac{d}{d\tau}\widetilde{U}_{\frac{1}{2}} &= \tfrac{\chi}{\sigma}\widetilde{U}_1 + \epsilon\left(-\tfrac{\hat{\mu}+\hat{\omega}+\hat{\nu}}{\sigma}\widetilde{U}_{\frac{1}{2}} + \tfrac{\hat{\nu}}{\sigma}\widetilde{T}_{\frac{1}{2}}\right) + \mathcal{O}(\epsilon^2)\\
\tfrac{d}{d\tau}\widetilde{T}_{\frac{1}{2}} &= -\tfrac{\sigma+\alpha}{\sigma}\widetilde{T}_{\frac{1}{2}} + \widetilde{U}_{\frac{1}{2}} + \tfrac{\hat{\beta}\hat{\nu}}{\sigma^2}(1-\widetilde{T}_{\frac{1}{2}}-\widetilde{T}_1)\widetilde{T}_1\\
&\quad + \epsilon\left(\tfrac{\hat{\mu}}{\sigma}\widetilde{U}_{\frac{1}{2}} - \tfrac{\hat{\mu}}{\sigma}\widetilde{T}_{\frac{1}{2}} - \tfrac{\hat{\beta}\hat{\nu}}{\sigma^3}(\hat{\nu}+\hat{\mu})(1-\widetilde{T}_{\frac{1}{2}}-\widetilde{T}_1)\widetilde{T}_1\right) + \mathcal{O}(\epsilon^2)\\
\tfrac{d}{dt}\widetilde{U}_1 &= -\tfrac{\chi}{\sigma}\widetilde{U}_1 + \epsilon\left(-\tfrac{\hat{\nu}+\hat{\mu}}{\sigma}\widetilde{U}_1 + \tfrac{\hat{\nu}}{\sigma}\widetilde{T}_1\right) + \mathcal{O}(\epsilon^2)\\
\tfrac{d}{d\tau}\widetilde{T}_1 &= -\widetilde{T}_1 + \widetilde{U}_1 + \tfrac{\alpha}{\sigma}\widetilde{T}_{\frac{1}{2}} - \epsilon\left(\tfrac{\hat{\mu}}{\sigma}\widetilde{T}_1 - \tfrac{\hat{\mu}}{\sigma}\widetilde{U}_1\right)
\end{aligned} \tag{2.12}$$

We now calculate the zero order approximation of the equilibrium, i.e., we look for the solution of:

$$\begin{aligned}
&\tfrac{\chi}{\sigma}\widetilde{U}_1 = 0\\
&-\tfrac{\sigma+\alpha}{\sigma}\widetilde{T}_{\frac{1}{2}} + \widetilde{U}_{\frac{1}{2}} + \tfrac{\hat{\beta}\hat{\nu}}{\sigma^2}(1-\widetilde{T}_{\frac{1}{2}}-\widetilde{T}_1)\widetilde{T}_1 = 0\\
&-\widetilde{T}_1 + \widetilde{U}_1 + \tfrac{\alpha}{\sigma}\widetilde{T}_{\frac{1}{2}} = 0
\end{aligned} \tag{2.13}$$

which is given by:

$$\begin{aligned}
&\widetilde{U}_1 = 0\\
&\widetilde{T}_1 = \tfrac{\alpha}{\sigma}\widetilde{T}_{\frac{1}{2}}\\
&\widetilde{T}_{\frac{1}{2}} = \tfrac{\sigma}{2}\left(-\tfrac{\sigma^2}{\alpha\hat{\beta}\hat{\nu}} + \tfrac{1}{\alpha+\sigma} + \sqrt{\left(-\tfrac{\sigma^2}{\alpha\hat{\beta}\hat{\nu}} + \tfrac{1}{\alpha+\sigma}\right)^2 + \tfrac{4\sigma^2}{(\alpha+\sigma)\hat{\beta}\hat{\nu}\alpha}\widetilde{U}_{\frac{1}{2}}}\right)
\end{aligned} \tag{2.14}$$

In the zero order approximation, $\widetilde{U}_1 = 0$ and we can conclude that both $\widetilde{U}_{\frac{1}{2}}$ and $\widetilde{U}_1$ change slowly. Up to first order in ϵ we have that

$$\widetilde{U}_1 = \epsilon\frac{\hat{\nu}}{\chi}\widetilde{T}_1 + \mathcal{O}(\epsilon^2) \tag{2.15}$$

Now substitute the equilibrium (2.14), a function of $\widetilde{U}_{\frac{1}{2}}$, and (2.15) into the differential equation for $U_{\frac{1}{2}}$. We obtain:

$$\begin{aligned}\frac{d}{dt}\widetilde{U}_{\frac{1}{2}} = \quad & -(\mu+\omega+\nu)\widetilde{U}_{\frac{1}{2}} - \frac{\sigma^2(\alpha+\sigma)}{2\alpha\beta} + \frac{\nu}{2} \\ & + \sqrt{\left(\frac{\nu}{2} - \frac{\sigma^2(\alpha+\sigma)}{2\alpha\beta}\right)^2 + \frac{\sigma^2(\alpha+\sigma)\nu}{\alpha\beta}\widetilde{U}_{\frac{1}{2}}}\end{aligned} \tag{2.16}$$

We have that $\frac{d}{dt}\widetilde{U}_{\frac{1}{2}} > 0$ as long as $0 \leq \widetilde{U}_{\frac{1}{2}} \leq \widetilde{U_{\text{eq}}}$, with $\widetilde{U_{\text{eq}}}$ the non-trivial equilibrium of (2.16). Therefore, after the introduction of the micoorganism into the population, the number of colonized individuals in the population will steadily increase towards its limiting value.

If $\widetilde{U}_{\frac{1}{2}}$ is small, for instance shortly after the introduction of the micoorganism, we can make a Taylor-approximation of equation (2.16). This results in:

$$\frac{d}{dt}\widetilde{U}_{\frac{1}{2}} = -\left(\mu+\omega+\nu - \frac{\sigma^2(\alpha+\sigma)\nu}{\alpha\beta\nu - \sigma(\alpha+\sigma)}\right)\widetilde{U}_{\frac{1}{2}} + \nu - \frac{\sigma^2(\alpha+\sigma)}{\alpha\beta}. \tag{2.17}$$

The solution of a differential equation of this form ($\frac{d}{dt}f = -c_1 f + c_2, f(0) = 0$) is:

$$U_{\frac{1}{2}}(t) = \frac{c_2}{c_1}\left(1 - e^{-c_1 t}\right). \tag{2.18}$$

It follows that the time scale of the initial rise of the prevalence in the extramural population is $\frac{1}{c_1}$, with

$$c_1 = \mu+\omega+\nu - \frac{\sigma^2(\alpha+\sigma)\nu}{\alpha\beta\nu - \sigma(\alpha+\sigma)} \tag{2.19}$$

2.2.4 Some numerical results

To compare the approximations with the exact solution we plot in Figure 2.3, for one choice of parameters and initial conditions, the numerical solution of (2.1), the numerical solution of the quasi-steady-state approximation and the Taylor approximation of the quasi-steady-state approximation. (The parameters choices were chosen for realism.) The Taylor approximation of the quasi steady state solution is very good if $U_{\frac{1}{2}}$ is not very large. The quasi steady state solution is almost identical to the exact solution, except for a shift in time, due to the longer period it takes before the equilibrium in the hospital is reached when we do

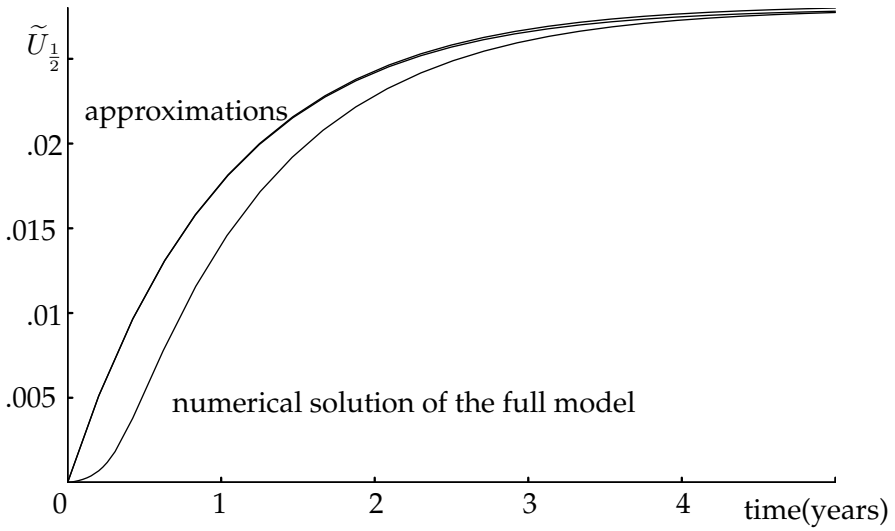
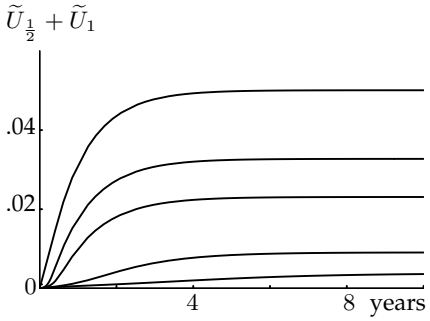


Figure 2.3: $U_{\frac{1}{2}}(t)$. For $\mu = \frac{1}{70}, \sigma = 26, \nu = \frac{1}{15}, \alpha = 26, \omega = 1, R_0 = 1.8$ the exact numerical solution, the numerical solution after the quasi steady state approximation and the solution of the equation based on the Taylor expansion of the quasi steady state are plotted (the last two almost overlap)

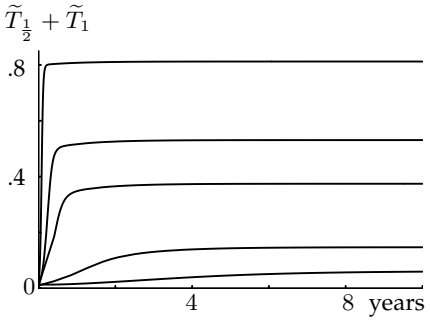
not make the quasi-steady-state approximation. The effects of different values of β and ω on the numerical solution of (2.1) are depicted in Figures 2.4 and 2.5 respectively for the level of colonized individuals in the hospital and in the community. If β is large, a high endemic equilibrium in the hospital is reached rather fast, while the presence of the strain in the community can slowly rise for a very long time, depending on ω. In contrast, when R_0 is only slightly larger than 1, admission of colonized patients becomes important. The shape of the curves is sigmoidal. In our model, all parameters are held constant which reflects constant antibiotic consumption in the sense that every individual has a constant probability per unit of time of receiving drugs (which stimulates the transition from $T_{\frac{1}{2}}$ to T_1).

2.2.5 Discussion and conclusions

It is very difficult to distinguish patients between being a spreader (U_1) or just a carrier $U_{\frac{1}{2}}$. Sometimes newly hospitalized patients are screened for the presence of the colonization [D'Agata et al., 2002a], but the distinction between patients being colonized or being a spreader

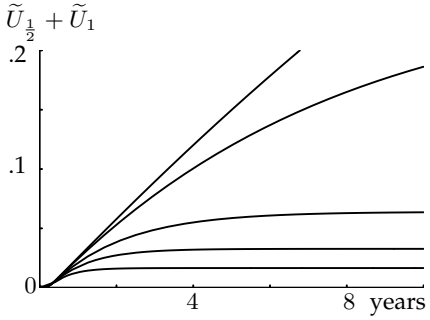


(a) $U_{\frac{1}{2}} + U_1$ as functions of the time in years since introduction of the pathogen for different R_A

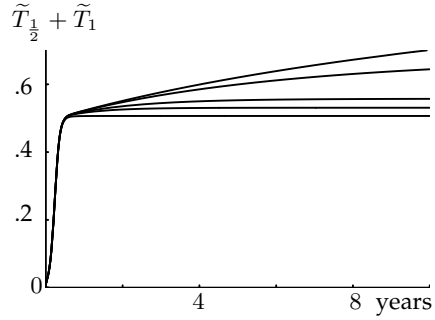


(b) $\widetilde{T}_{\frac{1}{2}} + \widetilde{T}_1$ as functions of the time in years since introduction of the pathogen for different R_A

Figure 2.4: For large R_A, a high endemic level is reached within hospitals in a short period of time. For values of R_A only slighter bigger than 1, readmission of carriers becomes important and within-hospital endemic levels are increasing over a longer period of time. $\chi \to \infty$, $\sigma = 26, \mu = \frac{1}{70}, \nu = \frac{1}{15}, \alpha = 26, \omega = 1, R_A = 1, 1.1, 1.5, 2, 5$ (up from below)



(a) $U_{\frac{1}{2}} + U_1$ as functions of the time in years since introduction of the pathogen for different decolonization rates



(b) $\widetilde{T}_{\frac{1}{2}} + \widetilde{T}_1$ as functions of the time in years since introduction of the pathogen for different decolonization rates

Figure 2.5: For a R_A significantly larger than 1, endemic levels within hospitals are not much influenced by the value of the decolonization rate. For the prevalence level in the community, the decolonization rate is very important. $\chi \to \infty$, $\sigma = 26, \mu = \frac{1}{70}, \nu = \frac{1}{15}, \alpha = 26, \omega = 0, 0.1, 0.5, 1, 5$ (from above), $R_A = 2$.

cannot be made. Furthermore we expect that people, who are discharged from the hospital, are rather healthy. Therefore it seems reasonable to assume that individuals are no longer spreaders when they leave the hospital. This holds true for enteral microorganisms (our assumption that transmission only takes place in the hospital is based on this). This also makes the category U_1 redundant. However, individuals may remain colonized after hospital discharge for prolonged periods of time [Byers et al., 2002].

If $R_0 > 1$ the fraction of infected individuals in the hospital will rise fast in the beginning, irrespective of the value of the decolonization rate (Figure 2.5) Also the final level of colonized patients in the hospital is not very sensitive to the value of the decolonization rate. Only if $R_0 \approx 1$ the precise value of the decolonization rate does influence the dynamics in the hospital substantially. For a relatively large decolonization rate, the fraction of carriers in the population is proportional to the level of infected patients in the hospital and inversely proportional to the decolonization rate. For small decolonization rates, the fraction of carriers in the population is mainly determined by the value of the decolonization rate provided $R_0 > 1$.

2.3 A Core Group

2.3.1 Basic assumptions and main conclusions

The model of the previous section describes the spread of a nosocomial microorganism in the population, but it ignores some factors that might have substantial impact. In this section we will investigate the influence of heterogeneity in the population (a core group) and in Section 2.5 we will investigate how the presence of more than one hospital can alter the results.

So far we assumed that the only relevant characteristics of individuals are their state with respect to colonization and infectiousness and whether they are in the hospital or not. However, some individuals will have a chronic disease and will visit the hospital more frequently. We will not concentrate on the dynamics within a designated unit [D'Agata et al., 2002b], but will look at the consequences for the dynamics in the population at large. Therefore, we use the extension of the basic model presented schematically in Figure 2.6 and represented mathematically by the differential equations (2.20). We now have two distinct groups

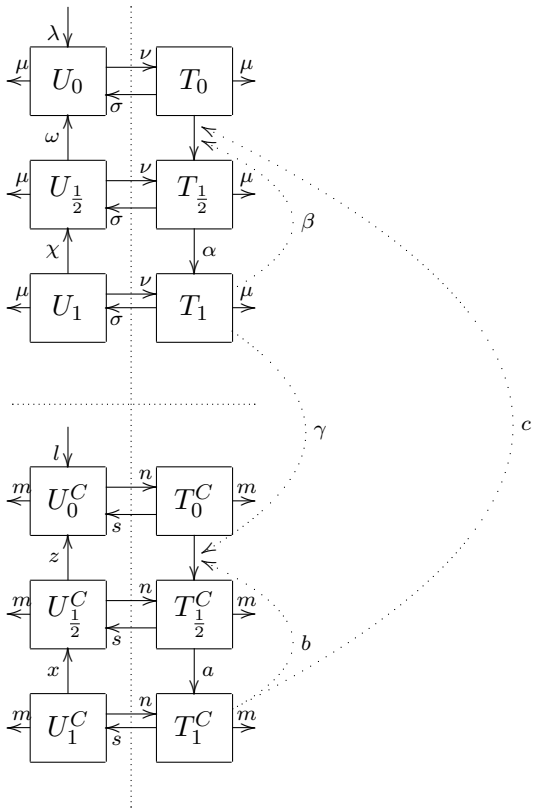


Figure 2.6: The full model with a core group

of individuals, members of a core group (C) that typically represent the chronically ill and frail segment of the population and 'normal' individuals (N). We assume that the fraction of core group patients in the population remains constant and that a core group patient remains in the core group the rest of his life, while a 'normal' individual will never belong to the core group. This last assumption is not very realistic, but it will hardly affect the results. Both groups obey the same dynamics, only with different values for the parameters. Specifically the parameter n is larger than ν, which corresponds to a more frequent admission to the hospital. Again we assume that transmission only takes place within the hospital. As patients in the core group may share medical equipment and may be cared for by a specialized group of health care workers, we assume that transmission of the microorganism happens

more likely between core group patients than from non-core group patients to other patients or vice versa. With the same kind of definitions and notation as before, (for instance the constant number of hospitalized core group patients is $T^{C*} = T_0^C + T_{\frac{1}{2}}^C + T_1^C$ and $\widetilde{T}_1^C = \frac{T_1^C}{T^{C*}}$), the analogue of (2.6) now becomes:

$$\begin{aligned}
\tfrac{d}{d\tau}\widetilde{U}_{\frac{1}{2}} &= -\tfrac{\mu+\omega+\nu}{\sigma}\widetilde{U}_{\frac{1}{2}} + \tfrac{\nu}{\sigma+\mu}\widetilde{T}_{\frac{1}{2}} + \tfrac{\chi}{\sigma}\widetilde{U}_1 \\
\tfrac{d}{d\tau}\widetilde{T}_{\frac{1}{2}} &= -\tfrac{\sigma+\mu+\alpha}{\sigma}\widetilde{T}_{\frac{1}{2}} + \tfrac{\mu+\sigma}{\sigma}\widetilde{U}_{\frac{1}{2}} + \tfrac{\beta}{\sigma}\tfrac{\nu}{\sigma+\mu+\nu}\widetilde{T}_0\widetilde{T}_1 + \tfrac{c}{\sigma}\tfrac{n}{s+m+n}\widetilde{T}_0\widetilde{T}_1^C \\
\tfrac{d}{d\tau}\widetilde{U}_1 &= -\tfrac{\mu+\chi+\nu}{\sigma}\widetilde{U}_1 + \tfrac{\nu}{\sigma+\mu}\widetilde{T}_1 \\
\tfrac{d}{d\tau}\widetilde{T}_1 &= -\tfrac{\sigma+\mu}{\sigma}\widetilde{T}_1 + \tfrac{\alpha}{\sigma}\widetilde{T}_{\frac{1}{2}} + \tfrac{\mu+\sigma}{\sigma}\widetilde{U}_1 \\
\tfrac{d}{d\tau}\widetilde{U}_{\frac{1}{2}}^C &= -\tfrac{m+z+n}{\sigma}\widetilde{U}_{\frac{1}{2}}^C + \tfrac{s}{\sigma}\tfrac{n}{s+m}\widetilde{T}_{\frac{1}{2}}^C + \tfrac{x}{\sigma}\widetilde{U}_1^C \\
\tfrac{d}{d\tau}\widetilde{T}_{\frac{1}{2}}^C &= -\tfrac{s+m+a}{\sigma}\widetilde{T}_{\frac{1}{2}}^C + \tfrac{m+s}{\sigma}\widetilde{U}_{\frac{1}{2}}^C + \tfrac{b}{\sigma}\tfrac{n}{n+m+s}\widetilde{T_0^C}\widetilde{T}_1^C + \tfrac{\gamma}{\sigma}\tfrac{\nu}{\nu+\mu+\sigma}\widetilde{T_0^C}\widetilde{T}_1 \\
\tfrac{d}{d\tau}\widetilde{U}_1^C &= -\tfrac{m+x+n}{\sigma}\widetilde{U}_1^C + \tfrac{s}{\sigma}\tfrac{n}{s+m}\widetilde{T}_1^C \\
\tfrac{d}{d\tau}\widetilde{T}_1^C &= -\tfrac{s+m}{\sigma}\widetilde{T}_1^C + \tfrac{a}{\sigma}\widetilde{T}_{\frac{1}{2}}^C + \tfrac{m+s}{\sigma}\widetilde{U}_1^C
\end{aligned} \tag{2.20}$$

We find that when β is large and the number of patients in T_1 is large, the influence of a core group will be small. The infection pressure within the hospital is already high and will only become slightly higher due to the core group. However, if $R_0 < 1$ for the basic model, the core group becomes important. They are a permanent source of resistant microorganisms and can maintain resistant microorganisms within the hospital setting that would disappear otherwise. When colonized patients remain colonized for a long period of time, the fraction of colonized patients in the population will increase, despite an R_0 smaller than 1 for the normal population when it would have been strictly separated from the core group population.

2.3.2 Computation of R_0

Suppose the microorganism is introduced in a virgin population. This can happen in two ways. Either a core group patient or a 'normal' patient can become a spreader. In both cases we can look at the expected number of individuals in both categories which become spreader due to the first spreader. By $R_0(N, C)$ we denote the expected number of core group patients who will become a spreader and who were infected by the initial 'normal' spreader and mutatis mutandis we define the

other three numbers. We obtain:

$$\begin{aligned}
R_0(N,N) &= \tfrac{\alpha\beta\lambda\nu(\omega+\nu+\mu)(\mu+\chi+\nu)}{\mu(\sigma+\mu+\nu)\{[(\alpha+\mu+\sigma)(\omega+\nu+\mu)-\sigma\nu][(\mu+\sigma)(\mu+\chi+\nu)-\sigma\nu\chi\alpha\}} \\
R_0(C,C) &= \tfrac{ablnz(z+m+n)(m+x+n)}{m(s+m+n)\{[(a+m+s)(z+m+n)-sn][(m+s)(m+x+n)-sn]-snax\}} \\
R_0(N,C) &= (a\gamma ln(m+z+n)(\mu+\omega+\nu)[(\alpha+\mu+\sigma)(\mu+\omega+\nu)-\sigma\nu]) / \\
&\quad (m(m+n+s)[(a+m+s)(m+z+n)-sn]) / \\
&\quad \{[(\alpha+\mu+\sigma)(\mu+\omega+\nu)-\sigma\nu][(\mu+\sigma)(\mu+\chi+\nu)-\sigma\nu(\mu+\omega+\nu)] \\
&\quad -\sigma\nu\chi\alpha(\mu+\omega+\nu)\} \\
R_0(C,N) &= (\alpha c\lambda\nu(\mu+\omega+\nu)(m+z+n)[(a+m+s)(m+z+n)-sn]) / \\
&\quad (\mu(\mu+\nu+\sigma)[(\alpha+\mu+\sigma)(\mu+\omega+\nu)-\sigma\nu]) / \\
&\quad \{[(a+m+s)(m+z+n)-sn][(m+s)(m+x+n) \\
&\quad -sn(m+z+n)]-snxa(m+z+n)\}
\end{aligned} \tag{2.21}$$

When all these reproduction rates between groups are strictly positive, the matrix

$$\begin{pmatrix} R_0(N,N) & R_0(C,N) \\ R_0(N,C) & R_0(C,C) \end{pmatrix}$$

is primitive and the largest eigenvalue is strictly dominant (see e.g., page 77 of [Diekmann and Heesterbeek, 2000]). The largest eigenvalue, which we call R_0, is given by:

$$\begin{aligned}
R_0 = \; & \tfrac{1}{2}\Big(R_0(N,N)+R_0(C,C) \\
& +\sqrt{(R_0(N,N)-R_0(C,C))^2+4R_0(C,N)R_0(N,C)}\Big)
\end{aligned} \tag{2.22}$$

2.3.3 Dynamics of the system

Two possible steady states of (2.20) exist, the trivial one (in which the resistant microorganism is not present and the non-trivial one (for which an awkward expression can be derived). To determine a quasi steady state approximation, we first exploit that members of the core group visit the hospital often, i.e. $n > \nu$ or $\nu = \epsilon\hat{\nu}$. We also assume that the decolonization rate ω satisfies $\omega = \epsilon\hat{\omega}$, which means that carriers in the community remain colonized for a substantial time. Finally we assume that the death rate is small and that transmission from T_1 to T_0^C is infrequent, i.e., $m = \epsilon\hat{m}$, $\mu = \epsilon\hat{\mu}$ and $c = \epsilon\hat{c}$.

There are two different cases. In the first one, we assume that $\beta = \frac{\hat{\beta}}{\epsilon}$, i.e., transmission is important in the normal population. In this case, up to zero order, the core group population and the normal population decouple and we re-obtain the results from Section 2.2. Intuitively, this means that there are already many spreaders in the hospital for

the normal population and the presence of some spreading core group members is of minor importance.

In the second case, we assume that transmission in the normal population is less important, i.e., $\beta = \hat{\beta}$. Up to zero order, we obtain for the equilibrium for $\widetilde{T}_1^C$:

$$\widetilde{T}_1^C = \frac{a(x+n)}{\sigma x + a(x+n)} + \frac{x(s+a)(s-\sigma)(n+s)}{bn(sx+a(x+n)} \tag{2.23}$$

which reduces to $\widetilde{T}_1^C = \frac{a(x+n)}{\sigma x + a(x+n)}$ if the discharge rate for core group and the non-core group is the same.

For the non-core group population, we then obtain the following system of differential equations (with slightly different coordinates)

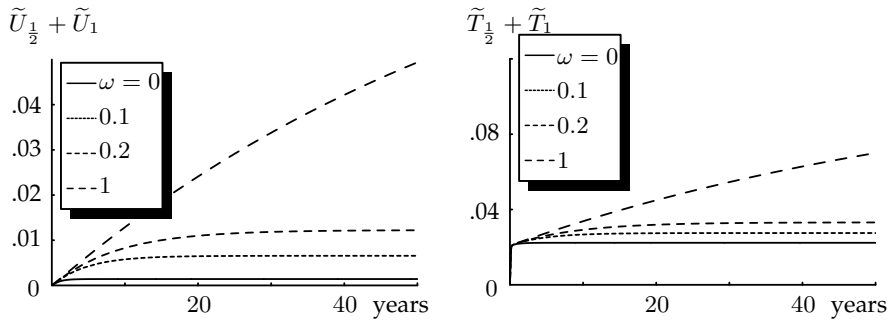
$$\begin{aligned}
\tfrac{d}{d\tau}\left(\widetilde{U}_{\frac{1}{2}} + \widetilde{U}_1\right) &= -\epsilon\tfrac{\hat{\mu}+\hat{\nu}+\hat{\omega}}{\sigma}\left(\widetilde{U}_{\frac{1}{2}} + \widetilde{U}_1\right) + \epsilon\tfrac{\hat{\nu}}{\sigma}\left(\widetilde{T}_{\frac{1}{2}} + \widetilde{T}_1\right) + \epsilon\tfrac{\hat{\omega}}{\sigma}\widetilde{U}_1 \\
\tfrac{d}{d\tau}\widetilde{U}_1 &= -\tfrac{\chi}{\sigma}\widetilde{U}_1 - \epsilon\tfrac{\hat{\mu}+\hat{\nu}}{\sigma}\widetilde{U}_1 + \epsilon\tfrac{\hat{\nu}}{\sigma}\widetilde{T}_1 + \mathcal{O}(\epsilon^2) \\
\tfrac{d}{d\tau}\left(\widetilde{T}_{\frac{1}{2}} + \widetilde{T}_1\right) &= -\left(\widetilde{T}_{\frac{1}{2}} + \widetilde{T}_1\right) + \left(\widetilde{U}_{\frac{1}{2}} + \widetilde{U}_1\right) - \epsilon\tfrac{\hat{\mu}}{\sigma}\left(\widetilde{T}_{\frac{1}{2}} + \widetilde{T}_1\right) \\
&\quad + \epsilon\tfrac{\hat{\mu}}{\sigma}\left(\widetilde{U}_{\frac{1}{2}} + \widetilde{U}_1\right) + \epsilon\tfrac{\beta\hat{\nu}}{\sigma^2}\widetilde{T}_1\left(1 - \left(\widetilde{T}_{\frac{1}{2}} + \widetilde{T}_1\right)\right) \\
&\quad + \epsilon\tfrac{\hat{c}n}{\sigma(s+n)}\widetilde{T}_1^C\left(1 - \left(\widetilde{T}_{\frac{1}{2}} + \widetilde{T}_1\right)\right) \\
\tfrac{d}{d\tau}\widetilde{T}_1 &= -\tfrac{\alpha+\sigma}{\sigma}\widetilde{T}_1 + \tfrac{\alpha}{\sigma}\left(\widetilde{T}_{\frac{1}{2}} + \widetilde{T}_1\right) - \epsilon\tfrac{\hat{\mu}}{\sigma}\widetilde{T}_1
\end{aligned} \tag{2.24}$$

These differential equations cannot be solved exactly. However, a lower-bound for the fraction of carriers can be obtained by putting $\beta = 0$, i.e., there is no transmission within the non-core group population. In this way, the differential equations become linear and can be solved exactly (but the expression is awkward).

If ω is large, the fraction of colonized patients in the population will be small. However, when ω is small, of the order of ν, the fraction of colonized patients can become high even if there is not enough transmission within the normal population to maintain the antibiotic resistant microorganism, see Figure 2.7.

2.4 Infection prevention strategies

We now return to the model without a core group. We want to investigate the influence of infection prevention measures of hospitals on the



(a) $U_{\frac{1}{2}} + U_1$ (b) $T_{\frac{1}{2}} + T_1$

Figure 2.7: The numerical solution of **(a)** $U_{\frac{1}{2}} + U_1$ and **(b)** $T_{\frac{1}{2}} + T_1$ as function of the time in years for several values of the decolonization rate $z = \omega$. We rescaled such that all values are fractions of the specific population (i.e. in or outside the hospital and member of the core group or not). For these graphs we used the following parameters: $R_A(N, N) = 0.5$, $R_A(F, F) = 1$, while transmission is 50 times less likely to occur from a core group patient to a normal patient than between normal patients if the prevalence in both groups would be the same. We assume there is no transmission from the non-core group to the core group. $\sigma = s = 26, \alpha = a = 52, \mu = m = \frac{1}{70}, \nu = \frac{1}{15}, n = 5, \chi = x = 26$ and $z = \omega = 1, 0.2, 0.1, 0$, where the lower the value of z and ω, the higher the prevalence in the hospital and the community.

dynamics of colonization. For simplicity we assume that within a hospital, patients are either non-colonized or spreaders, while in the community, individuals are either non-colonized or colonized, but never infectious. For notational reasons, we write $\widetilde{U}$ for the fraction of colonized individuals in the population and we write $\widetilde{T}$ for the fraction of infectious individuals in the hospital. In that case, we obtain the following two differential equations:

$$\begin{array}{l} \frac{d}{dt}\widetilde{U} = -(\mu + \omega + \nu)\widetilde{U} + \nu\frac{\sigma}{\sigma+\mu}\widetilde{T} \\ \frac{d}{dt}\widetilde{T} = -(\sigma + \mu)\widetilde{T} + (\sigma + \mu)\widetilde{U} + \frac{\beta\nu}{\sigma+\mu+\nu}(1 - \widetilde{T})\widetilde{T}. \end{array} \tag{2.25}$$

However, we now assume that every now and then, at rate ϵ, all colonization is eradicated from the hospital by a short but effective campaign, after which the prevalence will be rebuilding up slowly. The deterministic prediction of $\widetilde{U}$ and $\widetilde{T}$ is then impossible, so we switch to a stochastic description. We introduce the density function $f = f(U, T)$ such that the probability to find $(\widetilde{U}, \widetilde{T})$ to belong to a set ω is the integral

of f over ω.

We obtain the following partial differential equation for $f(\widetilde{U},\widetilde{T})$ as a function of time t:

$$\frac{\partial f}{\partial t}+\frac{\partial}{\partial \widetilde{U}}\left(\frac{d\widetilde{U}}{dt}f\right)+\frac{\partial}{\partial \widetilde{T}}\left(\frac{d\widetilde{T}}{dt}f\right)=-\epsilon f \tag{2.26}$$

with boundary condition:

$$\frac{d\widetilde{T}}{dt}(U,0)f(\widetilde{U},0)=\epsilon\int f(\widetilde{U},\widetilde{T})d\widetilde{T} \tag{2.27}$$

The equilibrium distribution should satisfy:

$$\frac{\partial}{\partial \widetilde{U}}\left(\frac{d\widetilde{U}}{dt}f\right)+\frac{\partial}{\partial \widetilde{T}}\left(\frac{d\widetilde{T}}{dt}f\right)=-\epsilon f \tag{2.28}$$

As before, we assume a quasi-steady state approximation applies, i.e. we assume that $\widetilde{U}$ is a constant. Hence we have that:

$$\tfrac{\partial f}{\partial t}+\tfrac{\partial}{\partial \widetilde{T}}\left(\tfrac{d\widetilde{T}}{dt}f\right)=-\epsilon f \tag{2.29}$$

This equation can be rewritten as:

$$\left(\epsilon+\frac{\partial}{\partial \widetilde{T}}\frac{d\widetilde{T}}{dt}\right)f=-\frac{d\widetilde{T}}{dt}\frac{\partial f}{\partial \widetilde{T}} \tag{2.30}$$

Hence we obtain by separation of variables,

$$\int\frac{df}{f}=-\int\frac{\frac{\partial}{\partial \widetilde{T}}\frac{d\widetilde{T}}{dt}+\epsilon}{\frac{d\widetilde{T}}{dt}}d\widetilde{T}\Rightarrow\log f(\widetilde{T})=-\log\frac{d\widetilde{T}}{dt}-\int\frac{\epsilon}{\frac{d\widetilde{T}}{dt}}d\widetilde{T}. \tag{2.31}$$

We rewrite $\frac{d\widetilde{T}}{dt}$ as:

$$\begin{array}{ll}\frac{d\widetilde{T}}{dt} & =-(\sigma+\mu)\widetilde{T}+(\sigma+\mu)\widetilde{U}+\frac{\beta\nu}{\sigma+\mu+\nu}(1-\widetilde{T})\widetilde{T}\\ & =-a\widetilde{T}^2+b\widetilde{T}+c=(a\widetilde{T}+\eta)(-\widetilde{T}+\xi)\end{array} \tag{2.32}$$

with $a=\frac{\beta\nu}{\sigma+\mu+\nu}$, $b=\frac{\beta\nu}{\sigma+\mu+\nu}-\sigma-\mu$, $c=(\sigma+\mu)\widetilde{U}$ and $-\frac{\eta}{a}=\frac{-\sqrt{b^2+4ac}+b}{2a}$ and $\xi=\frac{\sqrt{b^2+4ac}+b}{2a}$ the roots of $\frac{d\widetilde{T}}{dt}=0$.

We now write:

$$\begin{array}{c}\int \frac{\epsilon}{\frac{d\widetilde{T}}{dt}} d\widetilde{T} = \int \frac{\epsilon}{(a\widetilde{T}+\eta)(-\widetilde{T}+\xi)} d\widetilde{T} = \int \frac{\frac{\epsilon a}{a\xi+\eta}}{a\widetilde{T}+\eta} + \frac{\frac{\epsilon}{a\xi+\eta}}{-\widetilde{T}+\xi} d\widetilde{T} = \\ \frac{\epsilon a}{a\xi+\eta} \log\left(a\widetilde{T}+\eta\right) - \frac{\epsilon}{a\xi+\eta} \log\left(-\widetilde{T}+\xi\right) = \frac{\epsilon}{a\xi+\eta} \log\left(\frac{a\widetilde{T}+\eta}{-\widetilde{T}+\xi}\right).\end{array} \tag{2.33}$$

Now equation (2.31) can be written as:

$$\log f(\widetilde{T}) = -\log\left((a\widetilde{T}+\eta)(-\widetilde{T}+\xi)\right) - \frac{\epsilon}{a\xi+\eta} \log\left(\frac{a\widetilde{T}+\eta}{-\widetilde{T}+\xi}\right) + C \tag{2.34}$$

and we obtain:

$$f(\widetilde{T}) = \Omega\left(-\widetilde{T}+\xi\right)^{\frac{\epsilon}{a\xi+\eta}-1}\left(a\widetilde{T}+\eta\right)^{-1-\frac{\epsilon}{a\xi+\eta}} \tag{2.35}$$

with Ω a constant which we now determine from the condition that f is a density, i.e.

$$\int_0^{\xi} f(\widetilde{T})d\widetilde{T} = 1 \Rightarrow \Omega = \epsilon\left(\frac{\eta}{\xi}\right)^{\frac{\epsilon}{a\xi+\eta}} \tag{2.36}$$

Thus we deduce as formula for the density:

$$f(\widetilde{T}) = \epsilon\left(\frac{\eta}{\xi}\right)^{\frac{\epsilon}{a\xi+\eta}}\left(-\widetilde{T}+\xi\right)^{\frac{\epsilon}{a\xi+\eta}-1}\left(a\widetilde{T}+\eta\right)^{-1-\frac{\epsilon}{a\xi+\eta}}. \tag{2.37}$$

Note that $f(\widetilde{T})$ has a singularity at $\widetilde{T} = \xi$, i.e., at the steady state in the hospital. Typically, the density will look like the function in Figure 2.8, with almost all the mass of the density function being concentrated around $\widetilde{T} = 0$ and $\widetilde{T} = \xi$. (This coincides with the results of Section 2, where there is a rapid transition from low prevalence within the hospital to a high level prevalence.)

In a similar way, we can construct the density function $f(\widetilde{T})$ if the probability that a hospital takes infection prevention measures, such that the hospital becomes colonization free, depends on $\widetilde{T}$. If $\epsilon = \epsilon(\widetilde{T}) = k\widetilde{T}$, we have that:

$$f(\widetilde{T}) \propto \left(a\widetilde{T}+\eta\right)^{-1+\frac{k}{a}-k\frac{\xi}{a\xi+\eta}}\left(-\widetilde{T}+\xi\right)^{-1+k\frac{\xi}{a\xi+\eta}}. \tag{2.38}$$

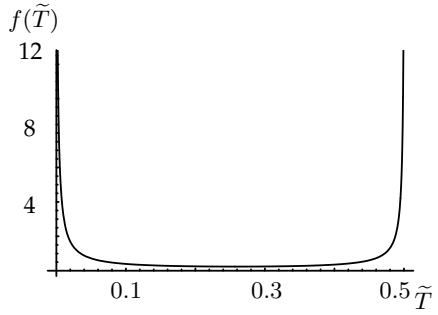


Figure 2.8: Typical example of the density $f(\widetilde{T})$

If $\epsilon = \epsilon(\widetilde{T}) = l\widetilde{T}^2$, we have that:

$$f(\widetilde{T}) \propto e^{\frac{l\widetilde{T}}{a}} \left(a\widetilde{T} + \eta\right)^{-1+l\frac{a\xi-\eta}{a^2}-l\frac{\xi^2}{a\xi+\eta}} \left(-\widetilde{T} + \xi\right)^{-1+l\frac{\xi^2}{a\xi+\eta}} . \qquad (2.39)$$

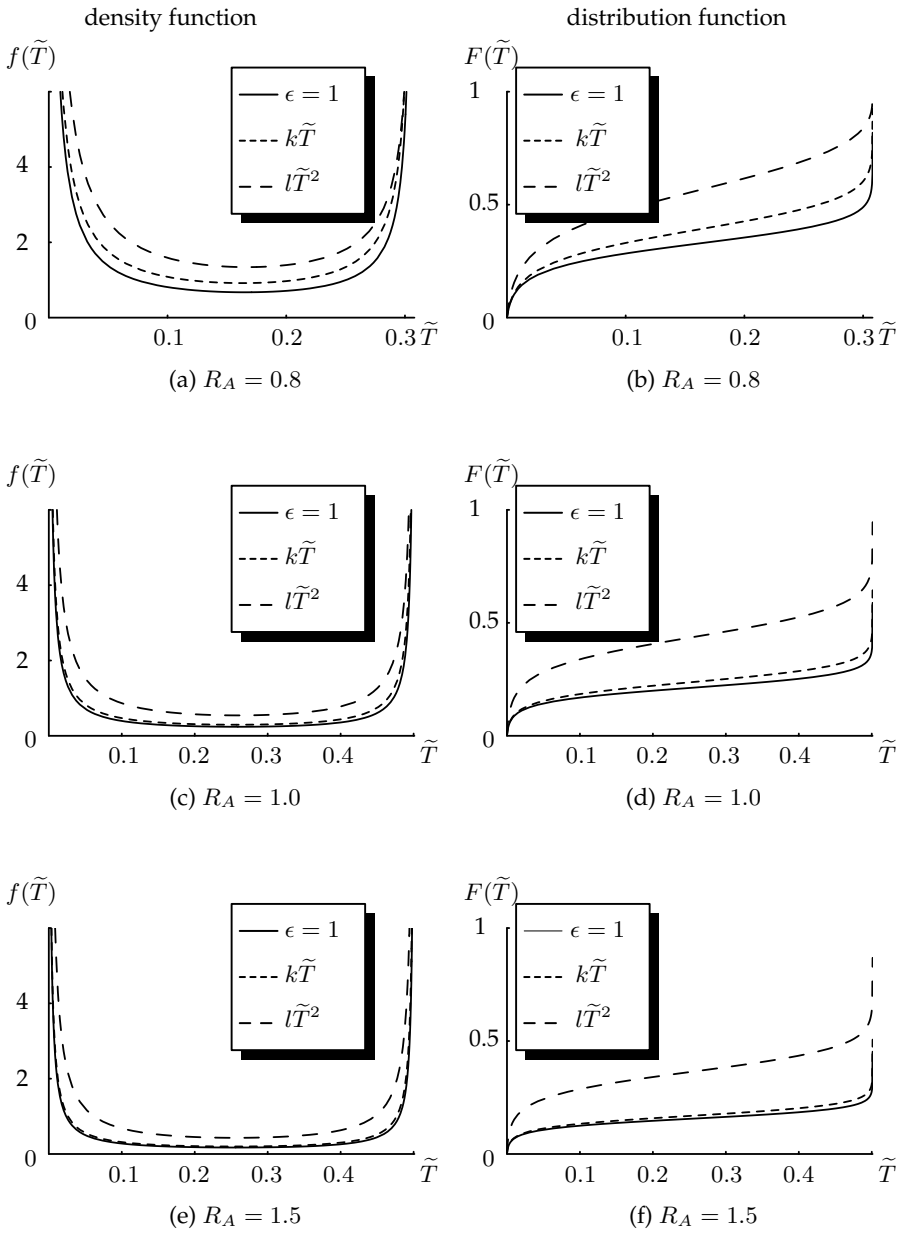
The constant of proportionality is such that $\int_0^\xi f(\widetilde{T})d\widetilde{T} = 1$. To compare the different functions for $\epsilon(\widetilde{T})$, we choose ϵ, k and l such that the total number of eradication episodes of hospitals is identical in the three situations, i.e. $\int_0^\xi f(\widetilde{T})\epsilon(\widetilde{T})d\widetilde{T}$ is the same for the three situations. As expected, if the total number of clearances remains constant, while relatively more clearances happen in hospitals with a high prevalence, as is the case when $\epsilon(\widetilde{T}) = l\widetilde{T}^2$, hospitals are less likely to have a high prevalence (see Figure (2.9)). However, the general picture, that the prevalence in a hospital is either low or high, does not change. We will use this observation in the next section.

2.5 Multi-center setting

In this section we will investigate the dynamics of colonization in a setting with many (smaller) instead of one (big) hospital. We restrict ourselves to the situation without direct contact between hospitals (i.e., we ignore transfer of a patient from one hospital to another).

2.5.1 Markov process

We will start this section with a description of a Markov process. On the basis of the previous section, it seems reasonable to assume that a hospital is either infected or infection-free. Therefore we assume the



(a) $R_A = 0.8$ (b) $R_A = 0.8$

(c) $R_A = 1.0$ (d) $R_A = 1.0$

(e) $R_A = 1.5$ (f) $R_A = 1.5$

Figure 2.9: Density function and distribution function for: $\epsilon = 1$, $\epsilon = k\widetilde{T}$, and $\epsilon = l\widetilde{T}^2$. k and l are chosen such that the expected number of eradication per year is 1. $\mu = 1/70$, $\nu = \frac{1}{15}$, $\sigma = 26$, $U = 0.001$. The distribution function is defined by: $F(\widetilde{T}) = \int_0^{\widetilde{T}} f(t)dt$.

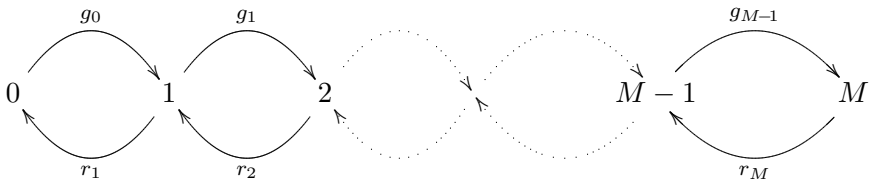


Figure 2.10: A Markov process

following: Let M denote the number of hospitals and p_i the probability that i hospitals are infected with the microorganism of interest.

We assume that the growth rate g_i, the rate at which the transition from i colonized to $i+1$ colonized hospitals occurs, is proportional to the number of colonization free hospitals and that the reduction rate r_i is proportional to the number of colonized hospitals (see Figure 2.10). In that case we obtain (with ϵ and δ constants (not related to the previous ϵ)):

$$\begin{gathered} r_i = \delta i \\ g_i = \epsilon(M-i) \\ \frac{d}{dt}p_i = r_{i+1}p_{i+1} + g_{i-1}p_{i-1} - (r_i + g_i)p_i \\ \sum_i p_i = 1 \end{gathered} \tag{2.40}$$

The steady state solution of these equations is given by the following binomial distribution [Kampen, 1981]: (with $x = \frac{\epsilon}{\delta}$)

$$p_i^s = \left(\begin{array}{c} M \\ i \end{array} \right) \left(\frac{x}{1+x} \right)^i \left(\frac{1}{1+x} \right)^{M-i} \tag{2.41}$$

The expectation and variance of this distribution are given by: (see also Figure 2.11)

$$E = M\frac{x}{1+x} \qquad Var = \frac{Mx}{(1+x)^2} \tag{2.42}$$

2.5.2 Dynamics outside the hospital

We now combine the model we described in Section 2.2 with the approach of the previous section. Therefore, we assume that the ϵ of Section 2.5.1 depends on the community reservoir of the microorganism.

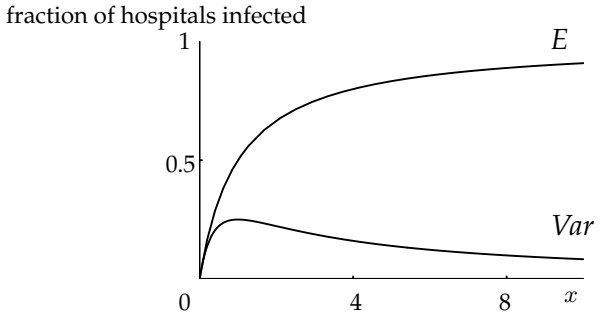


Figure 2.11: mean and variance of the binomial distribution as a function of $x = \frac{\epsilon}{\delta}$.

We would like to use the stationary distribution (2.41). The eigenvalues $\lambda_0, \lambda_1, \ldots, \lambda_M$ of system (2.40) are $0, -(\delta+\epsilon), -2(\delta+\epsilon), \ldots, -M(\delta+\epsilon)$. Therefore, the process will converge to the steady state solution (2.41). The typical time scale for this process is given by $-\frac{1}{\lambda_1} = \frac{1}{\delta+\epsilon}$. When hospitals are actively eradicating the microorganism, δ will be large and, therefore, we can use the stationary distribution.

First, we will investigate the situation when patient admission is not linked to any specific hospital and distribution over the hospitals occurs randomly.

In the Markov process model, ϵ is the rate at which a colonization free hospital becomes colonized. Therefore, ϵ is proportional to the number of admissions to the hospital, the probability that a newly hospitalized patient is colonized and the probability that a colonized patient initializes an outbreak in the hospital. We assume that $\epsilon = \hat{s}U_{\frac{1}{2}}$. The rate δ, at which a hospital eradicates colonization, is hard to determine. It depends on the efficacy of infection control measures, but also on the fraction of hospitalized patients which are colonized. For simplicity we assume that only the control mechanisms matter, i.e. δ does not depend on $U_{\frac{1}{2}}$. We have that $\frac{\epsilon}{\delta} = s\widetilde{U}_{\frac{1}{2}}$ with s a constant.

To construct a simultaneous model for hospital and community, we assume that in the colonized hospitals we can use the quasi steady state approximation (see (2.14)). This gives:

$$\frac{d}{d\tau}\widetilde{U}_{\frac{1}{2}} = -\frac{\mu+\omega+\nu}{\sigma}\widetilde{U}_{\frac{1}{2}} + \frac{\nu}{\sigma+\mu}\left(\widetilde{T}_{\frac{1}{2}}\left(\widetilde{U}_{\frac{1}{2}}\right) + \widetilde{T}_1\left(\widetilde{U}_{\frac{1}{2}}\right)\right)\frac{m}{M} \qquad (2.43)$$

with m the number of infected hospitals. The only difference with the

model with only one hospital, is the factor $\frac{m}{M}$ in the differential equation. If m is large, the results of both models will be similar. However, if m gets smaller, colonization of both patients and hospitals will slow down considerably (and the pathogen may even go extinct due to stochastic fluctuations).

For simplicity we next remove the stochastic influence, and we take for m the expected number of infected hospitals. $m = \frac{s\widetilde{U}_{\frac{1}{2}}}{1+s\widetilde{U}_{\frac{1}{2}}}M$, see equation (2.42). If we simplify the model even further by assuming that $\widetilde{U}_{\frac{1}{2}}$ is small, we can make a Taylor approximation of $\widetilde{T}_{\frac{1}{2}}+\widetilde{T}_1$ around $\widetilde{U}_{\frac{1}{2}} = 0$. Then we get as differential equation up to second order in $\widetilde{U}_{\frac{1}{2}}$:

$$\begin{aligned}\frac{d}{d\tau}\widetilde{U}_{\frac{1}{2}} = \quad & \left(-\tfrac{\mu+\nu+\omega}{\sigma} - \tfrac{(\alpha+\sigma)\sigma^2 s}{\alpha\beta(\sigma+\mu)} + \tfrac{\nu s}{\sigma+\mu}\right)\widetilde{U}_{\frac{1}{2}} \\ & + \left(\tfrac{(\alpha+\sigma)\sigma^2 s^2}{\alpha\beta(\sigma+\mu)} - \tfrac{\nu s^2}{\sigma+\mu} - \tfrac{\nu\sigma^2(\alpha+\sigma)s}{(\sigma+\mu)(\alpha\beta\nu-\sigma^2(\alpha+\sigma))}\right)\widetilde{U}_{\frac{1}{2}}^2\end{aligned} \tag{2.44}$$

The solution of a differential equation of the form $\frac{d}{dt}f = c_1 f + c_2 f^2$ is: $f(t) = \frac{Ae^{c_1 t}}{1-\frac{c_2}{c_1}Ae^{c_1 t}}$ with A a constant. When $c_1 > 0$ and $c_2 < 0$ (β large enough), this solution will behave similar to the function given in (2.18).

2.5.3 Spatial structure

In real life, people are usually admitted in a hospital in their own area. To investigate the influence of this phenomenon, we make the following assumptions:

- the community is divided in M sub-communities
- each sub-community has a hospital
- with probability q an individual is admitted, when the need arises, in its 'own' hospital.
- hospitalized patients that come from another community receive special care (for instance they are put into quarantine). Therefore, the probability that a colonized patient from another sub-community initializes an outbreak in hospital i is ξ_i $(0 \leq \xi_i \leq 1 \forall i)$ times the probability that a colonized patient from sub-community i initializes an outbreak in hospital i. Due to

the special care, patients may also be less likely to acquire colonization.) This is taken into account by introducing the factor η_i. The probability for a patient from a sub-community $j \neq i$ to acquire colonization during hospitalization in hospital i is η_i times the probability to acquire colonization for a patient from a sub-community i. Note that ξ_i and η_i only depend on the hospital that takes the intervention measures and not on the sub-community a non-local patient comes from.

- For simplicity we postulate that the probability that a hospitalized patient not receiving special care acquires colonization during the stay in a colonized hospital is y_i. We can justify this approximation by looking at the basic model in which the steady state in the hospital is not so strongly influenced by the fraction of colonized individuals in the community.

- We ignore that a hospital stay has a certain duration.

For the moment we assume that a hospitalized patient who does not go to the hospital in his own area, goes to some randomly chosen other hospital. With these assumptions we obtain the following equations:

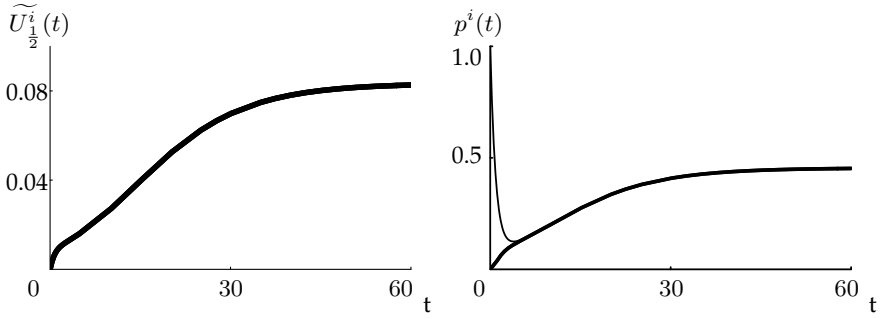
$$\frac{d}{dt}\widetilde{U}^i_{\frac{1}{2}} = -\omega\widetilde{U}^i_{\frac{1}{2}} + \nu y_i \left(qp^i + \frac{1-q}{M-1}\sum_{j\neq i}\eta_j p^j \right)\left(1-\widetilde{U}^i_{\frac{1}{2}}\right) \quad (2.45)$$

$$\frac{d}{dt}p^i = -\delta p^i + c_i \left(q\widetilde{U}^i_{\frac{1}{2}} + \xi_i \frac{1-q}{M-1}\sum_{j\neq i}\widetilde{U}^j_{\frac{1}{2}} \right)\left(1-p^i\right) \quad (2.46)$$

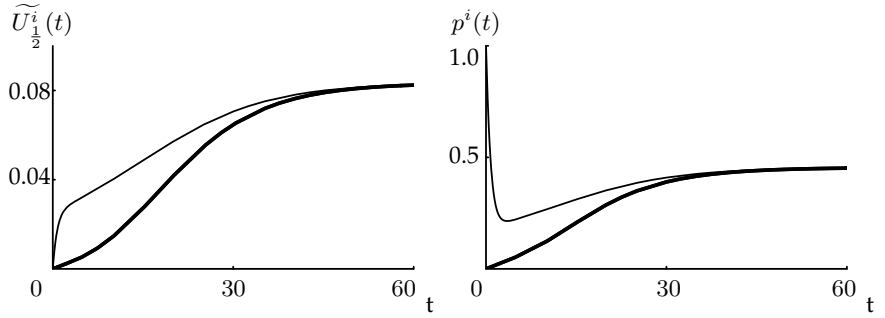
A disadvantage of this model is that people who have been in another hospital before, but are now in their own hospital, receive the same care as people who have only been in their own hospital. Therefore, this special care has only a limited effect.

The fact that there is no mortality in the model can be corrected by increasing the value of ω with μ. We choose the following parameter values: $\nu = \frac{1}{15}, y = 0.6, \omega = 0.2, g_i = 10 \forall i, \delta = 1$. We now solved the equations numerically for several combinations of the other parameters in the case of three regions, see Figures 2.12 and 2.13.

A hospital that implements special care to reduce nosocomial transmission, is able to keep the fraction of carriers in its extramural population at a lower level for some time, despite the higher prevalence rates

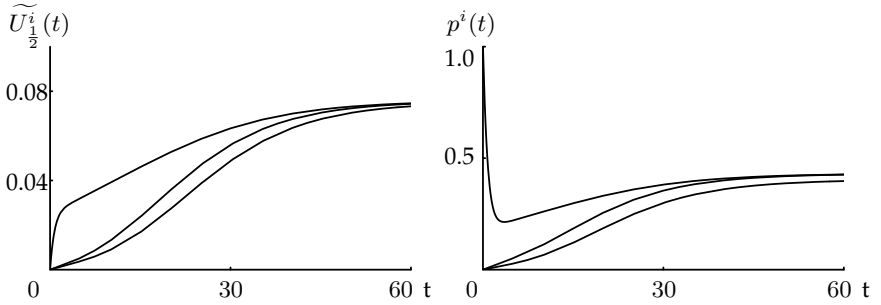


(a) No quarantine measures, no spatial structure: $\xi_i = 1$, $\eta_i = 1$, $q = \frac{1}{3}$.

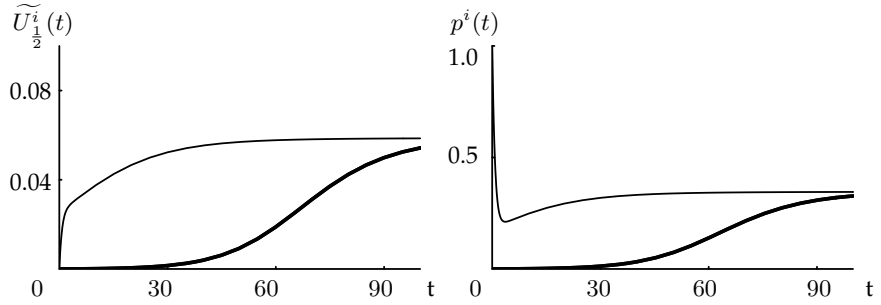


(b) No quarantine measures, 90% of the patients go to their own hospital. $\xi_i = 1$, $\eta_i = 1$, $q = 0.9$

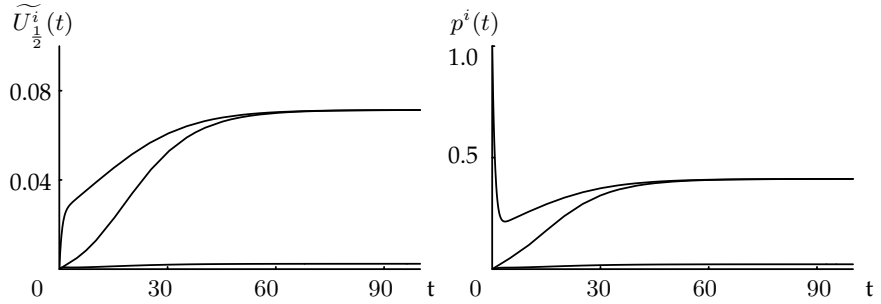
Figure 2.12: Changes over time in the fraction of carriers in sub-communities (left pane) and in the probability of a hospital to be infected (right pane). The population consist of three sub-communities with in each sub-community a hospital. $\nu = 1/15$, $y = 0.6$, $\omega = 0.2$, $g = 10$, $\delta = 1$, $p_1(0) = 1$, $p_2(0) = p_3(0) = 0$, $\widetilde{U^i_{\frac{1}{2}}}(0) = 0$.



(a) One hospital takes quarantine measures. $\xi_1 = \xi_2 = 1$, $\xi_3 = 0.01$, $\eta_1 = \eta_2 = 1$, $\eta_3 = 0.01$, $q = 0.9$.



(b) All hospitals have quarantine measures, $\xi_i = 0.01$, $\eta_i = 0.01$, $q = 0.9$.



(c) One hospital takes quarantine measures. $\xi_1 = \xi_2 = 1$, $\xi_3 = 0.01$, $\eta_1 = \eta_2 = 1$, $\eta_3 = 0.01$, $q = 0.9$, but the hospital with quarantine measures also has $\omega = 1$ as a consequence of a search and destroy policy.

Figure 2.13: Changes over time in the fraction of carriers in sub-communities (left pane) and in the probability of a hospital to be infected (right pane). The population consist of three sub-communities with in each sub-community a hospital. $\nu = 1/15$, $y = 0.6$, $\omega = 0.2$, $g = 10$, $\delta = 1$, $p_1(0) = 1$, $p_2(0) = p_3(0) = 0$, $\widetilde{U^i_{\frac{1}{2}}}(0) = 0$.

in the other sub-communities. However, in the long run, when a substantial fraction of the own community consists of carriers, the special care measures lose their effectiveness. The same applies for cohorting of patients to a hospital as captured by the parameter q. Although this will delay the spread of resistance to other communities substantially, in the long run, when a substantial fraction of the own community consists of carriers, cohorting is not effective anymore. Only measures such that R_0 becomes smaller than 1, i.e., an active search-and-destroy-policy, in the sense that colonization within known carriers is eradicated, or a reduction in the probability that the hospital will be infected when a colonized patient is admitted, can be effective in the long run.

So far we assumed that a 'foreign' hospital was chosen at random. This does not seem to reflect reality. We expect that in such a case a patient will go to a neighbouring hospital. For instance, define the following 1-dimensional model in which on each position corresponding to an integer number there is a hospital and there is a uniform distribution on the real axis of individuals. $U_{\frac{1}{2}}$ represents the fraction of individuals who carry the microorganism, U_0 is the fraction that doesn't carry the microorganism. S is the probability that a hospital is free of colonization, and P is the probability that colonization is endemic in the hospital. (We assume that there is no intermediate state.) Furthermore let $f(x-k)$ be the probability that when an individual at position x gets hospitalized, it is in the hospital at k and let $\eta(x)$ and $\xi(x)$ be the continuous analagon of η_i and ξ_i, i.e., when $\eta(0) = \xi(0) = 1$ the reduction in infectivity and susceptibility respectively when a patient is hospitalized in a hospital at distance x. Furthermore we assume that the likelihood of admission to a certain hospital only depends on the distance to that hospital, i.e., $f(a) = f(-a) \forall a \in \mathbb{R}$. In that case we obtain:

$$
\begin{aligned}
\tfrac{d}{dt}U_0(x,t) &= \mu - \mu U_0(x,t) + \omega U_{\frac{1}{2}}(x,t) - \nu y U_0(x,t) \sum_{l\in\mathbb{Z}} P(l,t) f(x-l)\eta(x-l) \\
\tfrac{d}{dt}U_{\frac{1}{2}}(x,t) &= -(\omega+\mu)U_{\frac{1}{2}}(x,t) + \nu y U_0(x,t) \sum_{l\in\mathbb{Z}} P(l,t) f(x-l)\eta(x-l) \\
\tfrac{d}{dt}S(k,t) &= \delta P(k,t) - gS(k,t) \int_{-\infty}^{\infty} U_{\frac{1}{2}}(k-s,t) f(s)\xi(s) ds \\
\tfrac{d}{dt}P(k,t) &= -\delta P(k,t) + gS(k,t) \int_{-\infty}^{\infty} U_{\frac{1}{2}}(k-s,t) f(s)\xi(s) ds
\end{aligned}
\tag{2.47}
$$

When we linearize around $U_{\frac{1}{2}}(x,t) = 0$ and $P(x,t) = 0$. We obtain:

$$\begin{aligned} \frac{d}{dt}U_{\frac{1}{2}}(x,t) &= -(\omega+\mu)U_{\frac{1}{2}}(x,t) + \nu y \sum_{l\in\mathbb{Z}} P(l,t)f(x-l)\eta(x-l) \\ \frac{d}{dt}P(k,t) &= -\delta P(k,t) + g \int_{-\infty}^{\infty} U_{\frac{1}{2}}(k-s,t)f(s)\xi(s)ds \end{aligned} \tag{2.48}$$

Integrating of these equations yields:

$$\begin{aligned} U_{\frac{1}{2}}(x,t) &= U_{\frac{1}{2}}(x,0)e^{-(\omega+\mu)t} + \nu y \int_0^t \sum_{l\in\mathbb{Z}} P(l,s)f(x-l)\eta(x-l)e^{(\omega+\mu)(s-t)}ds \\ P(k,t) &= P(k,0)e^{-\delta t} + g\int_0^t \int_{-\infty}^{\infty} U_{\frac{1}{2}}(x,t-r)f(x-k)\xi(x-k)e^{-\delta r}drdx \end{aligned} \tag{2.49}$$

We now substitute the equation for $U_{\frac{1}{2}}$ into the equation for P. This yields:

$$\begin{aligned} P(k,t) &= P(k,0)e^{-\delta t} + g\int_0^t dre^{-\delta r}\int_{-\infty}^{\infty} dxf(k-x)\xi(k-x)U_{\frac{1}{2}}(x,0)e^{-(\omega+\mu)(t-r)} \\ &+ g\nu y\int_0^t dre^{-\delta r}\int_{-\infty}^{\infty} dxf(k-x)\xi(k-x)\int_0^{t-r} dse^{(\omega+\mu)(s+r-t)}\sum_{l\in\mathbb{Z}} P(l,s)f(x-l)\eta(x-l) \end{aligned} \tag{2.50}$$

This can be written as:

$$P(k,t) = P_0(k,t) + g\nu y \int_0^t ds \sum_{l\in\mathbb{Z}} P(l,s)\gamma_1(k-l)\gamma_2(t-s) \tag{2.51}$$

where

$$P_0(k,t) = P(k,0)e^{-\delta t} + g\int_0^t dre^{-\delta r}\int_{-\infty}^{\infty} dxf(k-x)U(x,0)e^{-(\omega+\mu)(t-r)} \tag{2.52}$$

depends only on the initial conditions and

$$\begin{aligned} \gamma_1(k) &= \int_{-\infty}^{\infty} dxf(x+k)\xi(k+x)f(x)\eta(x) \\ \gamma_2(t) &= \int_0^t dre^{-(\omega+\mu)r}e^{-\delta(t-r)} = \tfrac{e^{-\delta t}-e^{-(\omega+\mu)t}}{\omega+\mu-\delta} \end{aligned} \tag{2.53}$$

To describe the evolution of the system from the beginning of time instead from $t = 0$, one has to solve the following differential equation: (see also [Diekmann, 1978], [Diekmann, 1979], [Rass and Radcliffe,

2003])

$$P(k,t) = \int_{-\infty}^{t} ds \sum_{l\in\mathbb{Z}} P(l,s)\gamma_1(k-l)\gamma_2(t-s) \tag{2.54}$$

If we look for a traveling wave solution $P(k,t) = w(k+ct)$, the solution w has to satisfy:

$$w(y) = \sum_{m\in\mathbb{Z}+y} \int_0^{\infty} d\tau w(m-c\tau)\gamma_1(y-m)\gamma_2(\tau) \tag{2.55}$$

An exponential solution $w(y) = e^{\lambda y}$ should satisfy

$$1 = \sum_{l\in\mathbb{Z}} \int_0^{\infty} \gamma_2(\tau)e^{\lambda(l-c\tau)}\gamma_1(l)d\tau \tag{2.56}$$

By using the definition of $\gamma_2(t)$ (2.53), this equation simplifies to:

$$(\delta+\lambda c)(\omega+\mu+\lambda c) = \sum_{l\in\mathbb{Z}} e^{\lambda l}\gamma_1(l) \tag{2.57}$$

The minimal value of c for which there is a solution is the wave speed.

We can also make a diffusion approximation of the system 2.47, where we assume that the fraction $1-q$ of the patients that it is not hospitalized in its own hospital go the next-neighbouring hospital. In one spatial dimension we obtain (with x now a continuous variable):

$$\begin{aligned}
&\tfrac{\partial}{\partial t}U(x,t) = -\omega U(x,t)+ \\
&\qquad \nu y\left(qp(x,t)+\eta(x)(1-q)\left[p(x,t)+\tfrac{1}{2}\tfrac{\partial^2}{\partial x^2}p(x,t)\right]\right)(1-U(x,t)) \\
&\tfrac{\partial}{\partial t}p(x,t) = -\delta p(x,t)+ \\
&\qquad g\left(qU(x,t)+\xi(x)(1-q)\left[U(x,t)+\tfrac{1}{2}\tfrac{\partial^2}{\partial x^2}U(x,t)\right]\right)(1-p(x,t))
\end{aligned} \tag{2.58}$$

This type of non-linear partial differential equation is difficult to handle, but again is possible to determine the propagation velocity of the colonizing microorganism. In an area where the microorganism appears for the first time, U and p are small, and, therefore, we can linearize the equations. If we assume that the speed of a traveling wave is found by looking for a travelling wave solution of the linearized equations we can find the minimal velocity of the wave. The linearized

equations (around $\widetilde{U}_{\frac{1}{2}} = p = 0$) are:

$$\begin{aligned}\frac{\partial}{\partial t}U(r,t) &= -\omega\widetilde{U(r,t)} + \nu p\left(qp(r,t) + \eta(r)(1-q)\left[p(r,t) + \tfrac{1}{2}\tfrac{\partial^2}{\partial r^2}p(r,t)\right]\right)\\ \frac{\partial}{\partial t}p(r,t) &= -\delta p(r,t) + g\left(qU(r,t) + \xi(r)(1-q)\left[U(r,t) + \tfrac{1}{2}\tfrac{\partial^2}{\partial r^2}U(r,t)\right]\right)\end{aligned} \tag{2.59}$$

If we now use the Ansatz of a traveling wave solution, $\begin{pmatrix} U \\ P \end{pmatrix} = e^{\lambda(x-\widetilde{c}t)}\begin{pmatrix} U_0 \\ P_0 \end{pmatrix}$ we obtain the following formula for the velocity $\widetilde{c}$:

$$\widetilde{c} = \tfrac{\omega+\delta}{2\lambda} - \frac{1}{2\lambda}\sqrt{(\omega-\delta)^2 + 4\nu y g\left\{q+\eta(1-q)+\tfrac{1}{2}\eta(1-q)\lambda^2\right\}\left\{q+\xi(1-q)+\tfrac{1}{2}\xi(1-q)\lambda^2\right\}} \tag{2.60}$$

As this formula is invariant under interchanging η and ξ, we conclude that they are equally important in reducing the spread between hospitals.

If $\widetilde{c} > 0$, we need that $\lambda < 0$ to consider a wave traveling to the right, so we look at the minimal $\widetilde{c}(\lambda)$ for which a solution with $\lambda < 0$ exists. In order to let $\widetilde{c}$ be positive we need that:

$$\omega + \delta - \sqrt{(\omega-\delta)^2 + 4\nu y g\left\{q+\eta(1-q)\right\}\left\{q+\xi(1-q)\right\}} < 0 \tag{2.61}$$

This follows from the following analysis. Write

$$\widetilde{c}(\lambda) = \frac{\alpha}{\lambda} - \frac{1}{\lambda}\sqrt{a + b\lambda^2 + d\lambda^4} \tag{2.62}$$

in which all parameters are positive (except λ). Taking the derivative with respect to λ, we find:

$$\frac{d\widetilde{c}}{d\lambda} = -\frac{\alpha}{\lambda^2} + \frac{1}{\lambda^2}\sqrt{a + b\lambda^2 + d\lambda^4} - \frac{b + 2d\lambda^2}{\sqrt{a + b\lambda^2 + d\lambda^4}} \tag{2.63}$$

Substitute $\lambda^2 = \mu$. Because we are only interested in negative λ, this substitution is one to one. If we put $\frac{d\widetilde{c}}{d\lambda} = 0$, we obtain the equation:

$$\frac{d^2}{\alpha^2}\mu^4 - \left(d + \frac{2ad}{\alpha^2}\right)\mu^2 - b\mu + \frac{a^2}{\alpha^2} - a = 0 \tag{2.64}$$

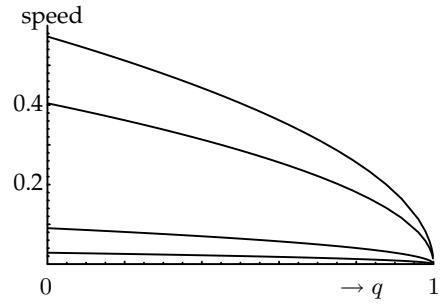
Condition (2.61) coincides with $\frac{a^2}{\alpha^2} - a < 0$. Therefore equation (2.64) is a polynomial of order ≥ 2, with a positive highest order coefficient and

for which all other coefficients are negative. Therefore there is a unique root of (2.64) with $\mu > 0$ which implies a unique minimal velocity of the wave. Equation (2.61) can be seen as a test whether the microorganism is able to spread to other hospitals. This becomes evident, when one realizes the following: When a colonized individual enters a fully susceptible population, colonization will remain for an average period of $\frac{1}{\omega}$. This individual visits a hospital on average $\frac{\nu}{\omega}$ times during this period. The probability that the hospital will become colonized during such a visit is given by g for the own hospital and $g\xi$ for a neighbouring hospital. If a hospital is colonized, it remains so for an average period of $\frac{1}{\delta}$. During this period, $y(q + (1 - q)\eta)$ new individuals become colonized. Therefore R_0 is given by:

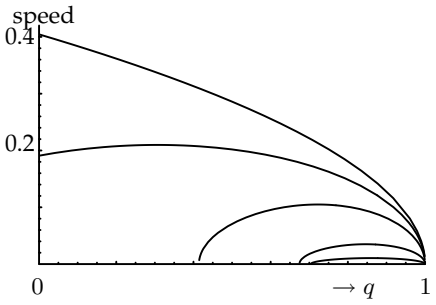
$$R_0 = \frac{\nu}{\omega\delta} g\{q + \xi(1 - q)\} y\{q + \eta(1 - q)\} \tag{2.65}$$

and inequality (2.61) is identical to the condition $R_0 > 1$. When each patient goes to its own hospital, $q = 1$, the infection cannot spread between hospitals. The speed of the traveling wave for other values of q is shown in Figure 2.14. The speed of the traveling wave as a function of the decolonization rate is shown in Figure 2.15. [t]

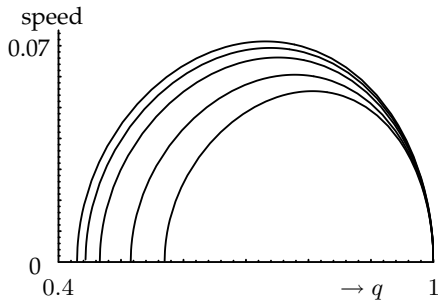
Another possible extension is the situation that when an individual visits its own hospital, (s)he gets the question whether (s)he has been in another hospital recently. In that case we have to introduce a group which answers "yes" to this question. (We denote this group with an asterix.) People leave this group at rate ϕ. We then get the following



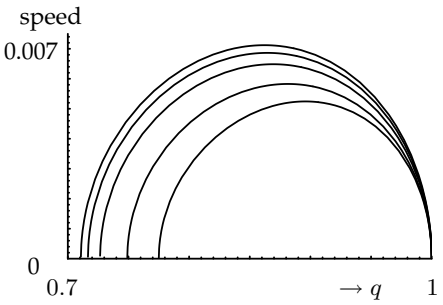
(a) basic model with $\omega = 0$, 0.2, 0.39 and 0.399 and $\xi = \eta = 1$



(b) basic model with $\omega = 0.2$ and $\xi = \eta = 0.01$, 0.1, 0.5, 0.8 and 1



(c) extended model with $\eta = \xi = 0.5$, $\omega = 0.2$ and $\phi = 0.1$, 0.2, 0.5, 1 and 2



(d) extended model with $\eta = \xi = 0.01$, $\omega = 0.2$ and $\phi = 0.1$, 0.2, 0.5, 1 and 2

Figure 2.14: Speed of the traveling wave as function of q, the fraction of hospitalized patients that go to their own hospital. $\nu = 1/15$, $g = 10$, $\delta = 1$,$y = 0.6$. (As to be expected, the curves are ordered according to the value of ω; a similar ordering occurs in the other pictures and in Figure 2.15).
Although the speed is plotted $\forall q \in [0, 1]$, we are mainly interested in large q. For small q only a minority of the hospital admissions will occur in the own hospital, which does not reflect reality. Moreover, as patients not admitted in the own hospital receive special treatment, a small q can prevent spread of the pathogen.

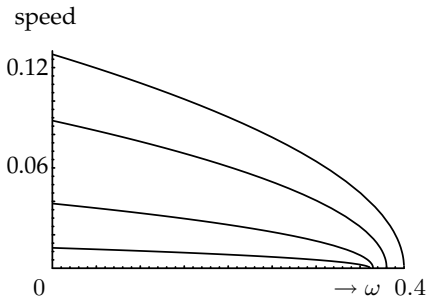


Figure 2.15: Wave velocity for the basic model as function of the decolonization rate ω. $\nu = 1/15, g = 20, \delta = 0.5, y = 0.6, q = 0.95$ and $\xi = \eta = 1, 0.5, 0.1$ and 0.01.

differential equations:

$$\begin{aligned}
\frac{d}{dt}U_0^i &= \phi U_0^{i*} + \omega U_{\frac{1}{2}}^i - q\nu y P^i U_0^i - (1-q)\nu U_0^i \\
\frac{d}{dt}U_{\frac{1}{2}}^i &= \phi U_{\frac{1}{2}}^{i*} - \omega U_{\frac{1}{2}}^i - \nu(1-q)U_{\frac{1}{2}}^i + q\nu y P^i U_0^i \\
\frac{d}{dt}U_0^{i*} &= -\phi U_0^{i*} + \omega U_{\frac{1}{2}}^{i*} - \nu q \eta_i p P^i U_0^{i*} - \nu(1-q)p \sum_{j\neq i} \frac{\eta_j P^j}{M-1} U_0^{i*} \\
&\quad + \nu(1-q) \sum_{j\neq i} \frac{1 - p\eta_j P^j}{M-1} U_0^i \\
\frac{d}{dt}U_{\frac{1}{2}}^{i*} &= -(\phi+\omega)U_{\frac{1}{2}}^{i*} + \nu(1-q)p \sum_{j\neq i} \frac{\eta_j P^j}{M-1} U_0^i \\
&\quad + \nu(1-q)U_{\frac{1}{2}}^i + \nu q p \eta_i P^i U_0^{i*} + \nu(1-q)p \sum_{j\neq i} \frac{\eta_j P^j}{M-1} U_0^{i*} \\
\frac{d}{dt}P^i &= -\delta P^i + g(1-P^i)\left\{ qU_{\frac{1}{2}}^i + qU_{\frac{1}{2}}^{i*}\xi_i + \xi_i \frac{1-q}{M-1} \sum_{j\neq i} \left(U_{\frac{1}{2}}^j + U_{\frac{1}{2}}^{j*} \right) \right\}
\end{aligned}$$

If we again replace the sum over all other hospitals by a sum over neighbouring hospitals, we can make a continuum approximation. Using a traveling wave-ansatz, we can determine the minimal speed numerically for this extended model, see Figure 2.14. Of course we are primarily interested in the region where q is close to one, because that is the case in reality.

2.5.4 Conclusions

We are interested in determining which precautions reduce the asymptotic velocity. As follows from Figure 2.14, cohorting (increasing q), can reduce the velocity substantially, but only if q is already quite high. If special care (quarantine) is effective, $\xi \ll 1$, cohorting becomes even more attractive. In that case, isolating people who have recently been in another hospital, decreases the velocity even more, but reaching substantial effect will be difficult (Figure 2.14).

Let us now focus on the influence of η and ξ. Decreasing η diminishes the probability that an individual acquires colonization when admitted in another colonized hospital. Decreasing ξ reduces the probability that a hospital acquires colonization when an individual from a neighbouring district is admitted. Both mechanisms will slow down the propagation speed and, as follows from equation (2.60), both intervention strategies are equally effective. Assuming $\xi = \eta$ we can obtain the influence of infection prevention measures as function of ω. From Figure 2.15, we can conclude that infection prevention measures can be a powerful tool in preventing the spread of a resistant pathogen. It can substantially reduce the spreading velocity, and can even prevent the spread if sufficiently effective. However, we have to stress that this method only gives a lower bound for the spreading velocity, because transfer of patients from one hospital to another is neglected.

2.6 Discussion

Interactions between levels of colonization within hospitals and within the community were analyzed. The main features of these analyses are (1) that the rate at which colonization is lost after hospital discharge is the major determinant for the ultimate level of colonization in the community, (2) that a small population of individuals characterized by frequent hospital admissions and high risk for colonization can maintain hospital endemicity, and thus create in the long run substantial community levels, even when the pathogen would disappear otherwise, and (3) that individual settings (e.g., hospitals or countries) can protect themselves against resistant pathogens by infection control measures, but that these are doomed to fail in the long term when resistance in the population builds up.

Modeling infectious diseases relies on making assumptions which

are open for discussion. First, we assumed that patients can only become colonized by cross-transmission within hospital settings. This holds true for many nosocomial pathogens, especially resistant forms of microorganisms that belong to the commensal flora of the intestinal tract. Examples of such pathogens are VRE and enteric Gram-negative bacteria. For pathogens colonizing the respiratory tract or skin, such as MRSA, within-hospital-transmission is most important, but the vicinity of these body surfaces to other people may allow cross-transmission even after hospital discharge.

Second, we assumed that spontaneous development of resistance was insignificant as compared to transmission. Again this holds true for VRE and MRSA, as resistance in these pathogens is based on the acquisition of large genetic elements that will not emerge through a few successive point mutations. The assumption implies that the model should be slightly adjusted for antibiotic agent & pathogen combinations for which single point mutations can lead to resistance, as in the case of quinolone antibiotics and Gram-negative bacteria.

Third, we assumed that antibiotic selection facilitated a transition from being colonized to becoming a potential spreader of resistant bacteria. There is strong clinical evidence for this process, especially for VRE, enteric Gram-negative bacteria and MRSA. Multiple risk factor analyses have identified broad-spectrum antibiotics as independent risk factors for colonization and spread of these pathogens. For VRE it was demonstrated that antibiotics not active against VRE increased the quantitative amounts of VRE per gram of feces and that with higher colonization density was associated an increased shedding of VRE in the inanimate surroundings of the patient, thereby increasing the risk for cross-transmission.

Fourth, the models used are mostly deterministic. We did consider the probability for a hospital to be infected, but we did not look at actual realizations. This prevents us from making a good prediction of, for instance, how long a hospital that takes quarantine measures can escape infection, see [Andersson and Britton, 2000].

Fifth, we have assumed that our parameters, such as antibiotic consumption, are constant in time.

And finally, when determining the wave speed for the spreading velocity we have neglected direct patient transmission between hospitals. This, of course, is not consistent with real practice, although infection control measures are usually stringent when a patient is trans-

ferred from another hospital.

Considering all the simplifications and assumptions our analyses identify three important points. The rate at which individuals lose colonization after hospital discharge is the major determinant for the ultimate level of carriers in the community. When this rate is low, the fraction of carriers can become quite high, even if the basic reproduction ratio is just slightly larger than 1. However, reliable estimates of ω, based on clinical experiments, are hardly available. Colonization with VRE can persist for at least five years [Byers et al., 2002; Baden et al., 2001], while reported average duration of colonization with MRSA ranges from three until 40 months [Mitsuda et al., 1995; Scanvic et al., 2001; Sanford et al., 1994]. As this parameter is of crucial importance for the prevalence of resistance in the community, studies determining this parameter for different pathogens are urgently needed. Moreover, these findings provide, at least on a theoretical basis, justification for the use of therapies that can increase ω in order to prevent emergence of multiple-resistant nosocomial pathogens in the community. One could think of antimicrobial eradication strategies or suppressing the resistant flora by competition with other microorganisms (e.g., probiotics).

Another finding is that a relatively small population, characterized by frequent admissions and high risks for colonization with resistant microorganisms, can maintain nosocomial endemicity even when a crude estimate of the basic reproduction ratio would suggest otherwise. Such populations usually consist of patients with some form of severe underlying illness, thereby increasing their risk of successful colonization with nosocomial pathogens and the subsequent development of infections. These infections must be treated with antibiotics and necessitate hospital admission. In a cycle of events, the individual components of the transmission processes are enhanced. This is comparable to the effect of a so-called core-group of individuals for sexually transmitted diseases. Examples of such small high-risk populations are haemodialysis patients for VRE, [D'Agata et al., 2002b], and MRSA and cystic fibrosis patients for *Pseudomonas aeruginosa*.

In the final analyses we investigated the influence of spatial structure on the spread of resistant pathogens in the community and in other (hospital) settings, and determined the effects of various infection control measures. From the traveling wave solution we determined the wave speed. One of the possible infection control measures is to put a newly hospitalized patient (usually a patient with an increased risk for

being colonized with a resistant strain, such as patients coming from a setting where colonization with resistant bacteria is endemic) into quarantine. This procedure will act as a double advantage, as it will decrease both the likelihood that this patient will spread bacteria to other patients and, vice versa, the risk for the quarantined patient to acquire colonization. In our simplified model of only three hospitals, we find that an individual hospital taking such measures is able to control the nosocomial spread of resistance for quite a long time. However, when a substantial part of the community population becomes colonized, due to discharges from other hospitals, quarantine measures will lose their effectiveness. In fact, in that situation patients admitted from the community, at high-risk for being colonized, need to be quarantined as well. Evidently, it will be more difficult to identify these patients than patients directly transported from other, suspected, settings. In addition, the hospitals' capacity for quarantine measures will be exhausted in due course. An example of this situation is the Dutch MRSA-policy. The Netherlands are surrounded by countries where colonization with MRSA is endemic within hospitals. Colonization and infection rates with MRSA in Dutch hospitals are still $< 1\%$. All patients that have been hospitalized abroad and that are transported to a Dutch hospital are put into quarantine until proven not to be colonized with MRSA. In case of colonization, the patient is kept into strict isolation until the microorganism has been eradicated (with antimicrobial agents) or until the patient is discharged from the hospital. This policy has been successful for the last twenty years. However, nowadays the number of patients seeking medical help abroad, especially in other countries of the European Union, is increasing and the financial possibilities for maintaining high effectiveness of detection, isolation and eradication are decreasing. In addition, a community-reservoir of MRSA seems to be building, further enlarging the risk of unnoticed introduction of MRSA into the hospital setting. With some modifications, the model can be used to determine the effects of these changes on the dynamics and to predict the success of this policy in coming years. The model could also be used to estimate the measures necessary to reverse a situation of nosocomial endemicity. In summary, our model describes how hospital settings can act as driving forces of antibiotic resistance in the community and identified some pivotal variables in the overall dynamics. With slight modifications the model could be used to analyze and, may be, predict changes in the dynamics for specific antibiotic-resistant pathogens.

normal	core.	
α	a	hospital transition rate from colonized to spreader
μ	m	death rate
ω	z	decolonization rate after discharge
χ	x	extramural rate of losing the high colonization load
ν	n	admission rate
σ	s	discharge rate
λ	l	birth rate
β, γ	b, c	transmission parameters

Chapter 3

The effect of intervention measures on the spread of MRSA

3.1 Introduction

Nosocomial infections with Methicillin-resistant *Staphylococcus aureus* (MRSA) have become a major problem in most European and American hospitals. MRSA-infections are associated with increased mortality compared with MSSA [Cosgrove et al., 2003] and in countries like the UK, USA and Greece, the fraction of MRSA isolates among the invasive *S. aureus* isolates has increased to more than 40% [Reacher et al., 2000]. In the Scandinavian countries and the Netherlands, in contrast, MRSA is still just a minor problem. In these countries a stringent policy of isolating colonized patients and screening high-risk patients is being followed (the so-called 'search and destroy' policy) [Werkgroep Infectie Preventie, 2003]. A recent paper [Cooper et al., 2004], provides an excellent overview of the nature of an MRSA epidemic and the major influence of re-admission of colonized patients. (Typically, per admission, a colonized patient transmits the pathogen insufficiently to other patients, to generate an outbreak. However, due to re-admission of patients who are still carriers, the prevalence within a hospital can increase to high levels.) In the setting analyzed in that article, however, only patients found to be colonized with MRSA by way of standard clinical cultures are isolated, which does not resemble the Dutch ap-

proach. In the Netherlands, all patients in the ward are screened for MRSA after a patient has unexpectedly been identified as MRSA carrier and, when positive, are put into isolation. Moreover, all high risk patients (patients who were colonized with MRSA on a previous hospitalization or patients recently hospitalized in a country with a high endemic level of MRSA in hospitals) are isolated on admission and only released from isolation when they are proven to be uncolonized. However, the relative contribution of each of these interventions is unknown [Voss, 2004].

To evaluate the effects of different interventions, we use two models, an analytical compartmental model of a hospital with two-level mixing and a more detailed simulation model. The majority of acquisitions occur within a ward, with low probability of MRSA-transmission to patients in other wards. As the depletion in the number of susceptibles is much faster in a small unit than in the hospital as a whole, the compartmental structure in itself already reduces the transmittability of MRSA. The key point, however, is that the compartmental structure allows us to analytically model the Dutch MRSA-policy. The main difference with the standard two-level mixing models (see e.g. [Ball et al., 1997]) is that the individuals within a unit can be replaced due to admission and discharge of patients. Therefore the size of an outbreak within a unit can be larger than the (finite) number of beds within the unit.

For the pathogen to persist, the expected size of an outbreak multiplied by the probability that a patient still carries the pathogen on a next admission, should be above 1. Our aim is to calculate the expected size on the basis of meaningful and plausible assumptions. Knowing how the size of an outbreak depends on transmission and other parameters, we can derive the critical values of these parameters. Based on these critical parameters, we can determine the effects of each of the components of the 'search and destroy' policy.

In the sequel the results of the analytical model will be analyzed by the simulation model. The advantage of this approach is that prevalence of MRSA in the community can be monitored as a function of time and that confidence intervals for the prevalence can be constructed.

The simulation model is also used to analyze all kinds of extensions of the analytical model, such as several hospitals with different intervention measures, heterogeneity between wards (inclusion of intensive care units (ICU's)), the effect of persistently colonized health

care workers, and the possibility of superspreaders. Both sections can be read separately according to the interest of the reader. Section 3.2 is probably more suitable for readers with some background in mathematics while Section 3.3 is also intended for readers without such a background.

3.2 An analytical model

3.2.1 A compartmentalized hospital

A single unit

Suppose a unit consists of N beds. We want to calculate the expected size of an outbreak. Due to the finite size of a unit and the stochastic nature of transmission, the expected size of an outbreak will be a finite number.

The dynamics in the ward is generated by two processes, transmission of the pathogen, and discharge and admission of patients. We assume that admission and discharge of patients happen at the same fixed time of the day. In particular, we assume that newly admitted patients are not colonized, i.e., the outbreak is triggered by just one exceptional case. (When the prevalence in the community is low and when the typical size of an outbreak is small, the probability that a second colonized individual in the community enters the hospital during the period of the outbreak is small.)

We also assume that the only route of transmission is cross transmission in which the rate for a susceptible patient to acquire colonization is proportional to the fraction of the beds that is occupied by colonized patients [Bonten and Weinstein, 1996; Ferrer et al., 2001; Grundmann et al., 2002]. This implies that the number of contacts during which colonization could be transferred is independent of the number of patients in the unit. For instance, when health care workers are the vectors of transmission and the per patient number of contacts with health care workers per day is independent of the size of the unit, the assumption is valid. Let $P(y|x)$ be the probability that there are y colonized patients in the unit at time $t + 1$ just before discharge and admission, given that x colonized patients were present in the unit at time t just after discharge and admission. For simplicity, we assume that uncolonized patients can acquire colonization only from those patients who

were already colonized after admission and discharge and not in a two step procedure during one and the same day. With these assumptions, the probability per susceptible patient per day to acquire colonization is given by $\left(1 - e^{-\beta \frac{x}{N-1}}\right)$ and the $P(y|x)$ is the binomial distribution, i.e., the probability $P(y|x)$ is given by ($0 \leq x, y \leq N$):

$$P(y|x) = \begin{cases} 0 & \text{if } y < x \\ \begin{pmatrix} N-x \\ y-x \end{pmatrix} \left(1 - e^{-\beta \frac{x}{N-1}}\right)^{y-x} \left(e^{-\beta \frac{x}{N-1}}\right)^{N-y} & \text{if } y \geq x \end{cases} \tag{3.1}$$

with β the transmission parameter. (We use the factor $N-1$ for later convenience to ensure that the expected number of acquisitions in the first day after admission to a 'virgin' ward is independent of the size of the ward for small β, see page 61 and [Ball et al., 1997].)

The number of occupied beds in the unit is assumed to be constant, i.e., we assume that all discharged patients are immediately replaced by newly admitted patients (who we have assumed to be uncolonized). We assume the length of stay to be independent of the colonization status of a patient and exponentially distributed with discharge rate $\frac{1}{d}$, where $d > 1$ is the mean length of stay in the unit. Finally we assume that colonization is persistent during the stay in the unit.

The probability $\hat{P}(z|y)$ that there are z colonized patients after discharge and admission, given that there were y colonized patients in the unit before discharge and admission, is given by ($0 \leq y, z \leq N$):

$$\hat{P}(z|y) = \begin{cases} 0 & \text{if } z > y \\ \begin{pmatrix} y \\ z \end{pmatrix} \left(\frac{1}{d}\right)^{y-z} \left(1 - \frac{1}{d}\right)^{z} & \text{if } z \leq y \end{cases} \tag{3.2}$$

Using a discrete convolution of transmission and discharge we find that the transition probabilities $f(b|a)$ that there are b colonized patients in the unit directly after discharge and admission at time $t+1$, given there were a colonized patients in the unit directly after discharge and admission at time t, are given by:

$$f(b|a) = \sum_{j=0}^{N-a} P(a+j|a)\hat{P}(b|a+j) \tag{3.3}$$

The expected number of acquisitions of colonization in a transition

from a colonized patients to b colonized patients is given by:

$$g(a,b) = \frac{\sum_{j=0}^{N-a} jP(a+j|a)\hat{P}(b|a+j)}{f(b|a)} \tag{3.4}$$

We can now calculate the expected size of an outbreak in a unit in terms of these quantities. Let $E(k)$ be the expected total number of new acquisitions in the unit when there are k colonized patients in the unit directly after admission and discharge. First step analysis then shows that the $E(k)$ satisfy the following set of equations:

$$\begin{aligned} &E(0) = 0 \\ &E(k) = \sum_{b=0}^{N} f(b|k)\left(E(b) + g(k,b)\right) \quad 1 \leq k \leq N \end{aligned} \tag{3.5}$$

We are especially interested in $E(1)$, as this is the expected number of acquisitions when a single colonized individual enters a 'virgin' unit. This number is an ingredient for computing the critical transmission parameter (related to R_0) as we will show later.

A hospital with many units

Apart from transmission within a unit, transmission between units is possible. From the expected size of an outbreak within a ward ($E(1)$ as defined in the previous section), we can calculate the expected size of an outbreak in the hospital as a whole (T) in which the outbreak may affect several units. To determine T, we consider a hospital with an infinite number of wards, each of size N. This assumption ensures that when two patients outside the initial unit are infected by patients in the initial ward, these two infected patients will never stay in the same ward. When a single colonized patient is hospitalized, we assume that the first acquisition occurs with probability p in the unit of the colonized patient and with probability $(1-p)$ in another unit, i.e., if we ignore depletion of susceptibles, a fraction p of all acquisitions occurs within the own unit of a patient. The expected number of acquisitions in the own unit due to the index patient during the first day is $(N-1)(1-e^{-\frac{\beta}{N-1}}) \approx \beta$ where the approximation is good when $\frac{\beta}{N-1}$ is small. The expected number of acquisitions in other units due to the index patient during the first day is then given by $\beta\frac{1-p}{p}$. (This coincides with mass action with transmission parameter $\beta\frac{1-p}{p}$.)

The expected size of the total outbreak T (excluding the index case) satisfies a consistency relation (equation (3.6)). The index patient infects on average $E(1)$ individuals within its own ward. The index patients can also infect patients in other wards. On average, one colonized individual infects $\beta\frac{1-p}{p}$ patients outside its own ward per day. The expected number of infection outside its own ward by the index patient is than given by $\beta\frac{1-p}{p}d$. The other $E(1)$ patients in the ward can also infect patients in other wards provided they are not immediately discharged after acquisitions (probability $\frac{d-1}{d}$). Therefore, the expected number of acquisitions due to patients in the initial ward is given by $\frac{(1-p)}{p}\beta d(1+\frac{d-1}{d}E(1))$. As we assumed an infinite number of wards, each of these acquisitions will occur in different wards. These patients are colonized themselves, and they can transmit the pathogen provided they are not immediately discharged after acquisition. Therefore, the consistency relation for T is given by:

$$T = E(1) + \frac{(1-p)}{p}\beta d(1+\frac{d-1}{d}E(1))(1+\frac{d-1}{d}T). \tag{3.6}$$

This leads to:

$$T = \frac{pE(1) + (1-p)\beta d(1+\frac{d-1}{d}E(1))}{p-(1-p)\beta(d-1)(1+\frac{d-1}{d}E(1))} \tag{3.7}$$

provided this expression is positive. Note that when the number of wards is finite the assumption that all acquisitions outside the own ward occur in virgin wards does not hold anymore. However, as the expected number of transmissions for a patient is maximal when all other patients within the unit are susceptible, the assumption will lead to an overestimation of the expected size of an outbreak T and this number can be regarded as a worst case scenario.

To determine the effect of the depletion of susceptibles within a unit, we first calculate the expected size of an outbreak in a hospital in which each ward is of infinite size. In one unit of infinite size, the expected number of patients which acquire colonization from the index patient is given by βd and the expected number of patients who acquire colonization from the index patient and are not discharged before they can spread is given by $\beta(d-1)(=R_A)$. Introduction of the pathogen can only lead to a major outbreak if $R_A > 1$. If $R_A < 1$, the expected size of an outbreak (excluding the index case) in wards of infinite size satisfies:

$$E(1) = \beta d + R_A E(1) \Leftrightarrow E(1) = \frac{\beta d}{1-\beta(d-1)} \tag{3.8}$$

which can be inserted in equation (3.7) to obtain the expected size of an outbreak in a hospital with an infinite number of wards, each of infinite size. However, the insertion of (3.8) in equation (3.7) can be omitted by realizing that the outbreak size in a hospital in which all wards are of infinite size, the transmission parameter in a unit is β and a fraction $1-p$ of the expected acquisitions occur between units, is the same as the outbreak size in a hospital in which all acquisitions occur within the unit and the transmission parameter is $\beta + \beta\frac{1-p}{p} = \frac{\beta}{p}$. Hence we have that the outbreak size T in a hospital with units of infinite size will be:

$$T = \frac{\frac{\beta}{p}d}{1-\frac{\beta}{p}(d-1)} = \frac{\beta d}{p-\beta(d-1)}. \tag{3.9}$$

Note that the p-dependency in this relation only arises due to our definition of the transmission parameters; p does not effect the number of transmissions within a unit but effects the number of transmissions between units.

Re-admission

When $T < \infty$, MRSA cannot persist in the hospital without re-introduction. However, colonized patients discharged from the hospital may carry MRSA for a long period of time. When such a carrier is re-admitted, a new outbreak in the hospital can occur. Suppose the probability that a patient leaving the hospital while being colonized will re-enter the hospital at some later time while still being a carrier is ξ (i.e., when the decolonization rate is χ and the hospitalization rate is h, $\xi = \frac{h}{h+\chi}$). When we assume that a person who dies is immediately replaced by an uncolonized individual, we can incorporate the probability that an individual dies while still being colonized in the decolonization rate ξ.

For the pathogen to be able to persist in the community as a whole (both extramural and intramural), the expected size of an outbreak T should exceed the critical value T_c, which is defined by the relation:

$$(1+T_c)\xi = 1. \tag{3.10}$$

In a hospital consisting of wards of infinite size, the critical value for β is given by:

$$\beta_c = \frac{p(1-\xi)}{d-1+\xi}.$$

With wards of smaller size, the recursion (3.5) has to be solved after which (3.7) and (3.10) can be used.

A small unit size can by itself already reduce transmission. The reason is that the structure works against large outbreaks as the pool of susceptibles declines faster in a small unit than in the hospital as a whole, see Figure 3.1(a) for the effect of the depletion of susceptibles in a more complex situation

3.2.2 Interventions for a single type of patients

We will now calculate the relative increase in the critical transmission parameter, β_c, for different infection prevention measures compared to a hospital not implementing any additional infection prevention measures.

Isolation of identified carriers

Samples for microbiological culturing are frequently obtained from patients with clinical suspicion of infection. Suppose that patients colonized with MRSA are detected as carriers in this way with probability ν per day. When patients known to be colonized with MRSA are transferred to an isolation room to prevent further spread, colonized patients are isolated at frequency ν. For modeling simplicity we assume that all cultures are performed at the same moment of the day, immediately before the moment of discharge and admission, that culture results are available at once, that cultures are 100% reliable and that isolated patients remain in isolation till discharge and, therefore, no longer spread MRSA to other patients. From a modeling perspective, we can, therefore, assume that patients who are found to be colonized are immediately discharged as they no longer participate in the transmission process. Note that we implicitly assume that the number of patients treated in isolation does not affect the number of patients not treated in isolation. With these assumptions, the effective discharge probability per day for a colonized patient becomes $\frac{1+(d-1)\nu}{d}$. Although the discharge probability for uncolonized patients remains $\frac{1}{d}$, we can assume it to be $\frac{1+(d-1)\nu}{d}$, as the replacement of an uncolonized patient by another uncolonized patient has no effect. This policy is incorporated in the model described by [Cooper et al., 2004] for a hospital consisting of only one unit. For a ward of infinite size, the relative increase in the

critical transmission parameter will be $1 + (d-1)\nu$ (provided the isolation capacity is sufficiently large). For a hospital consisting of wards of finite size, equations (3.5), (3.7) and (3.10) have to be solved in which d should be replaced by $\frac{d}{1+(d-1)\nu}$. The efficacy of this intervention relies mainly on ν, the detection probability of MRSA-colonization, as the average length of stay of hospitalized patients can hardly be altered.

Screening and isolation

The hospital structure can be used for an extension of the previous isolation policy: An active screening strategy as applied in the Netherlands. A patient identified as MRSA carrier will be transferred to an isolation room and all patients that may have had contact (directly or indirectly) are screened for colonization with MRSA. In our model, we replace screening of all contact patients by screening of all patients in the same unit as the index-patient, as the majority of transmissions occur within the unit. Again, identified carriers among these contact patients are transferred to an isolation room. This will, ideally, terminate the outbreak within the unit immediately, though spread may continue in other wards. This policy can therefore only be effective when the majority of acquisitions occur in the same unit.

Again we assume that patients are identified as MRSA carriers by standard clinical cultures with probability ν per patient per day and that these cultures are performed directly before the moment patients can be discharged. We can incorporate these clinical cultures in the 'discharge' probabilities. Without a clinical culture performed in any of the colonized patients (probability $(1-\nu)^y$, with y the number of colonized patients), the number of colonized patients can only decrease due to discharge. When at least one colonized patients is identified as carrier, all other patients in the unit will be screened and colonization is detected with probability θ per carrier. In this case the discharge function is given by:

$$\hat{P}(z|y) = \begin{cases} 0 & \text{if } z > y \\ \binom{y}{z} (1-\nu)^y \left(\frac{1}{d}\right)^{y-z} \left(1-\frac{1}{d}\right)^z + & \\ \sum_{j=1}^{y-z} \binom{y}{j} (1-\nu)^{y-j} \nu^j \binom{y-j}{z} \left(\frac{1+(d-1)\theta}{d}\right)^{y-j-z} \left(\frac{d-1}{d}(1-\theta)\right)^z & \text{if } 0 \le z \le y \end{cases} \tag{3.11}$$

When the tests are 100% reliable and all contact patients are screened ($\theta = 1$), culturing of at least one colonized patient leads to detection of all colonized patients. (In terms of the pathogen, this can be seen as a birth-death-catastrophe process.) As all colonized patients are transferred to isolation rooms, all remaining patients will be uncolonized. and the discharge function simplifies to:

$$\hat{P}(z|y) = \begin{cases} 0 & \text{if } z > y \\ \binom{y}{z} \left(\frac{1}{d}\right)^{y-z} \left(1-\frac{1}{d}\right)^{z} (1-\nu)^{y} & \text{if } 0 < z \leq y \\ (1-\nu)^{y} \left(\frac{1}{d}\right)^{y} + (1-(1-\nu)^{y}) & \text{if } z = 0 \end{cases} \tag{3.12}$$

Using (3.3) and (3.4) with the new 'discharge' function and (3.5), (3.7) and (3.10), the new critical value of the transmission parameter can be calculated.

Screening on admission

Another part of the Dutch policy is to isolate patients immediately after admission when it is known that they were colonized with MRSA on a previous admission. Only when colonization is excluded, isolation measures are withdrawn. We define S as the expected number of colonized patients, per outbreak in the hospital as a whole, who were identified as MRSA carriers during hospitalization (including the index patient). When no additional screening of contact patients is performed after demonstration of colonization, the number S satisfies the relation:

$$S = \frac{\nu}{\nu + (1-\nu)\frac{1}{d}}(1+T) \tag{3.13}$$

as each colonized patient has a probability $\frac{\nu}{\nu+(1-\nu)\frac{1}{d}}$ to be detected. Note that T should be calculated using an effective discharge probability of $\frac{1+(d-1)\nu}{d}$, see the previous section about isolation of identified carriers.

When the efficacy of this policy is g in the sense that a fraction $1-g$ of the previously colonized patients is missed by the screening, we find the following relation for criticality:

$$(1+T-S)\xi + S\frac{\xi(1-g)}{1-\xi g} = 1 \tag{3.14}$$

(The last term comes from the geometric series: Of the S patients, a fraction $(\xi g)^{j-1}\xi(1-g)$ will be able to spread at the j^{th} admission and could not spread at any previous admissions.)

The Dutch 'search and destroy' policy

Now suppose that contact patients are screened for colonization in case of an unexpected case of MRSA colonization and that previously colonized patients are screened on admission. To determine the efficacy of this policy, we not only need the expected size of an outbreak (T), but also the expected number of detected patients per outbreak (S). As the probability that colonization at a patient is detected depends on the total number of colonized patients in the unit (because of the search strategy), we introduce $S(k)$ as the expected number of detected cases in an outbreak in a single ward when the outbreak starts with k (undetected) patients. The $S(k)$ satisfy the set of equations:

$$\begin{aligned} S(0) &= 0 \\ S(k) &= \sum_{b=0}^{N} f(b|k)\,(S(b) + h(k,b)) \end{aligned} \tag{3.15}$$

where we should use equation 3.11 in the definition the transition probability f (equation 3.3) and where $h(x,y)$ denotes the expected number of detected patients in a transition from x colonized patients to y colonized patients. $h(x,y)$ is given by:

$$h(x,y) = \tfrac{1}{f(y|x)} \sum_{j=0}^{N-x} \sum_{i=1}^{x+j-y} P(x+j|x) \begin{pmatrix} x+j \\ i \end{pmatrix} (1-\nu)^{x+j-i}\nu^i \begin{pmatrix} x+j-i \\ y \end{pmatrix} \left(\tfrac{1+(d-1)\theta}{d}\right)^{x+j-i-y} \left(\tfrac{d-1}{d}(1-\theta)\right)^{y} \left(i + \tfrac{\theta}{\theta+(1-\theta)\frac{1}{d}}(x+j-i-y)\right) \tag{3.16}$$

When the screening will reveal every colonized patient $(\theta = 1)$, equation (3.15) simplifies to:

$$\begin{aligned} S(0) &= 0 \\ S(k) &= \sum_{b=1}^{N} f(b|k)S(b) + \sum_{j=0}^{N-k} P(k+j|k)(k+j)\left\{1-(1-\nu)^{k+j}\right\} \end{aligned} \tag{3.17}$$

where we can use (3.12) to determine the transition probabilities (3.3).

To determine the expected number of transmissions from the initial ward to other wards, we need to know the severity of the outbreak within a unit, i.e., the total number of patient days in the initial ward with patients being infectious. (We weight each day of the outbreak in the unit according to the number of patients in the unit that are infectious.) Let $D(k)$ be the severity of the outbreak within a unit, i.e.,

before the natural fade out of the outbreak or the detection of the outbreak, given there are initially k infectious patients in the ward. The $D(k)$ satisfy:

$$\begin{aligned} D(0) &= 0 \\ D(k) &= k + \sum_{b=0}^{N} f(b|k) D(b) \end{aligned} \tag{3.18}$$

The expected total number T of patients infected during an outbreak (which may spread over several wards) and the total number S of detected patients satisfies the equations:

$$\begin{aligned} T &= E(1) + \tfrac{(1-p)}{p} \beta D(1)(1 + \tfrac{d-1}{d} T) \\ S &= S(1) + \tfrac{(1-p)}{p} \beta D(1)(\nu + (1-\nu)\tfrac{d-1}{d} S) \end{aligned} \tag{3.19}$$

With these results, (3.19) and the relation (3.14), we can determine the effect of the Dutch policy.

Active decolonization

Finally we incorporate the possibility that patients treated in isolation are actively decolonized: Suppose that colonization is successfully eradicated in a fraction ϕ of the patients. This active decolonization can be incorporated in the critical relation:

$$\begin{gathered} (1 + (T - \phi S) - (1-\phi)S)\,\xi + S\tfrac{\xi(1-\phi)(1-g)}{1-\xi(1-\phi)g} = 1 \Rightarrow \\ (1 + T - S)\,\xi + S\tfrac{\xi(1-\phi)(1-g)}{1-\xi(1-\phi)g} = 1 \end{gathered} \tag{3.20}$$

3.2.3 A core group

In reality there is heterogeneity in the frequency of admission between individuals. Some individuals visit the hospital often, such as for example haemodialysis patients and elderly. To take heterogeneity to some extent into account, we repeat the analysis of section 3.2.2 for two types of individuals (core group, non-core group) which differ only in their probability to be still colonized with MRSA when on re-admitted. Individuals from the core group are hospitalized more frequently. If the spontaneous decolonization rate is the same for the core and the non-core group or if the decolonization rate is smaller for the core group, the probability to be colonized upon re-admission is higher for the core group. As the expected size of an outbreak within the hospital is not

altered by the introduction of two types of patients with identical behaviour within the hospital, the calculation of the outbreak size T is identical to the situation without a core group. (Use (3.3), (3.4), (3.5), (3.7) and (3.11), (3.12) or (3.2) depending on whether there is an active search policy in case of an unexpected case of MRSA-colonization or not, respectively.) However, the critical outbreak size T_c will change. A core group will enhance the importance of re-admission and, therefore, a lower R_A can lead to persistence in the population (both intramural and extramural). As for a lower R_A the average size of an outbreak will be smaller, the depletion of the number of susceptibles will also be smaller. Therefore, the effects of the compartmental structure of the hospital on the critical transmission parameter will be smaller in case of a core group. Moreover, relative effects of elements of the 'search and destroy' policy will be different.

Let ξ_i be the probability that a colonized patient of type i is still colonized when re-admitted and let γ be the expected fraction of hospitalized patients of type 1. When all other characteristics of the two types of patients are identical and previously colonized patients are not screened when re-admitted to the hospital, the pathogen can persist if and only if the largest eigenvalue of the next-generation matrix

$$\begin{pmatrix} (\gamma T+1)\xi_1 & \gamma T\xi_1 \\ (1-\gamma)T\xi_2 & ((1-\gamma)T+1)\xi_2 \end{pmatrix} \tag{3.21}$$

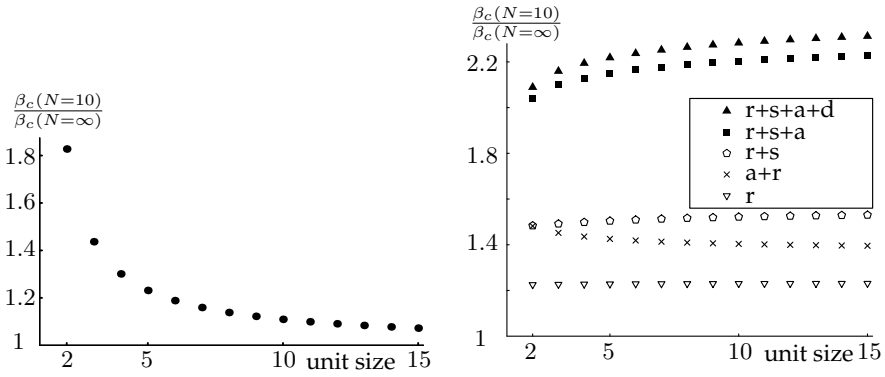
exceeds 1. The largest eigenvalue is equal to 1 when the outbreak size T equals:

$$T_c = \frac{(1-\xi_1)(1-\xi_2)}{\xi_2(1-\xi_1)+\gamma(\xi_1-\xi_2)} \tag{3.22}$$

The largest eigenvalue can be seen as the reproduction ratio between outbreaks in a hospital.

Isolation of identified carriers

We will now investigate the effect of different intervention measures in the presence of a core group. If the only intervention is to isolate patients known to be colonized with MRSA by clinical cultures, equation 3.22 still applies but T should be determined using an effective length of stay of $\frac{d}{1+(d-1)\nu}$ in (3.2) and (3.7), see Section 3.2.2.



(a) Effect compartmental structure alone as a function of the size of the unit

(b) Relative effect interventions as a function of the size of the unit

Figure 3.1: (a): Relative increase in critical transmission parameter β_c for different sizes of units without any interventions compared to a hospital consisting of wards of infinite size (i.e., without depletion of susceptibles) without any intervention. On average, $\frac{1}{10}$ of the colonized patients of type 1 and $\frac{1}{2}$ of colonized patients of type 2 is still colonized when re-admitted. 50% of the hospital is occupied by patients of type 2. 90% of the acquisitions occur within units. **(b)**: Relative increase in critical transmission parameter β_c for different intervention strategies compared to a hospital consisting of wards of identical size but without any intervention strategy. The symbols "r" stands for isolation of patients identified as MRSA-carriers upon clinical culture results, "s" stands for screening of all patients in a unit once one patient is found to be colonized, "a" stands for screening on admission of patients known to be colonized on previous admission with efficacy $g = 0.88$ and "d" stands for decolonization with efficacy of 20%.

Screening on admission

Now, screening of patients, known to be colonized during previous hospitalization with efficacy g, is added. Without additional screening of contact patients in case of an unexpected case of MRSA-colonization, each colonized patient has a probability to be detected of $\frac{\nu}{\nu+(1-\nu)\frac{1}{d}}$. Let $\eta(\nu, d, g, \xi_i)$ be the probability that a patient of type i who acquires colonization during hospitalization will ever return to the hospital while still being colonized but without being notified as such. When colonization was not detected during hospitalization (probability $1 - \frac{\nu}{\nu+(1-\nu)\frac{1}{d}}$), the patient is still colonized at the next admission with probability ξ_i. When colonization was detected during hospitalization, the patient has a probability $\xi^m g^{m-1}(1-g)$ to be able to transmit during the m^{th} admission without the ability to spread at any of the previous admissions. Therefore we have that:

$$\eta(\nu, d, g, \xi) = \xi \left(1 - \frac{\nu}{\nu + (1-\nu)\frac{1}{d}} + \frac{\nu}{\nu + (1-\nu)\frac{1}{d}} \frac{1-g}{1-\xi g} \right) \tag{3.23}$$

Note that $\eta(\nu, d, g, \xi) = \xi$ when the efficacy g of screening on admission of high-risk patients is 0. The next generation matrix is given by:

$$\begin{pmatrix} (\gamma T + 1)\eta(\nu, d, g, \xi_1) & \gamma T \eta(\nu, d, g, \xi_1) \\ (1-\gamma) T \eta(\nu, d, g, \xi_2) & ((1-\gamma T) + 1)\, \eta(\nu, d, g, \xi_2) \end{pmatrix} \tag{3.24}$$

where T again should be determined using an effective length of stay of $\frac{d}{1+(d-1)\nu}$. The pathogen can persist if and only if the largest eigenvalue of the next generation matrix exceeds 1.

The Dutch 'search and destroy' policy

With active search for MRSA in case of an unexpected identification of colonization with MRSA upon clinical cultures and screening on admission of previously colonized patients with efficacy g, the calculation is less straightforward as the number of detected patients of type i depends on whether the initial colonized patient was of type i or not. As the typical size of an outbreak is small, this cannot be neglected. Let $Q(k)$ denote the probability that colonization of the index patient is detected, given that there are k colonized patients in the unit (including the index patient) at the start of that day. We first consider the case when screening detects all colonized patients in the unit ($\theta = 1$).

Suppose that during that day j uncolonized patients in the unit acquire colonization. If the colonization of at least one of the $k+j$ colonized patients is detected (probability $1-(1-\nu)^{k+j}$), all patients in the unit will be screened and the colonization of the index patient is also detected. If colonization of none of the $k+j$ patients is detected by clinical cultures, colonization of the index patient can only be detected if the index patient is not discharged and the probability of detection will depend on the number of colonized patients that remain in the unit. Therefore, the $Q(k)$ satisfy the following relations for $1 \leq k \leq N$:

$$Q(k) = \sum_{j=0}^{N-k} P(k,k+j)\left(1-(1-\nu)^{j+k}\right) + \sum_{j=0}^{N-k} P(k,k+j)(1-\nu)^{j+k} \sum_{z=1}^{j+k} \binom{j+k}{z} \left(\tfrac{1}{d}\right)^{(j+k-z)} \left(1-\tfrac{1}{d}\right)^{z} \tfrac{z}{j+k} Q(z) \tag{3.25}$$

When $\theta \neq 1$, $Q(k)$ satisfies:

$$Q(k) = \sum_{j=0}^{N-k} P(k,k+j) \left(\sum_{z=0}^{k+j} \binom{k+j}{z} (1-\nu)^{k+j} \left(\tfrac{1}{d}\right)^{k+j-z} \left(1-\tfrac{1}{d}\right)^{z} \tfrac{z}{k+j} Q(z) + \sum_{z=0}^{k+j} \sum_{i=1}^{k+j} \binom{k+j}{i} (1-\nu)^{k+j-i} \nu^{i} \binom{k+j-i}{z} \left(\tfrac{1+(d-1)\theta}{d}\right)^{k+j-i-z} \left(\tfrac{d-1}{d}(1-\theta)\right)^{z} \tfrac{1}{k+j} \left\{ i + (k+j-i-z)\tfrac{\theta}{\theta+(1-\theta)\frac{1}{d}} + zQ(z) \right\} \right) \tag{3.26}$$

Let T_{ij} denote the expected number of new cases of type j when the initial case was of type i and let S_{ij} denote the expected number of detected cases of type j when the initial case was of type i. We obtain as next generation matrix:

$$\begin{pmatrix} \xi_1(1+T_{11}-S_{11})+S_{11}\frac{\xi_1(1-g)}{1-\xi_1 g} & \xi_1(T_{21}-S_{21})+S_{21}\frac{\xi_1(1-g)}{1-\xi_1 g} \\ \xi_2(T_{12}-S_{12})+S_{12}\frac{\xi_2(1-g)}{1-\xi_2 g} & \xi_2(1+T_{22}-S_{22})+S_{22}\frac{\xi_2(1-g)}{1-\xi_2 g} \end{pmatrix} \tag{3.27}$$

The pathogen can persist if and only if the largest eigenvalue of this matrix exceeds 1.

The calculation of T_{ij} is easy, as T is defined as the expected size of an outbreak excluding the index patient. Therefore, on average, a fraction γ of T patients will be of type 1 and a fraction $(1-\gamma)$ of type 2. The calculation of S_{ij} is more complex, as S is defined as the expected number of colonized patients for which colonization is detected. Therefore,

index patients also contributes to S and we have to use the probability $Q(1)$ that colonization is detected in the index patient. We obtain:

$$\begin{array}{cc} T_{i1} = \gamma T & i \in \{1,2\} \\ T_{i2} = (1-\gamma)T & i \in \{1,2\} \\ S_{11} = \gamma S + (1-\gamma)Q(1) & \\ S_{12} = (1-\gamma)(S - Q(1)) & \\ S_{21} = \gamma S - \gamma Q(1) & \\ S_{22} = (1-\gamma)S + \gamma Q(1) & \end{array} \tag{3.28}$$

Active decolonization

When colonized patients of type i treated in isolation rooms are actively decolonized with efficacy ϕ_i, the relations for T_{ij} and S_{ij} in system 3.28 change. The new values (denoted with a prime) are given by:

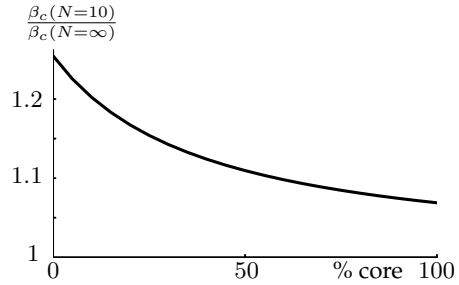
$$\begin{array}{c} S'_{ij} = S_{ij}(1-\phi_j) \\ T'_{ij} = T_{ij} - S_{ij}\phi_j \end{array} \tag{3.29}$$

The effects of different prevention strategies are shown in Figure 3.1. Note that the calculation in this section can be extended to a situation with more several core-groups, each with a different re-admission rate.

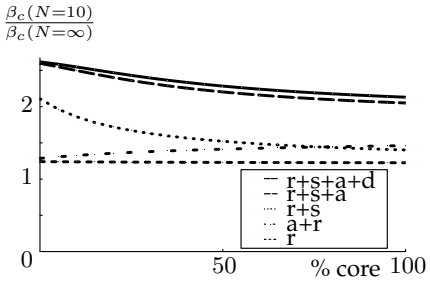
3.2.4 Results and conclusions of the analytical model

As shown in Figure 3.1, small units can be an effective way to prevent transmission as the number of susceptible patients decreases faster in small units than in large units. Also, the relative effects of interventions depend on the unit size. Screening of contact patients after identification of a MRSA-carrier is more effective for large units while screening on admission is more effective for small units. However this size-dependency is not very strong.

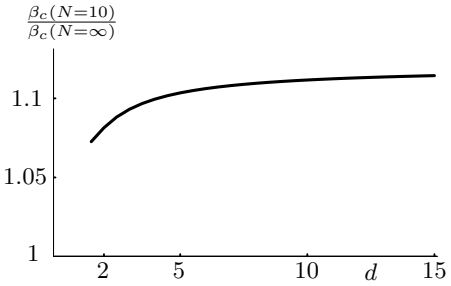
The R_0-value for MRSA is likely to be between 1 and 1.5. R_0 must be larger than 1 as endemicity has been established in hospitals in the US and UK. it should be above 1 and values larger than 1.5 are likely to lead to higher endemic levels. Based on this estimation and the results depicted in Figure 3.1, an intervention based on the isolation of patients found by clinical cultures alone is unlikely to be effective and should not be advised as policy. Combining this policy with screening on admission of patients who were colonized on a previous admission is more likely to be effective. Screening of contact patients in



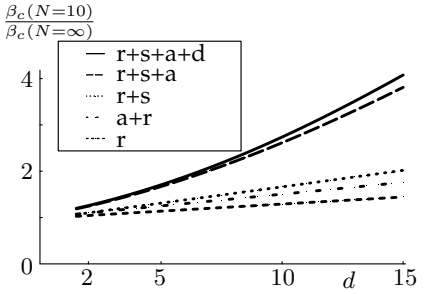
(a) Effect of core group size without intervention



(b) Effect of core group size

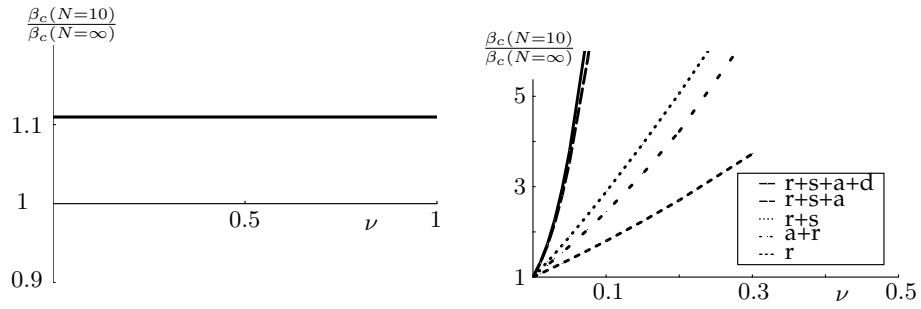


(c) Effect of the mean length of stay length without intervention



(d) Effect of the mean length of stay length with interventions

Figure 3.2: Left (a, c): The effects of parameters changes on the relative increase in the critical value of the transmission parameter β_c for a hospital with 10-bed units compared to a hospital with wards of infinite size, both in absence of infection control measures. Right (b, d): The relative increase in the critical value of the transmission parameter β_c for different intervention measures (see Figure 3.1 for definitions) in a hospital with wards of 10 beds compared to a hospital with wards of 10 beds without any intervention strategy. $p = 0.9$, $g = 0.88$, $\xi_1 = 0.5$, $\xi_2 = 0.1$. **(a)** and **(b)**: varying γ, (percentage of the hospitalized individuals that belong to the core group). Mean length of stay is 8 days, clinical cultures leading to identification of MRSA are performed once per month on average. **(c)** and **(d)**: Varying the mean length of stay in the hospital. $\gamma = 0.5$. Clinical cultures are performed once per month on average.



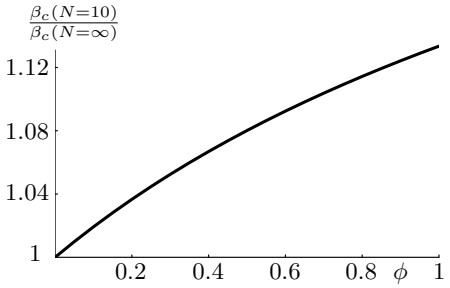
(a) No effect of the culture frequency without intervention

(b) Effect of the culture frequency with interventions

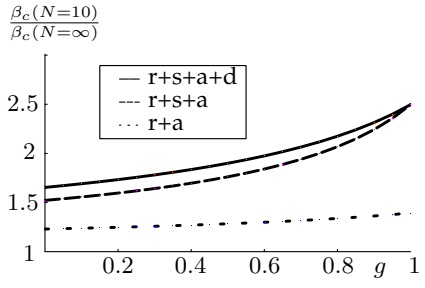
Figure 3.3: (a): Relative increase in the critical value of the transmission parameter β_c as a function of the detection rate of MRSA for a hospital with 10-bed units compared to a hospital with wards of infinite size, both in absence of infection control measures. **(b)**: Relative increase in the critical value of the transmission parameter β_c as a function of the detection rate of MRSA for different intervention measures (see Figure 3.1 for definitions) in a hospital with wards of 10 beds compared to a hospital with wards of 10 beds without any intervention strategy. $p = 0.9$, $g = 0.88$, $\xi_1 = 0.5$, $\xi_2 = 0.1$. Mean length of stay is 8 days, $\gamma = 0.5$. Without interventions, the culture frequency has no influence.

case of detection of MRSA in a unit, is more effective for typical unit sizes than screening on admission. However, combining both strategies, is far more effective and is 'guaranteed' to prevent endemicity of MRSA, even if screening of contact patients and screening on admission of high-risk patients is not 100% effective (see Figure 3.4). This explains the success of the Dutch 'search and destroy' policy. Although sporadic outbreaks occur [Vriens et al., 2002], they will not lead to the start of an epidemic. Additional decolonization of carriers is not very effective as the decolonization itself often fails.

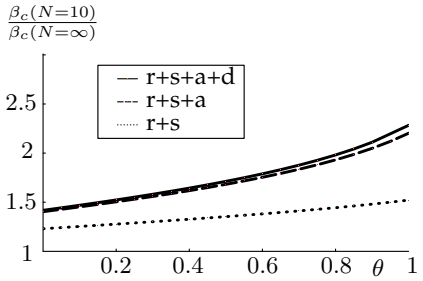
As can be seen in Figure 3.2, 3.3 3.4, effects of different intervention measures do not depend on the exact value of the parameters except on the frequency at which undetected carriers of MRSA in the hospital are detected. The conclusion in [Cooper et al., 2004], that isolation of patients identified carriers of MRSA by clinical cultures is effective, strongly depends on a sufficiently high frequency of clinical cultures. In contrast, according to our analysis, the 'search and destroy' policy is effective for very low values of the culture frequency (see Figure 3.3(b).)



(a) Additional effect of decolonization



(b) Effect of efficacy screening high-risk patients on admission



(c) Effect of efficacy screening contact patients

Figure 3.4: (a) Effects of the efficacy of decolonization on the critical transmission parameter β_c in the 'search and destroy' policy. **(b)** Effect of the efficacy of screening on admission, of patients known to be colonized with MRSA on previous admission, on the critical transmission parameter β_c for different interventions. **(c)** Effect of the efficacy of screening of contact patients in case of an detected MRSA-case in the hospital on the critical transmission parameter β_c for different interventions.

symbol	meaning
ν	probability of detection of colonization per colonized patient per day
β	transmission parameter
p	fraction of acquisitions in own unit (ignoring depletion of susceptible)
d	mean length of stay in a unit
ξ	P(still carrier on next admission)
g	efficacy screening high risk patients
ϕ	P(active decolonization successful)
θ	P(colonized contact patients is found by screening)
γ	fraction of hospitalized individuals of type 1
N	size of a unit
T	$\langle$size of an outbreak in the hospital as a whole$\rangle$ (excluding the index patient)
S	$\langle$number of detected patients during an outbreak$\rangle$
$E(1)$	$\langle$size of an outbreak in a single ward$\rangle$
$D(1)$	$\langle$infectious patient days per outbreak in a single unit$\rangle$
$Q(1)$	P(colonization of index patient detected)

Table 3.1: Parameters and symbols used in the analytical model. P stands for probability and $\langle\rangle$ stands for the average value

3.3 Simulation model

We use a discrete time stochastic simulation model for the spread of MRSA, both within hospitals and in the community at large. Our model has three hospitals of 693 beds, each with its own extramural community of 200,000 individuals. When a patients needs hospitalization, he/she is most likely to be admitted to the hospital of his/her community (p=0.95), but a small fraction (p=0.05) will be admitted in another hospital. However, after hospital discharge, all patients return to their own community. Furthermore, the population is divided into 2 categories with different health status. A small part of the extramural population (n=20,000) consists of a core-group, the majority (n=180,000) belongs to the non-core group. However, the core group represents on average half of the hospital population. (The only difference between the core group and the non-core group is the higher hospitalization rate for the core group population.) All deaths are immediately replaced by new-borns. These new-borns enter in the ex-

tramural population and belong to the same community and the same group (core group, non-core group) as the deaths, ensuring constant population sizes (see Figure 3.5). All new-borns are uncolonized with MRSA.

Within each hospital, we use a compartmentalized structure. Each hospital consists of multiple wards, each of fixed size. We distinguish between two types of wards. Each hospital has 5 intensive care units (ICU) and 36 'normal' wards consisting of 9 and 18 beds respectively. On admission, patients either go directly to an ICU or go to a 'normal' ward. The length of stay within a unit is exponentially distributed with a mean of 3 and 7 days for ICU's and 'normal' wards respectively. Patients treated in isolation have an expected length of stay of 20 days [Fitzpatrick et al., 2000]. Patients discharged from ICU's are transferred to a 'normal' ward, patients discharged from 'normal' wards return to their extramural population. Apart from these main flow of patients in the hospital, patients can be transferred between units (p=0.001 per patient per day) and between hospitals. (Patients transferred from another hospital constitute 5% of all hospital admissions, see Table 3.2.) After all these patient flows during a day have been simulated, patients are admitted from the extramural population to reach the right number of patients per unit, i.e., to compensate for the discharged patients or patients who died during hospitalization.

As most individuals carry MRSA asymptomatically, infections only represent the tip of the iceberg and, therefore, we focus on carriership of MRSA. Outside the hospital, we assume there are only two states of colonization: either an individual carries MRSA or an individual is uncolonized with MRSA (and is susceptible for colonization). Within hospitals, there are uncolonized patients and carriers of MRSA, but some hospitalized carriers are 'superspreaders', i.e., patients who are more likely to transmit MRSA to other patients and health care workers, e.g., because of infected wounds with heavy suppuration, scaling skin diseases or because patients need frequent endotracheal suction due to copious sputum production. Superspreaders are assumed to remain superspreaders during their hospitalization.

Apart from the actual colonization status, individuals are distinguished upon knowledge about their colonization status. In the extramural community individuals are distinguished upon the fact whether they were recognized as carrier during the previous hospital admission (as these patients are screened for colonization in the 'search and

destroy policy'). Within hospitals, colonized patients can be known to be colonized or not known to be colonized. In summary, the model distinguishes individuals on basis of four characteristics:

1. the colonization status of the patient and knowledge about the colonization status
2. the community to which the patient belongs
3. the health status of the patient (core group/non-core group)
4. the current location of the patient

We also include health care workers (HCW) in our model. We distinguish between two types of health care workers. The first type has only contact with patients in their own ward. The patient contacts of the second type of health care workers are not restricted to a single unit and can occur with all patients in the hospital. All health care workers stay in the unit for 8 hours and are then replaced by other health care workers. A health care worker of the first type is assumed to work always in the same unit. In the ICU's, the staff-patient ratio is 1-1, in the other wards 5 HCW are present. Apart from these HCW restricted to a single unit, there are 80 HCW's per hospital not restricted to any unit. The number of HCW's not working in the hospital is determined upon assuming that HCW work 8 hours per day for 220 days per year.

3.3.1 Dynamics of MRSA

In the extramural community, colonized individuals lose their colonization at a fixed rate (average duration of colonization in absence of readmission is 370 days). Individuals can acquire colonization in the hospital. Within the hospital, two routes of transmission exist (see Figure 3.6(b)). Patients can either transfer bacteria to other patients (via transiently colonized health care workers) or acquire colonization via persistently colonized health care workers. Patients in ICU have a higher risk for acquisition of MRSA, due to higher susceptibility for pathogen acquisition and a higher risk on transmission because of higher contact rates with HCW. Hence we assume that the transmission rates in ICU's are higher than in other units.

Transmission of MRSA is most likely to occur within a unit, but transmission to other units in the same hospital is also possible (see e.g.,

[Ball et al., 1997]). (The global mass action term is a factor 20 smaller than the mass action term within a unit (both for direct transmission (i.e., via transiently colonized HCW) and for transmission via persistently colonized HCW)).

There is no direct transmission from one hospital to another hospital, but hospitals can 'infect' each other by the transfer of a colonized patient from one hospital to another. Finally we assume that patients treated in isolation, cannot transmit or acquire MRSA.

For most strains of MRSA (and typically the hospital strains) there is little evidence of transmission in the open population and in most simulations, transmission in the open population was excluded. (As transmission of MRSA outside the hospital is rare ([Salgado et al., 2003]), we are mainly interested in the time between two admissions as in this period individuals may lose colonization. Therefore a realistic description of the open population (see e.g., [Halloran et al., 2002]) seems unnecessary.) We also performed simulations with transmission in the open population. In this case we assume homogeneous mixing within an extramural population belonging to a single hospital. Only a small fraction of acquisitions outside the hospital (p=0.05) occur between individuals of different communities (homogeneous mixing between all extramural individuals).

Patients transferred from foreign hospitals are a source of MRSA for countries with a low prevalence. We do not model foreign hospitals explicitly, but for each admission from an extramural community, there is a small probability that the admission is preceded by a stay in a foreign hospital during which MRSA may have been acquired. We neglect the duration of stay in a foreign hospital. With these assumptions, transfer of patients from foreign hospitals is a constant source of MRSA (but stochastic in nature) which will prevent extinction of MRSA.

3.3.2 Interventions

We explore several intervention scenarios. Within our model infection strategies can be targeted at different groups. General infection prevention measures like improved hand hygiene effect the risk of acquiring colonization for all patients within the unit. Other measures can be targeted at patients known to be colonized with MRSA (e.g., isolating) or at individuals who have a higher risk of being colonized, i.e., patients known to be colonized on a previous admission, patients who have

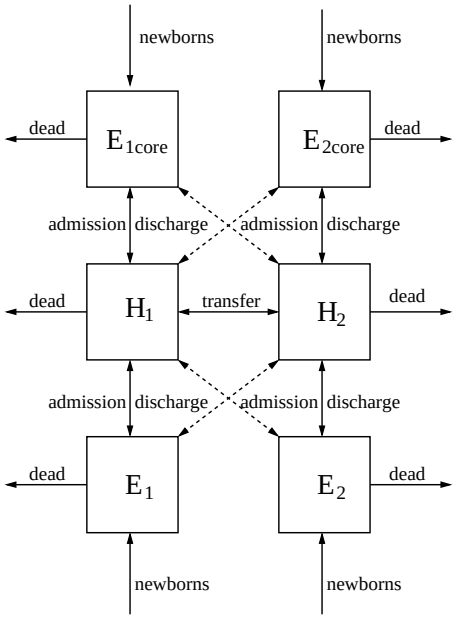
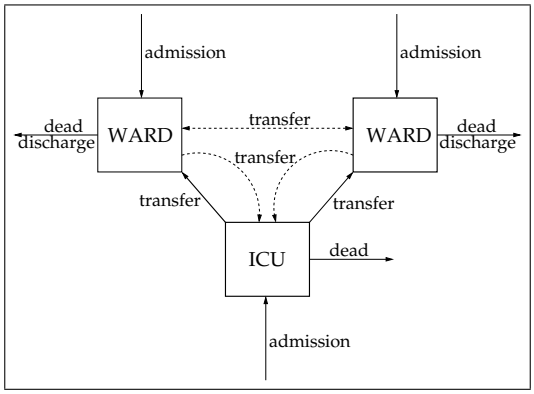


Figure 3.5: The model with 2 communities. Each hospital (denoted by an "H") has its own extramural community (denoted by an "E"). Each extramural community is divided into 2 groups: a core group who visits the hospital frequently and a non-core group with a much lower admission rate. The majority of hospitalizations are in the own hospital, however, a small fraction is hospitalized in another hospital (denoted by the dashed arrow). After discharge, patients return to their initial extramural community. For the structure within a hospital, see Figure 3.6(a)

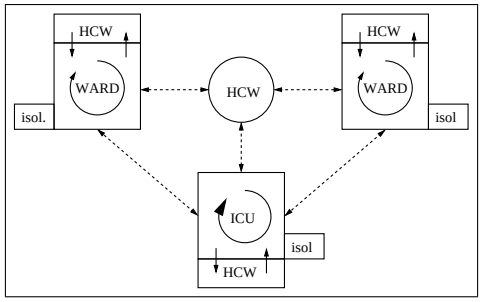
been hospitalized in a foreign hospital or members of the core group.

Without any intervention, we assume the per patient probability per day that colonization is detected by clinical cultures is constant (the constant ν of section 3.2). We consider the following intervention strategies.

- Identified MRSA carriers are treated in isolation rooms.
- High risk patients are screened for colonization on admission with efficacy 0.88, i.e., in a fraction 0.12 of the colonized patients that should be screened, either the screening is not performed or the screening does not lead to detection of the colonization.
- After detection of a colonized patient, all other patients in the



(a) patient flow in a hospital



(b) Transmission in a hospital

Figure 3.6: (a) Each hospital consists of 5 ICU's and 36 'normal' wards. Upon admission, patients either go to a 'normal' ward or to an ICU. Patients discharged from an ICU, go to a 'normal' ward, patients discharged from a 'normal' ward leave the hospital. There is also a small probability that patients are transferred from one unit (either intensive care or a 'normal' ward) to another unit. In each unit, both core-group patients and non-core group patients can be present as well as patients belonging to the community of different hospitals. Each unit has its own population of health care workers.
(b) There are two levels of transmission: within a unit and between units. Within a unit, there is transmission between patients (usually via temporarily contamined health care workers). Within a unit there is also transmission between persistently colonized health care workers and patients. Between units, transmission occurs at a lower rate and occurs again between patients (via temporarily contamined health care workers) and between patients and health care workers who do not belong to a specific unit. Colonized patients treated in isolation are assumed not to spread MRSA to other patients or health care workers. Solid arrows correspond to relatively frequent processes, dashed arrows to relatively infrequent ones.

same unit are screened for colonization.

- After detection of a colonized patients, all health care workers in the same unit are screened for colonization and positive HCW are send on leave.
- When there are two detected carriers of MRSA in an ICU, the ICU will be closed (with a maximum of 2 ICU's per hospital) and new patients are only admitted when at most one carrier is left in the unit.
- Active decolonization of patients known to be colonized is successfully performed in a fraction 0.25 of the patients.

3.3.3 Parameters

All parameters of the size and structure within the hospital and the structure in the population come from data from the UMCU and the Martini hospital Groningen (the Netherlands). The frequency of routine clinical cultures was calculated on the basis of the time between admission and detection of colonization of index-patients of (typically small) outbreaks in the UMCU (26.3 days). As some index-cases may not have been detected, we use a frequency of 0.03.

Transmission rates of MRSA within hospital settings have not been determined, specifically, transmission rates between wards and transmission rates between patients and health care workers are not well known. Although $R_0 > 1$ (as there is an endemic situation in many countries), the typical size of an outbreak in a hospital is small. Therefore the R_A-value, defined as the average number of secondary cases by one primary case when other patients are susceptible during a single hospital admission of the primary case will be less than 1 although staff deficiency may lead to an $R_A > 1$ [Grundmann et al., 2002]. We fitted values for different transmission parameters, with $R_0 > 1$ and $R_A < 1$, to the prevalence curve in the hospital and community as a function of time as observed in the US and the UK [Reacher et al., 2000]. We hypothesized that transmission between patients (via temporarily contamined health care workers) is 8 times more likely than that an uncolonized health care workers becomes a persistent carrier after contacting a colonized patient or that a persistently colonized health care worker transmits MRSA to an uncolonized patient and that transmission within a unit was three times more likely in an ICU compared

to a 'normal' ward and that superspreaders are 10 times as infectious as 'normal' colonized patients. Finally we assume that transmission within a ward is 20 times as likely as transmission between wards.

Efficacy of active decolonization is also not exactly determined. Data [Fitzpatrick et al., 2000] suggest that colonization is successfully eradicated in 25% of the cases. In the simulations with transmission in the open population, values for the basic reproduction number outside the hospital were chosen less than or equal. (Otherwise MRSA can persist without hospitals.)

With a high prevalence of MRSA, a shortage of isolation beds may well develop [Cooper et al., 2004]. In our main analysis we assume that the number of isolation beds will never be a restricting factor and wards will never be closed because of an outbreak of MRSA. Data from the Martini hospital Groningen, the Netherlands, show that the number of beds that could be used for isolation would be sufficient even if the prevalence of MRSA would reach the same level as in the US and the UK. However, in reality, the isolation capacity (especially for ICU-patients) might be less due to financial and/or staffing problems as there is an additional burden on HCW to treat patients in isolation rooms. Therefore, in countries with a high endemic level of MRSA, isolation capacity may be too small and we investigated the effect of a limited amount of isolation beds in ICU's on the efficacy of interventions.

3.3.4 Results

For each scenario, we performed 1000 simulations for a time period of 30 or 60 years. The transmission of MRSA and the HCW in the hospital are updated three times per day. All other model parts (intervention measures, discharge, dead and admission of patients, decolonization and transmission in the extramural population) are updated on a daily basis.

An initial low MRSA-prevalence

We focus on three communities with three hospitals. Indeed, without infection control measures, the epidemic curve ([Reacher et al., 2000]) as observed in the US and UK, results from our simulation (see Figure 3.7(b)). In a period of 10 years, the prevalence within the hospitals

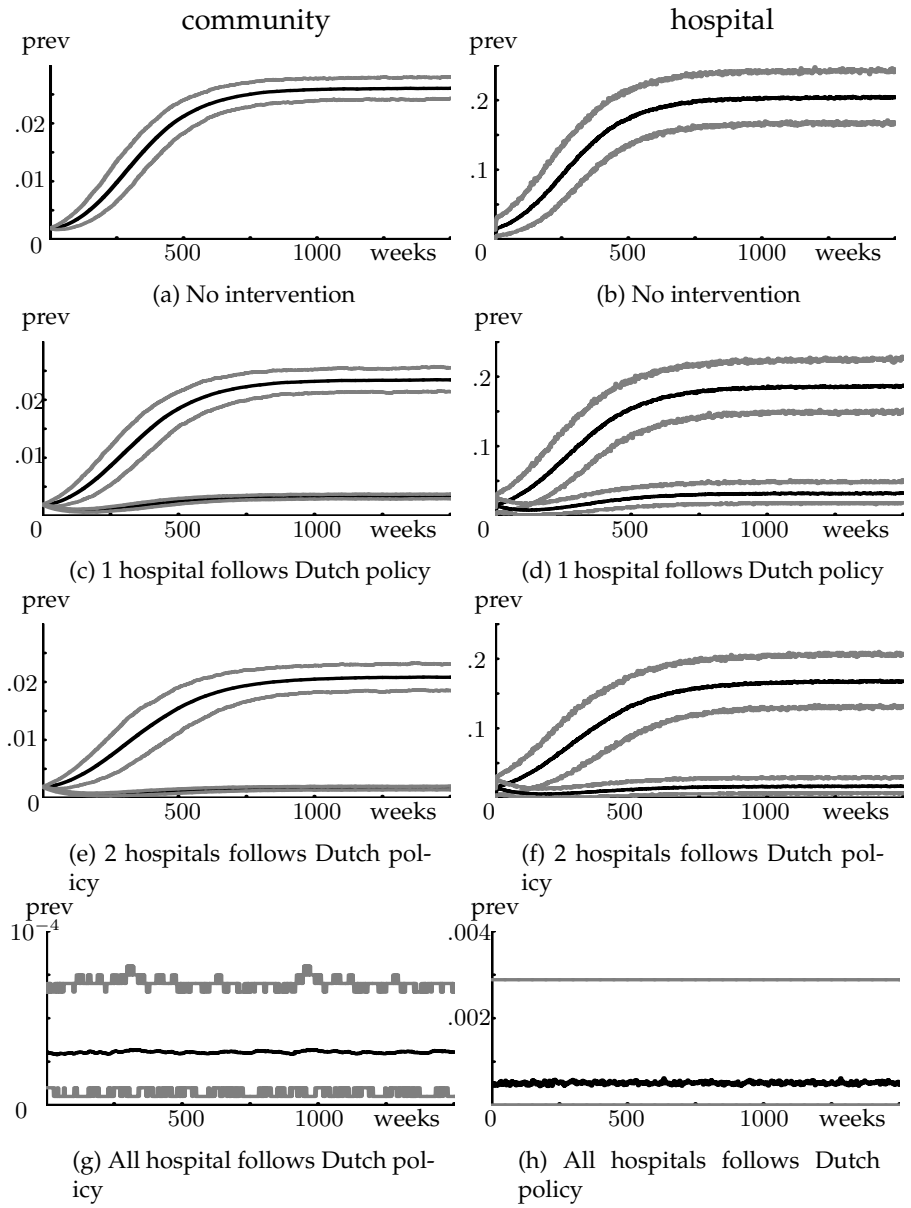
transmission parameter	(see text)[2]
fraction of transmission within unit	0.95
population size per hospital	200,000[1]
size hospital	693[1]
⟨size core group per open population⟩	20,000
⟨size core group in hospital⟩	364
#ICU per hospital	5[1]
# wards per hospital	36[1]
# beds per ICU	9[1]
# beds per ward	18[1]
# HCW per ICU	9[1]
# HCW per ward	5[1]
⟨ LOS in ICU⟩ (days)	7[1,3]
⟨ LOS in ward⟩ (days)	8[1,3]
⟨ LOS in isolation (days)⟩	20[3][Fitzpatrick et al., 2000]
recently admitted in foreign hospital	0.001[1,3]
extra infectivity superspreader	10 times higher[2,3]
P(superspreader—colonized)	0.001[2,3]
P(successful eradication)	0.25[3][Fitzpatrick et al., 2000]
efficacy screening on admission	0.88[1,3]
frequency routine cultures	0.03[1,3]
mean duration of colonization	1 year[3][Scanvic et al., 2001]
endemic level in foreign hospitals	0.1[Reacher et al., 2000]
fraction of admissions directly from other hospitals	0.05[1,3]
ratio transmission rate ICU vs transmission rate ward	3[2]
fraction dead in hospital	0.03[1,3]
isolation capacity of an ICU	1 bed
max. number ICU's closed	2

Table 3.2: Parameters used in the simulation model. With ⟨LOS⟩ we denote the average length of stay and P stands for probability. Parameters with a [1] are based on data from the UMCU and the Martini hospital Groningen (the Netherlands). For parameters with a [2], no values in literature could be found and these are chosen in such a way that without interventions, the prevalence in the hospital and in the community as a function of time resembles the curve as observed in the US and the UK [Reacher et al., 2000]. For parameters with a [3], the value is independent of whether the patient belongs to the core group or not.

increases from almost zero to more than 15% of the hospital population being colonized and the prevalence within the community at large remains low (see Figure 3.7(a)). However, within the core group a substantial percentage of the population becomes a carrier, see Figure 3.8. The fraction of colonized individuals in the extramural population is much larger in the core group than in the non-core group.

In Figures 3.7 to 3.9 the results of different intervention strategies are depicted. Applying the full Dutch 'search and destroy' policy (screening on admission of high-risk patients, isolating carriers of MRSA, screening of contact patients in case of an unexpected case of MRSA and active decolonization of carriers) ensures that the level of MRSA in the hospital remains below 5%. However, the colonization level within hospitals is influenced by the policy of other hospitals. If all hospitals use the 'search and destroy' policy, levels within the community and the hospital will be extremely low ($< 0.5\%$) with only sporadic outbreaks. With a single hospital using the 'search and destroy' policy, there will be MRSA carriers in the intervening hospital most of the time with a prevalence in the order of 3%, which is still much lower than the non-intervening hospitals. However, the effort to keep a low prevalence is large. The influx of colonized patients will be higher. Both the prevalence within the community of these neighbouring hospitals is higher and patients directly transferred from neighbouring hospitals have a higher probability of being colonized. hence. Hence more outbreaks will occur which all have to be controlled.

As execution of the 'search and destroy' policy is costly [Vriens et al., 2002], it is important to know which elements are vital. Now we start from a situation in which all three hospitals take the same intervention measures. When hospitals stop the screening of high-risk patients on admission , the prevalence both in the community (Figure 3.9(c)) and in the hospital (Figure 3.9(d)) will remain very low and the distinction with the full 'search and destroy' policy (Figures 3.7(g) and 3.7(h)) is small. If all hospitals stop the policy of screening contact patients in case of an unexpected carrier of MRSA in the hospital, levels of MRSA in the hospital and community will remain relatively low again, but are substantially higher than with the full 'search and destroy' policy due to more frequently occurring outbreaks of MRSA in the hospital. Discontinuation of active decolonization hardly effects colonization levels (data not shown) as active decolonization is not very effective anyway. The effects of transmission in the open population on the efficacy of the complete 'search and destroy' policy are shown in Figure



(a) No intervention

(b) No intervention

(c) 1 hospital follows Dutch policy

(d) 1 hospital follows Dutch policy

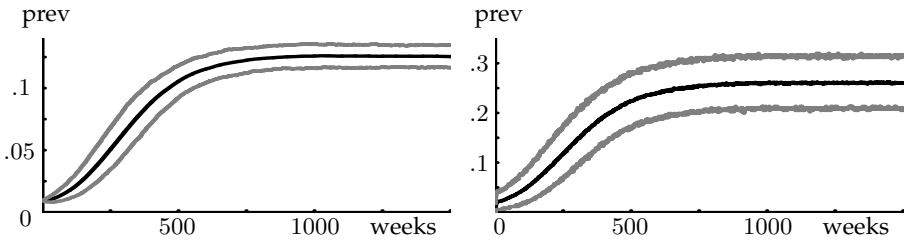
(e) 2 hospitals follows Dutch policy

(f) 2 hospitals follows Dutch policy

(g) All hospital follows Dutch policy

(h) All hospitals follows Dutch policy

Figure 3.7: An initial low prevalence. Left: the prevalence in the population. Right the prevalence in the hospitals. The black lines indicate the average behaviour of a hospital. The grey lines indicate the 90% confidence intervals. We only plotted the situation for different hospitals. The lower lines always correspond to hospitals that take intervention measures, the upper lines to a hospital that does not take any intervention measures.



(a) fraction of the core group in open population colonized

(b) fraction of the core group in the hospital colonized

Figure 3.8: The prevalence of carriership of MRSA in the core group population as function of the time when no hospitals take interventions and the initial prevalence is low. Left: the prevalence in the core population. Right the prevalence in the hospitals. The black line indicates the mean behaviour of any of the three hospitals. In 5% of the simulations, the prevalence was higher than the red line and in 5% of the simulations, the prevalence was lower than the green line.

3.10.

An initial high MRSA-prevalence

As a starting point for simulations in settings with an initial high prevalence, we used the prevalence level of a simulation after 1500 weeks starting with an initial low prevalence and without an intervention policy (see Figure 3.7(a) and 3.7(b)). With no interventions taken, the prevalence continues to fluctuate around the starting values. If all hospitals adopt the complete Dutch 'search and destroy' policy and the number of isolation beds would be sufficient, the prevalence, both in the hospitals and in the community , will decrease to extremely low values in a period of around 10 years time (see Figure 3.11(a) and 3.11(b). When there is only one isolation bed per ICU and wards are never closed in case of an MRSA outbreak, prevalence will still decrease fast (see Figure 3.11(c) and 3.11(d)). When patients known to be colonized on a previous admission (or coming from a hospital in a country with a high prevalence) are not screened on admission, the number of outbreaks will increase. However, due to the active search in case a carrier of MRSA is found in the hospital, outbreaks are unlikely to become large and still the prevalence will decrease roughly in the same time (see Figure 3.11(e) and 3.11(f)) as when all hospitals obeyed the full

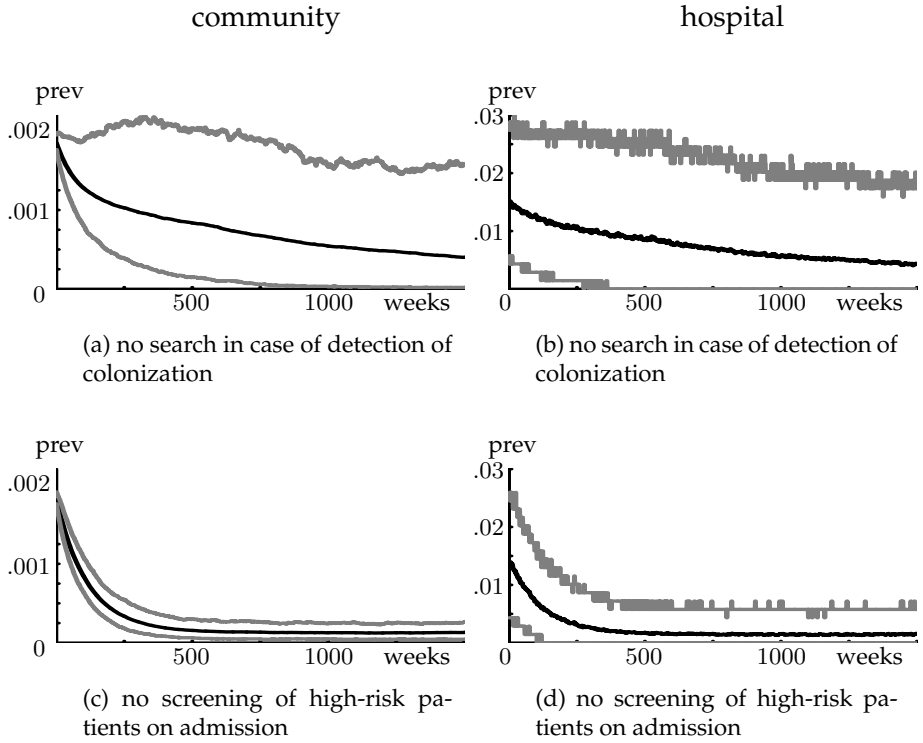
'search and destroy' policy. However, although the final prevalence of MRSA will be low, compared to the full 'search and destroy' policy the prevalence will be much higher as the number of small outbreaks will be higher. On the contrary, when patients known to be colonized on a previous admission are screened on admission but no active search is performed in case a carrier of MRSA is found in the hospital, the prevalence of MRSA will decrease but it will take considerably more time before low prevalence levels are reached (see Figure 3.11(g) and 3.11(h)).

Similarly to the case with an initial low prevalence, we are interested what will happen if only a limited number of hospitals obey a stringent policy, i.e., screening of contact patients and screening of high-risk patients on admission combined with isolation of identified carriers. Will these hospitals be able to bring down the prevalence within their hospitals to low levels despite the constant influx of carriers of MRSA from neighbouring hospitals? Results of simulations are shown in Figure 3.12). Non-intervening hospitals will hardly benefit from the intervening hospitals while the intervening hospitals can reach a relatively low prevalence of MRSA ($< 5\%$) in the hospital.

3.4 Discussion

Both the simulation model and the analytical model confirm that applying the Dutch 'search and destroy' policy is an effective way of maintaining a low prevalence of MRSA in hospitals and in the community. Either screening on admission of high-risk patients or screening of contact patients in case of a carrier of MRSA in the hospital appears to be sufficient to maintain the prevalence of MRSA at low levels. However, using a single intervention measure is far more vulnerable to changes in the dynamics, e.g., transmission in the open population, or possible simplifying assumptions in the model. Screening of contact patients appears to be more effective than screening of high-risk patients on admission. An intuitive explanation is that in case of a large outbreak in the hospital, the probability that a colonized patient is detected is far more likely than in case of a small outbreak. Therefore, screening on admission reduces the number of outbreaks while screening of contact patients stops transmission in the larger outbreaks. As active decolonization is successful in only a limited amount of the cases (p=0.2), the effect of active decolonization is only limited. However, the effect of a more

successful eradication would also be moderate. Patients identified as carrier will be recognized as such when re-admitted and screened. If decontamination has failed, they will remain isolated on their next admission and cannot spread MRSA. Identification of a core-group can be useful, as the prevalence within the extramural population is much higher in the core group compared to the non-core group. Therefore, screening of core group patients on admission, even if they are not known to be colonized on a previous admission, may be an effective infection prevention measure. An MRSA-clone that would also spread in the open population will not lead to a dramatic change in the prevalence when the R-value for transmission in the open population alone, remains below 0.5. The combined effect of spread in the open population and the spread within hospitals still leads to low endemic levels. For larger values, the prevalence in the extramural population will not remain extremely low anymore even when the Dutch 'search and destroy' policy is used while for an R_0 equal or larger than one, independent of the measures taken by hospitals, the prevalence in hospitals and in the community will reach high values. An R-value for transmission in the open population of 0.5, corresponds to an expected outbreak size of 2 in the population (1+1/2+1/4+1/8+1/16+...=2). This means that the number of MRSA-carriers in the open population doubles and the number of unidentified carriers would increase even more. As this has no major impact on the prevalence in hospitals using the Dutch 'search and destroy' policy, an increased number of patients seeking medical help abroad is unlikely to have a major influence.



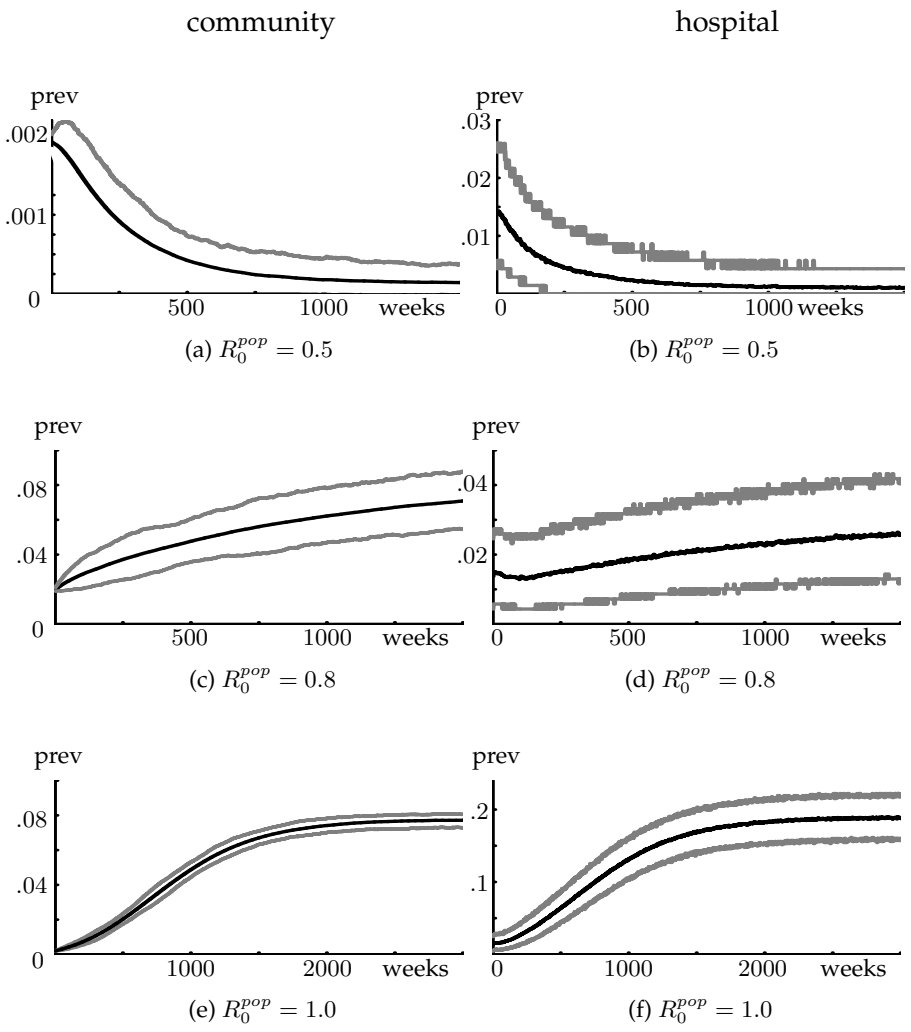
(a) no search in case of detection of colonization

(b) no search in case of detection of colonization

(c) no screening of high-risk patients on admission

(d) no screening of high-risk patients on admission

Figure 3.9: Left: the prevalence in the population. Right the prevalence in the hospitals. The initial prevalence in the hospital and the community is higher than the extreme low endemic prevalence in Figure 3.7(h) and 3.7(g). Hence we can determine the time scales to return to the extreme low prevalence values after a large outbreak has occurred. The blacks line indicate the average behaviour of any of the three hospitals. The grey lines indicate the 90% confidence intervals.

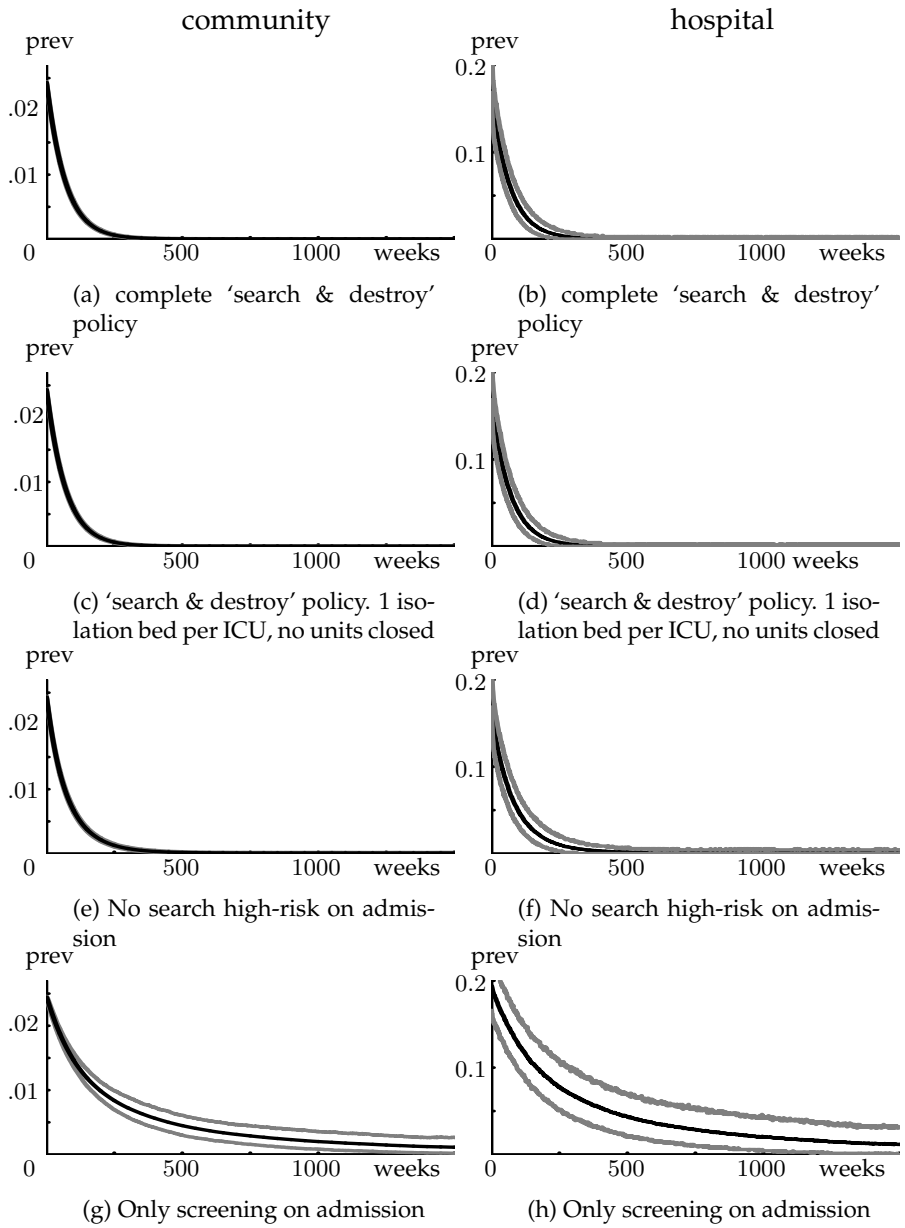


(a) $R_0^{pop} = 0.5$ (b) $R_0^{pop} = 0.5$

(c) $R_0^{pop} = 0.8$ (d) $R_0^{pop} = 0.8$

(e) $R_0^{pop} = 1.0$ (f) $R_0^{pop} = 1.0$

Figure 3.10: Transmission in the open population. Left: the prevalence in the population. Right the prevalence in the hospitals. The black line indicates the average behaviour of a hospital. The grey lines indicate the 90% confidence intervals. All hospitals obey the Dutch 'search & destroy' policy and the initial prevalence in the hospital and the community is low. The figures show the effect of different transmittability of MRSA in the open population.



(a) complete 'search & destroy' policy

(b) complete 'search & destroy' policy

(c) 'search & destroy' policy. 1 isolation bed per ICU, no units closed

(d) 'search & destroy' policy. 1 isolation bed per ICU, no units closed

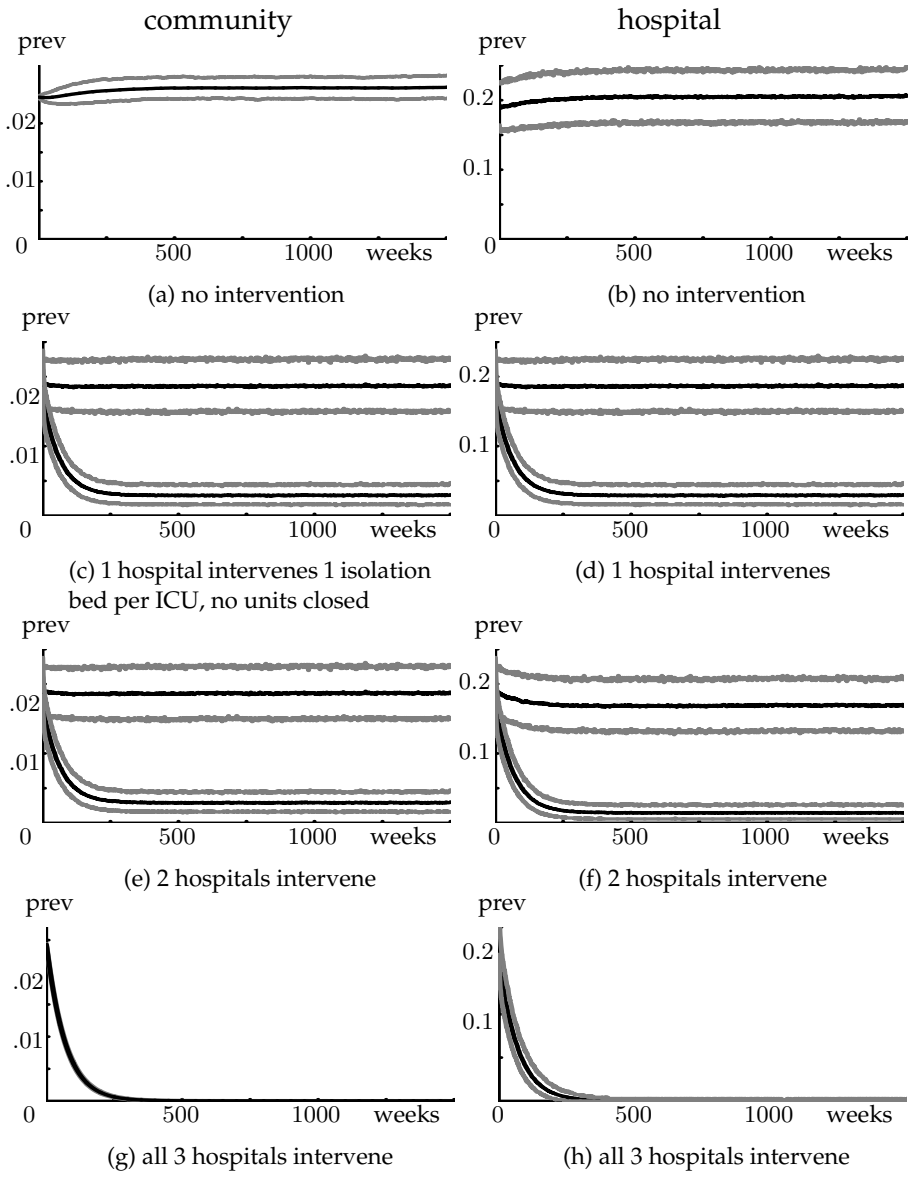
(e) No search high-risk on admission

(f) No search high-risk on admission

(g) Only screening on admission

(h) Only screening on admission

Figure 3.11: An initial high prevalence. Left: the prevalence in the population as a function of the time in weeks. Right the prevalence in the hospitals as a function of the time in weeks. The black line indicates the average behaviour of any of the three hospitals. The grey lines indicate the 90% confidence intervals.



(a) no intervention

(b) no intervention

(c) 1 hospital intervenes 1 isolation bed per ICU, no units closed

(d) 1 hospital intervenes

(e) 2 hospitals intervene

(f) 2 hospitals intervene

(g) all 3 hospitals intervene

(h) all 3 hospitals intervene

Figure 3.12: An initial high prevalence. Left: the prevalence in the population. Right: the prevalence in the hospitals. The intervening hospitals follow the Dutch 'search & destroy' policy but there is only 1 isolation bed per ICU and no units will be closed. A black line indicates the average behaviour of any of the three hospitals. The grey lines indicate the 90% confidence intervals. The lower lines always correspond to hospitals that take intervention measures, the upper lines to a hospital that does not take any intervention measures.

Part II

Data Analysis

Chapter 4

Independent patients

This work resulted from a brain storm session during which the observation was discussed that the probability for an event to happen to a patient depends on the length of stay of the patient in the unit. Therefore, even if acquisition of colonization with a certain microorganism is harmless for a patient, the mean length of stay of patients who acquired colonization will be longer than the mean length of stay of patients who did not acquire colonization. This can be shown by some very simplistic deterministic formulas (moment estimators). The simple deterministic formulas, however, are not optimal and a maximum likelihood procedure gives smaller asymptotic confidence intervals for the parameters. Testing the methods on real data led to all kind of problems. Assumptions like a constant discharge and infection rate, i.e., independent of the duration of stay in the unit, are not always satisfied. Moreover, taking into account the moments of detection of infection or colonization, could lead to different predictions. Patients were also assumed to be identical on admission, neglecting heterogeneity. The problems observed in this Chapter led to the model described in Chapter 5, in which no assumptions are made about the length of stay of patients but the observed lengths of stay are used and the culture moments are explicitly taken into account.

4.1 Introduction

Critically ill patients in intensive care units (ICU) often suffer from infections. In a one-day point-prevalence study including 10,038 patients

in 1,417 ICU in 17 European countries, 45% of the patients were found to suffer from one or more infections [Vincent et al., 1995], notably ventilator associated pneumonia (VAP) and bacteremia 64.7% and 18% respectively. For these two types of infections (as well as for others) one finds that, on average, the patients that become infected have a longer length of stay (LOS) in the ICU than patients that do not become infected [Heyland et al., 1999]. It is tempting to interpret this correlation as the infection causing an increased LOS. One should realize, however, that the risk of getting infected increases with LOS. So it may also be the other way around: a long LOS may have as an effect that one gets infected. Finally, there might be heterogeneity among patients such that some will, with higher probability, both stay longer and become infected. In order to decide about the potential effect of prevention measures one would like to disentangle how these various mechanisms contribute to the observed correlation (note that all three mechanisms may operate simultaneously).

Detailed case-control studies based on careful matching of patients on the basis of a multitude of traits [Fagon et al., 1993; Rello et al., 1996; Heyland et al., 1999; Timsit et al., 1996; Bonten et al., 1997; Schulgen, 1995] can lead to reliable conclusions. However, they are labour intensive (and therefore costly) and consequently not performed routinely. The aim of the present note is to propose a test that uses only standardly available data. The test ignores heterogeneity, so clinicians should separately wonder whether or not heterogeneity may play a big role. The outcome of the test provides us with a first rough estimate of attributable extra stay and as such should guide our thinking about the need for prevention measures and/or more costly statistical investigations.

We first analyzed data on intestinal colonization with Gram-negative bacteria resistant to third generation cephalosporins, which is considered harmless and for which we suspected that it will not prolong ICU-stay, and next for ICU-acquired bacteremia, for which we were quite confident that it prolongs ICU-stay. Finally, we tested the hypothesis that VAP prolongs ICU-stay.

4.2 Model assumptions and estimation

We adopt a Markov model for competing risks [Andersen et al., 1993] concerning a population of patients in an ICU. The event 'discharge'

happens to all of them (so in case of 'death' we also speak about 'discharge'; more about this in Appendix 4.6.3) and we know when. In other words, for every patient the LOS is recorded (say as an integer number of days). The event 'infection' may or may not happen, but if it happens, it happens before 'discharge'. In our basic model, we do not know when 'infection' happens, but we assume that at the time of 'discharge' it is known whether or not 'infection' has happened. We label a patient I if infection occurred and U otherwise. So the data consist of an integer (LOS) and a label (I or U) for every patient.

The Markov model involves three *unknown* parameters, d, d' and f. Their interpretation is as follows. While being 'uninfected' a patient has probability per unit $1/d$ (=λ) and f of being, respectively, discharged or infected. And an 'infected' patient has probability per unit of time $1/d'$ (=μ) of being discharged. In other words, in the absence of a force of infection (i.e., for $f = 0$) the sojourn times in the ICU is exponentially distributed with mean d and infected patients have an exponentially distributed remaining sojourn time with mean d'. Hence one can call $d' - d$ the attributable LOS. Our task is to estimate the parameters d, d' and f from the data. The statistical sound way to do so is Maximum Likelihood Estimation (MLE). In Appendix 4.6.1, we shall do exactly this. In Appendix 4.6.2, we introduce a 'deterministic' simplification leading to the explicit formulas for the estimators f^*, d^* and d'^*

$$f^* = \frac{P_I}{\overline{LOS(U)}} \tag{4.1}$$

$$d^* = \frac{\overline{LOS(U)}}{1 - P_I} \tag{4.2}$$

$$d'^* = \overline{LOS(I)} - \overline{LOS(U)} \tag{4.3}$$

where by definition

P_I	=	fraction of patients that leave the ICU infected
$\overline{LOS}(I)$	=	the average LOS of patients labeled I
$\overline{LOS}(U)$	=	the average LOS of patients labeled U

We emphasize that (4.1), (4.2) and (4.3) provide us with a recipe to calculate an estimate for attributable LOS, using readily available data as input and a (pocket) calculator as the only tool.

In Appendix 4.6.4 and 4.6.5 we substantiate that in replacing the far more laborious MLE by (4.1), (4.2) and (4.3), one doesn't make a big error. Moreover, we show how to derive confidence intervals for the various estimates, within the framework of the model (note that these

intervals don't tell anything about the reliability of the model itself; in particular they cannot serve to justify the neglect of heterogeneity or the assumption of constant discharge and infection rates).

Both the deterministic estimate and the MLE assume that the classification according to the categories I and U is perfect. When infection does not lead to visible changes in the status of a patient, misclassification can occur when the infection status of patients who are not known to be 'infected' is not checked at discharge. Moreover, when patients are cultured during their stay to determine the infection status, additional information about the moment of infection is available which can be used. The moments at which patients are cultured can be taken into account in either a deterministic way (section 4.6.9) or in a likelihood-based approach (section 4.6.10), although the deterministic way loses its main advantage as no explicit formulas can be obtained. The incorporation of possible misclassification, by taking the exact moments of culturing into account, can be important as it may lead to major changes in the prediction for the attributable length of stay.

4.3 Data

The model was used for three different infectious complications: Colonization with Gram-negative bacteria resistant to third generation cephalosporins, bacteremia and VAP. A cohort of 474 patients, consecutively admitted to two ICUs of the University Medical Center Utrecht between August 2001 and May 2002 [Nijssen et al., 2002], was used for colonization with Gram-negative bacteria resistant to third generation cephalosporins. During this period, on admission and every Monday and Thursday, patients were routinely screened for intestinal colonization with these bacteria by obtaining rectal cultures and using standard microbiological analyses. A finding of colonization with Gram-negative bacteria resistant to third generation cephalosporins, was not reported to the responsible physicians. For 17 of the 474 patients no culture results were available, typically patients with a short length of stay, and these patients were excluded from analysis.

For bacteremia, a cohort of 974 patients, consecutively admitted to the same two ICUs of the University Medical Center Utrecht between January 2000 and May 2001, was used. Bacteremia was defined as a blood culture growing with Gram-negative bacteria, *Staphylococcus aureus* or yeasts. There was no protocol for obtaining microbiological

surveillance blood cultures. All cultures were obtained for diagnostic purposes on demand of responsible physicians. Fever or other signs of infection usually are the incentive to obtain cultures for microbiological diagnosis.

For the analyses regarding VAP, three different patient cohorts were used. Two cohorts were from prospective randomized controlled studies on prevention of VAP [Bergmans et al., 2001; Fiore et al., 2000] (179 and 221 patients respectively) and the third cohort was similar to the patient cohort used for bacteremia. In the two intervention studies VAP was diagnosed according to a set of clinical criteria in combination with a positive quantitative culture of bronchoalveolar lavage fluid (cutoff point $\geq 10^4$ cfu/ml). In the non-experimental cohort VAP was diagnosed according to clinical criteria only. To guarantee that VAP is ventilator-associated, the clinical definition of VAP excludes episodes of VAP within the first two days of ICU-stay. Therefore, all patients who stayed at most two days in ICU or developed pneumonia during the first two days were excluded. This requires some changes in the parameter estimations (see Section 4.6.10).

In case of bacteremia or VAP, patients receive specific antimicrobial treatment. Often, the data allow for a distinction of patients on the basis of more categories than just I and U. Since infectious complications can increase mortality it may be important to distinguish in particular between surviving and non-surviving patients see Section 4.6.3. However, in general a more detailed description will both reduce the number of patients in each category and increase the number of parameters to be estimated, so the results will be statistically less reliable.

4.4 Results

4.4.1 Colonization with antibiotic resistant Gram-negative bacteria

In all, 424 patients were included in the analysis, of whom 37 developed colonization (33 patients where colonized on admission). The naive application of the deterministic estimates (equations (4.1), (4.2) and (4.3)) and the application of the MLE-estimates 4.20, gives almost identical

results. We find

$$\begin{array}{llll} d = 6.5 \pm 0.3 & d' = 13.5 \pm 2.2 & f = 0.015 \pm 0.002 & d' - d = 7.0 \pm 2.5 \\ d = 6.5 \pm 0.3 & d' = 13.3 \pm 2.4 & f = 0.015 \pm 0.002 & d' - d = 6.8 \pm 2.5 \end{array} \tag{4.4}$$

for the deterministic and the MLE-estimates respectively. We used Section 4.6.4 to obtain the confidence intervals. A maximum likelihood analysis, including the culture moments (see Section 4.6.10) and including the 33 patients colonized on admission, leads to different estimates:

$$d = 6.9 \pm 0.4 \quad d' = 9.0 \pm 0.8 \quad f = 0.030 \pm 0.004 \quad d - d' = 2.0 \pm 0.9 \tag{4.5}$$

(The inclusion of patients colonized on admission did not alter the estimates significantly.) Although acquisition of colonization with antibiotic resistant Gram-negative bacteria still seems to prolong length of stay, the estimated attributable length of stay due to colonization is much smaller than in the naive approach neglecting the moments of culturing.

Bacteremia

Of 974 patients included in the analysis for bacteremia, 67 had at least one positive blood culture (7%). As bacteremia leads to clinical symptoms like fever, which are the incentive for clinical cultures, misclassification will not be important. The basic method, not including the culture moments, gives:

$$\begin{array}{llll} d = 6.7 \pm 0.2 & d' = 16.6 \pm 2.0 & f = 0.011 \pm 0.0014 & d - d' = 11.1 \pm 2.2 \\ d = 6.7 \pm 0.2 & d' = 17.1 \pm 2.2 & f = 0.011 \pm 0.0014 & d - d' = 10.5 \pm 2.2 \end{array} \tag{4.6}$$

for the deterministic and the MLE method respectively. Including the information of the moment of detection of bacteremia (see Section 4.6.10), we find:

$$d = 5.9 \pm 0.2 \quad d' = 13.9 \pm 1.5 \quad f = 0.013 \pm 0.002 \quad d - d' = 8.0 \pm 1.5 \tag{4.7}$$

These methods indicate that d is significantly smaller than d', which confirms our hypothesis that bacteremia prolongs LOS. To test whether bacteremia also increases the death rate, we distinguished between death and discharge (see Section 4.6.3. We find:

$$d^{death} = 29 \pm 2 \quad d^{dis} = 7.3 \pm 0.3 \quad d'^{death} = 71 \pm 20 \quad d'^{dis} = 17 \pm 2 \tag{4.8}$$

colonization		no cult.		neg. cult		acq.col		pos adm.	
	#	#	$\overline{LOS}$	#	$\overline{LOS}$	#	$\overline{LOS}$	#	$\overline{LOS}$
all pat.	474	17	1.53	387	5.9	37	19.5	33	10.3

Table 4.1: Data for colonization with Gram-negative bacteria resistant to third generation cephalosporins. Patients are divided according to the results of the cultures, viz., patients for which no culture results was available, patients with only negative cultures, patients for which the first culture was negative but with at least one positive culture and patients for which the first culture was positive. $\overline{LOS}$ stands for the average length of stay of the patient group. # stands for the number of patients.

bacteremia		pos.cult.		neg. cult.	
	#	#	$\overline{LOS}$	#	$\overline{LOS}$
all pat.	974	67	22.9	907	6.2
surv	810	56	20.2	754	6.2
non-s.	164	11	36.2	153	6.1

Table 4.2: Data for bacteremia. All patients with a positive blood culture for Gram-negative bacteria, *Staphylococcus aureus* or yeasts were labeled as I. Positive blood cultures with other microorganisms or absence of positive cultures were labeled as U. $\overline{LOS}$ stands for the average length of stay of the patient group. # stands for the number of patients.

where d^{death} and d^{dis} are the inverse of the respective rates. (The infection rate remains $f = 0.013 \pm 0.02$.) The calculation of the mortality rates suggests that bacteremia decreases per diem mortality. However, overall mortality is independent of acquisition of bacteremia.

Ventilator-associated pneumonia

Incidence rates of VAP in the three studies were $48/179$ (27%), $21/190$ (11%) and $137/1090$ (13%). (The denominator is the number of patients with a length of stay of more than two days.) In all studies the mean LOS of infected patients was considerably longer (ranging from 22 to 28 days) than the mean LOS of non-infected patients (ranging from 10 to 13 days).

For the first and second study, per patients data about length of stay, whether or not the patient developed VAP and whether or not the

patient died during hospitalization were available. For study 3, also the day of acquisition of VAP was known. In the first study only patient who stayed more than 2 days in the ICU and were included. To compare results, in the other two studies patients who stayed at most 2 days in ICU were excluded from analysis. Furthermore, to prevent that a small number of patients with a long LOS (outliers) affect the results considerably, we restricted attention to those patients that stayed at most 65 days in the ICU. (The exact value does not influence the results as long as the outlier with a length of stay of 281 is excluded from analysis in study 2). We incorporated these restrictions on the LOS in the deterministic model. In principle our analysis remains the same, but the computations become more involved (see Section 4.6.8). These do not yield explicit formulas for d and d' and f, but they can be used to compute these quantities numerically. A bootstrap analysis (1000 simulations) was used to determine confidence intervals. We obtained:

$$\begin{array}{ll} d = 14.1(11.8; 16.9) & d' = 20.9(14.1; 43.0) \\ d = 12.4(10.5; 14.5) & d' = 8.6(2.5; 18.3) \\ d = 7.81(7.18; 8.52) & d' = 14.18(9.88 : 19.89) \end{array} \tag{4.9}$$

for study 1,2 and 3 respectively and the per diem probability to acquire VAP was 0.010, 0.015 and 0.017 for study 1, 2 and 3 respectively. The discharge rate for patients with VAP was higher than the discharge rate of patients without VAP in 90.1%, 17.5% and 100% of the simulations. When we distinguished between death and discharge (see Section 4.6.3), the MLE for the discharge rates λ and μ for uninfected and infected patients respectively are for study 1 and 2:

$$\begin{array}{llll} \lambda_{dis} = 0.057 & \lambda_{death} = 0.023 & \mu_{dis} = 0.088 & \mu_{death} = 0.027 \\ \lambda_{dis} = 0.044 & \lambda_{death} = 0.025 & \mu_{dis} = 0.029 & \mu_{death} = 0.019 \\ \lambda_{dis} = 0.099 & \lambda_{death} = 0.029 & \mu_{dis} = 0.047 & \mu_{death} = 0.023 \end{array} \tag{4.10}$$

(An analysis for study 3 based on a MLE using the moments of acquisition gave similar results.) d' appeared to be considerably larger than d in two studies. The per diem relative risk for developing VAP differed for the three studies, probably due to a different patient population (see Discussion). Calculation of the per diem mortality suggested that VAP decreases per diem mortality. However, due to the longer stay, the overall ICU-mortality is higher for patients who acquired VAP.

			VAP		No VAP	
study		#	#	$\overline{LOS}$	#	$\overline{LOS}$
1	all pat.	179	48	28.31	131	12.95
	surv	112	29	29.53	83	14.52
	non-s.	67	19	26.00	48	10.23
2	all pat.	190	21	21.86	169	10.6
	surv.	136	16	22.63	120	14.70
	non-s.	54	5	19.40	49	11.67
3	all pat.	1090	137	25.0	953	9.98
	surv.	474	43	25.23	431	10.59
	non-s.	150	21	17.43	129	7.91

Table 4.3: VAP data for the three studies. Only patients with $3 \leq LOS \leq 65$ were included. $\overline{LOS}$ stands for the average length of stay of the patient group while # stands for the number of patients.

4.5 Discussion

In principle, the method described allows us to distinguish on straightforward mathematical grounds between cause and effect regarding development of nosocomial infection on the one hand and LOS on the other hand. However, the simple deterministic formulas (4.1), 4.2) and (4.3) should be used with caution and serve primarily as a thought experiment. The MLE-method, based on the same assumptions, lacks the advantage of simple formulas and performs hardly better, see also Section 4.6.4. Apart from that, there are problems to prove uniqueness of the MLE estimates (see section 4.6.6 and 4.6.7). When information is available about the moment of acquisition during the stay, this information should be used. Specifically when misclassification can occur, incorporating the moments at which cultures are performed is essential.

To test the performance on real data, we first considered two infections for which the outcome is predictable. In ICU-patients, intestinal colonization with Gram negative bacteria resistant to third generation cephalosporins is generally believed to be in itself harmless. Moreover, presence or absence of colonization was not reported to physicians and did not lead to changes in treatment. In contrast, bloodstream infections are generally believed to prolong the LOS.

The results of the analysis agreed quite well with these presump-

tions and for VAP, for which no generally accepted view on its influence on LOS exists, our methods cannot solve this issue. Two of the three studies suggest that development of VAP prolongs LOS with approximately 7 days. In the other study development of VAP seems to shorten LOS although this result is not statistically significant (see below for discussion).

We consider these rough analyses as a first proof of principle of the applicability of the model, but further progress requires that we debate the underlying assumptions and do not close our eyes for the possibility that they are at variance with reality.

First, we have used exponential distributions. This means that we assume a constant discharge rate. This has the advantage that the parameters are few and have a clear medical interpretation (discharge rates), but the data themselves do not perfectly reflect the exponential distribution. Specifically, the number of patient with a short LOS and the number of patients with a long LOS is too high to fit an exponential distribution well. (A χ^2-test used for testing whether the LOS of uninfected patients followed an exponential distribution resulted in p-values of 0.02 and 0.005 for colonization with Gram-negative bacteria and bacteremia.)

Second, the distribution of the LOS of patients who died in the ICU is not equal to that of those who did not, while the simplest version of the model assumes they are. This might indicate that we should introduce some heterogeneity in the population. These two points are relatively easy to incorporate in the model, but at the expense of introducing new parameters which often do not have such a clear medical interpretation. The problem of different LOS of survivors and non-survivors can also be solved by introducing an extra state for patients. All patients arrive in a state with relatively high mortality and have a fixed probability per day to proceed to the second state.

Third, we have assumed a constant force of infection over time. This means that the method can only be used for infections where transmission between patients is unimportant see Section 5 (i.e. the model does not work in an outbreak setting, when the force of infection depends on the number of infected patients).

Fourth, we assume a constant vulnerability for infection during the stay. According to our data, this is a reasonable assumption for colonization with Gram-negative bacteria resistant to third generation cephalosporins and bacteremia (data not shown), but this may not

apply for VAP. In our patients, the probability to develop VAP was highest in the first week of ICU-stay (data not shown) which confirms data reported by [Cook et al., 1998].

In addition to uncertainties concerning the mathematical approach, there are potential imperfections in the data. We assume that directly after acquisition, the colonization can be detected. In reality colonization can only be detected with, an unknown, delay. In addition, all tests used for microbiological analysis have a lowest detection limit. Therefore some misclassification might have influenced our data. Moreover, diagnosing VAP is extremely difficult and there is no clinical useful gold standard for diagnosis. [Fagon et al., 1993]. In most ICUs VAP is diagnosed upon a set of clinical, microbiological and radiological criteria, which have a high sensitivity but a low specificity. These criteria were used in study 3. Addition of invasive techniques (e.g. bronchoscopy) with quantitative cultures of samples obtained increases specificity, reducing the number of false positive cases of VAP. These criteria were used in study 1 and 2. The different diagnostic criteria that were used may have influenced our findings. Moreover, the incidence rates are also influenced by patient selection. For example, diagnostic criteria were identical for study 1 and 2, yet incidence rates were much lower in study 2. In this study many patients admitted to ICU because of head trauma were excluded. As these patients have an extremely high risk for VAP, exclusion reduced the overall incidence (Table 4.3) of infection.

Whenever the test indicates that the infection is the cause of a prolonged stay, one should still interpret the word 'cause' with care, as there may exist confounding factors that are the true cause. Yet the possibility that the data are explained by the increased probability to contract the infection when LOS is long, has been ruled out.

Despite these potential imperfections, the method has advantages compared to a matched case-control study. The latter usually obtains cases and controls from large databases, but due to matching criteria relatively few patients are included in the final analysis. With the present model, complete databases can be used and good statistical power can be achieved. Moreover, no additional information is needed (except development of infection, LOS and, maybe, whether the patient survived or not), making it suitable for many existing databases. As the method is easy, fast and inexpensive, it provides an interesting tool to investigate the relationship between cause and effect in this part

of medicine.

4.6 Appendix

4.6.1 Maximum Likelihood Estimate

Suppose we look at a ward, and we are interested in a certain infection. If the force of infection is constant, and the discharge rate only depends on whether a patient is infected or not, we try to estimate the following parameters:

- infection rate=f.
- discharge rate of a non-infected patient=λ.
- discharge rate of an infected patient=μ.

given that we know for each patient the length of stay (LOS) and whether (s)he became infected or not. Let τ denote the time elapsed since a patient entered the ICU and let $\mathcal{G}_U(\tau)$ denote the probability that this patient is both still in the ICU and non-infected after time τ. According to the model, $\mathcal{G}_U(\tau)$ satisfies the following differential equation:

$$\begin{array}{rl} \frac{d}{d\tau}\mathcal{G}_U(\tau) & = -(\lambda+f)\mathcal{G}_U(\tau) \\ \mathcal{G}_U(0) = 1. & \end{array} \tag{4.11}$$

Hence we have:

$$\mathcal{G}_U(\tau) = e^{-(\lambda+f)\tau}. \tag{4.12}$$

The density for the length of stay, given that a patient did not become infected ($f_U(\tau)$), is proportional to $\lambda\mathcal{G}_U(\tau)$. The constant of proportionality serves to achieve the normalization $\int_0^\infty f_U(\tau)d\tau = 1$. Consequently we have:

$$f_U(\tau) = (\lambda+f)e^{-(\lambda+f)\tau}. \tag{4.13}$$

Let $\mathcal{G}_I(\tau)$ denote the probability that this patient is infected and still in the ICU after time τ. $\mathcal{G}_I(\tau)$ satisfies the following differential equation:

$$\begin{array}{rl} \frac{d}{d\tau}\mathcal{G}_I(\tau) & = f\mathcal{G}_U(\tau) - \mu\mathcal{G}_I(\tau) \\ \mathcal{G}_I(0) & = 0. \end{array} \tag{4.14}$$

The solution of this differential equation is given by:

$$\mathcal{G}_I(\tau) = \frac{f}{\lambda+f-\mu}\left(e^{-\mu\tau} - e^{-(\lambda+f)\tau}\right) \tag{4.15}$$

The density for the length of stay, given that a patient became infected ($f_I(\tau)$), is proportional to $\mu\mathcal{G}_I(\tau)$. The constant of proportionality serves to achieve the normalization $\int_0^\infty f_I(\tau)d\tau = 1$. Hence we have:

$$f_I(\tau) = \frac{\mu(\lambda+f)}{\lambda+f-\mu}\left(e^{-\mu\tau} - e^{-(\lambda+f)\tau}\right)\left(= \frac{\lambda+f}{f}\int_0^\tau f e^{-(\lambda+f)t}\mu e^{-\mu(\tau-t)}dt\right) \tag{4.16}$$

If we have $N+M$ patients, of which M became infected, and we order the patients such that the first N did not become infected, the likelihood becomes:

$$\begin{aligned} L = \quad & \binom{N+M}{M}\left(\tfrac{f}{f+\lambda}\right)^M\left(\tfrac{\lambda}{f+\lambda}\right)^N \\ & \prod_{j=1}^N(\lambda+f)e^{-(\lambda+f)T_j}\prod_{i=N+1}^{N+M}\tfrac{\mu(\lambda+f)}{\lambda+f-\mu}\left(e^{-\mu T_i} - e^{-(\lambda+f)T_i}\right) \end{aligned} \tag{4.17}$$

and the log likelihood $l = \log L$ can be written as:

$$\begin{aligned} l = \quad & \log\binom{N+M}{M} + M\log f + N\log\lambda + M\log\mu \\ & -\textstyle\sum_{j=1}^N(\lambda+f)T_j + \sum_{i=N+1}^{N+M}\log\left(\frac{e^{-\mu T_i}-e^{-(\lambda+f)T_i}}{\lambda+f-\mu}\right) \end{aligned} \tag{4.18}$$

To find the MLE-estimator, we solve the equations:

$$\begin{aligned} \tfrac{\partial}{\partial\lambda}l &= \tfrac{N}{\lambda} - \tfrac{M}{\lambda+f-\mu} - \textstyle\sum_{j=1}^N T_j + \sum_{i=N+1}^{N+M}\frac{T_i e^{-(\lambda+f)T_i}}{e^{-\mu T_i}-e^{-(\lambda+f)T_i}} \\ &= \tfrac{N}{\lambda} - \tfrac{M}{\lambda+f-\mu} - \textstyle\sum_{j=1}^{N+M} T_j + \sum_{i=N+1}^{N+M}\frac{T_i e^{-\mu T_i}}{e^{-\mu T_i}-e^{-(\lambda+f)T_i}} = 0 \\ \tfrac{\partial}{\partial f}l &= \tfrac{M}{f} - \tfrac{M}{\lambda+f-\mu} - \textstyle\sum_{j=1}^{N+M} T_j + \sum_{i=N+1}^{N+M}\frac{T_i e^{-\mu T_i}}{e^{-\mu T_i}-e^{-(\lambda+f)T_i}} = 0 \\ \tfrac{\partial}{\partial\mu}l &= \tfrac{M}{\mu} + \tfrac{M}{\lambda+f-\mu} - \textstyle\sum_{i=N+1}^{N+M}\frac{T_i e^{-\mu T_i}}{e^{-\mu T_i}-e^{-(\lambda+f)T_i}} = 0 \end{aligned} \tag{4.19}$$

This system of equations can be written as:

$$\begin{aligned} & fN = \lambda M \\ & \tfrac{M}{f} + \tfrac{M}{\mu} = \textstyle\sum_{j=1}^{N+M} T_j \\ & \tfrac{M}{\lambda+f-\mu} + \tfrac{M}{\mu} = \textstyle\sum_{i=N+1}^{N+M}\frac{T_i e^{-\mu T_i}}{e^{-\mu T_i}-e^{-(\lambda+f)T_i}} \end{aligned} \tag{4.20}$$

4.6.2 Deterministic estimates

As the system of equations characterizing the MLE-estimates cannot be solved analytically, we first investigate an estimate for which there is an analytical expression. It turns out to be easier to work with the inverses of the discharge rates:

- $d = \frac{1}{\lambda}$
- $d' = \frac{1}{\mu}$

We introduce the following observable quantities:

$$P_I = \frac{M}{M+N} \tag{4.21}$$

$$\overline{LOS}(U) = \frac{\sum_{i=1}^{N} T_i}{N} \tag{4.22}$$

$$\overline{LOS}(I) = \frac{\sum_{i=N+1}^{N+M} T_i}{M} \tag{4.23}$$

and compare them with their expectation.

The expectations can easily be expressed in the parameters f, d and d'. Let τ denote the time elapsed since a patient entered the ICU and let $\mathcal{G}_U(\tau)$ denote the probability that this patient is both still in the ICU and non-infected after time τ. Then according to the assumptions:

$$\mathcal{G}_U(\tau) = 1 - \int_0^\tau f_U(\sigma)d\sigma = e^{-\left(\frac{1}{d}+f\right)\tau} \tag{4.24}$$

As the probability per unit of time that the patient becomes infected (given that (s)he is still uninfected and in the ICU) equals $f\mathcal{G}_U(\tau)$ and $E(P_I)$ equals the integral with respect to τ we have:

$$E(P_I) = f\int_0^\infty \mathcal{G}_U(\tau)d\tau = \frac{f}{\frac{1}{d}+f} = \frac{fd}{1+fd} \tag{4.25}$$

By definition

$$E\left(LOS(U)\right) = \int_0^\infty \tau f_U(\tau)d\tau \tag{4.26}$$

Consequently

$$E\left(LOS(U)\right) = \int_0^\infty \tau\left(\frac{1}{d}+f\right)e^{-\left(\frac{1}{d}+f\right)\tau}d\tau = \frac{d}{1+fd} \tag{4.27}$$

To compute $E\left(LOS(I)\right)$, we first observe that the expected duration of the period in which the patient is not yet infected, given that (s)he is infected before being discharged, is equal to $E\left(LOS(U)\right)$. The reason is that the corresponding probability density function is proportional to $f\mathcal{G}_U$ and, therefore, equal to f_U. The second observation is that the stay of a patient that leaves infected is naturally subdivided into two

periods according to the moment of becoming infected. The second period is exponentially distributed with parameter $1/d'$, so with expected duration d'. We conclude that:

$$E\left(LOS(I)\right) = E\left(LOS(U)\right) + d' \tag{4.28}$$

We now have three relations, (4.25), (4.27) and (4.28), between the expectation of the quantities (4.21,4.22,4.23) and the parameters f, d and d'. These can easily be solved to obtain expressions for the parameters in terms of the three expressions. If we replace the expectations by the observed quantities, we obtain (4.1), (4.2) and (4.3) as estimates for the parameters:

4.6.3 Distinction between 'death' and 'discharge'

If an infection leads to worsening of the clinical condition of a patient, two effects for the average LOS of infected patients play a role. If the patient survives, the infection may attribute to a longer LOS because it takes the patient more time to recover, while if the patient dies (partly) as a result of the infection, the infection may have shortened the LOS of this patient. Therefore it remains difficult to interpret the attributable LOS found. To solve this problem, one can introduce a death rate λ_2 and μ_2 and a discharge rate (without 'death') λ_1 and μ_1 for the uninfected and infected patients respectively. If we put $\lambda = \lambda_1 + \lambda_2$ and $\mu = \mu_1 + \mu_2$, we obtain again (4.20) and (4.1), 4.2) and (4.3) for the MLE and the deterministic estimates respectively. If n of the N uninfected patients died in the ICU and m of the M infected patients, both (4.20) and equations (4.1), (4.2) and (4.3) can be extended by the following two equations:

$$m\lambda_1 = (M-m)\lambda_2 \qquad n\mu_1 = (N-n)\mu_2 \tag{4.29}$$

4.6.4 Asymptotic variance of the estimates

To obtain information about the accuracy of the estimate, we would like to calculate the standard deviation in the estimates. For the deterministic estimates, it is straightforward to calculate the standard deviation as the number of patients tends to infinity. For the MLE, we use the Fisher-information matrix to obtain asymptotic confidence intervals.

Asymptotic variance for deterministic estimates

We find the following asymptotic results:

$$\begin{aligned} \sigma_{d^*} &= d\sqrt{1+2fd}\frac{1}{\sqrt{N_{tot}}} \\ \sigma_{d'^*} &= \sqrt{\frac{d^2+d'^2+fdd'^2}{fd}}\frac{1}{\sqrt{N_{tot}}} \\ \sigma_{\frac{1}{f^*}} &= \sqrt{\frac{1+fd+(fd)^2}{f^3 d}}\frac{1}{\sqrt{N_{tot}}} \end{aligned} \tag{4.30}$$

and for the difference $d' - d$, which is a measure for the attributable length of stay, we find:

$$\sigma_{d'^*-d^*} = \sqrt{\frac{(1+fd)(d^2+d'^2+2fd^3)}{fd}}\frac{1}{\sqrt{N_{tot}}} \tag{4.31}$$

To obtain confidence intervals for d^*, d'^* and f^*, we replace d, d' and f in the expressions for the standard deviation by their estimates d^*, d'^* and f^*.

The derivation of the asymptotic variance in the estimate for d is shown below. The other asymptotic variances can be obtained in a similar way.

$$\begin{aligned} d - d^* = &\quad d - \frac{\overline{LOS(U)}}{1-P_I} = d - \frac{N+M}{N^2}\sum_{j=1}^{N} T_j \\ (d-d^*)^2 = &\quad d^2 - 2d\frac{N+M}{N^2}\sum_{j=1}^{N} T_j + \left(\frac{N+M}{N^2}\right)^2 \sum_{j=1}^{N} T_j^2 \\ &\quad +2\left(\frac{(N+M)^2}{N^4}\right)\sum_{j=1}^{N}\sum_{i>j}^{N} T_i T_j \end{aligned} \tag{4.32}$$

The expected value of $(d-d^*)^2$ is given by:

$$\begin{aligned} &E(d-d^*)^2 = \int_{[0,\infty)^N} \prod_{l=1}^{N} \left((\tfrac{1}{d}+f)e^{-(\frac{1}{d}+f)T_l} \right) \\ &\left\{ d^2 - 2d\frac{N+M}{N^2}\sum_{j=1}^{N} T_j + \left(\frac{N+M}{N^2}\right)^2 \sum_{j=1}^{N} T_j^2 + 2\left(\frac{(N+M)^2}{N^4}\right)\sum_{j=1}^{N}\sum_{i>j}^{N} T_i T_j \right\} dT_l \end{aligned} \tag{4.33}$$

Now evaluate the integrals separately for each of the four terms. Note first of all that:

$$\begin{aligned} \int_0^\infty (\tfrac{1}{d}+f)e^{-(\frac{1}{d}+f)T_j} dT_j &= 1 \\ \int_0^\infty (\tfrac{1}{d}+f)e^{-(\frac{1}{d}+f)T_j} T_j dT_j &= \frac{d}{1+fd} \\ \int_0^\infty (\tfrac{1}{d}+f)e^{-(\frac{1}{d}+f)T_j} T_j^2 dT_j &= \frac{2d^2}{(1+fd)^2} \end{aligned} \tag{4.34}$$

Therefore, the fourth term can be written as

$$\begin{aligned}&\int_{[0,\infty)^N}\prod_{l=1}^{N}\left((\tfrac{1}{d}+f)e^{-(\frac{1}{d}+f)T_l}\right)\left\{2\left(\tfrac{(N+M)^2}{N^4}\right)\sum_{j=1}^{N}\sum_{i>j}^{N}T_iT_j\right\}dT_l=\\&2\left(\tfrac{(N+M)^2}{N^4}\right)\tfrac{1}{2}N(N-1)\int_0^\infty\int_0^\infty\left((\tfrac{1}{d}+f)e^{-(\frac{1}{d}+f)T_l}\right)T_1T_2dT_1dT_2=\\&\tfrac{(M+N)^2}{N^3}(N-1)\left(\tfrac{d}{1+fd}\right)^2\end{aligned} \tag{4.35}$$

and the first, second and third term give respectively:

- d^2
- $-2d\frac{N+M}{N^2}\sum_{j=1}^{N}\int_0^\infty\left((\frac{1}{d}+f)e^{-(\frac{1}{d}+f)T_j}T_jdT_j\right)=-2d\frac{N+M}{N^2}N\frac{1}{\frac{1}{d}+f}=\frac{-2d^2}{1+fd}\frac{M+N}{N}$
- $\left(\frac{M+N}{N^2}\right)^2\sum_{j=1}^{N}\int_0^\infty\left(\frac{1}{d}+f\right)e^{-\left(\frac{1}{d}+f\right)T_j}T_j^2dT_j=2\frac{(M+N)^2}{N^3}\left(\frac{d}{1+fd}\right)^2$

Hence

$$\begin{aligned}E(d-d^*)^2=\;&d^2-\tfrac{2d^2}{1+fd}\tfrac{M+N}{N}+2\tfrac{(M+N)^2}{N^3}\left(\tfrac{d}{1+fd}\right)^2+\\&\tfrac{(M+N)^2}{N^3}(N-1)\left(\tfrac{d}{1+fd}\right)^2\end{aligned} \tag{4.36}$$

This is the expectation when N and M are given. Usually, in a clinical trial, only the total number of patients included is known beforehand. Therefore we average over all possibilities for N and M with appropriate weight. However, when everyone became infected, i.e., $N=0$, equation (4.22) is not defined and we have to determine $E(d-d^*)^2$ given that at least one patient did not acquire colonization. This leads to the following expression:

$$\frac{\sum_{N=1}^{N_{tot}}\binom{N_{tot}}{N}\left(\frac{1}{1+fd}\right)^N\left(\frac{fd}{1+fd}\right)^{N_{tot}-N}\left\{d^2-\frac{2d^2}{1+fd}\frac{N_{tot}}{N}+\left(\frac{d}{1+fd}\right)^2\left(\frac{N_{tot}^2}{N^2}+\frac{N_{tot}^2}{N^3}\right)\right\}}{1-\left(\frac{fd}{1+fd}\right)^{N_{tot}}} \tag{4.37}$$

For large N_{tot}, the binomial distribution can be approximated by a normal distribution. When we also rescale, i.e., $x=\frac{N}{N_{tot}}$, we obtain for the numerator (with $p=\frac{1}{1+fd}$):

$$\begin{aligned}&d^2\left(1-\left(\tfrac{fd}{1+fd}\right)^{N_{tot}}\right)+\\&\int_{\frac{1}{N_{tot}}}^1dx\sqrt{\tfrac{N_{tot}}{2\pi p(1-p)}}e^{-\frac{N_{tot}}{2p(1-p)}(x-p)^2}\left\{\tfrac{-2d^2}{1+fd}\tfrac{1}{x}+\tfrac{d^2}{(1+fd)^2}\tfrac{1}{x^2}+\tfrac{d^2}{(1+fd)^2N_{tot}}\tfrac{1}{x^3}\right\}\end{aligned} \tag{4.38}$$

For large N_{tot}, the contribution to the integral is maximal when $x \approx p$. Therefore we make a Taylor-expansion of the terms $\frac{1}{x^k}$, with $k \in \{1,2,3\}$ around $x = p$. Note that we only have to take into account even powers of the expansion because of the symmetry. Up to second order we obtain:

$$\begin{array}{l} d^2\left(1-\left(\frac{fd}{1+fd}\right)^{N_{tot}}\right)+\sqrt{\frac{N_{tot}}{2\pi p(1-p)}}\int_{\frac{1}{N_{tot}}}^{1} dx e^{-\frac{N_{tot}}{2p(1-p)}(x-p)^2} \\ \left\{-d^2+\frac{d^2(1+fd)}{N_{tot}}+\left(d^2(1+fd)^2+\frac{6d^2(1+fd)^3}{N_{tot}}\right)(x-p)^2\right\} \end{array} \tag{4.39}$$

For large N_{tot} we obtain:

$$-d^2\left(\frac{fd}{1+fd}\right)^{N_{tot}}+\frac{1}{N_{tot}}\left(d^2(1+2fd)\right)+\mathcal{O}\left(\frac{1}{N_{tot}}\right) \tag{4.40}$$

Therefore we have that for large N_{tot},

$$E(d^*-d)^2=\frac{1}{N_{tot}}\left(d^2(1+2fd)\right) \tag{4.41}$$

Asymptotic variance for the maximum likelihood estimates

It is less straightforward to calculate the variances for the estimates for the MLE, as we have no explicit solutions for the MLE. Therefore we use a variant of the Fisher Information Matrix. For notational convenience we define $\theta_1 = \lambda$, $\theta_2 = f$ and $\theta_3 = \mu$. The asymptotic covariance is given by:

$$Cov(\theta_i,\theta_j) := E((\theta_i^*-\theta_i)(\theta_j^*-\theta_j)) \approx I^{-1}(\theta_1,\theta_2,\theta_3)_{ij} \tag{4.42}$$

where the Fisher Information matrix I is given by:

$$I_{ij} = -E\left(\frac{\partial}{\partial\theta_i}\frac{\partial}{\partial\theta_j}l\right) \tag{4.43}$$

This approach does not lead to explicit expressions for the variances, but some approximations can be made.

If we define

$$\rho(\lambda,f,\mu)=\frac{M}{(\lambda+f-\mu)^2}-\sum_{i=N+1}^{N+M}\frac{T_i^2e^{-(\lambda+f+\mu)T_i}}{\left(e^{-\mu T_i}-e^{-(\lambda+f)T_i}\right)^2} \tag{4.44}$$

we can write:

$$I = -E\begin{pmatrix} -\frac{N}{\lambda^2}+\rho(\lambda,f,\mu) & \rho(\lambda,f,\mu) & -\rho(\lambda,f,\mu) \\ \rho(\lambda,f,\mu) & -\frac{M}{f^2}+\rho(\lambda,f,\mu) & -\rho(\lambda,f,\mu) \\ -\rho(\lambda,f,\mu) & -\rho(\lambda,f,\mu) & -\frac{M}{\mu^2}+\rho(\lambda,f,\mu) \end{pmatrix} \tag{4.45}$$

Let N_{tot} be the total number of patients, then $E(M) = \frac{f}{\lambda+f}N_{tot}$. The second part of $\rho(\lambda, f, \mu)$ is more interesting. The patients $N+1, \ldots, N+M$ became infected during their stay, therefore we have that (using (4.16)):

$$\begin{aligned} E\left(\frac{T_i^2 e^{-(\lambda+f+\mu)T_i}}{\left(e^{-\mu T_i}-e^{-(\lambda+f)T_i}\right)^2}\right) &= \frac{\mu(\lambda+f)}{\lambda+f-\mu}\int_0^\infty dt \frac{t^2 e^{-(\lambda+f+\mu)t}}{e^{-\mu t}-e^{-(\lambda+f)t}} \\ &= \frac{\mu}{(\lambda+f)^2(\lambda+f-\mu)}\int_0^\infty \frac{\tau^2}{e^\tau - e^{\frac{\mu}{\lambda+f}\tau}}d\tau \end{aligned} \tag{4.46}$$

This integral cannot be solved analytically, however the integral can be approximated. Let $\frac{\mu}{\lambda+f} = 1+\epsilon$, then

$$\int_0^\infty \frac{\tau^2}{e^\tau - e^{\frac{\mu}{\lambda+f}\tau}}d\tau = \int_0^\infty \frac{\tau^2}{e^\tau}\frac{1}{1-e^{\epsilon\tau}} \approx \frac{1}{\epsilon}+1+\frac{\epsilon}{2}-\frac{\epsilon^3}{6}+\frac{\epsilon^5}{6}+\mathcal{O}(\epsilon^7) \tag{4.47}$$

The inverse of I is given by:

$$\frac{-1}{N_{tot}(-fN_{tot}+(\lambda+f)(f(\lambda+f)+\mu^2)E(\rho))} \begin{pmatrix} \lambda(\lambda+f)(fN_{tot}-(\lambda+f)(f^2+\mu^2)E(\rho)) & \lambda f^2(\lambda+f)^2E(\rho) & \lambda(\lambda+f)\mu^2E(\rho) \\ \lambda f^2(\lambda+f)^2E(\rho) & f(\lambda+f)(fN_{tot}-(\lambda+f)(\lambda f+\mu^2)E(\rho)) & -f(\lambda+f)^2\mu^2E(\rho) \\ -\lambda(\lambda+f)^2\mu^2E(\rho) & -f(\lambda+f)^2\mu^2E(\rho) & (\lambda+f)\mu^2(N_{tot}-(\lambda+f)^2E(\rho)) \end{pmatrix} \tag{4.48}$$

If we take the first three terms in the expansion of the integral into account, the asymptotic variance in the estimates of λ, f and μ are given by:

$$\begin{aligned} Var(\lambda^*) &= \frac{1}{N_{tot}}\frac{\lambda(\lambda+f)(2\lambda(\lambda+f)(\lambda+2f)+f^2\mu-2(\lambda+f)\mu^2+\mu^3)}{2\lambda(\lambda+f)^2+f(\lambda+f)\mu-2(\lambda+f)\mu^2+\mu^3} \\ Var(f^*) &= \frac{1}{N_{tot}}\frac{f(\lambda+f)(2(\lambda+f)(f^2+\lambda f+\lambda^2)+\lambda f\mu-2(\lambda+f)\mu^2+\mu^3)}{2\lambda(\lambda+f)^2+f(\lambda+f)\mu-2(\lambda+f)\mu^2+\mu^3} \\ Var(\mu^*) &= \frac{1}{N_{tot}}\frac{(\lambda+f)^2\mu^2(2\lambda(\lambda+f)+f\mu)}{f(2\lambda(\lambda+f)^2+f(\lambda+f)\mu-2(\lambda+f)\mu^2+\mu^3)} \end{aligned} \tag{4.49}$$

Variance in the attributable LOS

The variance in the attributable LOS $(d' - d)$ can be determined with the covariance matrix obtained in the previous section. We have that:

$$E\left(\left((\frac{1}{\lambda^*}-\frac{1}{\mu^*})-(\frac{1}{\lambda}-\frac{1}{\mu})\right)^2\right) = \frac{Var(\lambda^*)}{\lambda^4}+\frac{Var(\mu^*)}{\mu^4}-2\frac{Cov(\lambda^*,\mu^*)}{(\lambda\mu)^2} \tag{4.50}$$

Proof.

$$E\left(\left((\tfrac{1}{\lambda^*}-\tfrac{1}{\mu^*})-(\tfrac{1}{\lambda}-\tfrac{1}{\mu})\right)^2\right) = E(\tfrac{1}{\lambda^{*2}})+E(\tfrac{1}{\mu^{*2}})-2E(\tfrac{1}{\lambda^*\mu^*}) + \left(\tfrac{2}{\mu}-\tfrac{2}{\lambda}\right)E(\tfrac{1}{\lambda^*})+\left(\tfrac{2}{\lambda}-\tfrac{2}{\mu}\right)E(\tfrac{1}{\mu^*})+\tfrac{1}{\lambda^2}-\tfrac{2}{\lambda\mu}+\tfrac{1}{\mu^2} \tag{4.51}$$

Now work out each of the terms separately:

$$\begin{aligned} E(\tfrac{1}{\lambda^{*2}}) = &\quad E\left(\tfrac{1}{\lambda^2}\tfrac{1}{1-\frac{\lambda^2-\lambda^{*2}}{\lambda^2}}\right) \\ = &\quad \tfrac{1}{\lambda^2}E\left(1+\tfrac{\lambda^2-\lambda^{*2}}{\lambda^2}+\tfrac{(\lambda^2-\lambda^{*2})^2}{\lambda^4}+\mathcal{O}(\lambda-\lambda^*)^3\right) \\ = &\quad \tfrac{1}{\lambda^2}-\tfrac{Var(\lambda^*)}{\lambda^4}+E\left(\tfrac{(\lambda+\lambda^*)^2(\lambda-\lambda^*)^2}{\lambda^4}\right)+\mathcal{O}(\lambda-\lambda^*)^3 \\ = &\quad \tfrac{1}{\lambda^2}+\tfrac{3Var(\lambda^*)}{\lambda^4}+\mathcal{O}(\lambda-\lambda^*)^3 \end{aligned} \tag{4.52}$$

In a similar way we have that

$$E(\frac{1}{\mu^{*2}}) = \frac{1}{\mu^2}+\frac{3Var(\mu^*)}{\mu^4}+\mathcal{O}(\mu-\mu^*)^3 \tag{4.53}$$

For the terms proportional to $E(\frac{1}{\lambda^*})$ and $E(\frac{1}{\mu^*})$ we find:

$$\begin{aligned} E(\tfrac{1}{\lambda^*}) = &\quad E\left(\tfrac{1}{\lambda}\tfrac{1}{1+\frac{\lambda^*-\lambda}{\lambda}}\right) = \tfrac{1}{\lambda}E\left(1-\tfrac{\lambda^*-\lambda}{\lambda}+\tfrac{(\lambda^*-\lambda)^2}{\lambda^2}\right)+\mathcal{O}(\lambda^*-\lambda)^3 \\ = &\quad \tfrac{1}{\lambda}+\tfrac{Var(\lambda)}{\lambda^3}+\mathcal{O}(\lambda^*-\lambda)^3 \\ E(\tfrac{1}{\mu^*}) = &\quad \tfrac{1}{\mu}+\tfrac{Var(\mu^*)}{\mu^3}+\mathcal{O}(\mu^*-\mu)^3 \end{aligned} \tag{4.54}$$

Finally, as

$$Cov(\lambda^*,\mu^*) = E\left((\lambda^*-\lambda)(\mu^*-\mu)\right) = E\left(\lambda^*\mu^*\right)-\lambda\mu) \tag{4.55}$$

and thus $E(\lambda^*\mu^*) = Cov(\lambda^*,\mu^*)+\lambda\mu$, we can write:

$$\begin{aligned} &E\left(\tfrac{1}{\lambda^*\mu^*}\right) = E\left(\tfrac{1}{\lambda\mu}\tfrac{1}{1+\frac{\lambda^*\mu^*-\lambda\mu}{\lambda\mu}}\right) \approx \\ &\tfrac{1}{\lambda\mu}-\tfrac{Cov(\lambda^*,\mu^*)}{(\lambda\mu)^2}+\tfrac{1}{(\lambda\mu)^3}E\left(\lambda(\mu^*-\mu)+\mu(\lambda^*-\lambda)+(\lambda^*-\lambda)(\mu^*-\mu)\right)^2 \approx \\ &\tfrac{1}{\lambda\mu}-\tfrac{Cov(\lambda^*,\mu^*)}{(\lambda\mu)^2}+\tfrac{1}{(\lambda\mu)^3}E\left(\lambda^2(\mu^*-\mu)^2+\mu^2(\lambda^*-\lambda)^2+2\lambda\mu(\mu^*-\mu)(\lambda^*-\lambda)\right) \\ &= \tfrac{1}{\lambda\mu}-\tfrac{Cov(\lambda^*,\mu^*)}{(\lambda\mu)^2}+\tfrac{1}{(\lambda\mu)^3}\left(\lambda^2Var(\mu^*)+\mu^2Var(\lambda^*)+2\lambda\mu Cov(\lambda^*,\mu^*)\right) \\ &= \tfrac{1}{\lambda\mu}+\tfrac{Cov(\lambda^*,\mu^*)}{(\lambda\mu)^2}+\tfrac{Var(\mu^*)}{\lambda\mu^3}+\tfrac{Var(\lambda^*)}{\mu\lambda^3} \end{aligned} \tag{4.56}$$

Combining the expressions above, gives expression (4.50). □

This last expression can be compared with $\sigma^2_{d'^*-d^*}$ from formula (4.31)

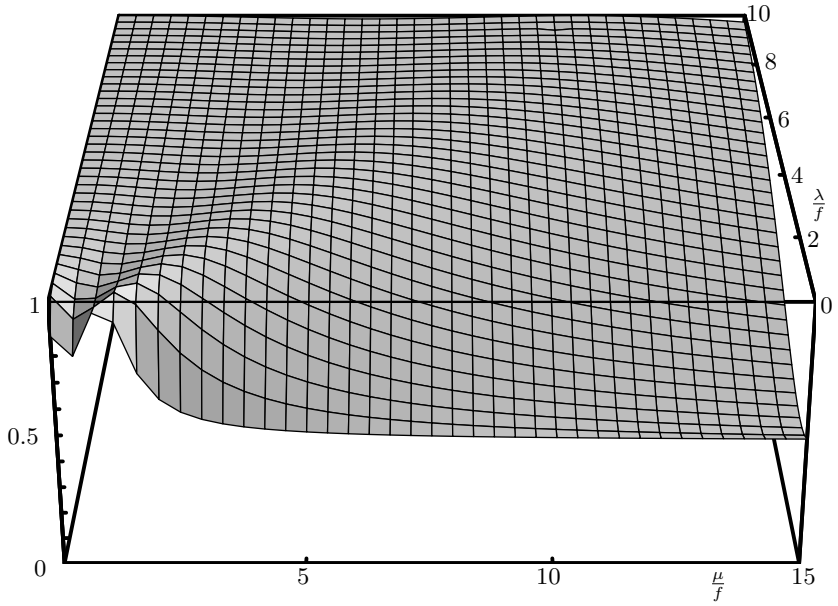


Figure 4.1: ratio between the Variance in the attributable LOS for the MLE and the deterministic estimate. $0 < \frac{\mu}{f} < 15$, $0 < \frac{\lambda}{f} < 10$

4.6.5 Numerical estimates of confidence intervals

As the MLE is asymptotically efficient, the variance in the estimates should be smaller or equal to the variance in the 'deterministic' estimates. This is indeed the case. However, the ratio between the two variances can be close to 1 for some parameter values. For the ratio between the variances for the estimate of the attributable LOS, we find the following results (see figure 4.6.5): When $\mu > \lambda + f$, which corresponds to the situation that the infection is beneficially for the patient, the variance in th MLE-estimate is clearly smaller than the variance in the deterministic estimate. However, when the infection is neutral or harmful for the patient, the deterministic estimate does not do much worse than the MLE.

It is also very interesting to know how well the asymptotic approximations for the variances are if the number of patients is finite. To investigate this, for given λ, f and μ , we generated 10000 sets of N_{tot} patients. We estimated the parameters back from the data in two ways, with the deterministic estimate and with the MLE. The standard deviations found in the estimate of the attributable LOS are shown in Figure 4.2. As can be seen from Figure 4.2, even for small number of patients,

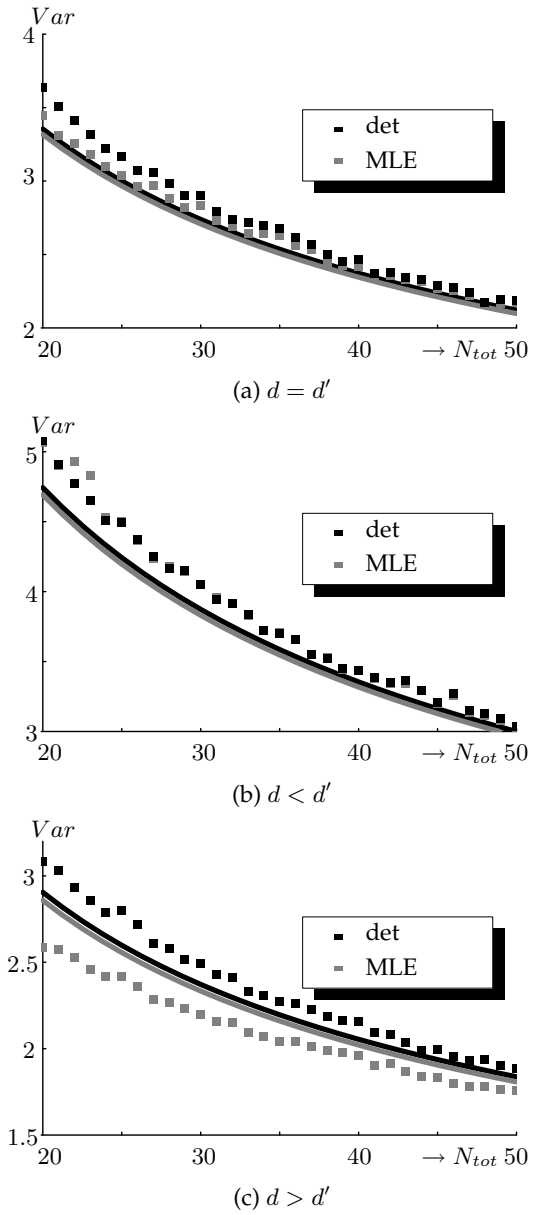


Figure 4.2: Variance in the attributable LOS for the MLE and the deterministic estimate. The lines are the asymptotic approximations. The dots are the results of the simulations. (a): $d = 5$, $d' = 5$, $f = 0.1$. (b): $d = 5$, $d' = 10$, $f = 0.1$. (c): $d = 5$, $d' = 2.5$, $f = 0.1$. The variance in the MLE is smaller than the variance in the deterministic estimates although the variances for both methods are almost equal when $d < d'$. The the simulation results converge to the asymptotic lines as the number of patients increases.

the asymptotic approximations are quite good. The variance in the attributable LOS of the MLE is smaller than the deterministic estimate. However, the difference is small when $d' \geq d$. Therefore if the deterministic estimate indicates that $d' \geq d$, it is probably not necessary to perform the MLE.

4.6.6 Existence and positivity of MLE-estimates

If we eliminate λ and f from the equations (4.20) for the ML-estimates, we obtain:

$$\frac{M(M+N)}{\mu(N+2M-\mu\sum_{j=1}^{N+M}T_j)} = \sum_{i=N+1}^{N+M} \frac{T_i e^{-\mu T_i}}{e^{-\mu T_i} - e^{-\frac{N+M}{\sum_{j=1}^{N+M}T_j - \frac{M}{\mu}}T_i}} \tag{4.57}$$

There are two interesting values for μ: $\mu_1 = \frac{M}{\sum_{j=1}^{N+M}T_j}$ and $\mu_2 = \frac{N+2M}{\sum_{j=1}^{N+M}T_j}$.

The left hand side (LHS) of (4.57) in $\mu = \mu_1$ is given by $\sum_{j=1}^{N+M}T_j$. The right hand side of (4.57) is discontinuous in $\mu = \mu_1$.

$$\lim_{\mu\uparrow\mu_1} RHS = 0 \qquad \lim_{\mu\downarrow\mu_1} RHS = \sum_{i=N+1}^{N+M} T_i \tag{4.58}$$

For $\mu \to \mu_2$, both RHS and LHS have a pole. Therefore we expand them in powers of $(\mu - \mu_2)$.

$$\begin{aligned}
\frac{M(M+N)}{\mu(N+2M-\mu\sum_{j=1}^{N+M}T_j)} &= -\frac{M(M+N)}{2M+N}\frac{1}{(\mu-\mu_2)} \\
&\quad + \frac{M(M+N)}{(2M+N)^2}\sum_{j=1}^{N+M}T_j + \mathcal{O}((\mu-\mu_2)^2) \\
\sum_{i=N+1}^{N+M} \frac{T_i e^{-\mu T_i}}{e^{-\mu T_i} - e^{-\frac{N+M}{\sum_{j=1}^{N+M}T_j - \frac{M}{\mu}}T_i}} &= -\frac{M(M+N)}{2M+N}\frac{1}{(\mu-\mu_2)} - \frac{M^2}{(2M+N)^2}\sum_{j=1}^{N+M}T_j \\
&\quad + \tfrac{1}{2}\sum_{i=N+1}^{N+M}T_i + \mathcal{O}((\mu-\mu_2)^2)
\end{aligned} \tag{4.59}$$

Notice that:

$$\begin{aligned}
RHS(\mu_2) - LHS(\mu_2) &= -\tfrac{M}{2M+N}\sum_{j=1}^{N+M}T_j + \tfrac{1}{2}\sum_{i=N+1}^{N+M}T_i \\
&= \tfrac{MN}{2M+N}\left(-LOS(U) + \tfrac{1}{2}LOS(I)\right) \\
\lim_{\mu\downarrow\mu_1} RHS(\mu) - LHS(\mu) &= -\sum_{j=1}^{N}T_j < 0
\end{aligned} \tag{4.60}$$

If $(-LOS(U)+\frac{1}{2}LOS(I)) > 0$, there is a solution in the interval (μ_1, μ_2), otherwise there is a solution on (μ_2, ∞) as the LHS converges from below to 0, proportional to $\frac{1}{\mu^2}$ when $\mu \to \infty$, while the RHS converges

to 0 from below proportional to $e^{-\min\limits_{N+1\leq i\leq N+M}\mu T_1}$. There cannot be a solution for $\mu < \mu_1$, as the sign of LHS and RHS differs.

From the second equation of the system (4.20), we observe that the estimate for f is positive when $\mu > \mu_1$. With the first equation of the system (4.20) we obtain that the estimate of λ is positive. This is an advantage of the MLE, as the deterministic estimate of μ can be negative.

4.6.7 Uniqueness of the MLE estimates

$$
\begin{aligned}
l(\lambda+\epsilon_\lambda, f+\epsilon_f, \mu+\epsilon\mu) = & \log\binom{N+M}{M} + M\log(f+\epsilon_f) + N\log(\lambda+\epsilon_\lambda) \\
& + M\log(\mu+\epsilon_\mu) - M\log(\lambda+f+\mu+\epsilon_\lambda+\epsilon_f-\epsilon_\mu) \\
& - \sum_{j=1}^{N}(\lambda+f+\epsilon_\lambda+\epsilon_f)T_j + \sum_{i=N+1}^{N+M}\log\left(e^{-(\mu+\epsilon_\mu)T_i} - e^{-(\lambda+f+\epsilon_\lambda+\epsilon_f)T_i}\right) \\
= & \, l(\lambda,f,\mu) + \tfrac{M}{f}\epsilon_f + \tfrac{N}{\lambda}\epsilon_\lambda + \tfrac{M}{\mu}\epsilon_\mu - \tfrac{M}{\lambda+f-\mu}(\epsilon_\lambda+\epsilon_f-\epsilon_\mu) - (\epsilon_\lambda+\epsilon_f)\sum_{j=1}^{N}T_j \\
& + \sum_{i=N+1}^{N+M} T_i \tfrac{(\epsilon_\lambda+\epsilon_f)e^{-(\lambda+f)T_i} - \epsilon_\mu e^{-\mu T_i}}{e^{-\mu T_i} - e^{-(\lambda+f)T_i}} - \tfrac{M}{2f^2}{\epsilon_f}^2 - \tfrac{N}{2\lambda^2}{\epsilon_\lambda}^2 - \tfrac{M}{2\mu^2}{\epsilon_\mu}^2 \\
& + \left(\tfrac{M}{2(\lambda+f-\mu)^2} - \sum_{i=N+1}^{N+M}\tfrac{T_i^2}{2\left(e^{\frac{1}{2}(\lambda+f-\mu)T_i} - e^{-\frac{1}{2}(\lambda+f-\mu)T_i}\right)^2}\right)(\epsilon_\lambda+\epsilon_f-\epsilon_\mu)^2 + \mathcal{O}(\epsilon^3) \\
= & \, l(\lambda,f,\mu) - \tfrac{M}{2f^2}{\epsilon_f}^2 - \tfrac{N}{2\lambda^2}{\epsilon_\lambda}^2 - \tfrac{M}{2\mu^2}{\epsilon_\mu}^2 + \tfrac{M}{2(\lambda+f-\mu)^2}(\epsilon_\lambda+\epsilon_f-\epsilon_\mu)^2 \\
& - \sum_{i=N+1}^{N+M}\tfrac{T_i^2}{2\left(e^{\frac{1}{2}(\lambda+f-\mu)T_i} - e^{-\frac{1}{2}(\lambda+f-\mu)T_i}\right)^2}(\epsilon_\lambda+\epsilon_f-\epsilon_\mu)^2 + \mathcal{O}(\epsilon^3)
\end{aligned}
\tag{4.61}
$$

In this derivation we used the system of equations (4.20) to show that all terms proportional to $\epsilon_\lambda, \epsilon_f$ and ϵ_μ cancel out. If for small ϵ_λ, ϵ_f and ϵ_μ, the remaining expression is smaller than $l(\lambda, f, \mu)$, both evaluated in the MLE, each critical points of the likelihood would be a maximum, which would imply uniqueness. However, this could not be proven for arbitrary lengths of stay. However, in all numerical examples, only one extremum could be found.

4.6.8 Restriction on LOS

Another point is the following: In reality, it is usually hard to determine whether an infection is ICU-acquired or not. To exclude non-ICU acquired infections, ICU-acquired infections are usually defined as infections diagnosed after a fixed number of days after entering the ICU.

Furthermore, to prevent that a small number of patients with a long LOS (outliers) affect the results considerably, one often restricts the attention to those patients who stay at most a fixed number of days in the ICU. As a second extension, we incorporate these restrictions on the LOS in the model. In principle our analysis remains the same, but the computations become more involved.

We use the basic model of section 4.2 but we now only look at patients who stay at least α and at most β days in ICU. This gives the following relations. These do not yield explicit formulas for d and d' and f, but they can be used to compute these quantities numerically.

$$\begin{aligned}
\overline{LOS}(NI) &= \alpha + \frac{\int_0^{\beta-\alpha} \tau \mathcal{G}_{NI}(\tau) d\tau}{\int_0^{\beta-\alpha} \mathcal{G}_{NI}(\tau) d\tau} \\
\overline{LOS}(I) &= \frac{\int_\alpha^\beta d\tau \int_\alpha^\tau dx \tau f \mathcal{G}_{NI}(x) \frac{1}{d'} e^{-\frac{1}{d'}(\tau - x)}}{\int_\alpha^\beta d\tau \int_\alpha^\tau dx f \mathcal{G}_{NI}(x) \frac{1}{d'} e^{-\frac{1}{d'}(\tau - x)}} \\
P_I &= f \int_\alpha^\beta d\tau \mathcal{G}_{NI}(\tau) \int_0^{\beta-\tau} dx \frac{1}{d'} e^{-\frac{x}{d'}} / \\
&\quad \left(\frac{1}{d} \int_\alpha^\beta \mathcal{G}_{NI}(\tau) d\tau + \right. \\
&\quad \left. f \int_\alpha^\beta d\tau \mathcal{G}_{NI}(\tau) \int_0^{\beta-\tau} dx \frac{1}{d'} e^{-\frac{x}{d'}} \right)
\end{aligned} \tag{4.62}$$

4.6.9 When screening is done at fixed time intervals

Assume patients are screened for a microorganism on admission and at fixed days of the week, say after alternatingly w_1 and w_2 days. In this setting there is an extra possibility of misclassification. Patients who become colonized, but are discharged before the next screening, will never have a positive test. To correct for this phenomenon, we do the following: Define m as the time at which the first culture is performed after admission. We assume for the moment that cultures are performed at times 0, m, $m + w_1$, $m + w_1 + w_2$, $m + 2w_1 + w_2$, ... We now label the cultures such that screening 0 is performed at $t = t_0 = 0$, screening 1 at $t = t_1 = m$ and so on. Again we only look at patients with a length of stay in the interval (α, β). As we know whether a patient was colonized on admission, we could take $\alpha = 0$. However, usually an infection is classified ICU-acquired only if detection takes place after a fixed number of days. We now define $n_{\min}(m)$ as the lowest culture number n, such that $t_n \in (\alpha, \beta)$. $n_{\max}(m)$ is defined as the largest culture number such that $t_n \in (\alpha, \beta)$.

Next, we calculate the fraction $\mathcal{G}_{ND}$ of the individuals who are colonized, but not yet tested positive as a function of time t elapsed since

admission and given the value of m. Directly after a test, this fraction will be zero. Therefore $\mathcal{G}_{ND}$ will obey the following differential equation:

$$\begin{array}{l} \frac{d\mathcal{G}_{ND}(t)}{dt} = f\mathcal{G}_{NI}(t) - \frac{1}{d'}\mathcal{G}_{ND}(t) \\ \mathcal{G}_{ND}(t_n) = 0 \end{array} \tag{4.63}$$

with $n \geq 0$. The solution of this differential equation is:

$$\frac{fdd'}{d' + fdd' - d}\left(e^{-\left(\frac{1}{d}+f-\frac{1}{d'}\right)t_n}e^{-\frac{1}{d'}t} - e^{-\left(\frac{1}{d}+f\right)t}\right) \tag{4.64}$$

Directly after a screening, $\mathcal{G}_{ND}$ will drop to zero, while $\mathcal{G}_I$ increases with the same amount. With $\mathcal{G}_{ND}((t_n)_-)$ we denote this amount, which is mathematically defined as $\lim_{t' \uparrow t_n} \mathcal{G}_{ND}(t')$.

Now we are able to calculate P_I, $LOS(NI)$ and $LOS(I)$ as a function of m. ($\mathcal{G}_{NI}$ is the same as in 4.6.2.)

$$\begin{array}{l} P_I(m) = \sum_{n=n_{\min}(m)}^{n_{\max}(m)} \mathcal{G}_{ND}((t_n)_-)(1 - e^{-\frac{1}{d'}(\beta - t_n)})/ \\ \left(\frac{1}{d}\int_\alpha^\beta \mathcal{G}_{NI}(\tau)d\tau + f\int_\alpha^\beta d\tau \mathcal{G}_{NI}(\tau)\int_0^{\beta-\tau} dx \frac{1}{d'}e^{-\frac{x}{d'}}\right) \end{array} \tag{4.65}$$

$$\begin{array}{l} LOS(NI)(m) = \frac{1}{d}\int_\alpha^\beta d\tau\tau\mathcal{G}_{NI}(\tau) \\ +\frac{1}{d'}\sum_{n=n_{\min}(m)}^{n_{\max}(m)-1}\int_{t_n}^{t_{n+1}} d\tau\tau\mathcal{G}_{ND}(\tau) \\ +\int_{t_{n_{\max}(m)}}^{\beta} d\tau\tau\frac{1}{d}\mathcal{G}_{NI}(\tau)\int_0^{\beta-\tau} dx\frac{1}{d'}e^{-\frac{x}{d'}} \\ +\int_{t_{n_{\min}(m)-1}}^{\alpha} d\tau\tau f\mathcal{G}_{NI}(\tau)\int_\alpha^{t_{n_{\min}(m)}} dx\frac{1}{d'}e^{-\frac{x-\tau}{d'}} \\ +\int_\alpha^{t_{n_{\min}(m)}} d\tau\tau f\mathcal{G}_{NI}(\tau)\int_\tau^{t_{n_{\min}(m)}} \frac{1}{d'}e^{-\frac{x-\tau}{d'}} \end{array} \tag{4.66}$$

$$LOS(I)(m) = \frac{1}{d'}\sum_{n=n_{\min}}^{n_{\max}(m)} \mathcal{G}_{ND}((t_n)_-)\int_{t_n}^{b} e^{-\frac{1}{d'}(t-t_n)} \tag{4.67}$$

where we omitted the normalization for $LOS(NI)(m)$ and $LOS(I)(m)$, which is obtained by dividing by the same expressions without the τ in the integral.

Until now we assumed that cultures were performed at $t = 0$, m, $m + w_1$, $m + w_1 + w_2$, ... with m in the interval $(0, w_2)$. However, if

$w_1 \neq w_2$, it is also possible that cultures are performed at $t = 0$, m, $m + w_2$, $m + w_1 + w_2$, ... where m is now in the interval $(0, w_1)$. If we assume that m is uniformly distributed, we obtain our final formulas:

$$\begin{array}{l}\overline{LOS}(NI) = \frac{1}{w_1+w_2} \\ \left(\int_0^{w_2} \overline{LOS}(NI)(m, w_1, w_2)dm + \int_0^{w_1} \overline{LOS}(NI)(m, w_2, w_1)dm \right) \\ \overline{LOS}(NI) = \frac{1}{w_1+w_2} \\ \left(\int_0^{w_2} \overline{LOS}(I)(m, w_1, w_2)dm + \int_0^{w_1} \overline{LOS}(I)(m, w_2, w_1)dm \right) \\ P_I = \frac{1}{w_1+w_2} \left(\int_0^{w_2} P_I(m, w_1, w_2)dm + \int_0^{w_1} P_I(m, w_2, w_1)dm \right)\end{array} \tag{4.68}$$

4.6.10 MLE including moments of culturing

For each patient, we divide the period of stay into at most three periods: the period in which the patient is certainly uninfected, a period in which the patient is colonized (period after a positive culture) and the period between the last negative culture preceding the first positive culture and the moment of the positive culture itself.

If the patient is colonized on admission, the likelihood is given by

$$\mu e^{-\mu T} \tag{4.69}$$

If the patient never had a positive culture, the length of stay of the patient T can be written as $T = T_u + T_q$ with T_u the period that the patient is certainly uninfected and T_q is the period after the last culture during which the infection status is not known for sure. In this case the likelihood is given by:

$$\begin{array}{l} e^{-(\lambda+f)T_u} \left(\lambda e^{-(\lambda+f)T_q} + \int_0^{T_q} dt f e^{-(\lambda+f)t} \mu e^{-\mu(T_q-t)} \right) = \\ e^{-(\lambda+f)T_u} \left(\lambda e^{-(\lambda+f)T_q} + \frac{\mu f}{\lambda+f-\mu} \left(e^{-\mu T_q} - e^{-(\lambda+f)T_q} \right) \right) \end{array} \tag{4.70}$$

Finally, when a patient was not 'infected' on admission but became infected during the stay, the the length of stay of the patient T can be written as $T = T_u + T_q + T_c$ with T_u the period that the patient is certainly uninfected, T_c the period after the first positive culture and T_q the period for which the 'infection' status is not known for sure. In this

case the likelihood is given by:

$$\begin{aligned} &e^{-(\lambda+f)T_u}\mu e^{-\mu T_c}\int_0^{T_q} dt f e^{-(\lambda+f)t}e^{-\mu(T_q-t)} = \\ &e^{-(\lambda+f)T_u}\mu e^{-\mu T_c}\frac{f}{\lambda+f-\mu}\left(e^{-\mu T_q}-e^{-(\lambda+f)T_q}\right) \end{aligned} \tag{4.71}$$

The total likelihood is the product of all individual likelihoods.

Suppose patients who stayed at most α days in the ICU were excluded as well as patients who developed 'infection' within the first α days of stay. The likelihood that a patient did not acquire 'infection' and is till in the ICU after α days is $e^{-(\lambda+f)\alpha}$. For the patients remaining in analysis, the likelihoods in equations (4.71) and (4.70) should be multiplied by $e^{(\lambda+f)\alpha}$.

Chapter 5

Dependency between patients

5.1 Introduction

Within health-care settings, antibiotic resistance increasingly hampers successful treatment of infections, especially in intensive care units (ICUs) [Kollef and Fraser, 2001]. For some pathogens (e.g., vancomycin-resistant *Staphylococcus aureus* and pan-resistant *Pseudomonas aeruginosa* and *Acinetobacter* species) the post-antibiotic era is approaching. With a limited armamentarium of antibiotics remaining available for treatment, infection prevention becomes more and more important. The epidemiology of antibiotic resistance in hospital settings, however, is complex and quantitative understanding of the dynamics is essential for designing optimal infection control strategies.

As only a fraction of colonized patients will develop clinically apparent infections [Bonten and Weinstein, 1996], the true volume of antibiotic resistance is best represented by asymptomatic carriage (i.e., colonization). Changes in the prevalence of colonization with antibiotic-resistant microorganisms within hospital settings may occur through different processes: admission and discharge of colonized and non-colonized patients; mutations, changing susceptible bacteria into resistant ones, followed by selection due to antibiotic pressure; and cross-transmission, usually via temporarily contaminated hands of health care workers [Bonten et al., 2001]. A key characteristic of cross-transmission is dependency among patients. The risk of acquisition

(also called 'colonization pressure') is influenced by the colonization status of other patients [Bonten et al., 1998]; also see [Merrer et al., 2000] for MRSA, [Puzniak et al., 2002] for VRE and [Man et al., 2001] for Enterobacteriaceae.

Because of the typically small patient populations in ICUs (usually $<$20) and the rapid patient turnover, large fluctuations in proportions of colonized patients in time occur naturally [Bonten et al., 2001]. In relatively short time series, with large fluctuations in numbers of colonized patients, the dependency created by cross transmission leads to overdispersion and autocorrelation [Cooper and Lipsitch, 2004]. This means that in periods of high prevalence the likelihood that colonization resulted from cross-transmission is higher than in periods of low prevalence. So, the variation in time in the average number of patients colonized each day will be skewed.

Processes in which patients interact are usually called 'non-linear'. In contrast, mutations, selection of resistant flora and admission of colonized patients occur independently of the colonization status of other patients. For these processes there is no autocorrelation and when analyzed in time, the numbers of patients colonized each day will be binomially distributed. In the following we call such processes 'linear', to distinguish them from the non-linear cross-transmission. This distinction between linear and non-linear processes has implications for the design of infection control strategies, as well as for the interpretation of the observed effects of interventions [Cooper and Lipsitch, 2004; Harris et al., 2004]. Barrier precautions, for instance, can only prevent cross-transmission. And with regard to interpretation, generally used statistical tests assume independence of data, which is violated when patient-to-patient transmission is involved.

Two previous studies attempted to derive estimates of the relative importance of both processes underlying the dynamics of antibiotic resistance. [Pelupessy et al., 2002] proposed a Markov chain model to analyze transmission dynamics in small hospital settings [Pelupessy et al., 2002]. The model used data on the number of admitted patients, number of patients colonized per day, and the duration of length of stay. The daily changes in the numbers of patients that were colonized were used to estimate the rates of linear and non-linear processes and to estimate relative contributions of cross-transmission and endogenous acquisition. Limitations of this model included

- that two linear processes (i.e., admission and endogenous selec-

tion rates) could not be distinguished

- the assumption of a constant and complete bed occupancy
- a presumed constant discharge rate (or, in other words, sojourn times are exponentially distributed)
- uncertainty of patient status in-between culture sampling moments.

More recently Cooper and Lipsitch, building on the work of [Pelupessy et al., 2002] proposed a 'hidden Markov model', in which they used infection data, instead of longitudinal colonization data, to determine the relative contributions of linear and non-linear processes of acquisition [Cooper and Lipsitch, 2004]. Using three data sets, they concluded that patient-to-patient transmission was relevant for VRE and MRSA, but not for Gram-negative bacteria. Even though the conclusions were not validated by way of reference data, they elegantly demonstrated that their model offers marked improvement over generally used approaches when patient-to-patient transmission is relevant. However, as infection rates only represent the tip of the iceberg, uncertainty about underlying transmission processes remains and long surveillance periods would be needed to derive reliable predictions in settings with low infection incidence.

In the present study, we first present an extension of the Markov model proposed by [Pelupessy et al., 2002] for the interpretation of longitudinal colonization data. The adaptations are

- that admission rates are explicitly distinguished from endogenous selection rates
- that actual changes in bed occupancy are used
- the use of a stochastic model to estimate the status of patients in-between culture sampling moments.

So the model formulation is data driven from the very beginning and incorporates all the information that is available. The model is first used to interpret the epidemiology of cephalosporin-resistant *Enterobacteriaceae* (CRE) in two ICUs, using extensive surveillance and genotyping as reference. Next an extended version of the model, with one extra infection route and additional patient characteristics, is used for

studying the epidemiology of *Staphylococcus aureus* in a burn wound centre. Finally we use the framework to determine optimal culture frequencies.

5.2 The model

We need a mechanistic model that incorporates the different infection routes and that, preferably, requires the specification of only few parameters. Based on this mechanistic model, we will perform the processing of the available data, leading to estimates of the parameters in the mechanistic model and thus to estimates for the relative importance of the infection routes.

For the moment we assume for the mechanistic model that:

- patients can be in two states: either a patient carries the infective agent of interest or a patient is uncolonized, i.e., does not carry the infective agent at a detectable level
- uncolonized patients can acquire the agent and thus become colonized
- colonized patients can transmit the infective agent to uncolonized patients
- when we know the colonization status of all patients at a certain time, we know the transition rates. (This is the so-called Markov property.)
- once a patient becomes colonized, he/she remains colonized during the rest of the stay. In this way we know that, as long as there is no change in the population, the prevalence can only increase.
- the colonization status of a patient is determined on admission

Suppose we know for each patient:

- day of admission
- day of discharge
- days at which a sample is taken (which is cultured)

- results of cultures (assumed, for the time being, to be 100 % reliable)

and, possibly, we also know individual characteristics of the patients, like

- susceptibility per patient (e.g., proportional to the burned body surface)
- infectivity per patient (e.g., proportional to quantity of bacteria found in cultures)

We divide the period of stay of a patient retrospectively into (at most) 3 periods based on the results of the culturing:

1. patient is certainly uncolonized $[t_0, t_1]$
2. patient may or may not be colonized $[t_1, t_2]$
3. patient is certainly colonized $[t_2, t_e]$

Per day, we have three categories of patients, uncolonized patients, patients whose colonization status is uncertain and colonized patients. We label these three categories $\mathcal{U}$, $\mathcal{Q}$ (for 'questionable') and $\mathcal{C}$ respectively. The number of patients in the categories are represented by respectively, u, q and c . For later convenience, for every time t, we order the patients in category $\mathcal{Q}$ in increasing order of the time they are already in category $\mathcal{Q}$, i.e., a patient entering category Q will be the first one in the ordering.

5.3 Data processing

The up-dating (on a day by day basis) consists of 4 parts:

1. Evolution according to the mechanistic model

2. Use results of culturing
3. Removal of the patients that leave state $\mathcal{Q}$
4. Incorporate the new patients in $\mathcal{Q}$

We assume that culturing, discharge and admission take place at the same fixed time of the day. In the bookkeeping, we only incorporate those population states that are compatible with the data.

By definition of the categories, we are certain about the colonization status of the patients in $\mathcal{U}$ and $\mathcal{C}$. For each day, the number of patients in these two categories can be determined from the data directly. Hence, u and c are treated as (time-dependent) parameters and are not included in our definition of the state space. As each of the patients in $\mathcal{Q}$ can be colonized or not, the remaining state space is $(\mathbb{Z}_2)^q = \{0,1\}^q$. Each state is denoted by a vector $\mathbf{v} = (v_1, v_2, \ldots, v_q)$, with $v_k \in \{0,1\}$, where v_k denotes whether the k^{th} patient in $\mathcal{Q}$ is colonized or not. The state $(v_1, v_2, \ldots, v_q)$ can also be represented by the binary number $v_1 v_2 \ldots v_q$ and therefore we have a natural labeling j of each of the 2^q states ($0 \leq j \leq 2^q - 1$).

For notational convenience, we would like to switch between a state represented as a finite sequence of 0's and 1's, i.e., as an element of $(\mathbb{Z}_2)^q$, and its corresponding number. Therefore we introduce the numbering function defined as:

$$\begin{array}{rc} N: & (\mathbb{Z}_2)^q \to \mathbb{Z}_{2^q} \subset \mathbb{N} \\ & (v_1, v_2, \ldots, v_q) \mapsto \sum_{i=1}^{q} v_i 2^{(q-i)} \end{array} \tag{5.1}$$

The inverse of the numbering function, N^{-1}, relates a state number m ($0 \leq m \leq 2^q - 1$) to the colonization status of the individuals in $\mathcal{Q}$. Specifically, the element $N^{-1}(m)_k$ denotes whether individual k ($1 \leq k \leq q$) in state m is colonized or not. We also introduce a notion of ordering on $(\mathbb{Z}_2)^q$.

$$\mathbf{v} \geq \tilde{\mathbf{v}} \Leftrightarrow v_i \geq \tilde{v_i} \quad \forall \quad 1 \leq i \leq q \tag{5.2}$$

and a L^1-norm:

$$|v| = \sum_{i=1}^{q} v_i \tag{5.3}$$

As the state is actually uncertain, we want to employ a stochastic description and assign to each state a probability that it is the actual (unknown) state. So we introduce the probability vector $\mathbf{p} =$

$\{p_0, p_1, \ldots, p_{2^q-1}\}$ of length 2^q in which p_j denotes the likelihood that the system is in state j.

For the description of the time evolution of the vector $\mathbf{p}$, we need the mechanistic model. The mechanistic model should give probabilities A_{mn}, $(0 \leq m, n \leq 2^q - 1)$, which describe how likely state m is at time $t + 1$ just before culturing, discharge and admission, given that the system was in state n at time t just after the culturing, discharge and admission. The evolution can then be defined in terms of matrix multiplication:

$$\begin{aligned} E: \quad & \mathbb{R}^{2^q} \to \mathbb{R}^{2^q} \\ & \mathbf{p} \mapsto A\mathbf{p} \quad \text{with } A = (A_{mn}) \end{aligned} \tag{5.4}$$

Note that the matrix A has a special structure when no additional individual characteristics play a role. Let $\pi(k)$ be the probability that an uncolonized patient acquires colonization during a day, given that there are k colonized patients in the ward. Each transition rate in column m is either zero, when the transition to state m is not allowed by the mechanistic model, or it can be written as a product of powers of $\pi(c+j)$ and $(1-\pi(c+j))$ with c the number of colonized patients in $\mathcal{C}$ and j the number of patients in $\mathcal{Q}$ that are colonized when the system is in state n. Explicitly,

$$\begin{cases} A_{mn} = 0 & \text{if } N^{-1}(m) \not\geq N^{-1}(n) \\ A_{mn} = (1-\pi(k))^u \pi(k)^l (1-\pi(k))^{q-l} & \text{if } N^{-1}(m) \geq N^{-1}(n) \\ \text{with } l = |N^{-1}(m)| - |N^{-1}(n)| & \text{and } k = c + |N^{-1}(n)| \end{cases} \tag{5.5}$$

For instance, in the case that there are 2, c and u patients in $\mathcal{Q}$, $\mathcal{C}$ and $\mathcal{U}$ respectively, the matrix A becomes:

$$\begin{pmatrix} (1-\pi(c))^{2+u} & 0 & 0 & 0 \\ (1-\pi(c))^{1+u}\pi(c) & (1-\pi(c+1))^{1+u} & 0 & 0 \\ (1-\pi(c))^{1+u}\pi(c) & 0 & (1-\pi(c+1))^{1+u} & 0 \\ (1-\pi(c))^{u}\pi(c)^2 & (1-\pi(c+1))^{u}\pi(c+1) & (1-\pi(c+1))^{u}\pi(c+1) & (1-\pi(c+2))^{u} \end{pmatrix}$$

Note that the matrix A does not preserve the norm of the vector $\mathbf{p}$ when $u \neq 0$. (This is due to the fact that we leave out all transitions that could in principle have happened to the $\mathcal{U}$ category.)

We now will use the culture results, the discharge data and the admission data. Suppose that the k^{th} patient in $\mathcal{Q}$ is cultured. By the definition of the category $\mathcal{Q}$, this culture will be positive. Therefore only the states m, $0 \leq m \leq 2^q - 1$, with $N^{-1}(m)_k = 1$ are allowed by the

data and the other states have zero a posteriori likelihood. Mathematically, culturing of patient k in $\mathcal{Q}$ amounts to projecting the vector $\mathbf{p}$ on a linear subspace isomorphic to $\mathbb{R}^{2^{q-1}}$.

$$\begin{aligned} C: \quad & \mathbb{R}^{2^q} \to \mathbb{R}^{2^q} \\ & \mathbf{p} \mapsto C\mathbf{p} \quad \text{with } C = (C_{mn}) \end{aligned} \tag{5.6}$$

and the diagonal matrix C is given by:

$$C_{mn} = \begin{cases} 0 & \text{if } n \neq m \text{ or } N^{-1}(n)_k = 0 \\ 1 & \text{if } n = m \text{ and } N^{-1}(n)_k = 1 \end{cases} \tag{5.7}$$

Example: In the case that $q = 3$ and we culture the second patient in $\mathcal{Q}$ and before the culturing the state vector is $\mathbf{p} = (p_0, p_1, \ldots, p_7)$, then after the culturing, the vector will be $(0, 0, p2, p3, 0, 0, p6, p7)$.

When a category $\mathcal{Q}$ patient 'leaves' $\mathcal{Q}$, either because he/she was cultured or because he/she leaves the unit without being cultured, the number of possible states is reduced by a factor 2.

For $1 \leq k \leq q$ we can define the operator O_k that removes the k^{th} patient in Q via:

$$\begin{aligned} O: \quad & \mathbb{R}^{2^q} \to \mathbb{R}^{2^{q-1}} \\ & \mathbf{p} \mapsto \mathbf{p}' \end{aligned} \tag{5.8}$$

where the components of $\mathbf{p}'$ are defined by:

$$p'_{N(v_1,\ldots,v_{q-1})} = p_{N(v_1,\ldots,v_{k-1},0,v_k,\ldots v_{q-1})} + p_{N(v_1\ldots,v_{k-1},1,v_k,\ldots v_{q-1})} \tag{5.9}$$

This operator O_k just adds the probabilities of the states for which the colonization status of the remaining patients is identical.

If several category $\mathcal{Q}$ patients 'leave' $\mathcal{Q}$ at the same time, we either have to generalize the operator O_k or we have to apply the operator O_k several times for different k (and different q). In the last case, to avoid confusion about which of the patients in $\mathcal{Q}$ 'leaves' $\mathcal{Q}$, we should order the operators such that we do the removal in decreasing order of the patient number in $\mathcal{Q}$.

Suppose now that l patients enter category $\mathcal{Q}$ at a certain time t. By the definition of the category $\mathcal{Q}$, patients enter category $\mathcal{Q}$ directly after their last negative culture, so we know that these patients enter category $\mathcal{Q}$ uncolonized. As we ordered the patients in category $\mathcal{Q}$ according to the day they entered this category, these l patients correspond to

the first l digits in the binary expansion. Due to this ordering, the function I_l that deals with the admission of l new patients to $\mathcal{Q}$ is defined by:

$$\begin{array}{rc} I_l: & \mathbb{R}^{2^q} \to \mathbb{R}^{2^{q+l}} \\ & \mathbf{p} \mapsto \mathbf{p}' \end{array} \tag{5.10}$$

where the elements in the vector $\mathbf{p}'$ are given by $(0 \leq k \leq 2^{q+l}-1)$:

$$p'_k = \begin{cases} 0 & \text{if } k \geq 2^q \\ p_k & \text{if } k < 2^q \end{cases} \tag{5.11}$$

Note that O_k and I_l involve a change of the dimension of the state space. Indeed, we 'glue' together state spaces of different size according to the need as exposed by observed events.

The likelihood of the observed events during 1 day is the norm of the final state vector $\mathbf{p}$ (assuming that the initial state vector had norm 1). More precise, the likelihood is given by

$$\frac{|CA\mathbf{p}|}{|\mathbf{p}|} \tag{5.12}$$

The likelihood of the observed events over several days is the product of the relevant 1 day likelihoods.

5.4 Estimating parameters and confidence intervals

In this section we will describe the statistical methods used in the analysis of section 5.5. The value of the parameter that maximizes the likelihood of the totality of the observed events is called the Maximum Likelihood Estimator (MLE) and is denoted with $\hat{\theta}$. To find confidence bounds for MLE, we use the likelihood ratio test. This test is based on the likelihood ratio:

$$\frac{L(\theta)}{L(\hat{\theta})} \tag{5.13}$$

Note that $\hat{\theta}$ and L are functions of the data. Suppose a 'true' value for θ exists. This value is denoted by θ_0. The function

$$\begin{array}{c} \tilde{l}: \Omega \to \mathbb{R} \\ \tilde{l}(\theta) = -2\log\frac{L(\theta)}{L(\hat{\theta})} \end{array} \tag{5.14}$$

converges in distribution to the χ^2-distribution with k degrees of freedom, with k the dimension of the parameter θ, see e.g., [Cox and Hinkley, 1979]. Therefore:

$$P\left(\theta_0 \in \left\{\theta | \tilde{l}(\theta) \leq \chi^2_{\alpha;k}\right\}\right) \to \alpha \tag{5.15}$$

when the amount of data increases (see for instance [Andersen et al., 1993]) where $\chi^2_{\alpha;k}$ is the size $1-\alpha$ critical value of the chi-square distribution with k degrees of freedom. The volume

$$\{\theta | \tilde{l} \leq \chi^2_{\alpha;k}\} \tag{5.16}$$

is an α-confidence region which is asymptotically correct. Intuitively, this is a good approximation when the data show that all states have been visited several times. To find confidence intervals for a component of the parameter, we treat the other components as nuisance parameters, see [Venzon and Moolgavkar, 1988]. Let $\Omega_i(x) = \{\theta \in \Omega | \theta_i = x\}$. The profile likelihood for $\theta_i, 1 \leq i \leq k$, is given by

$$L_i(\theta_i) = \max_{\theta \in \Omega_i(\theta_i)} L(\theta) \tag{5.17}$$

and the approximate α-confidence interval for the i^{th} component is given by:

$$\{\theta_i | 2(\log L(\hat{\theta}) - \log L_i(\theta_i)) \leq \chi^2_{\alpha;1}\} \tag{5.18}$$

The profile likelihood method can also be used to construct confidence regions for a function $g : \Omega \to \mathbb{R}$. The probability

$$P\left(g(\theta_0) \in \left\{x | \min_{\{\theta | g(\theta) = x\}} \tilde{l}(\theta) \leq \chi^2_{\alpha;1}\right\}\right) \tag{5.19}$$

converges to α when the amount of data increases and so

$$\left\{x | \min_{\{\theta | g(\theta) = x\}} \tilde{l}(\theta) \leq \chi^2_{\alpha;1}\right\} \tag{5.20}$$

is an asymptotically correct α-confidence region for the function g. To obtain a one-sided confidence region, we first construct a $(-1+2\alpha)$ two-sided confidence region, which will be, as a rule, an interval $I = (i_0, i_1)$. The function

$$\tilde{g}(x) = \min_{\{\theta | g(\theta) = x\}} \tilde{l}(\theta) \tag{5.21}$$

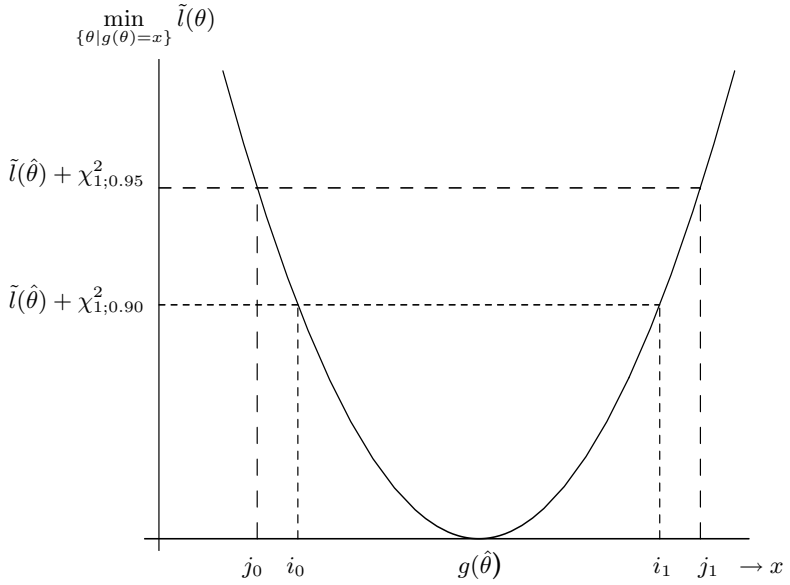


Figure 5.1: Construction of one-sided 95% confidence intervals for a function g using the profile likelihood method (see equation 5.21) The intervals (i_0, i_1) and (j_0, j_1) are respectively 90% and 95% confidence intervals. The intervals $(-\infty, i_1)$ and (i_0, ∞) are one-sided 95% confidence intervals.

will become symmetrical around $g(\theta_0)$ when the amount of data increases. Therefore, $(-\infty, i_1)$ and (i_0, ∞) are one-sided α-confidence intervals for the function g (see Figure 5.1).

For a few calculation we also used a different, more Bayesian heuristic method, to construct confidence intervals. (This method is only used when stated explicitly, otherwise the previous method is used.) Let $f(x)$ be the density function of the χ^2-distribution with k free parameters. We assume that $f(\tilde{l}(\theta))$ is the distribution for the parameter θ, i.e., the probability that $\theta \in \omega$ (given that the model is correct and that the χ^2-distribution is a good approximation) is given by:

$$P(\theta \in \omega) = \int_{\omega} d\theta \frac{f(\tilde{l}(\theta))}{V^{k-1}\left(\tilde{l}^{-1}(\tilde{l}(\theta))\right)} \tag{5.22}$$

where the k-dimensional volume of a set ξ is denoted by $V^k(\xi)$ and $\tilde{l}^{-1}$ is the set-valued map assigning to a number the set of all points in Ω that are mapped to this number by $\tilde{l}$. (The following intuitive reasoning shows that $\frac{f(\tilde{l}(\theta))}{V^{k-1}\left(\tilde{l}^{-1}(\tilde{l}(\theta))\right)}$ is the correct density. For each

interval $I = (i_0, i_1) \subset \mathbb{R}^+$, we have by definition that $P(\tilde{l}^{-1}(I)) = F(i_1) - F(i_0) = \int_{i_0}^{i_1} f(t)dt$ with F the cumulative distribution function of the χ^2-distribution with k degrees of freedom. Suppose $g(\theta)$ is the density we are looking for. Then we also have that $P(\tilde{l}^{-1}(I)) = \int_\Omega d\theta g(\theta) 1_{\tilde{l}^{-1}(I)}(\theta) = \int_0^\infty dt \int_{\tilde{l}^{-1}(t)} d_{k-1}g(x) 1_{\tilde{l}^{-1}(I)}(x) = \int_{i_0}^{i_1} dt \int_{\tilde{l}^{-1}(t)} d_{k-1}g(x)$. When we compare the two expressions, we notice that $f(t) = \int_{\tilde{l}^{-1}(t)} d_{k-1}g(x)$. $g(x) = \frac{f(\tilde{l}(\theta))}{V^{k-1}\left(\tilde{l}^{-1}(\tilde{l}(\theta))\right)}$ is a solution of this equation.)

For a general function $g : \Omega \to \mathbb{R}$, the expected value of g is given by:

$$< g >= \int_\Omega d\theta \frac{f(\tilde{l}(\theta))g(\theta)}{V^{k-1}\left(\tilde{l}^{-1}(\tilde{l}(\theta))\right)} \tag{5.23}$$

We call an interval $I = (y, z)$ an α-confidence interval for the function g, if $P(\{\theta | g(\theta) \in I\}) \geq \alpha$. The choice of the interval depends on the problem, for instance one can choose for 1-sided or 2-sided tails.

None of the integrals in this section can be evaluated exactly, as the likelihood function defined in Section 5.2 is a complicated function. However, as long as the number of parameters remains small and the calculation of the likelihood is relatively fast, we can evaluate the integrals numerically with a Monte Carlo approach. Assume Ω is bounded (otherwise restrict the domain to the physically relevant domain). Randomly pick $\theta^1, \theta^2, \ldots, \theta^N \in \Omega$ and calculate $\tilde{l}(\theta^i)$ $(1 \leq i \leq N)$. Let I_j $(1 \leq j \leq M)$ be the interval $(\chi^2_{\frac{j-1}{M};k}, \chi^2_{\frac{j}{M};k})$ with $M \in \mathbb{N}$ (By construction, the probability $P(\{\theta | \tilde{l}(\theta) \in I_j\}) = 1/M \quad \forall j$.) We define N_j as $\{\theta^i | \tilde{l}(\theta^i) \in I_j\}$. Now we can approximate the probability (5.22) that $\theta \in \omega$ by

$$\sum_{j=1}^{M} \frac{1}{M} \frac{\#(N_j \cap \omega)}{\#N_j} \tag{5.24}$$

where $\#S$ is the cardinality of the set S and (5.23) can be approximated by

$$\sum_{i=1}^{N} \frac{g(\theta^i)}{M \# N_{j(i)}} \tag{5.25}$$

where $j(i)$ is the number j of the set N_j to which i belongs.

The mean prevalence in a unit is defined as the quotient of the number of patient days for which patients are colonized divided by the total number of patient days. In case the exact days of acquisitions are

known, the prevalence can easily be calculated. When there is uncertainty in the days of acquisitions due to the fact that patients are not cultured on a daily basis, the expected prevalence can be calculated by calculating the prevalence for all of the realizations that are allowed by the culture data and afterwards averaging over these prevalences with appropriate weight (which will depend on the parameters in the mechanistic model). However, typically, the number of realizations of the acquisition days will be extremely large which makes an exact calculation impossible. Therefore we will use an approximation which is reasonable in case the dependency between patients is not very large. In this approximation, we will approximate the days that a patients is actually colonized while belonging to category $\mathcal{Q}$. The expected number of days that a patient is positive depends on the value of the parameters in the mechanistic model and the mean prevalence will be a function of these parameters.

The days for which it is demonstrated that a patient is uncolonized (patient belongs to $\mathcal{U}$) do not contribute to the numerator (the number of patient days for which patients are colonized), the days that a patient has demonstrated colonization fully contribute to the numerator. The contribution to the numerator for the days that the colonization status of a patient is uncertain is more subtle. If a patient leaves the unit while the colonization status is still uncertain after discharge, the contribution to the numerator for a day of stay in $\mathcal{Q}$ is the sum of the elements of $\mathbf{p}$ for which the patient is positive in the corresponding state.

If a patient became colonized during stay, we have to calculate, for all days that the patients belonged to category $\mathcal{Q}$, the probability that the patient was colonized at a day given that the patient had become colonized on the day of culturing. To do so, we determine the likelihood that the culture result was indeed positive at the day of detection of colonization, i.e., we determine $\sigma = \frac{\sum_i p'_i}{\sum_i p_i}$ (see equation (5.6). The contribution to the numerator per day that the patient stayed in $\mathcal{Q}$ is the sum of the elements of $\frac{\mathbf{p}}{\sigma}$ for which the patient is positive in the corresponding state (i.e., the probability that the patient is colonized that day given that the patient became colonized on the day of culturing.)

5.5 Application to data

The method described in the previous sections was applied to two data sets. For the first data set, section 5.5.1, we use a model with only two colonization routes and without additional patient characteristics. For the second data set, section 5.5.2, we use a slightly more complicated model for the colonization dynamics of *Staphylococcus aureus* in a burn wound center. We hypothesized three colonization routes exist and we also included additional patient characteristics. For both data sets, independently, genotyping was performed, which serves to provide a gold standard.

In both cases, the aim was to determine the relative importance of different infection routes and not necessarily to obtain the best possible description of the data by a non-parametric approach. However, we use a χ^2-test to avoid the possibility that we fit to the data a model which does not reflect any of the properties of the data.

5.5.1 Gram-negative bacteria resistant to third generation Cephalosporins

One data set was collected in two ICU's of the UMC Utrecht, The Netherlands, during an eight-month period. Rectal cultures were analyzed and the presence of Gram-negative bacteria resistant to third generation Cephalosporins was determined. Cultures were performed on admission and afterwards twice weekly [Nijssen et al., 2005]. The aim of this study was to determine the relative importance of two different colonization routes, i.e., cross-transmission and the endogenous route (outgrowth of already pre-existing colonization in undetectable quantities due to selective advantage when competitors suffer from antibiotic treatment). In the gold standard (genotyping), epidemiological linkage was defined as two patients having an overlap in stay in the unit. Because of the possibility of low-level colonization directly after acquisition, a maximum time window of 7 days was accepted between periods that did not overlap. Cross-transmission was defined as acquired colonization with a species identical to one in a patient that was earlier found to colonized and which was epidemiologically linked. Identity of species was determined on the base of high genetic relatedness of different isolates.

In the method described in section 5.3, the per diem probabil-

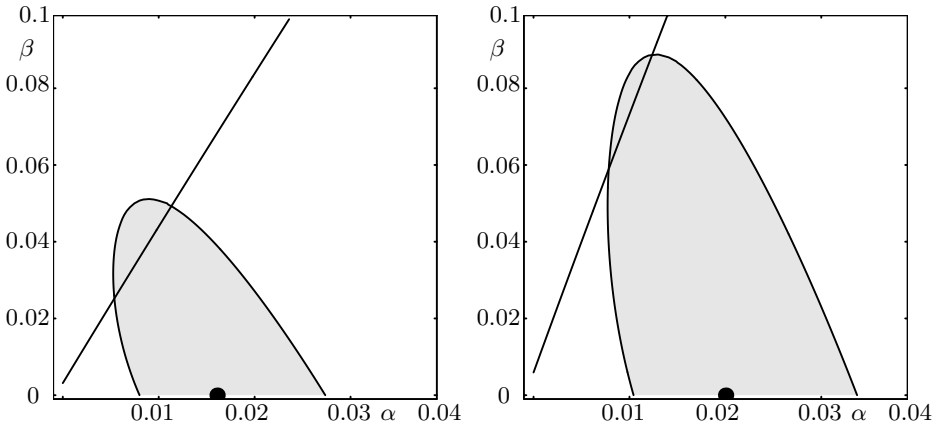
ity per susceptible patient to acquire colonization was set to be $1 - e^{-(\alpha+\beta I/N)}$ where α represents the endogenous term and $\beta I/N$ the cross-transmission term, with I the number of colonized patients in the ICU ($c \leq I \leq q + c$) and N the total number of patients in the ICU. Both α and β should be positive. The 95% confidence set for the parameters is shown in Figure 5.2 combined with the MLE-estimator for the parameters and the line $\alpha = \beta p(\alpha, \beta)$, for which both routes are equally important, where $p(\alpha, \beta)$ is the estimated prevalence for given α and β. The 95% confidence intervals for each of the parameters are shown in Table 5.1.

The probability that cross-transmission is more important than the endogenous route was calculated using the profile likelihood method. Although the line of equal importance intersects the 95%-confidence sets with the smallest area in Figure 5.2, we can predict with more than 95% accuracy that the endogenous route is more important than the exogenous route for both ICU's, see Table 5.1. To determine the expected relative importance of cross-transmission we evaluated the integral (5.23) with $g(\alpha, \beta) = \frac{\beta p(\alpha,\beta)}{\alpha+\beta p(\alpha,\beta)}\}$ (see Table 5.1). These values differ considerably from the MLE, which predicts that the influence of cross-transmission is 0. Note that the MLE lies on the boundary of the domain. This is because β was required to be non-negative. Otherwise the MLE for β would be negative ($|\beta| < 0.01$ for both ICU's).

The model predicts cross-transmission to be responsible for 18.3% and 16.4% of the acquisitions in IC1 and IC2 respectively. This coincides reasonably well with the results of the gold standard based on the genotyping. Six out of 21 (28.6%) and four out of 19 (21.1%) acquisitions were likely to have resulted from cross transmission in IC1 and IC2 respectively (see Table 5.1 and Figure 5.2). See also the discussion for an explanation of the difference. A goodness of fit χ^2-test (with two free parameters) based on the MLE did not give an indication to reject the model for any of the ICU's. (The χ^2-statistics has a value larger than the obtained value in 32% and 18% of the realization for IC1 and IC2 respectively.)

5.5.2 *Staphylococcus aureus* infections in burn wound patients

The second data set was collected in the burn wound center of the Martini hospital Groningen, the Netherlands, during a three-year period.



(a) IC1 (b) IC2

Figure 5.2: 95% confidence area for the parameters. The per diem probability per susceptible patient to acquire colonization was set to be $1 - e^{-(\alpha+\beta I/N)}$ where α represents the endogenous term and $\beta I/N$ the cross-transmission term, with I the number of colonized patients in the ICU and N the total number of patients in the ICU. The lines indicate equal importance of the endogenous route and cross transmission.

	IC1	IC2
α	0.017[0.008,0.035]	0.021[0.011,0.040]
β	0.0[0,0.039]	0.0[0,0.068]
prev.	24.42%[24.33,24.57]	14.89%[14.69,15.04]
P(exo>endo)	2.3%	1.5%
% exo	18.3%[0,49]	16.4%[0,43.0]

Table 5.1: Estimates and 95%-confidence intervals for Gram-negative bacteria resistant to third generation Cephalosporins. Rows 1,2 and 4 are calculated using the profile likelihood method. Rows 3 and 5 are calculated using the heuristic method of section 5.4.

	IC1	IC2	Total
Admitted patients	277	180	457
Rectal swabs	753	490	1243
Patients colonized (%)	41 (14.7)	29 (16.1)	70 (15.3)
Patients colonized on admission (%)	19 (6.9)	7 (3.9)	26 (5.7)
Patients with acquired colon. (%)	21 (7.6)	19 (10.6)	40 (8.7)
Observed endemic prev., mean (%)	17.5± 13.5	14.2± 12.6	16± 9.4
range (%)	0-67	0-60	0-67
Median time to acq. in days (range)	7 (2-48)	10 (4-52)	8 (2-52)

Table 5.2: Basic data for Gram-negative bacteria resistant to third generation Cephalosporins.

Burn wound patients lack the normal skin barrier which make them vulnerable for colonization with *Staphylococcus aureus* which is associated with delayed wound healing [Kooistra-Smid et al., 2004]. Nasal and pharyngeal colonization with *Staphylococcus aureus* is present in 30% of all patients on admission. The aim of the study was to investigate the colonization dynamics of *Staphylococcus aureus* in this 10 bed unit.

During the three-year period cultures were performed on admission of the nose, throat and burn wounds and afterwards weekly of the burn wounds. The study period was divided in two periods, a baseline period from January 1999 till June 2000 and an intervention period from July 2000 till December 2001. In the intervention period, on admission all patients were nasally treated with mupirocin to eradicate and prevent nose colonization with *Staphylococcus aureus*.

To evaluate the effect of the intervention, we distinguish three types of patients in the mechanistic model:

- non-colonized patients
- patients who are colonized in the nose/throat but not in their wounds
- patients with wound colonization

(Although it seems logical to distinguish patients with wound colonization between those with and without colonized nares, we choose not to do so as nose cultures were not obtained after admission in many patients and nose cultures were frequently positive for the nose as well as

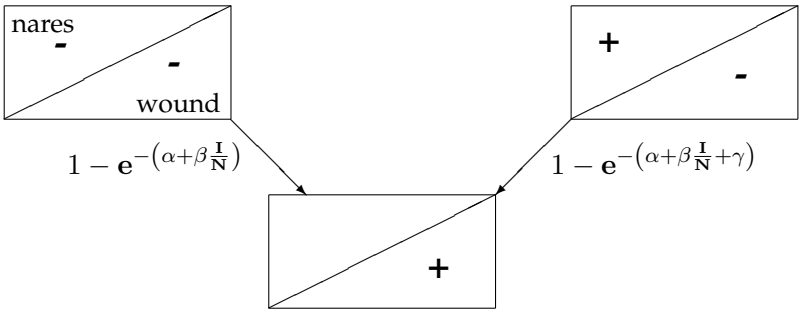


Figure 5.3: Acquisition routes for *Staphylococcus aureus*. The upper triangle denotes the nose and throat. The lower triangle denotes the burn wounds. A $+/-$ sign denotes whether that site is colonized or not. We hypothesized (see text) that once a patient is colonized in the burn wounds, the colonization status of that patient is no longer important. We define three routes of bacterial spread: First, due to a constant background force of infection (rate α), patients can acquire wound colonization. Second, due to cross transmission (rate is proportional to the number of patients with wound colonization (I) divided by the total number of patients (N) and the constant of proportionality is β) patients can acquire wound colonization. Third: Patients carrying *Staphylococcus aureus* in the nose/throat can colonize their own wound at rate γ. Id est, patient colonized in the nose/throat can acquire wound colonization in three ways, patients without colonization in the nares/throat can acquire wound colonization in two ways.

for the wounds when cultures were obtained.

We hypothesized that three routes of acquisition existed (see Figure 5.3):

1. a constant background force of infection α, for instance due to health care workers who are persistent carriers in the nose or due to visitors.

2. cross transmission from wounds of patients with wound colonization to wounds of patients who are not colonized in the wounds. The rate is proportional to the fraction of patients in the unit with wound colonization and equals $\beta\frac{I}{N}$ with I the number of patients with wound colonization and N the total number of patients.

3. a constant colonization pressure γ on the wounds when the nose/throat of the patient are colonized.

	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$
baseline	0.054(0,0.084)	0.012(0,0.15)	0.035(0,0.13)
mup	0.032(0.013-0.046)	0(0,0.041)	0.093(0.026,0.197)
whole period	0.042(0.017-0.054)	0.001(0-0.05)	0.067(0.020-0.129)

Table 5.3: Maximum likelihood estimates for the baseline period, the mupirocin period and the whole period for the full model. (95%-confidence intervals are based on the profile-likelihood method.)

	baseline	mup	total
α	0.061[0.042,0.083]	0.032[0.021,0.043]	0.049[0.033,0.054]
γ	0.035[0,0.121]	0.093[0.026,0.197]	0.067[0.022,0.128]

Table 5.4: Estimates and 95%-confidence intervals for the model for colonization with *Staphylococcus aureus* without cross-transmission ($\beta = 0$).

We perform the likelihood analysis of Sections 5.2, 5.3 and 5.4. However, the evolution matrix A (5.4) not only depends on the colonization status of the wounds of patients, but also on the colonization status of the nose of patients in $\mathcal{U}$ and $\mathcal{Q}$.

To determine the importance of the parameters, we calculated MLE's for the full model (see Table 5.3). The calculated prevalence using the MLE-estimates was 62%, 58% and 60% for the baseline period, the intervention period and the whole period respectively. We also determined MLE's for the models in which one of the parameters was set to zero. For $\beta = 0$, there was hardly any difference in MLE compared to the MLE of the full model (loglikelihood decreases with 0.012 and 0.001 for the baseline and the whole period respectively). Therefore we assume $\beta = 0$ in the rest of our analysis. (This can be formalized by the likelihood ratio test for nested models (see e.g. [Wasserman, 2004]) and it was supported by the genotyping which also indicated that patient-to-patient transmission hardly plays a role.) The results of the model without cross transmission are shown in Figure 5.4 and Table 5.4.

As can be seen from Figure 5.4 and Table 5.4, the confidence intervals for γ are wide which indicates a large uncertainty in the importance of acquiring colonization from the patients own nose. To test whether there is evidence that the parameters α and γ changed due to

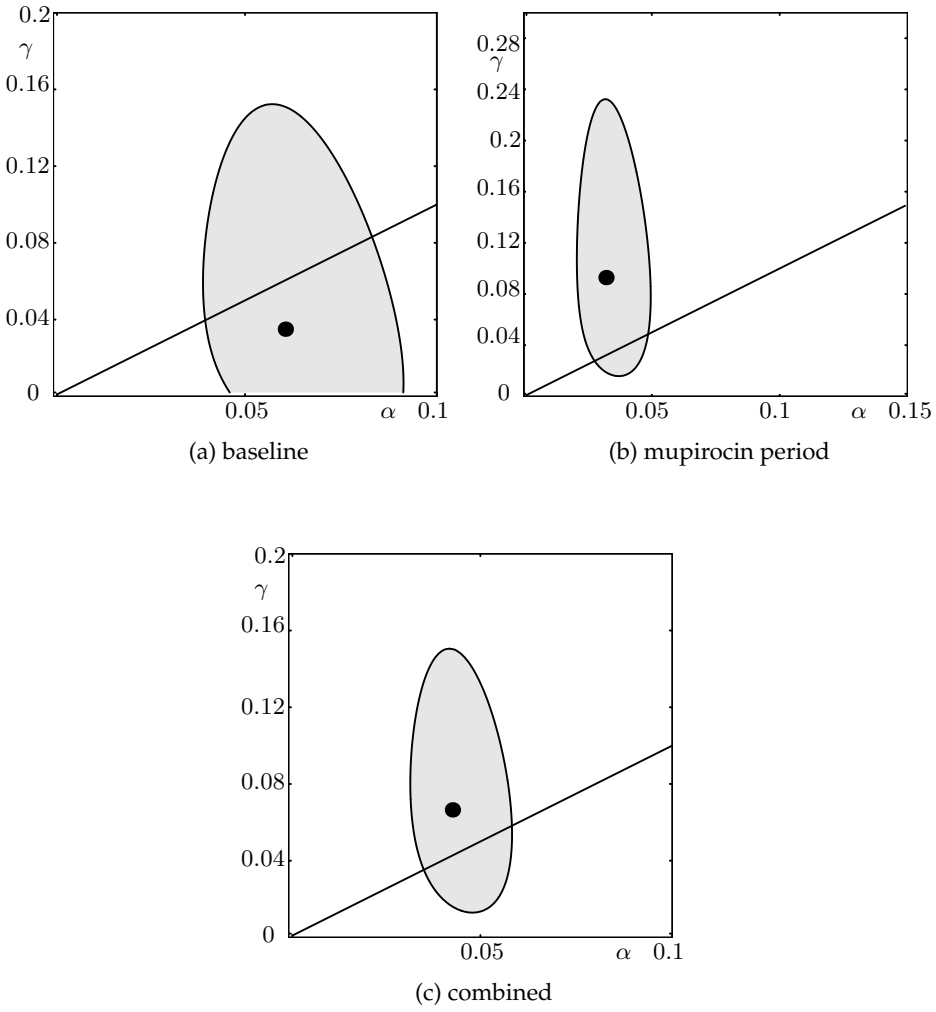
the intervention, we use the likelihood ratio test for nested models. In one model, α and γ are assumed to be constant over the whole period (both baseline and intervention period). In the other model, the parameters α and γ were allowed to be different in the two periods. With this likelihood ratio test, α and γ were constant over the whole period with a p-value< 0.05. That there is evidence for a change in the parameters between the two periods could also be deduced from Table 5.4 as the 95%-confidence intervals for α do not overlap for the baseline period and the intervention period. The intervention seems to protect patients who are uncolonized in the nose/throat on admission. (For patients who are colonized in the nose/throat on admission, it is less clear, as the MLE of γ increases in the intervention period although they were nasally treated with mupirocin.) This could be explained if the nose would serve as an intermediate step in the transmission process, in the sense that acquired colonization in the nose precedes colonization of wounds.

We also tested whether the vulnerability for wound colonization was linearly correlated to the percentage of burned body surface. Id est, whether the model in which the per diem probability for an uncolonized patient with a fraction burned body surface of ρ would be $1-e^{-(\alpha+\gamma n)\rho}$, with $n \in \{0,1\}$ representing the status of the nose, would give a higher value of the likelihood in the MLE. (This assumption again increases the complexity of the evolution matrix A.) For all periods, this assumption decreased the MLE. (Including cross-transmission in the model in which vulnerability for wound colonization is linearly correlated to the percentage of burned body surface did not change this observation.) Therefore, there is no linear correlation between acquisition rate and the percentage burned body surface.

A goodness of fit χ^2-test (with two estimated parameters, α and γ) based on MLE's did not give an indication to reject the model for any of the two periods. (The χ^2-statistics has a value larger than the obtained value in 12.3%, 28.3% and 36.4% of the realizations for the baseline, the intervention period and the total period respectively.)

5.6 Optimal culturing frequency

In this section, we focus on the optimal culture frequency for distinction between the endogenous route (fixed probability per patient per day to acquire colonization) and the exogenous route (probability per patient



(a) baseline

(b) mupirocin period

(c) combined

Figure 5.4: 95% confidence area for the parameters α, the per diem likelihood of acquisition of colonization due to the background force of infection, and γ the additional per diem likelihood of acquisitions of colonization in the wounds for a patient colonized in the throat/nares. **(a)** baseline period, **(b)** the mupirocin period (nasal treatment with mupirocin on admission to eradicate and prevent nasal colonization), **(c)**: total period. The lines indicate equal importance of the constant background force of infection and the infection pressure from the throat/nares to the own wounds for patients colonized in the throat/nares.

per day to acquire colonization depends linearly on the fraction colonized patients in the unit). When data about length of stay, moments of cultures and results of the cultures are known, the model provides estimates on the relative importance of both colonization routes.

However, which culture frequency should be chosen? The optimal culture frequency will depend on the balance between the expected time to establish the dominant infection route and the cost and effort involved in the culture frequency. The answer depends mainly on two variables. First, the relative importance of both routes and second the endemic prevalence in the unit.

With an endemic prevalence of almost 100%, nearly all patients acquire colonization shortly after admission and distinction between the routes becomes impossible without genotyping. On the other hand, if acquisition of colonization is very rare, there is hardly any information about acquisition in the data and again very long study periods are required to establish the dominant route.

As an experiment, we performed 1000 simulations for different culture frequencies. For a period of 10 years, we simulated an ICU with 10 beds which are always occupied and with two routes transmission; the endogenous route and the exogenous route. A database of patients with an exponential distributed length of stay (mean 9 days) was constructed. The probability to be discharged was not influenced by colonization. A fraction 0.05 of the patients was colonized on admission. Colonization dynamics were run with parameters α and β for the endogenous and exogenous route respectively, In this way we obtained a database with patients for which we knew the exact day of colonization. For each simulation the maximum likelihood algorithm as described in Sections 5.2, 5.3 and 5.4 was performed. In each simulation and for each culture frequency, we determined each month whether we could correctly predict with at least 0.95% confidence which of the two routes is most important (using the profile likelihood method) leading to a fraction of the simulations in which the dominant acquisition route is predicted correctly.

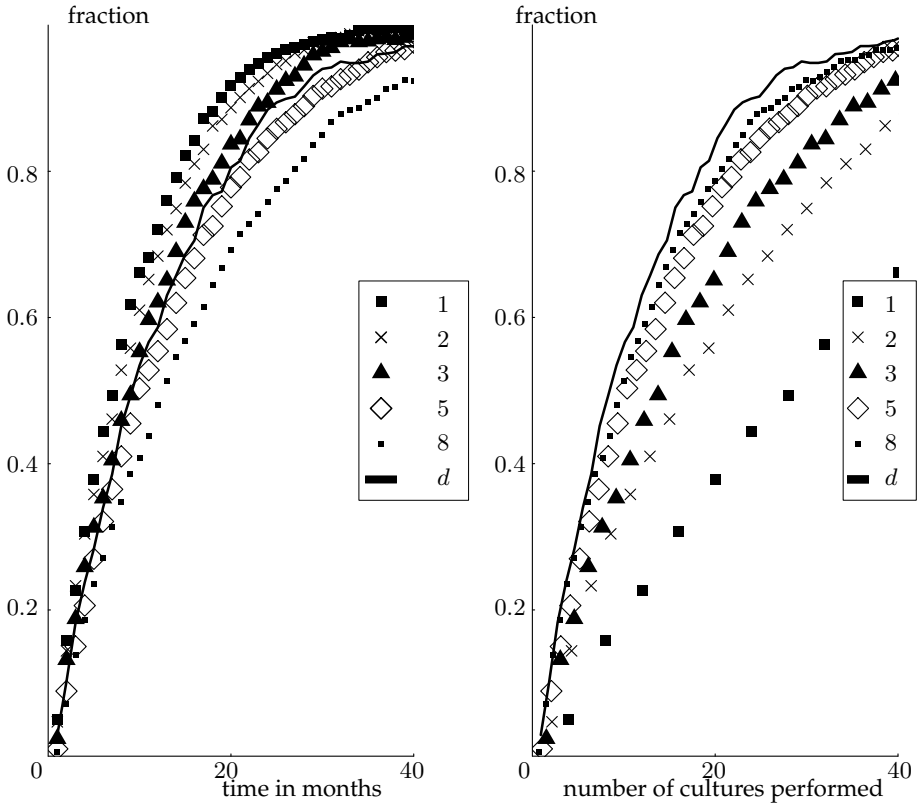
Culture frequency is associated with the time until the dominant transmission route is established (see Figure 5.5(a)). The culture regime in which patients are only cultured on the days of admission and discharge performs slightly worse than culturing on admission and afterwards every 4 days. To determine the maximal information gained per culture performed, we rescale the horizontal axis of Figure 5.5(a). The

horizontal axis now represents the number of cultures performed (with as unit the number of cultures performed per month if all patients are cultured on a daily basis independent of their colonization status.) (see Figure 5.5(b)). In this rescaling, we assume no additional cultures are performed after a positive culture.

Performing daily cultures now performs very poor and culturing only on admission and discharge gives the most information. However, the difference in information gained per performed culture are small for all culture regimes except culturing on a daily or two-daily basis.

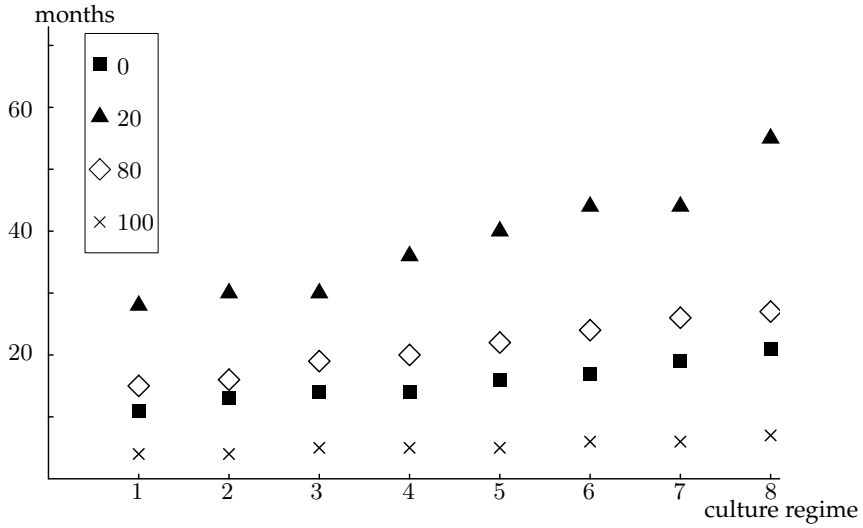
5.7 Discussion

According to the gold standard provided by genotyping and epidemiological linkage, the Markov chain model accurately quantified acquisition routes of colonization with cephalosporin-resistant Enterobacteriaceae in two intensive care units and correctly established predominance of endogenous over exogenous acquisition and correctly predicted that cross transmission of *Staphylococcus aureus* in a burn wound centre was not important. This method, therefore, seems a promising tool to provide essential information on the dynamics of microorganisms in hospital settings on the basis of fluctuations in the prevalence, without the necessity of labour-intensive and costly genotyping procedures. Antibiotic resistance is emerging in hospital settings worldwide and with diminishing antibiotics remaining available for treatment, prevention of spread will become more important. Up till now, genotyping of multiple isolates in combination with interpretation of epidemiological data have been the method of choice to reliably determine dynamics of antibiotic resistance, especially in endemic settings. However, extensive genotyping is costly and time-consuming and, therefore, hardly feasible on a daily basis. The Markov model, as proposed in this study, fulfills the need for an easy and reliable tool to evaluate the dynamics of antibiotic resistance and is able to disentangle the relevance of patient-dependent and non-dependent acquisition routes on the basis of longitudinal colonization data only. As the results of cultures usually determine whether a patient is classified as colonized or uncolonized, it seems logical to take these results as the basis of the analysis and take into account the uncertainty in the moment of acquisition due to the fact that patients are not cultured continuously. Also, the real days of admission and days of discharge of the

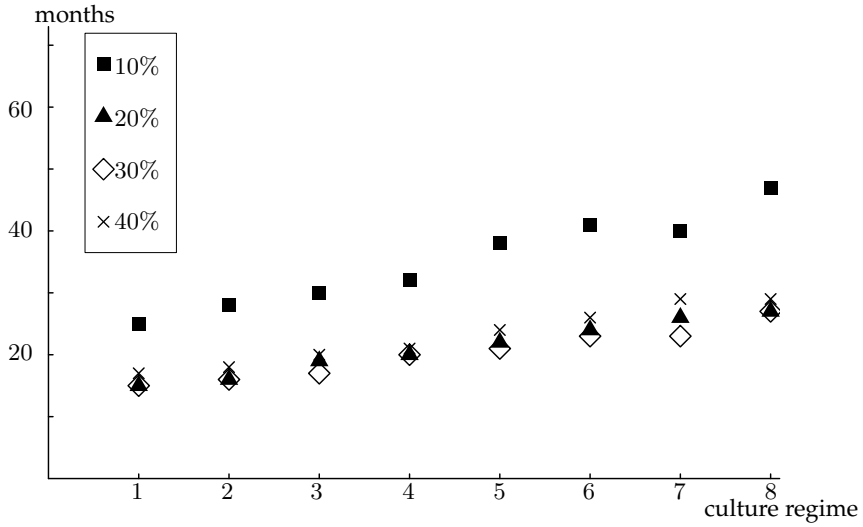


(a) time in months (b) time rescaled

Figure 5.5: Fraction of the 1000 simulation of an ICU with 10 beds in which the dominant infection route could be correctly predicted with 95% confidence for difference culture regime. **(a)** The horizontal axis denotes time in months **(b)** The horizontal axis denotes the number of performed cultures with as unit the number of cultures performed per month if patients are cultured on a daily basis. Regimes labeled with number j denote culturing on admission and afterwards every j^{th} day. The culture regime labeled with a 'd' denotes culturing on admission and on discharge. 5% of the patients is colonized on admission. The true parameters are $\alpha = 0.00515152$ and $\beta = 0.10303$ leading to a mean prevalence of 20% in the unit in which 80% of the acquisitions is due to cross transmission.



(a) mean prevalence is 30%



(b) relative importance of cross transmission is 80%

Figure 5.6: Time in months before in 80% of the 1000 simulations the dominant route could be predicted with a p-value of 0.95 as a function of the frequency. Cultures are performed on admission and afterwards every j^{th} day. An ICU with 10 beds, 5% of the patients is colonized on admission. **(a)** For each simulation the mean prevalence is 30% and the relative importance of cross transmission is 0, 20, 80 and 100%. **(b)** For each simulation the relative importance of cross transmission is 80%, the mean prevalence is 10, 20, 30 and 40%.

patients are used, avoiding extra assumptions like exponentially distributed lengths of stay. This framework allows for all kinds of Markovian transmission models and also all kind of individual patient characteristics can be used. Therefore patients need not be assumed to be identical apart from their colonization status (see the analysis of the data of the burn wound center in which we incorporated the status of the nose and the percentage of burned body surface). Moreover, spatial effects, such as whether patients share rooms or distance between beds, can be incorporated. Our model may offer three important advantages for clinical practice. First, it allows quantifying of the relative importance of exogenous and endogenous acquisition routes, which is relevant for designing infection control strategies. Exogenous transmission, usually occurring via temporarily contaminated hands of staff, depends on health-care worker-related variables, such as contact rates, level of cohorting and compliance with hand hygiene, as well as on patient (body site and bacterial load) and microbial characteristics (such as survival time of microorganisms on hands) [Bonten et al., 2001]. Modulation of these variables is warranted when exogenous transmission is an important acquisition route. Endogenous selection is driven by selective antibiotic pressure and does not depend on colonization pressure in the unit. Transformation from susceptible to resistant microorganisms can occur through mutations, upregulation of resistance genes or horizontal gene transfer. In fact, the term 'acquired' may not always be correct, as selection of pre-existing, but undetectable, flora on admission may only become apparent after some time in ICU. Reducing selective antibiotic pressure is warranted when endogenous selection is an important acquisition route. Second, the Markov methodology may improve the reliability of the interpretation of interventions, taking patient dependency into account. Many infection control interventions (such as improving hand hygiene, use of gloves and gowns and antibiotic cycling) have been analyzed in quasi-experimental designs, such as before-after studies [Puzniak et al., 2002; Harris et al., 2004]. Results were evaluated by standard statistical tests, such as χ^2-test, T-test and regression analysis that neglect dependence among patients. Therefore, if cross-transmission is relevant, differences between baseline and intervention period, considered to be statistically significant according to these statistical tests, do not necessarily prove causality between intervention and outcome. Third, this method allows quantification of infection control practices. A central concept in infectious disease dynamics is the basic reproduction number R_0,

which corresponds to the average number of secondary infected cases in a wholly susceptible population [Anderson and May, 1991; Diekmann and Heesterbeek, 2000]. Within hospital settings, R_0 represents the number of secondary cases through cross-transmission generated by a primary case in a pathogen-free ward. (Note that the concept of R_0 is not entirely correct when it is possible to attain the disease without transmission (endogenous route). However, the value of R_0 predicts whether cross transmission on itself can maintain endemicity.) Infection prevention aims to reduce R_0 to an effective R (R_N) value below unity. In our study the R_N values for cephalosporin-resistant Enterobacteriaceae were close to zero in both wards. These findings suggest that an intervention aimed at reducing cross-transmission can hardly reduce resistance prevalence any further. In other settings or for other pathogens, where R_N is > 1, calculation of R_N after an intervention allows quantification of its effects.

Our model has some limitations. First, we also need that patients are cultured on admission, as otherwise we have to make additional assumptions about the probability that an admitted patient was colonized on admission. Second, the role of environmental contamination is not explicitly incorporated. When environmental contamination can persist even when the patient who was the source of the contamination is no longer in the unit, the Markov property is not satisfied and the model is not applicable. In theory, colonization status of a patient might determine the likelihood of contamination of the inanimate environment and with discharge of the colonized patient, environmental contamination might disappear as well. In that case, the inanimate environment could be considered as a functional part of the patient and the Markov model still would apply. Third, the role of persistently colonized health care workers has not been incorporated. Such health care workers might act as a continuous source for transmission, though contacts with different patients is unlikely to be randomly distributed. However, despite the fact that multiple examples of outbreaks caused by health care workers exist, persistently colonized health care workers are, in general, not considered relevant sources for most nosocomial pathogens. Moreover, permanently colonized health care workers would impose a colonization pressure that would not depend on the prevalence of colonized patients and would, therefore, be part of the so-called endogenous process. Fourth, the number of acquisitions per unit of time is associated with the time needed for obtaining reliable results. When the prevalence is very low, there is not much information about

acquisitions in the data. On the other hand, the mean endemic prevalence in the unit should not be very high as the relative fluctuations in the daily endemic prevalence will then be small, making distinction between a constant force of infection and a force of infection that is proportional to the daily prevalence difficult. Another point is that the numerical algorithm becomes slow when the number of patients for whom the colonization status is unknown becomes large. However, the actual unit size can be very large, as long as the number of patients in $\mathcal{Q}$ does not become much larger than 10. If it does, the method can still be performed but techniques to approximate the likelihood (e.g., EM algorithm, [Wasserman, 2004]) have to be used. A hidden Markov chain approach [Cooper et al., 2004] for these glued Markov chains would be possible, but as the colonization status of most patients would be unknown, the size of a unit has to be relatively small. If we incorporate that cultures are not 100% reliable (false positive and false negative cultures), we are uncertain of the colonization status of all patients (although the likelihood of colonization depends on the outcome of the culture tests). Therefore all patients are in $\mathcal{Q}$ and again, the size of the unit should be relatively small. Another problem, not specific for this approach, is that with a very realistic model with many parameters, often little can be said about the values of the parameters while a more simplistic model ignores elements that are relevant in real life. However, the advantage of a mechanistic model is that the parameters have a clear medical/biological meaning compared to non-parametric approaches. Note also that this framework is unnecessary when the patients are independent of each other (no transmission). Nevertheless, this framework could serve as a test to determine whether dependency between patients is relevant. The confidence intervals are based on the assumption that the model, as used, is correct. Although the models are not rejected by a χ^2-test and, therefore, seem to behave reasonably well, the models do not capture the whole reality. Therefore, the true confidence intervals might be wider than the reported ones. However, this problem is inevitable as distinguishing between different infection routes is the aim of this study and a model is needed to describe the different routes.

In the data set of the burn wound center, cross transmission from wound to wound appeared to be unimportant. Interventions, based on improving the already high level of hand hygiene during wound care, are, therefore, unlikely to be effective. Colonization of the nose however, seemed to be involved in acquiring wound colonization, although

the exact route could not be determined from this dataset. Reduction of the constant background force, possibly by reducing nose colonization among HCW and preventing nose colonization of patients might be effective measures to reduce wound colonization. More detailed studies in this direction are currently performed.

Finally, simulations show that culturing on a daily basis is the fastest way to obtain results, however, the information gained per performed culture is minimal. The practice of culturing on admission and afterwards twice a week seems to provide a reasonable balance between rapid results and avoiding too many cultures.

We thank Richard Gill for useful advice.

Bibliography

Andersen, P., Borgan, Ø., Gill, R., and Keiding, N. (1993). *Statistical Models Based on Counting Processes.* Springer-Verlag, New York.

Anderson, R. M. and May, R. M. (1991). *Infectious Diseases of Humans: Dynamics and Control*. Oxford University Press.

Andersson, D. I. and Levin, B. R. (1999). The biological cost of antibiotic resistance. *Current Opinion in Microbiology*, 2:489–493.

Andersson, H. and Britton, T. (2000). *Stochastic Epidemic Models and Their Statistical Analysis*, volume 151 of *Lecture Notes in Statistics*. Springer.

Austin, D. and Anderson, R. (1999). Transmission dynamics of epidemic methicillin-resintant *Staphylococcus aureus* and vancomycin-resitant enterococci in england and wales. *The Journal of Infectious Diseases*, 179:883–891.

Austin, D., Kakehashi, M., and R.M., A. (1997). The transmission dynamics of antibiotic-resistant bacteria: the relationship betwee resistance in commensal organisms and antibiotic consumption. *Proceedings of the Royal Society, Series B*, 264:1629–1638.

Austin, D., Kristinsson, K., and Anderson, R. (Feb. 1999a). The relationship between the volume of antimicrobial consumption in human communities and the frequency of resistance. *Proceedings of the National Academy of Sciences of the United States of America*, 96:1152–1156.

Austin, D. J., Bonten, M. J., Weinstein, R. A., Slaughter, S., and Anderson, R. M. (June 1999b). Vancomycin-resitant enterococci in intensive-care hospital settings: Transmission dynamics, persistence, and the impact of infection control programs. *Proceedings of the National Academy of Sciences of the United States of America*, 96:6908–6913.

Baden, L., Thiemke, W., Skolnik, A., Strymish, J., Gold, H., Moellering, J., and Eliopoulos, G. (2001). Prolonged colonization with vancomycin-resistant *Enterococcus faecium* in long-term care patients and the significance of 'clearance'. *Clinical Infectious Diseases*, 33:1654–1660.

Ball, F., Mollison, D., and Scalia-Tomba, G. (1997). Epidemics with two levels of mixing. *The Annals of Applied Probability*, 7, No.1:46–89.

Bergmans, D. C., Bonten, M. J., Gaillard, C. A., Paling, J. C., Geest, S. v. d., Tiel, F. H., Beysens, A. J., Leeuw, P. W. d., and Stobberingh, E. E. (2001). Prevention of ventilator-associated pneumonia by oral decontamination. *American Journal of Respiratory and Critical Care Medicine*, 164:382–388.

Bergstrom, C. T., Lipsitch, M., and Levin, B. R. (August 2000). Natural selection, infectious transfer and the existence conditions for bacterial plasmids. *Genetics*, 155:1505–1519.

Bergstrom, C. T., Lo, M., and Lipsitch, M. (2004). Ecological theory suggest that antimicrobial cycling will not reduce antimicrobial resistance in hospitals. *Proceedings of the National Academy of Sciences of the United States of America*, 101, no.36:13285–13290.

Bonhoeffer, S., Lipsitch, M., and Levin, B. R. (Oct. 1997). Evaluating treatment protocols to prevent antibiotic resistance. *Proceedings of the National Academy of Sciences of the United States of America*, 94:12106–12111.

Bonten, M., Austin, D., and Lipsitch, M. (2001). Understanding the spread of antibiotic resistant pathogens in hospitals: mathematical models as tools for control. *Clinical Infectious Diseases*, 33:1739–1746.

Bonten, M., Froon, A., Gaillard, C., and et al. (1997). The systemic inflammatory response in the development of ventilator-associated pneumonia. *American Journal of Respiratory and Critical Care Medicine*, 156:1105–1113.

Bonten, M., Slaughter, S., Ambergen, A., Hayden, M., Voorhis, v. J., Nathan, C., and et al. (1998). The role of colonization pressure in the spread of vancomycin-resistant enterococci. an important infection control variable. *Archives of Internal Medicine*, 158:1127–32.

Bonten, M. and Weinstein, R. (1996). The role of colonization in the pathogenesis of nosocomial infections. *Infection Control and Hospital Epidemiology*, 17:193–200.

Byers, K., Anglim, A., Anneski, C., and Farr, B. (2002). Duration of colonization with vancomycin-resistant enterococcus. *Infection Control and Hospital Epidemiology*, 23(4):207–11.

Cook et al., D. J. (15 September 1998). Incidence of and risk factors for ventilator-associated pneumonia in critically ill patients. *Annals of Internal Medicine*, 129:433–440.

Cooper, B. and Lipsitch, M. (2004). The analysis of hospital infection data using hidden markov models. *Biostatistics*, 5,2:223–237.

Cooper, B., Medley, G., Stone, S., Kibbler, C., Cookson, B., Roberts, J., Duckworth, G., Lai, R., and Ebrahim, S. (2004). Methicillin-resistant *Staphylococcus aureus* in hospitals and the community: Stealth dynamics and control catastrophes. *Proceedings of the National Academy of Sciences of the United States of America*, 101(27):10223–10228.

Cosgrove, S. E., Sakoulas, G., Perencevich, E. N., Schwaber, M. J., Karchmer, A. W., and Carmeli, Y. (2003). Comparision of mortality associated with methicillin-resistant and methicillin-susceptible *Staphylococcus aureus* bacteremia: A meta-analysis. *Clinical Infectious Diseases*, 36:53–59.

Cox, D. and Hinkley, D. (1979). *Theoretical Statistics.* Chapman & Hall/CRC.

D'Agata, E. M., Gautam, S., Green, W. K., and Tang, Y.-W. (2002a). High rate of false-negative results of the rectal swab culture method in detection of gastrointestinal colonization with vancomycin-resistant enterococci. *Clinical Infectious Diseases*, 34:167–172.

D'Agata, E. M. C., Horn, M. A., and Webb, G. F. (2002b). The impact of persistent gastrointestinal colonization on the transmission dynamics of vancomycin-resistant enterococci. *The Journal of Infectious Diseases*, 185:766–773.

Diekmann, O. (1978). Thresholds and travelling waves for the geographical spread of infection. *Journal of Mathematical Biology*, 6:109–130.

Diekmann, O. (1979). Run for your life. A note on the asymptotic speed of propagation of an epidemic. *Journal of Differential Equations*, 33:58–73.

Diekmann, O. and Heesterbeek, J. (2000). *Mathematical Epidemiology of Infectious Diseases*. Wiley Series in Mathematical and Computational Biology.

Diekmann, O., Heesterbeek, J., and Metz, J. (1990). On the definition and the computation of the basic reproduction number R_0 in models for infectious diseases in heterogeneous populations. *Journal of Mathematical Biology*, 28:365–382.

Fagon, J.-Y., Chastre, J., Hance, A. J., Montravers, P., Novara, A., and Gibert, C. (1993). Nosocomial pneumonia in ventilated patients: A cohort study evaluating attributable mortality and hospital stay. *The American Journal of Medicine*, 94:281–288.

Ferrer, R., Bauer, T., and Torres, A. (2001). *Ventilator-associated pneumonia: Is Prevention Cost Effective? European Pespective*, volume 5 of *Perspectives on Critical Care Infectious Diseases*, chapter 13, pages 155–168. Kluwer Academic Publishers, Boston.

Fiore, M., Nieuwenhoven, M., and Ramsay, G. (2000). Decreased incidence of ventilator-associated pneumonia caused by *Pseudomonas aeruginosa* over 3 years. *British Journal of Surgery 87*, 7:931.

Fitzpatrick, F., Murphy, O., Brady, A., Prout, S., and Fenelon, L. (2000). A purpose built MRSA cohort unit. *Journal of Hospital Infection*, 46:271–279.

Grundmann, H., Hori, S., Winter, B., Tami, A., and Austin, D. J. (2002). Risk factors for the transmission of methicillin-resistant *Staphylococcus aureus* in an adult intensive care unit: Fitting a model to the data. *The Journal of Infectious Diseases*, 185:481–488.

Halloran, M. E., Longini, I. M., Nizam, A., and Yang, Y. (2002). Containing bioterrorist smallpox. *Science*, 298:1428–1432.

Harris, A., Bradham, D., Baumgarten, M., Zuckerman, I., Fink, J., and Perencevich, E. (2004). The use and interpretation of quasi-experimental studies in infectious diseases. *Clinical Infectious Diseases*, 38:1586–1591.

Hayden, M. K. (2000). Insights into the epidemiology and control of infection with vancomycin-resitant enterococci. *Clinical Infectious Diseases*, 31:1058–1065.

Heyland, D., Cook, D., L., G., Keenan, S., and Brun-Buisson, C. (1999). The attributable morbidity and mortality of vantilator-associated pneumonia in the critically ill patient. *American Journal of Respiratory and Critical Care Medicine*, 159:1249–1256.

Jonge, E. d., Schultz, M. J., Spanjaard, L., Bossuyt, P. M., Vroom, M. B., Dankert, J., and Kesecioglu, J. (2003). Effects of selective decontamination of digestive tract on mortality and acquisition of resistant bacteria in intensive care: a randomised controleed trial. *Lancet*, 362:1011–1016.

Kampen, N. v. (1981). *Stochastic Processes in Physics and Chemistry*. North-Holland Personal Library.

Kermack, W. and McKendrick, A. (1927). A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London. Series A*, 115:700–721.

Kollef, M. and Fraser, V. (2001). Antibiotic resistance in the intensive care unit. *Annals of Internal Medicine*, 134:298–314.

Kooistra-Smid, M., Dijk, S. v., Beerthuizen, G., Vogels, W., Zwet, T. v., Belkum, A. v., and Henri, V. (2004). Molecular epidemiology of *Staphylococcus aureus* colonization in a burn center. *Burns*, 30:27–33.

Lipsitch, M., Bergstrom, C. T., and Levin, B. R. (Feb. 15 2000). The epidemiology of antibiotic resistance in hospitals: Paradoxes and prescriptions. *Proceedings of the National Academy of Sciences of the United States of America*, 97, No. 4:1938–1943.

Man, P. d., Veeke, E. v. d., Leemreijze, M., Leeuwen, M. v., Vos, G., Anker, J. v. d., Verbrugh, H., and Belkum, A. v. (2001). Enterobacter species in a pediatric hospital: horizontal transfer or selection in individual patients? *The Journal of Infectious Diseases*, 184(2):211–214.

Merrer, J., Santoli, F., Appéré-De Vecchi, C., Tran, B., de Jonghe, B., and Outin, H. (2000). Colonization pressure and risk of acquisition of methicillin-resistant *Staphylococcus aureus* in a medical intensive care unit. *Infection Control and Hospital Epidemiology*, 21:718–23.

Mitsuda, T., Arai, K., Fujita, S., and Yokota, S. (1995). Epidemiological analysis of strains of methicillin-resistant *Staphylococcus aureus* (MRSA) infection in the nursery: prognosis of MRSA carrier infants. *Journal of Hospital Infection*, 31:123–134.

Nijssen, S., Bootsma, M., Bonten, M., and Hoepelman, A. (2002). Dynamics of colonization with 3rd generation cephalosporin-resistant enterobacteriaceae in icu patients. abstract no: 2179 ICAAC.

Nijssen, S., Bootsma, M., Fluit, A., Diekmann, O., and Bonten, M. (2005). Determining the relative importance of microbial acquisition routes in hospital settings. *Submitted*.

Pelupessy, I., Bonten, M. J., and Diekmann, O. (2002). How to assess the relative importance of different colonization routes of pathogens within hospital settings. *Proceedings of the National Academy of Sciences of the United States of America*, 99, No. 8:5601–5605.

Puzniak, L., Leet, T., Mayfield, J., Kollef, M., and Mundy, L. (2002). To gown or not to gown: The effect on acquisition of vancomycin-resistant enterococci. *Clinical Infectious Diseases*, 35:18–25.

Rass, L. and Radcliffe, J. (2003). *Spatial Deterministic Epidemics*, volume 102 of *Mathematicl Surveys and Monographs*. American Mathematical Society.

Reacher, M., Shah, A., Livermore, D., Wale, M., Graham, C., and Johnson, A.P., e. a. (2000). Bacteremia and antibiotic resistance of its pathogens reported in england and wales between 1990 and 1998: trend analysis. *British Medical Journal*, 320:213–216.

Rello, J., Jubert, P., Valles, J., Artigas, A., Rué, M., and Niederman, M. (1996). Evaluation of outcome for intubated patients with pneumonia due to *Pseudomonas aeruginosa*. *Clinical Infectious Diseases*, 23:973–978.

Salgado, C. D., Farr, B. M., and Calfee, D. P. (2003). Community-acquired methicillin-resistant *Staphylococcus aureus*: A meta-analysis of prevalence and risk factors. *Clinical Infectious Diseases*, 36:131–139.

Sanford, M., Widmer, A., Bale, M., Jones, R., and Wenzel, R. (1994). Efficient detection and longterm persistence of the carriage of methicillin-resistant emphS. aureus. *Clinical Infectious Diseases*, 19:1123–1128.

Scanvic, A., Denic, L., Gaillon, S., Giry, P., Andremont, A., and Lucet, J. (2001). Duration of colonization by methicillin-resistant *Staphylococcus aureus* after hospital discharge and risk factors for prolonged carriage. *Clinical Infectious Diseases*, 32:1393–1398.

Schulgen, G. (1995). *Multistate models for event-history data: Modelling occurence and impact of nosocomial infections*. PhD thesis, University of Freiburg.

Stewart, F., Antia, R., Levin, B., Lipsitch, M., and Mittler, J. (1998). Population genetics of antibiotic resistance ii: Analyic theory for sustained populations of bacteria in a community of hosts. *Theoretical Biology*, 53:152–165.

Timsit, J., Chevret, S., J., V., and et al. (1996). Mortality of nosocomial pneumonia in ventilated patients: influence of diagnostic tools. *American Journal of Respiratory and Critical Care Medicine*, 154:116–123.

Venzon, D. and Moolgavkar, S. (1988). A method for computing profile-likelihood-based confidence intervals. *Applied Statistics*, 37, No. 1:87–94.

Verhoef, J., Beaujean, D., Blok, H., Baars, A., Meyler, A., and Werken, C. v. d. (1999). A dutch approach to methicillin-resistant *Staphylococcus aureus*. *European Journal of Clinical Microbiology and Infectious Diseases*, 18:461–466.

Vincent, J., Bihari, D., Suter, P., and et al. (1995). The prevalence of nosocomial infection in intensive care units in europe. results of the european prevalence of infection in intensive care (EPIC) study. *Journal of the American Medical Association*, 274:639–644.

Vollaard, E. (1991). *The concept of colonization resistance*. PhD thesis, Katholieke Universiteit Nijmegen.

Voss, A. (2004). Preventing the spread of MRSA. common sense and observational studies are of benefit. *British Medical Journal*, 329:521.

Vriens, M., Blok, H., Fluit, A., Troelstra, A., Werken, C. V. D., and Verhoef, J. (2002). Costs associated with a strict policy to eradicate methicillin-resistant *Staphylococcus aureus* in a dutch university medical center: A 10-year survey. *European Journal of Clinical Microbiology and Infectious Diseases*, 21(11):782–786.

Wallinga, J. and Teunis, P. (2004). Different epidemic curves for severe acute respiratory syndrome reveal similar impacts of control measures. *American Journal of Epidemilogy*, 160(6):509–516.

Wasserman, L. (2004). *All of Statistics*. Springer Verlag, New York.

Werkgroep Infectie Preventie (2003). Ziekenhuizen: MRSA algemeen.

Willems et al., R. (2001). Variant esp gene as a marker of a distinct genetic lineage of vancomycin-resistant *Enterococcus faecium* spreading in hospitals. *Lancet*, 357.

Samenvatting

Dit proefschrift gaat over de dynamiek van antibioticaresistentie. Ziekenhuizen, en binnen ziekenhuizen intensive care afdelingen, fungeren als epicentra van resistentieproblemen. Hiervoor zijn twee redenen. In ziekenhuizen, en in het bijzonder op intensive care afdelingen, liggen patiënten met verhoogde vatbaarheid voor infecties. Het gebruik van antibiotica op intensive care afdelingen is dan ook substantieel hoger dan in de rest van het ziekenhuis en in de open populatie. Dit antibioticagebruik leidt ertoe dat bacteriestammen die resistent zijn voor de gebruikte antibiotica een selectievoordeel krijgen ten opzichte van niet-resistente stammen. In hoofdstuk 2 wordt het belang van ziekenhuizen als epicentra onderzocht en wordt er specifiek ingegaan op de interactie tussen dragerschap van resistente bacteriën in de open populatie en verspreiding van deze stammen in het ziekenhuis. In hoofdstuk 3 worden twee modellen behandeld voor de gevreesde ziekenhuisbacterie Methicilline-resistente *Staphylococcus aureus* (MRSA). In het eerste model wordt naar de effectiviteit van verschillende infectiepreventiestrategiën gekeken uitgaande van de huidige Nederlandse situatie. In het tweede model worden verschillende scenario's gesimuleerd, zowel voor de Nederlandse als de Amerikaanse situatie. Ook wordt het effect van transmissie in de open populatie onderzocht.

De laatste twee hoofdstukken zijn data-georiënteerd. In hoofdstuk 4 bekijken we twee methodes om onderscheid te maken tussen de twee fenomenen dat de kans op infectie toeneemt naarmate een patiënt langer op een intensive care afdeling ligt en dat de toestand van een patiënt verslechtert als gevolg van de infectie waardoor de patient langer op de intensive care afdeling ligt.

In het laatste hoofdstuk wordt een kader geïntroduceerd waarmee op basis van longitudinale data (opnamegegevens en kweekgegevens) het belang van verschillende infectieroutes bepaald kan worden. Binnen deze Markov-keten benadering wordt getracht zo dicht mogelijk bij de beschikbare data te blijven door bij de wiskundige beschrijving van de 'toestand' in, bijvoorbeeld, een intensive care afdeling, de werkelijk aanwezige gegevens als uitgangspunt te nemen.

Dankwoord

Nu de resultaten van een jarenlang onderzoek zijn neergeschreven in dit proefschrift is het moment aangebroken om een aantal mensen te bedanken.

Mijn dank gaat in de eerste plaats uit naar mijn beide promotores. Odo, het was uitermate prettig om met je samen te werken. Je kritische lezen van manuscripten en je wens om een gedachtengang nauwkeurig te formuleren hebben de leesbaarheid van de stukken aanmerkelijk vergroot. Daarnaast wil ik je bedanken voor de prettige omgang. De wekelijkse werkbesprekingen waren niet alleen nuttig maar ook gezellig.

Marc, je interesse in de toepasbaarheid van wiskunde en de relevantie voor de medische wereld was een welkome aanvulling op de theoretische aanpak zoals gebruikelijk binnen wiskunde. Bedankt voor je gedrevenheid en enthousiasme, voor het uitgebreid lezen van de manuscripten, en ook voor de aangename uurtjes in de kroeg.

I am also grateful to the members of the reading comittee:

1. Tom Britton (Stockholm University, Sweden)
2. Neil Ferguson (Imperial College, London, Great Britain)
3. Richard Gill (Utrecht University, the Netherlands)
4. Hajo Grundmann (Rijksinstituut voor Volksgezondheid en Milieu, Bilthoven, the Netherlands)
5. Hans Heesterbeek (Utrecht University, the Netherlands)

for the time and expertise they have invested in reading the manuscript and for their useful comments and suggestions.

I would like to thank my current and former colleagues in room 610, Johnny Edwards, Oliver Lorscheid, Tobias Mühlenbruch Hamid Rahmati and Joost Rommes, the other AIO's of Odo, Barbara Boldin, Natalia Davydova, Philipp Getto and Yunxin Huang but also the other PhD-students and other members of the Mathematical Institute for the nice atmosphere.

Ook wil ik de medewerkers van de afdeling infectieziekten van het UMCU bedanken voor de prettige sfeer. Ik ben altijd met veel plezier

naar de werkbesprekingen en congressen gegaan. In het bijzonder wil ik Andy, Saskia, Helen, Jan Jelrik, Karin, Stefan, Fieke, Irene, Marianne, Jannetta, Rob, Willem, Ad en Alice noemen. Aan de ziekenhuishygiène ben ik dank verschuldigd voor het aanleveren van data, evenals aan Constance Schultz en Anneke van der Zee.

Ook wil ik het brandwondencentrum van het Martini-ziekenhuis in Groningen graag vermelden. De prettige en langdurige samenwerking heeft zeker bijgedragen aan het aanscherpen van de probleem formulering en de oplossing in hoofdstuk 5. In het bijzonder wil ik Mirjam Kooistra en Evert van Zanten noemen voor het aanleveren van de data.

De deelnemers aan het epidemiologie-leesclubje wil ik bedanken voor de genomen moeite om steeds naar Utrecht te komen. Ik hoop dat dit initiatief een lang leven beschoren zal zijn.

Naast het werk was er gelukkig ook nog tijd voor ontspanning. Ik wil Robbert bedanken voor het tennissen, alles en iedereen van de studentenbridgeclub Dombo voor de gezelligheid, in het bijzonder René, Hendrik, Jannes, Daniël, Nira en Remco. De leden van de schaakvereniging Almelo hebben ervoor hebben gezorgd dat ik nog steeds met veel plezier een paar partijtjes per jaar schaak, waarvoor dank. Verder dank aan de nog niet eerder genoemde vrienden Evert, Sylvie, Edward, Germen, Aliza, Florine, Beernd, Joanna en Stefan.

Inti, mijn paranimf, bedankt voor je belangstelling in de sterrenkunde waardoor deze promotieplaats voor mij beschikbaar kwam. Ook wil ik mijn andere paranimf, mijn zus Janet, bedanken evenals mijn broer Huub en de overige familieleden. In het bijzonder wil ik mijn ouders, Leni en Chris Bootsma bedanken voor hun onvoorwaardelijke steun.

Martin Bootsma
Utrecht, 2005

Curriculum vitae

Op 18 december 1977 ben ik geboren te Almelo. Van 1989 tot 1995 bezocht ik het Pius X College te Almelo alwaar ik het Gymnasium diploma behaalde. In september 1995 ging ik studeren aan de Universiteit Utrecht en in augustus 1996 haalde ik mijn propedeuse Wiskunde (cum laude) en Natuurkunde. In januari 2001 studeerde ik af in de theoretische natuurkunde (cum laude) bij Prof. Dr. H. van Beijeren met de scriptie getiteld "The influence of impurities on equilibrium crystal shapes". Tevens behaalde ik in januari 2001 een vrij doctoraal wiskunde. Tijdens mijn studie gaf ik als student-assistent werkcolleges.

In februari 2001 werd ik aangesteld als assistent in opleiding (AIO) bij het Mathematisch Instituut van de faculteit Wiskunde en Informatica van de Universiteit Utrecht onder begeleiding van Prof. Dr. O. Diekmann en Prof. Dr. M.J.M. Bonten wat geresulteerd heeft in dit proefschrift. In deze periode heb ik diverse congressen en scholen bezocht, o.a. in Bielefeld en Oberwolfach (Duitsland), Mariefred (Zweden), Venetië (Italië), Lissabon (Portugal), San Diego, Chicago (Verenigde Staten), Woudschoten en Wageningen (Nederland). Tevens heb ik als AIO diverse werkcolleges gegeven.