MATHEMATICAL MODELING OF DISEASES TO INFORM HEALTH POLICY

A Thesis
Presented to
The Academic Faculty

by

Daniel Mello Faissol

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
H. Milton Stewart School of Industrial and Systems Engineering

Georgia Institute of Technology
Aug 2008

MATHEMATICAL MODELING OF DISEASES TO INFORM HEALTH POLICY

Approved by:

Thomas L. Gift
Division of STD Prevention
Centers for Disease Control and Prevention

David M. Goldsman
H. Milton Stewart School of Industrial and Systems Engineering
Georgia Institute of Technology

Paul M. Griffin
H. Milton Stewart School of Industrial and Systems Engineering
Georgia Institute of Technology

Pinar Keskinocak
H. Milton Stewart School of Industrial and Systems Engineering
Georgia Institute of Technology

Julie L. Swann
H. Milton Stewart School of Industrial and Systems Engineering
Georgia Institute of Technology

Date Approved: Aug 2008

To my father,

Sergio Zarur Faissol,

whose inspiration made this possible.

ACKNOWLEDGEMENTS

I would like to thank many people for their contributions that led to this dissertation. I would first like to thank my advisors, Julie Swann and Paul Griffin, with whom I had the very good fortune of working. I am grateful for the generous amount of time and effort they both contributed. I consider myself to be one of the more fortunate students for having advisors of such good nature and genuine interest in my success as a graduate student and beyond.

I would like to thank the other members of my committee, especially Thomas Gift who collaborated with us on Chapter 2, as well as David Goldsman and Pinar Keskinocak for their helpful comments and patience. Additionally, I would like to express my gratitude to many of the ISyE faculty and staff, several of whom went well beyond the call of duty, particularly Steven Hackman for his entertaining classes from which I learned a great deal.

I am also very grateful for the financial support that the AT&T Labs Fellowship Program and the NASA Harriet G. Jenkins Predoctoral Fellowship Program have provided me during my time as a graduate student.

I have had the pleasure of knowing and working with many fellow graduate students of whom there are too many to list. I would however, like to thank H. Eser Kirkizlar for his collaboration in Chapter 3, Claudio Santiago and Helder Inacio for their help in Chapter 5, Charles Leslie Wardell for allowing me to carry him on my shoulders during our course work and as we prepared for the comprehensive exam, and each of them for their invaluable friendship and support during my entire graduate career.

Finally, I would like to thank my family: Sergio, Mirian, Cecilia and Patricia

Faissol, in decreasing order of very old age. My parents' hard work, dedication, patience, support and love are more than anybody could ever ask for, without which I would not have even begun writing this document. Lastly, I owe a special debt of gratitude to my fiancée, Dorothy, who is arguably my better half. Her patience, commitment, support and love are cherished, without which I would not have finished writing this document.

TABLE OF CONTENTS

[1]This chapter is a collaborative effort with H. Eser Kirkizlar.

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Disease affects virtually everyone in society. Our increasing interconnectedness makes understanding infectious diseases ever more critical. Coupled with the expanding amount of health data currently available, the complex nature of many diseases make them ideal candidates for mathematical modeling. In light of the rising costs of health care, mathematical modeling provides a relatively inexpensive way of helping us better understand how infectious diseases spread, how to better allocate funds, and how to pick the most effective interventions aimed at preventing and treating disease. In this dissertation we present mathematical models that help answer health policy questions relating to HIV and Hepatitis C, and analyze bias in Markov models of general disease progression.

In the next chapter, we study a specific question in HIV policy. During the initial HIV outbreak in the 1980's public health agencies around the country actively closed down bathhouses because they were identified as a venue in which high risk behavior was taking place and were perceived to be one of the major factors contributing to the spread of HIV. Many of these venues have reopened today, and some blame them for the increasing HIV incidence that is currently taking place. The debate as to how public health agencies should react to this phenomenon lacked a mathematical analysis of the issue. To study the problem, we develop a Bernoulli process transmission model where, for a given individual, each risky person-to-person contact is treated as an independent Bernoulli trial with an associated probability of HIV transmission. We extend the model to include a heterogeneous population with multiple risk groups and add the effect of co-infection with other diseases, such as Syphilis, which increase

the probability of transmission when present. We show that the HIV attack rate is concave as a function of the proportion of the bathhouse patrons' contacts that are with other bathhouse patrons. We use this fact to draw conclusions on the effect of closing bathhouses under certain assumptions.

We populate the model of HIV transmission with data from a survey of four major cities in the US. A key finding is that the impact on HIV incidence from the disproportionate mixing of the population due to the presence of bathhouses is small compared to the impact from changes in some key parameter values, such as condom usage. The effect that closing bathhouses will have on these parameter values is not clear; however, the results suggest that alternative interventions targeted at individuals in bathhouse venues could have greater effects on the spread of HIV than closing bathhouses.

In Chapters 3 and 4, we build a mathematical model to examine the timing of testing and treatment for diseases, particularly Hepatitis C. Hepatitis C is the leading cause for liver transplants and the 10^{th} leading cause of death in US adults. It is typically asymptomatic for decades yet still infectious during this period. Many papers in the medical literature analyze the cost-effectiveness of screening by simulating the disease and a limited number of *a priori* testing policies. However, this may be insufficient to determine the best timing of the tests or incorporate changes over time. In Chapters 3 and 4, we study this problem with a dual approach, both analytical and simulation. We develop a Markov Decision Process (MDP) model for diseases where our goal is to determine the best timing for testing (and treatment) decisions when the presence of the disease is not known in advance; our model allows for the awareness of a disease to change behavior.

Using medical data, we arrive at a dynamic policy of testing and treating for the case of Hepatitis C in Chapter 3. In Chapter 4 we also simulate all policies of up to five tests in a lifetime to examine both optimal and practically implementable policies.

A key finding in both chapters is that the current policy recommendations on testing for Hepatitis C are too restrictive, and that it is cost-effective to test the overall population if done at the appropriate times. We also demonstrate the importance of including behavior changes in the model and analyzing the optimal timing of tests by comparing the results to previous studies.

The Markov models used in the study of Hepatitis C in Chapters 3 and 4 motivated the topic in Chapter 5 where we examine bias in Markov models of diseases, including the one studied in Chapter 3 and 4. We examine two classes of diseases and the associated Markov models commonly used to model them: ones in which the transition probabilities are state dependent (that is, they vary by severity of the disease), and ones in which the transition probabilities are time dependent. We find the behavior of these Markov models in steady state and arrive at sufficient conditions for bias to exist in models with aggregated transition probabilities when compared to models with state/time dependent transition probabilities. We also find that when aggregating data to compute transition probabilities, the bias increases with the degree of data aggregation. We apply the results to Hepatitis C, Alzheimer's disease and lung cancer and find that the bias is significant depending on the method used to aggregate the data.

In the final chapter we present possibilities for extensions of the research topics discussed. We plan to expand the models in Chapters 3 and 5 to larger state spaces so that we can model more complicated diseases. We aim to use the future results to model diseases such as cancer, diabetes, and cardiovascular disease.

CHAPTER II

A BERNOULLI PROCESS TRANSMISSION MODEL FOR HIV TRANSMISSION: THE ROLE OF BATHHOUSES IN HIV TRANSMISSION

2.1 Introduction

In the 1980s and early 1990s, sexually transmitted disease (STD) rates in men who have sex with men (MSM) fell following the adoption of safer sex behavior among those who were impacted most strongly in the early years of the HIV/AIDS epidemic [49]. However, in the US, there continues to be a substantial amount of high risk sexual behavior and HIV infection among MSM [15, 117, 17, 44, 30]. Although most demographic categories have seen decreases in HIV incidence, MSM remain at high risk and studies have shown increasing STD rates and risk behavior among MSM [49, 27]. They make up approximately 45% of newly reported HIV/AIDS diagnoses [18, 24] and annual incidence ranges from high levels of 1.2% to 8.0% [22]. In addition, in several large US cities where approximately 25% of MSM have HIV, nearly 50% are unaware of their infection [22].

Research has documented that commercial sex venues, such as bathhouses, provide opportunities for casual and anonymous sex between MSM [140, 129, 45, 134]. Bathhouses have been associated with behaviors that increase risk for STDs and HIV, including sex with multiple partners and unprotected sex [129, 81, 10].

In the early 1980s, bathhouses were identified as a nexus of HIV transmission. As a result, bathhouses were the topic of heated debate among public health officials, government leaders and the gay community [111, 39, 13, 99, 8]. Many states developed policies to regulate bathhouse behavior; the most extreme of these was advocating

bathhouse closure [111, 39, 13, 141]. Alternatively, opponents of bathhouse closure argued that bathhouses could be used as venues to promote HIV and STD prevention interventions among both high risk and hard to reach populations [137].

By the late 1980s, much of the public debate over bathhouses ended because many of the bathhouses were shut down [72]. However, many bathhouses are currently operating around the U.S. in major U.S. cities [72, 142].

Recent research indicates that men who attend bathhouses have been shown to engage in more sexually risky behavior than those who do not [143]. A case-control study in New York City showed that MSM with syphilis were more likely to have visited bathhouses than controls [90]; surveillance data have indicated high rates of bathhouse patronage among MSM with repeat syphilis infections [32]. A second case control study from Los Angeles demonstrated that MSM who attend bathhouses and sex clubs have an increased risk of contracting syphilis as compared to those MSM who do not patronize bathhouses and sex clubs [73]. Other US and international studies have shown that bathhouse patronage was associated with hepatitis A, syphilis, and lymphogranuloma venereum [49].

Some jurisdictions have existing regulations that would provide a basis for the closure of bathhouses, and doing so has been suggested as a public health intervention to limit STD and HIV transmission [13, 46]. Historically, most of the policies aimed at regulating bathhouse behavior have been based on little, if any, data [141]. Alternatively, because they provide access to a population engaging in sexual risk behavior, they have been and could be used as venues to promote HIV and STD prevention interventions [137]. Although numerous bathhouse interventions to reduce high risk sexual behavior have been suggested, developed and implemented, most of these have not been evaluated for effectiveness [137, 9]. One exception is a study by Woods et al [141], which assesses how changes in bathhouse architecture, mandated by city-wide policies, affect the sexual behavior of MSM.

Thompson [126] considered the role that a small, yet very high-risk, subset of a population has on HIV transmission by comparing it to a population with uniform risk. Since there was very little data regarding HIV available at that time, the model was constructed so that it would rely on very few parameters. The availability of behavioral data that enables risk behavior to be differentiated among MSM enabled us to develop an HIV transmission model that explicitly considers behaviors in various subpopulations.

We sought to estimate the potential impact of bathhouse closure on HIV transmission in comparison to an alternative of a reduction in risk behavior by those attending bathhouses. To accomplish this, we constructed a transmission model based on the Bernoulli process model of HIV infection originally developed by Steven Pinkerton and Paul Abramson to determine how the closure of bathhouses would affect HIV transmission [95]. We extend the model to include subpopulations with different behaviors and the presence of syphilis, populate the model with data from a survey of MSM, and estimate the implications of various bathhouse policies on HIV transmission.

2.2 Methods

2.2.1 The Mathematical Model

In the Bernoulli process model each sex act is treated as an independent Bernoulli trial. That is, the probability of infection from one act is a constant and is independent of all other acts. In our model, we defined the probability of HIV transmission by sex act and not by partnership due to the detail of the data that we used.[1]

Let π_T be the HIV prevalence of the total population, and α_u (α_p) be the probability an individual acquires HIV from one unprotected (protected) sex act with an

[1]Note that we also modeled the probability of HIV transmission by partnership, and the conclusion of whether closing a bathhouse would increase or decrease HIV prevalence does not change. Readers interested in the details of this analysis can contact the authors.

HIV infected partner. Note that α_p can be represented as $\beta\alpha_u$ where $(1-\beta)*100\%$ is the percent effectiveness of condoms. Pinkerton and Abramson [95] established that the probability (P_1) of an individual acquiring HIV after m acts where $x\%$ of them are protected with a single ("main") partner of unknown HIV status is:

$$P_1 = \pi_T(1-(1-\alpha_u)^{(1-x)m}(1-\alpha_p)^{xm}). \tag{1}$$

Similarly, they established that for a person who has n total acts of which $y\%$ are protected each with *different* (or "non-main") partners of unknown HIV status, the probability (P_2) of becoming infected is:

$$P_2 = 1-(1-\pi_T\alpha_u)^{(1-y)n}(1-\pi_T\alpha_p)^{yn}. \tag{2}$$

We can combine Equations 1 and 2 to find the probability an individual does *not* acquire HIV after m contacts with one (main) partner (x percent of which are protected) and n additional contacts each with different (non-main) partners (y percent of which are protected).

In this paper, we have extended the Bernoulli transmission framework to account for different subpopulations, each with different behavioral characteristics and HIV prevalences. This development will provide the machinery to analyze the effect that bathhouses have on HIV transmission.

In order to compare different bathhouse policies, we will need to distinguish between sexual behaviors of those who visit bathhouses and those who do not, as indicated by the data. Indeed, it is useful to consider several different categories of behavior. To this end, we define "types" of individuals based on a variety of criteria, and effectively divide the population into subpopulations where all individuals in a subpopulation have the same type. The data indicate that there may be significant differences across groups including whether an individual visits bathhouses, has a

"main" partner, or is HIV+.

Consequently, we first divide the population into individuals who visit bathhouses and those that do not. Call these subpopulations BH and NB, respectively. Let π_{BH} (π_{NB}) be the HIV prevalence of BH (NB). For each of BH and NB, we define two subpopulations based on the presence or absence of a main partner. For those men who currently have a main partner, their contacts are split between a main partner and others from the population of MSM at large. Those who do not have a main partner draw all of their contacts from the population at large. Finally, we further divide each of these 4 subpopulations by HIV+ and HIV-, for a total of 8 subpopulations (four of which are BH and four of which are NB).

For a given individual in subpopulation i we define:

- m_i as the number of sex acts with a main partner
- x_i as the percentage of main partner sex acts that are protected
- n_i as the number of sex acts with non-main partners
- y_i as the percentage of non-main partner sex acts that are protected

Because the effect bathhouses have on the spread of HIV is the main topic of this study, we are also concerned with the proportion of the sex acts that are with members of BH. We denote this proportion as z. We define z_{BH} as the proportion of non-main contacts of members of BH that are with other members of BH. Similarly, we define z_{NB} as the proportion of non-main contacts of members of NB that are with members of BH. We can extend the Bernoulli model to express the probability of a member of BH not acquiring HIV with a partner of unknown HIV status in a time window of a year by:

$$\begin{aligned}\overline{P_{xm,ynz_{BH}}} &= \underbrace{[1-\pi_T[1-(1-\alpha_p)^{xm}(1-\alpha_u)^{(1-x)m}]]}_{A} \\ &\times \underbrace{(1-\pi_{BH}\alpha_p)^{yz_{BH}n}}_{B}\underbrace{(1-\pi_{BH}\alpha_u)^{(1-y)z_{BH}n}}_{C} \\ &\times \underbrace{(1-\pi_{NB}\alpha_p)^{y(1-z_{BH})n}}_{D}\underbrace{(1-\pi_{NB}\alpha_u)^{(1-y)(1-z_{BH})n}}_{E}, \end{aligned} \tag{3}$$

where we removed the indices on m, x, n, and y for simplicity. Each labeled term in Equation 3 above represents the probability of not acquiring HIV given a set of partners and acts. The partners and acts corresponding to each term are described in words below:

- A: m sex acts with main partner (x percent protected)
- B: $yz_{BH}n$ protected sex acts with non-main partners who are bathhouse patrons
- C: $(1-y)z_{BH}n$ unprotected sex acts with non-main partners who are bathhouse patrons
- D: $y(1-z_{BH})n$ protected sex acts with non-main partners who are not bathhouse patrons
- E: $(1-y)(1-z_{BH})n$ unprotected sex acts with non-main partners who are not bathhouse patrons

Now, let $P_{xm,ynz_{BH}} := 1 - \overline{P_{xm,ynz_{BH}}}$ be the probability that one individual from BH acquires HIV given the set of behavior parameters m, x, n, y, and z_{BH}. A similar equation exists for the individuals from NB.

If we assume everyone in a given subpopulation has the same behavior parameters, then 1 - Equation 3 gives the expected number of new HIV cases for that subpopulation. For clarity, let us index the BH subpopulations by i and the NB subpopulations

by j. Summing over all subpopulations gives us the expected number of new HIV cases for the population as a whole as

$$\sum_{i=1}^{M_{BH}} (P_{x_i m_i, y_i n_i z_{BH}}) * N_i + \sum_{j=1}^{M_{NB}} (P_{x_j m_j, y_j n_j z_{NB}}) * N_j, \tag{4}$$

where N_k represents the number of people in subpopulation k, and there are four M_{BH} bathhouse subpopulations indexed by i and four M_{NB} nonbathhouse subpopulations indexed by j[2].

Given a set of parameters, Equation 4 makes it possible to compare the average total number of new infections in a year given that a bathhouse is open or closed. It is clear that the parameters z_{BH} and z_{NB}, indicating the degree of mixing between BH and NB, will be different depending on whether the bathhouse is open or closed; we estimated these values from the data. We also examined the impact of other differences in behavioral parameters to indicate changes brought on by policies, system changes, or interventions.

For analysis on the subpopulations to be accurate, certain identities must hold. For example, the number of sex acts that the members of BH have with NB can be calculated from the data as the number of total acts by a BH individual times the percentage of those acts outside the bathhouse times the proportion of those acts with the NB population. The number of sex acts that the members of NB have with BH can correspondingly be calculated from data as the number of total acts by an NB individual time the proportion of those acts with the BH population. Clearly these two number of sex acts should be the same. Similarly, the number of protected and unprotected acts each must also match. However, survey data on sexual behavior will for a variety of reasons contain some inaccuracies, and as result, the identities will not precisely hold. Even if every member of the population is sampled in the survey,

[2]For bathhouse and nonbathhouse, the subpopulations are {main/HIV+, main/HIV-, no main/HIV+, no main/HIV-}.

it is possible that someone reports having a sex act in the survey, but his partner does not report it. It is also possible that they recalled their condom usage during the act differently. Additionally, we are using averages within each of the subpopulations, i.e., we do not explicitly model each individual's behavior, and not all individuals within a community were included in the survey. In our results we therefore used an error tolerance of 33%. That is, we only considered data points of our parameters if each of these described identities had an error of less than 33%.

2.2.2 Including Syphilis

The presence of primary or secondary syphilis (P&S) has been shown to dramatically increase the probability of HIV acquisition and transmission. Estimates of this increase range from 3-15 times the probability without syphilis, though most studies have focused on heterosexual transmission and many have combined all genital ulcerative diseases together (including chancroid and genital herpes simplex virus infection) [105, 125, 132]. Because of the increase in syphilis rates among MSM in major US cities [92], the model may more accurately reflect transmission under varying scenarios if syphilis can be incorporated.

Because we included the effect of syphilis into the model by using the same process as the previously-described model, and the equations become much more cumbersome, we will simply provide a qualitative explanation and provide mathematical details in the Appendix.

In the most specific consideration, syphilis can change the rate of infectivity of an individual in different ways depending on whether the individual without HIV has syphilis, whether his partner has syphilis, or whether both have the disease. To calculate the probability of infection for a susceptible individual, we use a similar method as in Equation 3 but extend the equation to incorporate changes in infectivity rates depending on whether one of the partners, both, or none have syphilis.

To extend the model, we take into the account the susceptible individual's syphilis status and the proportion of their contacts that are with syphilis positive individuals. The former implies that the total number of subpopulations under consideration will double (positive or negative for syphilis). The latter is determined simply by the syphilis prevalence of the corresponding subpopulation. When neither individual with sexual contact has syphilis, the same infectivity rates are used as before. When the individual in question has syphilis but his partner does not, the appropriate transmission probability (α_u or α_p) is multiplied by some constant, a. When the individual in question does not have syphilis but his partner does, the appropriate transmission probability is multiplied by a different constant, b. When both individuals in question have syphilis, the appropriate transmission probability is multiplied by yet another constant, c. See Equations 12 and 13 in the Appendix for the full model. This general model can also be simplified for the case when the infectivity rates are not different for all of these combinations by making the constants $a = b = c$.

2.2.3 Analytics

We show that the HIV attack rate is concave as a function of z_{BH}. This result is useful because the data in the subsequent section indicate that z_{BH} is greater than where the maximum of the concave function occurs regardless of whether bathhouses are open or closed (given that the number of contacts remains the same in both scenarios). Consequently, a decrease in z_{BH} (resulting from bathhouse closure) causes an increase in the HIV attack rate. The intuition is that the increase in z_{BH} benefits the low risk group more than it harms the high risk group causing an overall decrease in the number of HIV cases. Data cited in the literature (cite Woods letter) supports the hypothesis that the number of contacts will remain near constant in both scenarios, and hence we can use the concavity result to conclude that the HIV attack rate will increase with bathhouse closure.

Equation 3 defined the probability of an individual of the bathhouse population characterized by the parameters x, m, y, n and z_{BH} not acquiring HIV as

$$\begin{aligned}\overline{P_{xm,ynz_{BH}}} &= \underbrace{[1-\pi_T[1-(1-\alpha_p)^{xm}(1-\alpha_u)^{(1-x)m}]]}_{A} \\ &\times \underbrace{(1-\pi_{BH}\alpha_p)^{yz_{BH}n}}_{B}\underbrace{(1-\pi_{BH}\alpha_u)^{(1-y)z_{BH}n}}_{C} \\ &\times \underbrace{(1-\pi_{NB}\alpha_p)^{y(1-z_{BH})n}}_{D}\underbrace{(1-\pi_{NB}\alpha_u)^{(1-y)(1-z_{BH})n}}_{E}, \end{aligned} \tag{5}$$

but it can easily be re-written as

$$\begin{aligned}P_{xm,ynz_{BH}} &= [1-\pi[1-(1-\alpha_p)^{xm}(1-\alpha_u)^{(1-x)m}]] \\ &\times [(\frac{(1-\pi_B\alpha_p)}{(1-\pi_B\alpha_u)} * \frac{(1-\pi_N\alpha_u)}{(1-\pi_N\alpha_p)})^y * \frac{(1-\pi_B\alpha_u)}{(1-\pi_N\alpha_u)}]^{z_{BH}n} \end{aligned} \tag{6}$$

If we define

$$a = (\frac{(1-\pi_B\alpha_p)}{(1-\pi_B\alpha_u)} * \frac{(1-\pi_N\alpha_u)}{(1-\pi_N\alpha_p)} \tag{7}$$

$$b = \frac{(1-\pi_B\alpha_u)}{(1-\pi_N\alpha_u)} \tag{8}$$

and

$$c = [1-\pi[1-(1-\alpha_p)^{xm}(1-\alpha_u)^{(1-x)m}]] \tag{9}$$

then we can re-write this equation as

$$P_{xm,ynz_{BH}} = (a^y * b)^{z_{BH}n} * c \tag{10}$$

which is a convex function of z_{BH} since $(a^y * b)$ is always positive.

We can arrive at a similar expression for the NB population. Since z_{NB} is a linear function of z_{BH}, $P_{xm,ynz_{NB}}$ for the nonbathhouse patrons is also a convex function of z_{BH}. Finally, since the negative of a convex function is concave, and the sum of concave functions is concave, equation 4 is a concave function of z_{BH}.

2.2.4 Data

The primary data source for behavioral data was the Urban Men's Health Study (UMHS), which was conducted in 1997 and has been described elsewhere [15]. Briefly, the survey was conducted via telephone in four urban centers (San Francisco, Los Angeles, New York City, and Chicago) and focused on geographical areas that were believed to contain the majority of MSM in each city. Survey participants were MSM and were asked about the number of sex partners in the last year overall, and about the number of sex partners with whom they had engaged in receptive or insertive anal intercourse, as well as other behaviors. They were also asked about venues in which they met sex partners and about condom usage. For their four most recent sex partners, they were asked about behavior, condom usage, and the venue where they met the specific partner. All behavioral data were taken from this survey, as were data for the HIV prevalence in the BH and NB MSM populations.

Data from UMHS were used to define the subpopulations (BH and NB, HIV-infected and non-infected) and partnerships (main and non-main) described in the model section. Where variable values for the subpopulations were not different at $p < 0.05$, we used the overall mean for the variable in question. Because the data for the last four partners were specific by partner and less likely to be subject to recall error, we used the last four partners to determine the rates of condom usage, number of sex acts, and for BH MSM, the proportion of sex acts with non-main partners that occurred with a partner met at a bathhouse venue. The means for each variable were calculated using all non-missing values for the respective field in the data, rather than discarding observations that had missing values for some fields. Only anal sex acts were considered as risky sex acts for the model. See Table 1 for a summary of the survey data and significance across subpopulations.

Table 1: Summary of data from the 1997 Urban Men's Health Study. Any differences in values shown are significant at the 0.05 level.

Total Population (11,646 MSM)							
Bathhouse patrons				Non-bathhouse patrons			
29.4%				70.6%			
w/ main partner		w/out main partner		w/ main partner		w/ out main partner	
41%		59%		49%		51%	
HIV+	HIV-	HIV+	HIV-	HIV+	HIV-	HIV+	HIV-
25.4%	74.6%	25.4%	74.6%	16.5%	83.5%	16.5%	83.5%
Number of non-main sex acts in last 12 months							
128.98	58.59	141.22	70.83	16.64	11.18	28.87	33.99
Percentage condom usage with non-main partners							
76.0%				83.0%			

The distinction between those that have a main partner and those who do not was determined by the survey respondent citing he had a special commitment to or was in love with a particular person. The number of acts that are with main partners is 34.65, and the estimated condom usage with main partners is 58%. Neither of these are significantly different across subpopulations. Based on data from the last four partners, approximately 22% of the bathhouse patrons' non-main sex acts occur in the bathhouse. This implies that 78% of their non-main sex acts occur outside of the bathhouse which are distributed among bathhouse patrons and non-bathhouse patrons. The way in which we distribute the acts outside of the bathhouse is as follows: If there are 2 bathhouse patrons who have 10 acts each outside of the bathhouse and 10 non-bathhouse patrons who have 5 acts each outside of the bathhouse then there are 2*10 + 10*5 = 70 total acts to be distributed outside of the bathhouse. 2*10/70 * 100% is the percentage of the acts that any given member of either population has with bathhouse patrons while outside of the bathhouse when the bathhouse is open. The true value using data from Table 1 is 54.4%. Multiplying the true value with the percentage of BH acts that take place outside the bathhouse (78%), we obtain the

percentage of BH acts outside the bathhouse that take place with other BH patrons. This also gives us the number of BH acts with BH patrons outside the bathhouse, and by adding the number of acts within the bathhouse (22% of the BH total acts), we obtain the number of acts of BH with BH. By converting this to percentages, we estimate that 64.3% of the bathhouse patrons' non-main sex acts are with other bathhouse patrons. That is, the total acts between BH patrons includes the non-main acts that are *in* the bathhouse plus the remainder of their sex acts outside the bathhouse but divided among the overall population, some of which are with other bathhouse patrons and some of which are not. Given a percentage of the bathhouse patrons non-main sex acts that are with other bathhouse patrons, we can derive the percentage of sex acts that non-bathhouse patrons have with bathhouse patrons in order to account for the total sex acts of the population.

When bathhouses are closed, we assume that all non-main sex acts are drawn uniformly from the total number of sex acts in the population, which is based on the population sizes of BH and NB and the number of sex acts of each group. In this case, an estimated 60% of the (would be) bathhouse patrons' non-main sex acts are with other (would be) bathhouse patrons. That is, in the absence of bathhouses, there is a "natural" mixing of the entire population, whereas the presence of the bathhouse skews this mixing so that bathhouse patrons have relatively more sex acts with each other. Therefore z_{BH} is always smaller when bathhouses are closed than open.

Data for the probability of HIV transmission, protective effect of condoms, prevalence of P&S syphilis, and syphilis multiplier effect were taken from the literature. See Table 2 for the variables, their abbreviation in the model, baseline values, ranges, and sources.

[3]In our calculations, the same value for increased likelihood of HIV transmission was used regardless of whether either partner or both partners had syphilis. Hence a = b = c = 3.

Table 2: Epidemiologic variables used for base case and sensitivity analysis along with corresponding references.

Variable	Value	Range	References
Per-act HIV transmission risk (α)	0.01	0.005-0.03	[37, 131]
Protective effect of condoms ($1-\beta$)	0.9	0.90-0.95	[94]
Syphilis prevalence (P&S) (Π_s)	0.005	0.001-0.05	[20]
Syphilis multiplier effect (a, b, c)[3]	3	3-15	[105, 125, 132]

2.3 Results

Results are reported by computing the annual HIV attack rates. The HIV attack rate is defined as the number of new HIV cases in a year divided by the number of susceptible people (i.e., people currently without HIV) in the populations under consideration. It is useful in demonstrating relative degrees to which the epidemic is spreading under different scenarios.

Figure 1 shows the HIV attack rate given the number of acts as listed in the data section with 80% condom usage for the non-bathhouse population, 75% condom usage for the bathhouse population, an overall syphilis prevalence of 0.5%, and a syphilis multiplier of 3. The attack rate is shown as a function of z_{BH}, that is, a function of the percentage of the bathhouse patrons' non-main sex acts that are with other bathhouse patrons. In this figure, we make all calculations under the assumption that the total number of sex acts is the same under both scenarios of bathhouses closed and open. We will also examine the sensitivity of the results to this assumption in later analysis.

In all results that we have seen, the attack rate can be closely approximated by a linear relationship, as depicted in Figure 1. In general, the attack rate can be shown to be a concave function with respect to z_{BH}. In addition, for the given data, the attack rate is a strictly *decreasing* function of z_{BH}. This monotonicity is significant since it means that the HIV attack rate will always be higher when bathhouses are

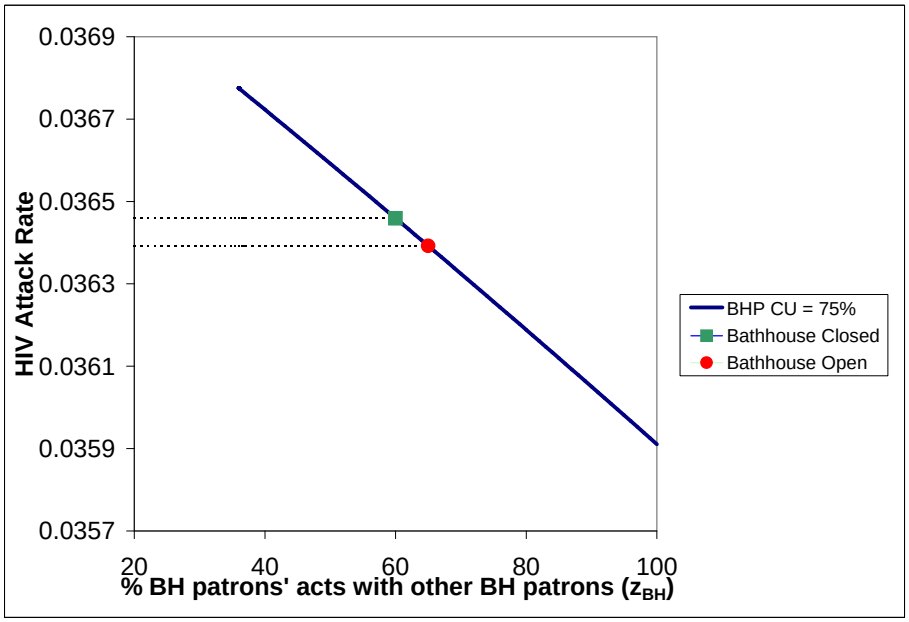


Figure 1: HIV attack rate as a function of z_{BH} when the Bathhouse Patron Condom Usage (BHP CU) is 75%, Non-bathhouse patron condom usage is 80%, overall syphilis prevalence is 0.5%, and the syphilis multiplier is 3.

closed. This conclusion is a consequence of the attack rate being decreasing and that z_{BH} will always be larger when the bathhouses are open than when the bathhouses are closed given that the number of acts remain the same. Consequently, it follows that under the stated assumptions, the attack rate will increase with bathhouse closure. In experiments using other values for behavioral parameters, we have found that the decreasing nature of the function is driven by the relatively larger number of non-main acts of the bathhouse patron population and the higher HIV prevalence of the bathhouse patrons. Single points on Figure 1 indicate the points corresponding to bathhouses being open or closed.

Figure 2 compares the effects of varying rates of condom usage and the number of acts to the effect of bathhouse closure on the attack rate. The top curve is calculated with the same parameter values as Figure 1 and is again shown as a function of z_{BH}.

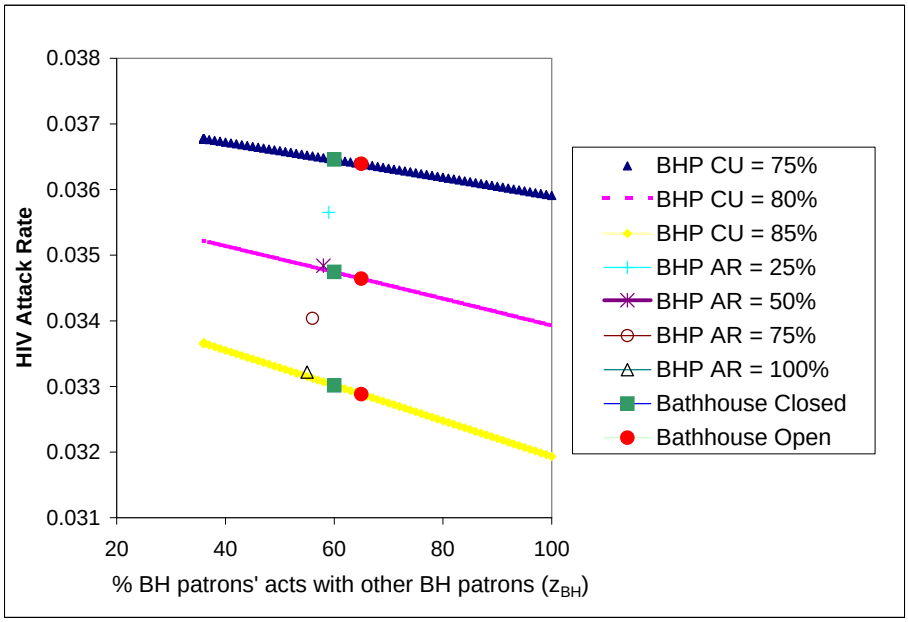


Figure 2: HIV attack rate for varying levels of Bathhouse Patron Condom Usage (BHP CU) and Bathhouse Patron Act Reduction due to bathhouse closure (BHP AR), with other parameters set as in Figure 1.

The two curves below it represent the attack rates when the condom usage of the bathhouse population increases by 5 percentage points (to 80%) and 10 percentage points (to 85%), respectively. Each of these three curves are calculated under the assumption that there is no change in the number of sex acts in the event of bathhouse closure.

The three individual scatter points, on the other hand, represent the attack rates calculated with the parameter values for Figure 1 with the bathhouse open but where the number of sex acts for the bathhouse population *does* decrease in the event of bathhouse closure. The analysis is done for several values of the potential percentage reduction in the bathhouse patrons' non-main partner sex acts that took place in the bathhouse. The reduction in acts is referred to as Bathhouse Patron Act Reduction (BHP AC) in the figure. The remainder of the acts (including non-main partner acts

outside of the bathhouse and main partner acts) are assumed to be unchanged. An entire curve is not traced in this case since the bathhouse is assumed closed making it unnecessary to plot it as a function of z_{BH}; we only consider the non-skewed "natural" mixing of the population which occurs when the bathhouse is closed corresponding to one value of z_{BH} and hence one point on the graph. Note that the "natural" percentage of bathhouse patrons' acts with other bathhouse patrons with bathhouses closed decreases as the number of sex acts is scaled down since they are having fewer acts.

Under the stated assumptions, this result suggests that increasing BH condom usage from 75% to 80% has a much larger effect on the attack rate than a bathhouse closure. Similarly, a change in the number of sex acts also has a much larger effect on attack rate than that of closure alone, that is, if the reduction in sex acts with the bathhouse closed were sufficiently large, then closing the bathhouse could be effective in reducing HIV transmission. This would assume that the bathhouse patrons were *not* able to replace their bathhouse sex acts by meeting sex partners in other venues. Not shown on the graphs, but also noteworthy, is that the effect of bathhouse closure is roughly equal to a 0.2% change in the condom usage of the bathhouse patrons.

Figure 3 demonstrates the effects that our assumptions regarding syphilis have on the results. We consider syphilis prevalence values of 0.5%, 5%, and 10% (the bottom three curves, respectively). Additionally, whereas all previous results assumed that the presence of syphilis caused the probability of infection per sex act to increase by a factor of three, Figure 3 considers the case where syphilis causes the probability of infection to increase by a factor of 15 (top curve). We consider this sensitivity since there is variation in the literature on the effect of syphilis.

It should be noted that while the graph in Figure 3 will vary with different values of the syphilis multiplier effect and the underlying syphilis prevalence, the overall conclusion will not change. In other words, we expect that the attack rate will still

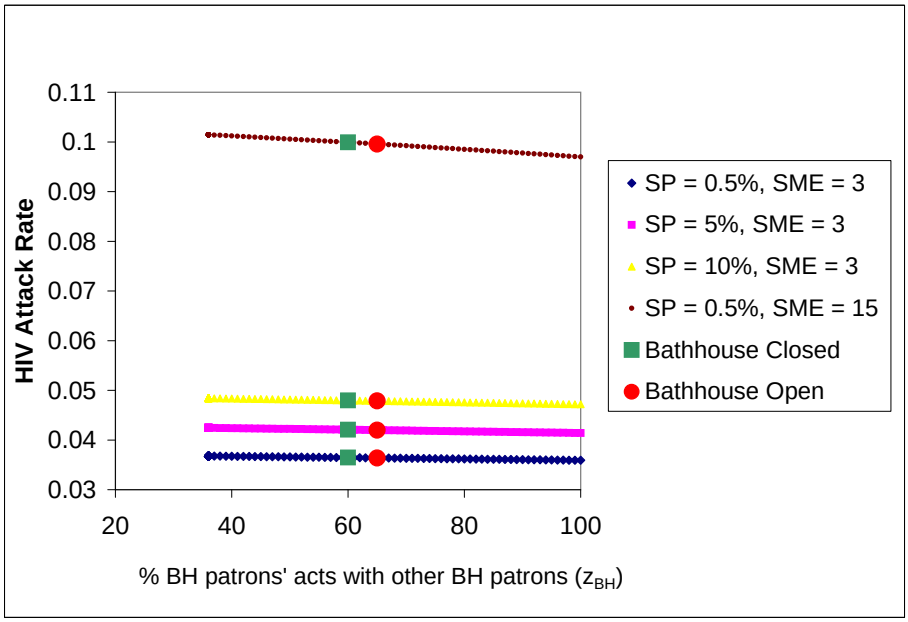


Figure 3: HIV attack rate for different Syphilis Prevalence (SP) and Syphilis Multiplier Effect (SME) values, where Bathhouse patron condom usage is 75% and Non-bathhouse patron condom usage is 80%.

be a decreasing function of z_{BH}, and as a result bathhouse closures are expected to create an increase in the attack rate unless the reduction in acts is large enough.

Table 3 shows sensitivity on attack rate for different parameter values. In this table, the base case corresponds to the base case parameter values as listed in Table 2. All other attack rates correspond to the base case parameter values except where explicitly stated (e.g. alpha = 0.005 corresponds to the base case parameter values except that alpha is changed from 0.010 to 0.005). The best case scenario corresponds to the parameter values that will produce the lowest attack rate, namely $\alpha = .005$, (1-β) = 0.95, $\pi_S = .001$, and syphilis multiplier = 3. The worst case scenario corresponds to the values that will produce the highest attack rate, namely $\alpha = .030$, (1-β) = 0.900, $\pi_S = .050$, and syphilis multiplier = 15. Notice that although the magnitude of the attack rate values change, it remains higher for the case of open versus closed

Table 3: Annual attack rate for different parameter values. The base case corresponds to the values reported in Table 2.

	BH closed	BH open	50% BHP AR after BH closure	80% BHP CU BH closed	80% BHP CU BH open
Best case scenario	0.0165	0.0165	0.0158	0.0156	0.0156
Alpha = .005	0.0187	0.0186	0.0178	0.0178	0.0177
95% condom effectiveness	0.0328	0.0328	0.0314	0.0310	0.0309
Base case	0.0365	0.0364	0.0348	0.0347	0.0346
Alpha = 0.030	0.0997	0.0995	0.0960	0.0958	0.0954
Worst case scenario	0.1685	0.1678	0.1614	0.1611	0.1603

bathhouses.

We have assumed that sexual activity remains the same during bathhouse closure. It is possible that sexual activity would decrease after bathhouse closure, though we have no data on this issue. In Figure 4 we show the resulting HIV attack rate from bathhouse closure as a function of percentage decrease in sexual activity. The base case for an open bathhouse from Figure 1 is shown as a horizontal dashed line for comparison. It is assumed that the base case of BH condom usage is 75%, non-bathhouse condom usage is 80%, overall syphilis prevalence is 0.5%, and the syphilis multiplier is 3. From this figure we see that if bathhouse closure leads to a reduction of bathhouse patrons' sexual activity within the bathhouse of at least 2%, HIV transmission would be reduced as compared to keeping the bathhouse open. Note that part of the reason for the decreased HIV attack rate with decreased sexual activity is that the relative percentage of sexual acts with condom usage increases.

2.4 Discussion

The results of the model suggest that, given the characteristics of the MSM population of the four survey cities (and the assumption that these characteristics do not change) bathhouse closure would result in an increase in the HIV attack rate. However, the

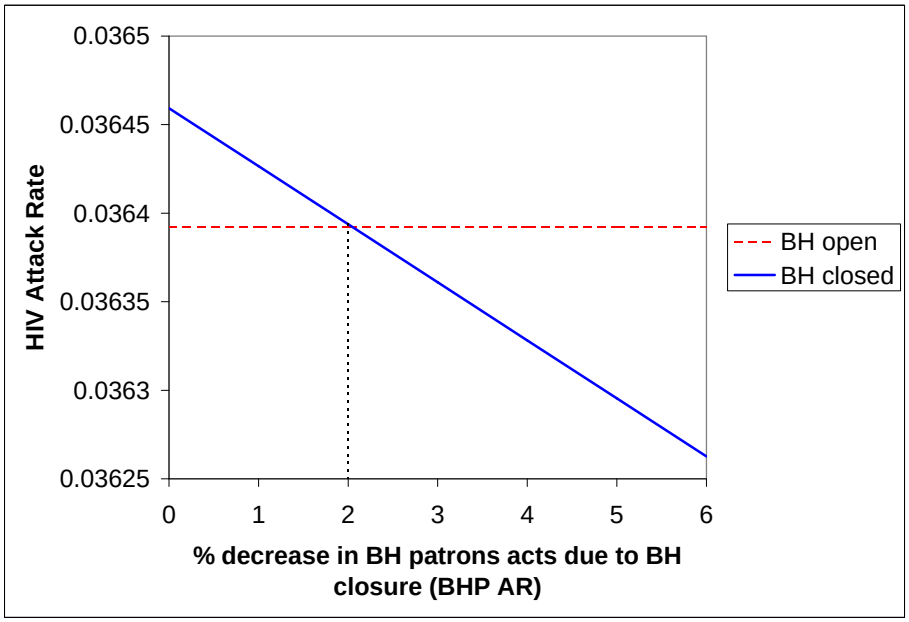


Figure 4: HIV attack rate for bathhouse closure as a function of the percentage decrease in BH sexual activity within the bathhouse due to bathhouse closure. BH condom usage is 75%, non-bathhouse condom usage is 80%, overall syphilis prevalence is 0.5%, and the syphilis multiplier is 3. The horizontal dashed line is for the base case of keeping the bathhouse open.

magnitude of the effect of the closure would be small compared to the effect of a change in condom usage or the number of sex acts. Therefore, these results need to be coupled with other studies. That is, if one can show that condom usage will not change, but the number of sex acts will decrease dramatically as a result of closure, then Figure 2 suggests that bathhouse closure would lead to a decrease in the attack rate. Conversely, if by leaving the bathhouses open one can show that condom usage will increase (by, for example, condom usage enforcement or promotion/education at the bathhouse venue) then the policy of keeping bathhouses open may result in a lower HIV attack rate. Indeed, leaving bathhouses open facilitates such intervention programs since it makes access to the MSM group with the highest prevalence and highest number of sex acts much more reachable. However, one study has found little

effect on risk behavior of bathhouse policies focused on regulating the amount of public versus private space in the facilities. The authors suggested that such policies may have moved risk behavior of BH patrons to different venues, rather than causing it to decline [139]. The impact of these and other potential interventions remains speculative, because there has been little evaluation [10, 137].

The results must be interpreted under consideration of the model's limitations. First and foremost, as a single stage model it only serves as a short term prediction as it ignores the effects of secondary infections. It does not differentiate between acute and chronic HIV infection. Evidence in the literature suggests that HIV infectivity differs over time. Wawer et al. [132] state that HIV infectivity is much higher during the acute stage (3 month period following initial infection) than during the chronic stage which is compounded by the fact that during the first three months the virus has likely not been detected. On the other hand, Rapatski et al. [100] find that most of the transmission occurs in the later stage of the disease. Either way, it could be important to have a model that can capture temporal differences in HIV infectivity.

If the aforementioned factors were included, it is possible that the graph in Figure 1 could become steeper. This steepness, or lack thereof, however, is what determines how the effect of bathhouse closure compares with a change in condom usage and number of sex acts. A change in steepness could possibly result in a different conclusion regarding the role of bathhouses in HIV transmission when changes in behaviors are considered.

An additional limitation of the model is that some of the data estimators for the parameters included in the model rely on the survey respondent's last four partners. The assumption that a respondent's condom usage behavior over the course of an entire year is accurately reflected by their condom usage with their last four partners, for example, may not always be valid. Moreover, the survey only includes city residents of four US cities and does not reach the surrounding areas, which may contain MSM

who visit city bathhouses. The UMHS survey, however, is the most comprehensive to date.

Another factor that we did not consider was club (recreational) drug use. The use of club drugs such as crystal methamphetamine has been shown to increase sexual risk-taking behavior and is associated with increased HIV/STD transmission [33, 34, 70]. Research has documented the prevalence of club drugs at specific gay events such as circuit parties. It is possible that many bathhouse patrons use substances prior to arriving at the venue. However, it may also be possible that bathhouse patrons acquire substances from other bathhouse patrons. Environment may influence club drug usage, and by extension risk behavior, suggesting that if bathhouses were closed, club drug usage preceding sex could diminish. We should mention that although we did not have accurate estimates of the prevalence of drug use in bathhouses, the model could be easily extended to determine the impact of drug use on transmission by using the same approach we used for including syphilis, and it is possible that drug use is already incorporated in the survey responses, since the condom usage was lower among bathhouse patrons.

The Centers for Disease Control and Prevention estimate that 25% of the US population who are HIV positive does not know it [26] Unfortunately, the data available from the UMHS survey did not allow us to model knowledge of infection. If the data were available, we could create new classes of those that know they are infected and those that do not, in the same way that syphilis was modeled.

Lastly, an implicit underlying assumption in our model is that the disproportionate mixing of the bathhouse and non-bathhouse populations will be eliminated as a result of bathhouse closure. That is, those who were bathhouse patrons prior to closing would no longer disproportionately have more sex with each other than the rest of the population after closure. This assumption seems quite intuitive in the short term. However, the bathhouse patrons would still remain high activity members of

the population in terms of partner numbers. As time passes, it is unpredictable the degree to which the high activity members would organize private venues similar to the bathhouse through, for instance, coordination via the Internet. Recent literature already suggests that the Internet is serving as a "cyber-bathhouse" in the sense that many high activity or high risk MSM can easily arrange private environments similar to bathhouses, with even less regulation than in bathhouses [12].

In future work, we plan to extend this model to a multi-stage setting. One of the benefits of this extension is that, it will eliminate some of the aforementioned limitations. A multi-stage model will also be useful in determining the importance of modeling subpopulations when considering HIV transmissions, since the impact of transmissions is compounded over time.

2.5 Appendix

In order to incorporate the effects of syphilis, we divided each subpopulation above further by whether the individuals are positive or negative for syphilis. The probabilities of HIV infection will be different for each of the subpopulations. When populating our earlier HIV transmission model, we had data that indicated the number of sex acts by subpopulation. However, when we expand the model to include syphilis status, our data does not specify which partners have syphilis, so we take an expectation over the syphilis prevalence in the corresponding population to determine the number of acts across each of the subpopulations.

For an individual from BH who is positive for syphilis, his probability of not acquiring HIV with the behavior parameters x, m, y, n, and z_{BH} is given by:

$$\begin{aligned}\overline{PP_{xm,ynz_{BH}}} &= [1-\pi_T[\pi_S[1-(1-c\alpha_p)^{xm}(1-c\alpha_u)^{(1-x)m}] \\ &\underbrace{+(1-\pi_S)[1-(1-a\alpha_p)^{xm}(1-a\alpha_u)^{(1-x)m}]]]}_{A+B} \\ &\times \underbrace{(1-\pi_{BH}c\alpha_p)^{yz_{BH}\pi_S n}}_{C}\underbrace{(1-\pi_{BH}c\alpha_u)^{(1-y)z_{BH}\pi_S n}}_{D} \\ &\times \underbrace{(1-\pi_{NB}c\alpha_p)^{y(1-z_{BH})\pi_S n}}_{E}\underbrace{(1-\pi_{NB}c\alpha_u)^{(1-y)(1-z_{BH})\pi_S n}}_{F} \\ &\times \underbrace{(1-\pi_{BH}a\alpha_p)^{yz_{BH}(1-\pi_S)n}}_{G}\underbrace{(1-\pi_{BH}a\alpha_u)^{(1-y)z_{BH}(1-\pi_S)n}}_{H} \\ &\times \underbrace{(1-\pi_{NB}a\alpha_p)^{y(1-z_{BH})(1-\pi_S)n}}_{H}\underbrace{(1-\pi_{NB}a\alpha_u)^{(1-y)(1-z_{BH})(1-\pi_S)n}}_{I},\end{aligned} \tag{11}$$

where the indices are suppressed again for clarity. For an individual from BH who is syphilis negative, his probability of not acquiring HIV with the behavior parameters x, m, y, n, and z_{BH} is given by:

$$\begin{aligned}\overline{PN_{xm,ynz_{BH}}} &= [1-\pi_T[\pi_S[1-(1-b\alpha_p)^{xm}(1-b\alpha_u)^{(1-x)m}] \\ &\underbrace{+(1-\pi_S)[1-(1-\alpha_p)^{xm}(1-\alpha_u)^{(1-x)m}]]]}_{A+B} \\ &\times \underbrace{(1-\pi_{BH}b\alpha_p)^{yz_{BH}\pi_S n}}_{C}\underbrace{(1-\pi_{BH}b\alpha_u)^{(1-y)z_{BH}\pi_S n}}_{D} \\ &\times \underbrace{(1-\pi_{NB}b\alpha_p)^{y(1-z_{BH})\pi_S n}}_{E}\underbrace{(1-\pi_{NB}b\alpha_u)^{(1-y)(1-z_{BH})\pi_S n}}_{F} \\ &\times \underbrace{(1-\pi_{BH}\alpha_p)^{yz_{BH}(1-\pi_S)n}}_{G}\underbrace{(1-\pi_{BH}\alpha_u)^{(1-y)z_{BH}(1-\pi_S)n}}_{H} \\ &\times \underbrace{(1-\pi_{NB}\alpha_p)^{y(1-z_{BH})(1-\pi_S)n}}_{I}\underbrace{(1-\pi_{NB}\alpha_u)^{(1-y)(1-z_{BH})(1-\pi_S)n}}_{J}.\end{aligned} \tag{12}$$

As before, each labeled term in the equations above represents the probability of not acquiring HIV given a set of partners and acts. The partners and acts corresponding to each term are described in words below:

- A + B: m sex acts with main partner who is syphilis positive x percent of which are protected (weighted with probability π_S), and m sex acts with main

partner who is syphilis negative x percent of which are protected (weighted with probability $(1 - \pi_S)$)

- C: $yz_{BH}\pi_S n$ protected sex acts with non-main partners who are bathhouse patrons and syphilis positive
- D: $(1 - y)z_{BH}\pi_S n$ unprotected sex acts with non-main partners who are bathhouse patrons and syphilis positive
- E: $y(1 - z_{BH})\pi_S n$ protected sex acts with non-main partners who are not bathhouse patrons and are syphilis positive
- F: $(1 - y)(1 - z_{BH})\pi_S n$ unprotected sex acts with non-main partners who are not bathhouse patrons and are syphilis positive
- G: $yz_{BH}(1-\pi_S)n$ protected sex acts with non-main partners who are bathhouse patrons and syphilis negative
- H: $(1 - y)z_{BH}(1 - \pi_S)n$ unprotected sex acts with non-main partners who are bathhouse patrons and syphilis negative
- I: $y(1 - z_{BH})(1 - \pi_S)n$ protected sex acts with non-main partners who are not bathhouse patrons and are syphilis negative
- J: $(1 - y)(1 - z_{BH})(1 - \pi_S)n$ unprotected sex acts with non-main partners who are not bathhouse patrons and are syphilis negative

We can combine Equations 12 and 13 to represent the total number of new HIV cases for all BH subpopulations by:

$$\sum_{i=1}^{M_{BH}} [1 - (1 - \pi_S)\overline{PN_{x_i m_i, y_i n_i z_{BH}}} - (\pi_S)\overline{PP_{x_i m_i, y_i n_i z_{BH}}})] * N_i, \tag{13}$$

where there are 8 M_{BH} BH subpopulations indexed by i, and N_i members in subpopulation i. With an identical development we can include the NB subpopulations in Equation 13 and it will have the same form as Equation 3, the only difference being each population is split into two groups by syphilis status, and each corresponding probability of infection is more complicated. The total number of new HIV cases for the entire population is then represented by:

$$\begin{aligned} & \sum_{i=1}^{M_{BH}} [1-(1-\pi_S)\overline{PN_{x_i m_i, y_i n_i z_{BH}}} - (\pi_S)\overline{PP_{x_i m_i, y_i n_i z_{BH}}})] * N_i \\ + & \sum_{j=1}^{M_{NB}} [1-(1-\pi_S)\overline{PN_{x_j m_j, y_j n_j z_{NB}}} - (\pi_S)\overline{PP_{x_j m_j, y_j n_j z_{NB}}})] * N_j, \end{aligned} \tag{14}$$

where there are 8 M_{NB} NB subpopulations indexed by j, and N_j members in subpopulation j.

CHAPTER III

TIMING OF TESTING AND TREATMENT FOR ASYMPTOMATIC DISEASES

3.1 Introduction[*]

Recommended testing protocols to determine the presence of a disease have two major components. First they must specify the population to be tested (i.e., universal testing or targeting). Second, they must specify the timing and frequency of testing, that is, one-time testing, routine testing, or intermittent testing. If it is the latter, they must also specify the time interval between tests and whether the interval should be constant or varying. Although current guidance typically (though not always) specifies the population to be tested, it is often silent as to the frequency and the timing of testing. Timing can be important because it can impact whether testing is cost-effective in a population.

Several criteria are considered when determining the optimal testing protocol. These include the prevalence of the disease, accuracy of the test, whether awareness of disease status reduces costs, the associated costs of the disease and test, and how the disease progresses. As disease prevalence decreases, more persons must be tested to identify one case and thus targeting tests to at-risk persons may be more efficient. However, if the cost of the condition is high relative to the cost of the test, universal testing may be more cost-effective. Awareness of disease status will influence costs if an effective intervention to treat the disease is available or if changes in personal behavior can reduce future morbidity, mortality, or probability of transmission to others. For example, a person who learns he has Hepatitis C (HCV) may reduce

[*] ***This chapter is a collaborative effort with H. Eser Kirkizlar.***

his alcohol intake, which would reduce the probability of future liver damage. In addition, if he were an injection drug user (IDU), he could participate in a needle exchange program, which would reduce the likelihood of transmission to others.

Our motivation is to study testing and treatment protocols from a societal perspective. We specifically focus on the *time* to test an individual one or more times for a disease. We use a Markov Decision Processes (MDP) approach to model the disease progression and testing decision, where the reward function (or utility) is based on testing and treatment costs, quality adjusted life years (QALY) defined in different stages of disease, and the cost of infecting other individuals. We also allow the awareness of the presence of the disease to affect behavior, which can change the transition probabilities and secondary infection costs after testing. We analyze the model for structural results and find sufficient conditions to establish when testing (and treating) the disease is cost-effective, and provide insights for healthcare practice. We demonstrate the results for HCV using medical data, and compare our findings to current HCV screening recommendations. We also solve for the threshold incidence for a population group beyond which testing is cost-effective.

The Chapter is organized as follows. A brief literature review is given in the next section. In Section 3.3, we describe the Markov model for a general disease setting, and present structural results for the MDP. Then in Section 3.4, we perform a numerical study for HCV using data obtained from patient studies and health databases. Finally, we discuss the implications of our results and several directions for future research in the last section.

3.2 Literature

We briefly summarize past work in HCV screening and treating as it relates to our work; this is in no way a complete review of the literature. Several researchers have

modeled *treatment* of HCV and have in general found it to be cost-effective for populations who are known to have the disease [62, 107]. The literature on *testing* for HCV is much less clear. For example, Castelnuovo [14] find based on a decision analytic model that testing is cost-effective in injection drug users; Gordon [54] also reports that testing can be cost-effective for high-risk groups. However Plunkett [96] finds screening of asymptomatic pregnant women for HCV infection is not cost-effective. In addition, Singer [114] found that routine HCV testing was not cost-effective in asymptomatic, average-risk adults. Pereira [91] used a Markov simulation model to show that although testing blood donors for HCV was cost-saving for the health care system, screening of post-transfusion patients was not; this is likely because most HCV-infections due to transfusions have already been identified. Note that the literature on cost-effectiveness for screening of HCV does not explicitly consider the age as part of the decision process, which is a factor we will consider in this Chapter.

Testing recommendations from public health agencies can also differ. For example, the US Public Health Service (USPHS) and Infectious Diseases Society of America (IDSA) recommend testing of all HIV positive persons for HCV [16]. The US Preventive Services Task Force (USPSTF) recommends against routine testing for HCV in the general population and makes no recommendation for populations with high risk for infection [128]. The Centers for Disease Control and Prevention (CDC), on the other hand, does recommend routine testing for population groups with a high risk for acquiring HCV such as drug users or commercial sex workers [19]. However, none of these agencies make recommendations about the frequency or timing of testing. Explicit consideration of timing could change the recommendations.

Our work in HCV testing and treating differs from previous work in that we explicitly model the timing decision of the testing and explicitly consider behavior change as a result of knowledge of infection from testing. The former is possible because of the Markov decision process model that we develop and study. In addition,

we compare different strategies for the number of tests.

There have been numerous papers in the medical literature using Monte Carlo simulation of Markov models to study disease progression or screening for other diseases (e.g., [53, 103, 116]) In most of these papers, the progression of the disease is modeled as Markovian, and cost-effectiveness of a specified testing policy is calculated using simulation across a population group with particular risk characteristics. The simulation medical papers often address *whether* to screen the risk group in question, although fewer address *when* the screening should be performed, which can affect cost-effectiveness. The definition of the risk group may include an age range (which is an implicit way of capturing timing), but this may not be sufficient to capture the progression of the disease and behavior over time. Example papers that examine cost-effectiveness of repeated screenings where a limited number of testing policies are specified *a priori* include [11, 74, 130, 84]. Diehl [38] studies a large number of screening policies for disease where over 1000 testing policies for breast cancer are evaluated.

Analytical approaches may complement the simulations and help to provide additional insight on the characteristics of the testing policies, and there are also relevant papers that have used analytics for screening decisions. Some of these papers also relate to scheduling examinations or replacements for machines in a production system, although papers in this area may have different assumptions than those that focus on medical decisions. Early operations research approaches to this problem include Smallwood [115] and Kirchklein [63], both of which have perfect testing information and stationary parameters; the first is an early example of POMDPs with medical implications. Lee [65] finds that the optimal screening is equally spaced if tests have perfect reliability, although other papers have found that spacing may not be equal if parameters vary over time. Examples of papers that primarily focus on inspection of production systems include [135, 75, 76, 82, 144]. One key assumption in these papers

is that the testing procedure does not impact the performance of the machine unless a corrective action is taken. An important aspect of our problem is the behavioral change brought on by awareness of the disease gained through testing, i.e., we can have "partial" treatment at no cost.

One of the more relevant papers from the inspection literature is Ozekici [83], which uses a simplified version of a POMDP that the authors then transform into an MDP with complete information. The authors include false positives and negatives for the test but allow no death from causes other than disease and no recurrence of disease (thus no testing after disease has been treated). Houshyar [58] allows death from causes other than disease and formulates a screening problem where the disease progression can be modeled with a discrete-time Markov chain. He gives guidelines for calculating the costs as well as the transition probabilities and applies the model to a disease but does not study structural results of the problem.

Zelen [147] focuses on medical screening timing along with follow-on papers [67] and [66]. These papers focus on a weighted utility function that is linear in the probability of finding a case and being incident between tests; they focus on testing when probabilities are stationary over time, or not age-dependent, while in our case risk behaviors or disease progression may depend on age. Other papers that study screening problems but have stationary parameters include Monahan [77] and Parmogiani [88]. An interesting approach is taken in Kaplan [61], where the authors use an analytical model similar to inventory modeling to show that the interval between screenings depends on the prevalence in the population, and apply the model to HIV.

Parmigiani [86] uses a non-Markovian stochastic model to solve for test timing and takes into account the effect of age on disease progression. Parmigiani [87] finds an exact solution with fallible tests and two disease states. As in our case, the latter paper only gains information about disease presence by use of a test, although in the second paper tests take a random amount of time and do not alter the state of the

system. We also use a higher number of disease states and allow recurrence of the disease and death from causes other than disease. Special cases with two (or even three) states have also been considered in other papers including some above but also others, e.g., Eddy [43] and Grosfeldnir [55], although under different assumptions than the ones we use. Specifically, Grosfeldnir [55] consider a production system and the only states are "good" or "bad", and the actions available are to replace to machine or not. Others have considered the special case of one or two tests over the time horizon, such as in [86, 89, 77]. In our case we are able find explicit conditions for a dynamic testing and/or treating strategy to be beneficial. That is, we do not restrict the number of tests beforehand.

The subject of using MDPs and POMDPs to model medical screening problems is discussed in Hauskrecht [56], where POMDPs may be used to capture informational aspects. As the authors state, even the definition of the POMDP may be difficult for a disease, and the number of transitions and probabilities to define can become "practically impossible". They use a hybrid POMDP with an MDP and use approximations to solve it with data from ischemic heart disease, but they point out that other structural refinements are possible to make the models reasonable to define and solve.

To summarize, our research contributes to the literature on medical screening by developing and analyzing a special case of a POMDP model for the timing of a screening test for a disease that may have secondary infections. Key aspects include a transition probability matrix that changes with the actions and states to model behavioral changes. We find some analytical results on the timing of disease screening and interpret these for policy implications.

3.3 Model and Structural Results

3.3.1 The Mathematical Model

We formulate the progression of the disease as well as the testing and treatment decisions as a finite-horizon, discrete time discounted MDP model. We allow the health states to be partially observable, with the test providing the only update in information regarding the current health state. In the description of the model we will focus on the maximization of utility. The utility is based on QALYs for the different states of health of a person as well as the cost of testing, treating, and corresponding complications. Note that costs may be converted to QALYs using a cost-effectiveness threshold, e.g., \$50,000 per QALY [29]. We assume that false positive tests can result but a second test is available with higher accuracy, which reflects common practice. In our base model we will analyze the model with a test that has high sensitivity but we will describe how the model and results extend to tests with false negatives.

Let $S \subset (\{(h,i)|h \in \{0,\ldots,H\}, i \in \{0,1\}\}$ be the state space of our stochastic process. We use $h \in \{0,\ldots,H\}$ to denote the health state of the individual (note that in the remainder of the Chapter we will use "health state" to refer to the first component of the state space). 0 is the state individual is healthy and H is the state individual is dead, with the other states representing disease states that may have different utilities, transition probabilities, or other characteristics. An individual's belief about whether or not he has the health condition is denoted by i; 0 is used for the case that he thinks he is healthy and 1 for the case that he knows that he has the disease. The probability of having the disease (e.g., determined by the prevalence in the population to which the individual belongs) is only updated by the use of a screening test. This simpler representation of a partially-observable MDP is used to capture the issue of awareness of disease, and it reflects testing as the primary way to determine presence of a disease.

$A = \{NT, T_1, T_2, T_1T_2\}$ denotes the action set and A_s denotes the feasible actions

for every $s \in S$, where $A_s \subseteq A$. The action NT denotes "do not test" option, T_1 denotes "test but do not treat if the individual is sick" option, T_2 denotes "treat the individual if he is already known to be sick" option, and T_1T_2 is used for "test and treat if the individual is sick" option. In our model, T_1 is treated differently than T_1T_2, since in some cases the treatment may not exist, is expensive, is not very effective, or may have side effects (in which case the patient might choose not to be treated). Testing without treatment (T_1) is a reasonable option (i.e., an improvement over not testing) only if the awareness of the disease implies benefits other than treatment and cure. These may include changes in utility, changes in probability of infecting others, or changes in progression rates (see below). This also allows for partial treatments. Furthermore, the action T_2 allows the patient to delay his treatment. This can occur when the patient cannot be treated immediately due to an existing health condition or when a patient who did not want the treatment just after the test result changes his mind later. We assume that the test results are immediately available and the length of the treatment is negligible compared to the decision epoch. Hence, we allow transitions to occur at the beginning of the decision epoch when a test and/or treatment action is taken.

$T = \{0, \ldots, N\}$ is the set of (finite) decision epochs. Decision epochs can be years, months or even days depending on the disease. In addition, the number of decision epochs might change with respect to the problem. For example, when modeling certain diseases, decisions are made every year (e.g., annual exams) and the number of decision epochs is chosen to be an age after which an individual's utility is negligible while in other cases shorter time periods are desirable.

$r_t(s, a)$ is defined to be the utility of taking decision a in state s at decision epoch t. It includes the QALYs of different health states and the cost associated with the likelihood of infections to other people (if the disease under consideration is infectious). The cost associated with infections would be primarily for secondary

infections, since for many diseases this captures the majority of the reduction of QALYs. It is also a standard way of including QALYs for diseases like HCV [114] and HIV [84]. Finally, we assume that the costs of a test are only immediate costs and that the test does not cause future harm to the individual. $p(s'|s,a)$ denotes the probability of going to state s' from state s, when decision a is taken. The non-stationary probability transitions are considered in the numerical results section.

Let $\pi = \{\pi_0, \ldots, \pi_N\} \in \Pi$ be a policy where Π denotes the set of all policies and π_t is the action at time $t \in \{0, \ldots, N\}$. We assume $\pi_N = NT$ for all $\pi \in \Pi$. We are interested in finding a policy that maximizes the total discounted expected utility over the horizon. The objective function is thus

$$\max_{\pi \in \Pi} \mathbb{E}_\pi \Big\{ \sum_{t=1}^{N} \lambda^t r_t(s, \pi_t) \Big\},$$

where $0 \leq \lambda \leq 1$ is the discount factor. Let $u_t^\pi(s) = r_t(s, \pi_t) + \sum_{j \in S} \lambda p(j|s, \pi_t) u_{t+1}^\pi(j)$. In other words, $u_t^\pi(s)$ is the total utility from decision epoch t onwards under policy π if the system is in state s at that time. Let $b_t^\pi(s)$ denote the probability that the system is in state s at time t when policy π is used. Then, $\mathbb{E}[u_t^\pi]$ can be calculated as $\sum_s b_t^\pi(s) u_t^\pi(s)$. Furthermore, it should be noted that $\mathbb{E}[u_0^\pi] = \mathbb{E}_\pi \Big\{ \sum_{t=1}^{N} \lambda^t r_t(s, \pi_t) \Big\}$.

In the remainder of this Chapter we study a special class of diseases with two disease stages (in addition to healthy and death states). By considering this simpler class of diseases, we are able to find closed-form expressions in terms of the problem parameters that determine whether or not it is beneficial to test and/or treat the person at a specific time. This type of Markov model might be appropriate for a disease such as Chlamydia, which has earlier asymptomatic stage and a later symptomatic stage in some patients.

The state space in this case is $S = \{(0,0), (1,0), (1,1), (2,1), (3,0)\}$. State (0,0) denotes the healthy state, state (3,0) denotes the death state, and the other states represent the different stages of the disease as well as the awareness of the patient as described previously. We assume that once the disease reaches stage 2, the individual

is aware of it and the condition causes some irreversible damage (for example when glaucoma causes blindness). The feasible actions are as follows:

$$A_s = \begin{cases} \{NT, T_1, T_1T_2\} & \text{if } s \in \{(0,0),(1,0)\}, \\ \{NT, T_2\} & \text{if } s = (1,1), \\ \{NT\} & \text{if } s \in \{(2,1),(3,0)\}. \end{cases}$$

We assume that $r_t(s,a) = r(s,a)$ for all $t \in \{0,\ldots,N\}$. We let $r(s,NT) = R_s$, and we subtract the cost of testing and/or treatment from this value whenever any action other than NT is chosen. In other words, R_s denotes the utility of being in state s when there are no testing and treatment costs. We let $R_{(3,0)} = 0$, hence $u_t^\pi((3,0)) = 0$, for all $t \in \{0,\ldots,N\}$ and $\pi \in \Pi$. Let c_0 be the cost of testing, c_1 be the cost of treating the patient (in disease stage 1), v_1 be the success probability of the treatment (in disease stage 1). The elements of the probability transition matrix under action NT are as follows:

$$P_{s,s'} = \begin{cases} p_{ij} & \text{if } s = (i,0), s' \in \{(j,0),(j,1)\}, \\ q_{ij} & \text{if } s = (i,1), s' = (j,1). \end{cases}$$

We assume that there is no direct transition to disease stage 2 from the healthy state.

The following lemma is used throughout the Chapter and its proof is immediate from backwards induction.

Lemma 1. *Let π and π' be two policies that might differ after time t and agree otherwise. If $\mathbb{E}[u_t^\pi] \geq \mathbb{E}[u_t^{\pi'}]$, then $\mathbb{E}[u_0^\pi] \geq \mathbb{E}[u_0^{\pi'}]$. In other words, the policy that is better at time t is a better policy.*

Let us define $f_i(x,y)$ for $i \in \{1,\ldots,6\}$ and $x_1, x_2 \in \{0,\ldots,N\}$ as follows:

$$
\begin{aligned}
f_1(x_1,x_2) &= -c_0\Big(\frac{p_{00}^{x_2-x_1}(p_{11}-p_{00}-p_{01})+p_{11}^{x_2-x_1}p_{01}}{p_{01}(p_{11}^{x_2-x_1}-p_{00}^{x_2-x_1})}\Big) \qquad (15)\\
&\quad +(R_{11}-R_{10})\Big(\frac{1-(\lambda p_{11})^{N-x_2+1}}{1-\lambda p_{11}}\Big),\\
f_2(x_1,x_2) &= -c_1+v_1R_{00}-v_1R_{11}+v_1\Big(R_{00}-R_{10}\frac{p_{01}}{p_{11}-p_{00}}\Big)\Big(\frac{\lambda p_{00}-(\lambda p_{00})^{N-x_2+1}}{1-\lambda p_{00}}\Big)\\
&\quad -v_1\Big(R_{11}-R_{10}\frac{p_{01}}{p_{11}-p_{00}}+R_{21}\frac{p_{12}}{q_{22}-p_{11}}\Big)\Big(\frac{\lambda p_{11}-(\lambda p_{11})^{N-x_2+1}}{1-\lambda p_{11}}\Big),\\
f_3(x_1,x_2) &= f_1(x_1,x_2)+f_2(x_1,x_2),\\
f_4(x_1,x_2) &= -c_0\Big(\frac{p_{00}^{x_2-x_1}(p_{11}-p_{00}-p_{01})+p_{11}^{x_2-x_1}p_{01}}{p_{01}(p_{11}^{x_2-x_1}-p_{00}^{x_2-x_1})}\Big)\\
&\quad +R_{21}(\frac{p_{12}}{q_{22}-p_{11}}-\frac{q_{12}}{q_{22}-q_{11}})\Big(\frac{\lambda q_{22}-(\lambda q_{22})^{N-x_2+1}}{1-\lambda q_{22}}\Big)\\
&\quad +\Big(R_{10}-R_{21}(\frac{p_{12}}{q_{22}-p_{11}}-\frac{q_{12}}{q_{22}-q_{11}})\Big)\\
&\quad \Big(\frac{\lambda q_{11}-(\lambda q_{11})^{N-x_2+1}}{1-\lambda q_{11}}-\frac{\lambda p_{11}-(\lambda p_{11})^{N-x_2+1}}{1-\lambda p_{11}}\Big),\\
f_5(x_1,x_2) &= -c_1+v_1R_{00}-v_1R_{11}+v_1\Big(R_{00}-R_{10}\frac{p_{01}}{p_{11}-p_{00}}\Big)\Big(\frac{\lambda p_{00}-(\lambda p_{00})^{N-x_2+1}}{1-\lambda p_{00}}\Big)\\
&\quad +v_1R_{10}\Big((\frac{p_{01}}{p_{11}-p_{00}})\Big(\frac{\lambda p_{11}-(\lambda p_{11})^{N-x_2+1}}{1-\lambda p_{11}}\Big)-\frac{1-(\lambda q_{11})^{N-x_2+1}}{1-\lambda q_{11}}\Big)\\
&\quad -v_1R_{21}(\frac{q_{12}}{q_{22}-q_{11}})\Big(\frac{\lambda q_{22}-(\lambda q_{22})^{N-x_2+1}}{1-\lambda q_{22}}-\frac{\lambda q_{11}-(\lambda q_{11})^{N-x_2+1}}{1-\lambda q_{11}}\Big),\\
f_6(x_1,x_2) &= f_4(x_1,x_2)+f_5(x_1,x_2).
\end{aligned}
$$

The following theorems provide conditions for cost-effective testing and treatment strategies when false negatives do not occur.

Theorem 1. *If the individual was tested and/or treated at t_i (where $t_0 = 0$) for the last time and awareness of the disease only affects the immediate utilities, then the following strategy is cost-effective:*

1. *If he is healthy, or treated successfully at time t_i, then it is cost-effective to test him (and treat him if he is sick) at time t_{i+1} if $f_1(t_i,t_{i+1}) \geq 0$ ($f_3(t_i,t_{i+1}) \geq 0$).*

2. *If he is tested to be in disease stage 1 and the treatment is not successful (or*

he is not treated) at time t_i, then it is cost-effective to treat him at time t_{i+1} if $f_2(t_i, t_{i+1}) \geq 0$.

The proof is in the Appendix to this Chapter.

Theorem 2. *If the individual was tested and/or treated at t_i (where $t_0 = 0$) for the last time and awareness of the disease only affects the disease's progression, then the following strategy is cost-effective:*

1. *If he is healthy, or treated successfully at time t_i, then it is cost-effective to test him (and treat him if he is sick) at time t_{i+1} if $f_4(t_i, t_{i+1}) \geq 0$ ($f_6(t_i, t_{i+1}) \geq 0$).*

2. *If he is tested to be in disease stage 1 and the treatment is not successful (he is not treated) at time t_i, then it is cost-effective to treat him at time t_{i+1} if $f_5(t_i, t_{i+1}) \geq 0$.*

The proof is similar to the proof of Theorem 1, hence it is omitted. Theorems 1 and 2 show that if there exists a time t_1 such that it is cost-effective to test(and/or treat) the individual, assuming that there will be no other interventions later on, then this action should be taken at time t_1. Furthermore, if there exists a time t_2 after time t_1 such that the intervention at time t_2 is better than doing nothing, then this intervention should be done as well. Similarly, at every t_{i+1}, the decision is based on the health state at the time of previous intervention and the remaining time until N.

Note that the previous theorems assumed that awareness of the disease either affects the immediate rewards or the progression of the disease, and this is correct for most of the diseases. We could simultaneously incorporate both changes into our model, but for the sake of simplicity, these expressions were not provided here. The next section considers a disease (HCV) where both progression and the immediate rewards depend on the awareness of the disease.

Tests for diseases may result in false negatives or false positives. In this study, we assume that a false positive is detected by a followup-up test with greater accuracy,

which is the case in practice. False negatives are more complicated since the individual is led t believe that he is healthy, when in fact he has the disease.

3.4 Application of the Model to HCV

Hepatitis C virus (HCV) is a blood-borne virus that typically leads to a slow progression of chronic liver disease. The majority of people are asymptomatic for decades before the negative health effects first become noticeable. In the US, an estimated 3.9 million people are currently infected [25], making it the most common chronic blood-borne infection in the country. HCV can cause liver cell damage, cancer, and cirrhosis, and is the leading cause for liver transplants. Most people are unaware they have the disease until they develop end stage liver disease but may spread the disease to others even when they are asymptomatic. There is currently no vaccine for HCV, although treatments exist that can cure with a 54% rate if applied early enough. The high cost of treating the advanced disease, combined with the infectivity and long asymptomatic period make HCV a candidate for screening programs.

Recall that Theorems 1 and 2 make the assumption that awareness of HCV infection effects only the disease progression *or* the rewards. These restrictions are only imposed for the sake of brevity. The theorems can be joined into one where the progression and rewards can change simultaneously. However, since the expressions become much more cumbersome, we omit them. For the computational results that follow we use the results from Section 3.3.1 allowing disease progression and rewards to change simultaneously. We populate the model with parameters specific to HCV. The state space in this case is defined as:

- (0,0) = Healthy
- (1,0) = Infected (unaware)
- (1,1) = Infected (aware)

- (2,1) = Decompensated cirrhosis including associated complications (hepatocellular carcinoma, liver transplant)
- (3,0) = Death

This is a simplified natural history of the disease obtained from [7, 48, 57, 114]. We develop a Monte Carlo simulation of the disease progression in MATLAB. Individuals enter the system uninfected at age 15. We then implement the following testing strategy: from ages 15-35, in each period, we check whether taking any action is cost-effective. Whenever it is cost-effective to take an action, the respective action is taken at that time (i.e. test, test and treat, treat, treat again). Ages 15-35 constitute the decision epochs. After age 35, we add a final reward which represents the continued disease progression (or lack thereof) until age 80 (or death) based on modified model parameters that reflect a higher natural death rate and a lower incidence.

The model is based on QALYs and costs, which are associated with each health state as summarized in Table 4. The utility in each period is calculated by converting the QALY associated with that period to a cost where one QALY corresponds to \$50,000, and then subtracting the dollar costs of the screening test, treatment, and annual HCV related health costs. The values for the transition probabilities were taken from the literature. Some of the transitions depend greatly on the population considered. For example, injection drug users (IDUs) are at higher risk for initially acquiring HCV, while those who drink more than 50 grams of alcohol per day have faster progression rates to decompensated cirrhosis. See Table 4 for the details on the probabilities and rates. Whenever simplifications were made due to the nature of the model contained herein, we used conservative estimates so that the results would not be biased towards increased testing.

Due to limited HCV incidence data in the literature, it is difficult to assess the age period during an individual's life when they are susceptible to HCV. Therefore, since

the HCV risk groups are very similar to that of HIV, we examined HIV incidence data to extrapolate the ages at which individuals are at most risk for the overall population. According to the Centers for Disease Control and Prevention (2005a), children under age 13 constitute less than 1% of persons living with HIV through the year 2000 (mostly acquired through mother-to-infant transmission, which is not studied in this model), and after age 55 HIV infection rates drop dramatically. According to the Department of Health and Human Services (2005), percent drug use rises dramatically from age range 12-13 to age range 14-15 and remains high until age range 50-54. Consequently, we assume individuals are susceptible to HCV infection during the entire time horizon of ages 15-35, making it a conservative assumption that does not bias towards testing. Table 4 also indicates the probabilities of infection for each risk group.

The cost of infecting others with HCV is calculated using the given data as the *additional* total discounted lifetime cost of an individual as a result of acquiring HCV. We do this by subtracting the total discounted cost of an individual with probability of acquiring HCV equal to zero from the total discounted cost of an individual who acquires HCV at age of 23 (since this data does not exist for HCV, the value is based on the average age of HIV infection from [102] because they have similar behavioral factors). Both calculations are made assuming no screening. The cost of infecting others is thus calculated to be $\$50,939$, which is equivalent to a reduction of 1.1 QALYs in our model. The expected cost of a secondary infection must be multiplied by the probability of infecting someone else at each time step, which will depend greatly on the behavioral characteristics of the individual in question and his age. Hence, we will assume that the probability of infecting others is identical to the patient's probability of acquiring HCV himself.

Most of the existing screening papers have not taken into account varying progressions due to alcohol, although medical studies have found there can be a significant

Table 4: Parameter values for model

Parameter	Value	Reference
Time Horizon of Decision	ages 15-35	model assumption
Probability of Infection for IDU population	0.014	[28]
Probability of Infection for general population	0.0004	[25]
Probability of Infection after age 33 (all pops)	0	model assumption
Progression to decomp cirrhosis (alcohol)	0.0115	[136]
Progression to decomp cirrhosis (no alcohol)	0.0025	[136]
Death rate in decomp cirrhosis	0.22	[48]
Death rate due to other causes (ages 13-33)	0.0016	[23]
Death rate due to other causes (age $>$ 33)	0.015	[23]
Probability of Treatment Success	54%	[69]
% population heavy drinkers	4.9%	[78]
Discount factor for QALYs	0.97	[114]
Discount factor for costs	0.97	[68]
Costs		
Decompensated Cirrhosis	25,691	[122]
Test	24.42	[118]
Treatment	22,896	[40]
Secondary Infections	50,939	calculated from model
QALYs		
Infected Aware	0.98	[114]
Disease Complications	0.48	[57]

effect. We define a person as a heavy drinker if he has 2 or more drinks per day (greater than 50 grams of alcohol). We assume that once a person becomes aware of his infection, he reduces his drinking below the 50 grams threshold (we also study sensitivity to this factor), and he reduces his probability of infecting others by half [114]. Behavioral change, then, not only affects the progression of the disease in a patient, but also the rate of transmission to others.

Figure 5 and Table 5 show the average rewards obtained through the Monte Carlo simulations using the aforementioned testing strategy. Several scenarios are considered including testing the overall and IDU populations under different alcohol and test acceptance assumptions. In the case of the overall and IDU populations we

assume 4.9% of the population drink excessively (estimated using the 2001-2004 National Health and Nutrition Examination Survey) as the base case. Since it may not be feasible to conduct an HCV test at the date specified by our testing strategy, we also consider cases where there is a 70% chance of the test occurring when it is cost-effective to do so. We refer to this scenario in the results as 70% test acceptance, where we assume 4.9% of the population consumes alcohol excessively. We consider cases where individuals who drink excessively and become aware of an HCV infection reduce alcohol consumption to less than 50 grams/day only 50% of the time, rather than the base case of always reducing alcohol consumption. Finally, in the last column, we consider only individuals who do not consume alcohol excessively.

Testing is never cost-effective for members of the overall population who do not consume alcohol excessively or when there is a only 50% alcohol reduction, and consequently the mean number of tests for those is 0 (as is the mean QALY gain and cost). Testing occurs at age 20 and 25 for the overall population base case. When we use a 70% acceptance rate and the first test is missed, the testing strategy then recommends tests at age 21 and 26. If the test at age 21 is also missed, then the testing recommends testing at age 22 only. For IDU population, testing is cost-effective every year of the decision epochs (i.e., ages 16-35), regardless of whether any tests are missed. It is clear from the figure that the largest QALY gains are a result of high incidence and alcohol consumption.

The threshold incidence for testing (and treating when necessary) is 0.021% when 4.9% of the population consumes alcohol excessively. In that case testing becomes cost-effective for age 27. When no excessive alcohol consumption is assumed, the threshold incidence is 0.066% and testing becomes cost-effective for age 24.

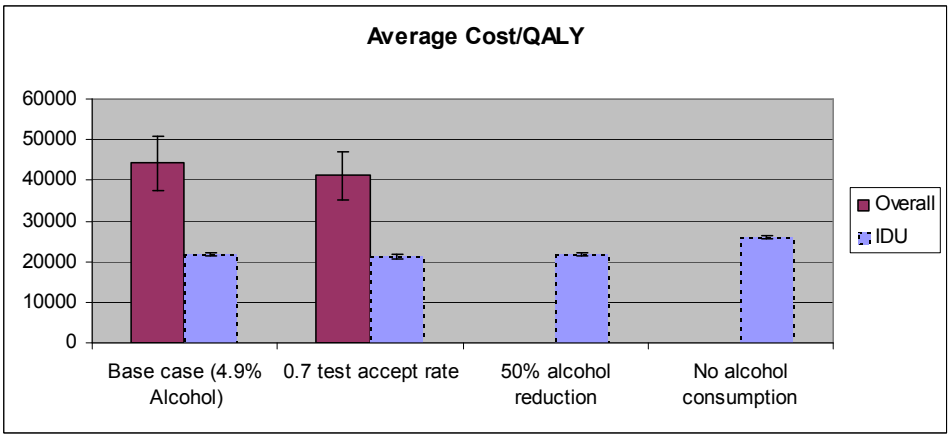


Figure 5: Mean cost/QALY for the overall and IDU populations from implementing the dynamic policy. Confidence intervals calculated using standard errors.

3.5 Conclusions

In this Chapter we have developed a Markov Decision Process (MDP) model for examining the timing of testing with treatment as an option where the action can result in disease awareness that changes the progression of the disease, and the MDP has partial updating of disease presence based on testing. In particular, we find sufficient conditions for testing (and treating) to be cost beneficial from a societal perspective. We consider dynamical screening and treatment policies that are determined by the result of the individual's previous test result as well as disease's progression and infectivity characteristics.

We use the MDP model in the case of Hepatitis C to study the timing of test and treatment actions for various populations. We use medical data to estimate the progression of the disease, prevalence, health costs, and infectivity. We find that both test-only and test-and-treat are cost-effective for IDU populations. The additional QALYs gained as compared to no testing can be as high as 0.15 for test-and-treat for IDUs. Regarding the general population, our recommendations find that testing and treating meets the \$50,000/QALY cost-effectiveness threshold, albeit not by much.

Table 5: Results for the overall and IDU populations from implementing the dynamic testing policy (standard errors are shown in parenthesis)

Overall population			
	Mean QALYs gained	**Mean Cost**	**Mean number of tests**
Base case (4.9% Alcohol)	0.0026 (0.00027)	116.82 (3.1)	1.9826 (0.00026)
0.7 test accept rate	0.0028 (0.00026)	119.39 (3.2)	1.9701 (0.00032)
IDU Population			
Base case (4.9% Alcohol)	0.1625 (0.0028)	3548.9 (42.4)	17.591 (0.0045)
0.7 test accept rate	0.1503 (0.0029)	3214.5 (40.9)	12.269 (0.0062)
50% alcohol reduction	0.1622 (0.0028	3551.1 (42.5)	17.591 (.0045)
No alcohol consumption	0.1401 (0.0023)	3663.6 (42.0)	17.591 (.0045)

We also find that incorporating behavior has an impact on recommendations, but that the IDU population should be tested even if there are no behavioral changes from awareness of having HCV. We also find that the number of actions tends to decrease with the effectiveness of behavioral changes in drinking, and that the effectiveness of alcohol behavior change can impact whether test-and-treat is better.

Our analysis also supports the CDC recommendations to test and treat groups with high risk of acquiring HCV. The overall population group is not currently addressed by CDC recommendations since most analysis has not incorporated the effect of alcohol behavior change on progression of Hepatitis C. We also add to the literature by specifying ages for testing of other populations. We have also studied cost-effectiveness using cost minimization (including productivity losses) as the objective or examining the ratio of cost paid to utility gained, and we find similar recommendations on which groups should be tested and how often.

For future work, it would be useful to apply our MDP for determining when to test or treat to other diseases with different characteristics. The research also suggests that examining dynamic screening policies (e.g., as in [38]) could be beneficial for some diseases. Further analysis of MDPs could also suggest other types of policies to consider.

3.6 Appendix

Proof of Theorem 1

Let π be a policy that satisfies $\pi_{t_{i+1}} = NT$ for $t > t_i$. Backwards induction shows that

$$\begin{aligned}
u^{\pi}_{t_{i+1}}((0,0)) &= R_{(0,0)} \sum_{i=t_{i+1}}^{N} (\lambda p_{00})^{N-i} + R_{(1,0)} p_{01} \sum_{i=t_{i+1}}^{N-1} \lambda^{N-i} \sum_{j=0}^{N-i-1} p_{00}^{j} p_{11}^{N-i-j-1} \\
&\quad + R_{(2,1)} p_{01} p_{12} \sum_{i=t_{i+1}}^{N-2} \lambda^{N-i} \sum_{j=0}^{N-i-2} p_{00}^{j} \sum_{k=0}^{N-i-j-2} p_{11}^{k} q_{22}^{N-i-j-k-2}, \\
u^{\pi}_{t_{i+1}}((1,0)) &= R_{(1,0)} \sum_{i=t_{i+1}}^{N} (\lambda p_{11})^{N-i} + R_{(2,1)} p_{12} \sum_{i=t_{i+1}}^{N-1} \lambda^{N-i} \sum_{j=0}^{N-i-1} p_{11}^{j} q_{22}^{N-i-j-1}, \\
u^{\pi}_{t_{i+1}}((1,1)) &= R_{(1,1)} \sum_{i=t_{i+1}}^{N} (\lambda q_{11})^{N-i} + R_{(2,1)} q_{12} \sum_{i=t_{i+1}}^{N-1} \lambda^{N-i} \sum_{j=0}^{N-i-1} q_{11}^{j} q_{22}^{N-i-j-1}, \\
u^{\pi}_{t_{i+1}}((2,1)) &= R_{(2,1)} \sum_{i=t_{i+1}}^{N} (\lambda q_{22})^{N-i}.
\end{aligned}$$

1. First, assume that the individual is healthy at time t_i (either his test result shows that he is healthy or he is treated successfully). We will compare the policies π and π', where $\pi'_t = \pi_t$ for $t \leq t_i$, $\pi'_{t_{i+1}} = T$, $\pi'_t = NT$ for $t > t_i$ and $t \neq t_{i+1}$. When policy π' is used, we obtain the following:

$$\begin{aligned}
u^{\pi'}_{t_{i+1}}((0,0)) &= -c_0 + u^{\pi}_{t_{i+1}}((0,0)) \\
u^{\pi'}_{t_{i+1}}((1,0)) &= -c_0 + u^{\pi}_{t_{i+1}}((1,1)) \\
u^{\pi'}_{t_{i+1}}((2,1)) &= u^{\pi}_{t_{i+1}}((2,1)).
\end{aligned}$$

We then find the expected values of utility-to-go functions as

$$\begin{aligned}
\mathbb{E}[u^{\pi}_{t_{i+1}}] &= \Big(p_{00}^{t_{i+1}-t_i}\Big)u^{\pi}_{t_{i+1}}((0,0)) + \Big(p_{01}\sum_{j=0}^{t_{i+1}-t_i-1} p_{00}^{j}p_{11}^{t_{i+1}-t_i-j-1}\Big)u^{\pi}_{t_{i+1}}((1,0)) \\
&\quad + \Big(p_{01}p_{12}\sum_{j=0}^{t_{i+1}-t_i-2} p_{00}^{j}\sum_{k=0}^{t_{i+1}-t_i-j-2} p_{11}^{k}p_{22}^{t_{i+1}-t_i-j-k-2}\Big)u^{\pi}_{t_{i+1}}((2,1)), \\
\mathbb{E}[u^{\pi'}_{t_{i+1}}] &= \Big(p_{00}^{t_{i+1}-t_i}\Big)u^{\pi'}_{t_{i+1}}((0,0)) + \Big(p_{01}\sum_{j=0}^{t_{i+1}-t_i-1} p_{00}^{j}p_{11}^{t_{i+1}-t_i-j-1}\Big)u^{\pi'}_{t_{i+1}}((1,1)) \\
&\quad + \Big(p_{01}p_{12}\sum_{j=0}^{t_{i+1}-t_i-2} p_{00}^{j}\sum_{k=0}^{t_{i+1}-t_i-j-2} p_{11}^{k}p_{22}^{t_{i+1}-t_i-j-k-2}\Big)u^{\pi'}_{t_{i+1}}((2,1)).
\end{aligned}$$

Some algebra shows that $\mathbb{E}[u^{\pi'}_{t_{i+1}}] \geq \mathbb{E}[u^{\pi}_{t_{i+1}}]$ if $f_1(t_i, t_{i+1}) \geq 0$. Since π and π' agree before time t_{i+1}, Lemma 1 shows that $\mathbb{E}[u_0^{\pi'}] \geq \mathbb{E}[u_0^{\pi}]$.

Similarly, we compare two policies π and π'', where $\pi''_t = \pi_t$ for $t \leq t_i$, $\pi''_{t_{i+1}} = T_1T_2$, $\pi''_t = NT$ for $t > t_i$ and $t \neq t_{i+1}$. When policy π'' is used, we obtain the following:

$$\begin{aligned}
u^{\pi''}_{t_{i+1}}((0,0)) &= -c_0 + u^{\pi}_{t_{i+1}}((0,0)) \\
u^{\pi''}_{t_{i+1}}((1,0)) &= -c_0 - c_1 + v_1 u^{\pi}_{t_{i+1}}((0,0)) + (1-v_1)u^{\pi}_{t_{i+1}}((1,1)), \\
u^{\pi''}_{t_{i+1}}((2,1)) &= u^{\pi}_{t_{i+1}}((2,1)).
\end{aligned}$$

We then find the expected values of utility-to-go functions as above and some algebra shows that $\mathbb{E}[u^{\pi''}_{t_{i+1}}] \geq \mathbb{E}[u^{\pi}_{t_{i+1}}]$ if $f_3(t_i, t_{i+1}) \geq 0$. Since π and π'' agree before time t_{i+1}, Lemma 1 shows that $\mathbb{E}[u_0^{\pi''}] \geq \mathbb{E}[u_0^{\pi}]$.

2. We compare two policies π and π''', where $\pi'''_t = \pi_t$ for $t \leq t_i$, $\pi'''_{t_{i+1}} = T_2$, $\pi'''_t = NT$ for $t > t_i$ and $t \neq t_{i+1}$. When policy π''' is used, we obtain the following:

$$u^{\pi'''}_{t_{i+1}}((1,1)) = -c_1 + v_1 u^{\pi}_{t_{i+1}}((0,0)) - v_1 u^{\pi}_{t_{i+1}}((1,1)),\ u^{\pi'''}_{t_{i+1}}((2,1)) = u^{\pi}_{t_{i+1}}((2,1)).$$

Some algebra shows that $\mathbb{E}[u^{\pi'''}_{t_{i+1}}] \geq \mathbb{E}[u^{\pi}_{t_{i+1}}]$ if $f_2(t_i, t_{i+1}) \geq 0$. Since π and π''' agree before time t_{i+1}, Lemma 1 shows that $\mathbb{E}[u_0^{\pi'''}] \geq \mathbb{E}[u_0^{\pi}]$. $\square$

Next, we generalize our results to include the case when false negatives may occur Specifically, let n_1 denote the probability that the test result is negative when the person's health state is 1. In this case, closed form expressions can be obtained but they are quite complicated, hence we will simply give the solution methodology for different cases.

1. If the person's test result is negative at time t_i and he has not been tested since time t_{i-1}, we can find the real probabilities of being at each state under policy π as follows:

$$\begin{aligned} b^{\pi}_{t_i}((0,0)) &= \frac{p_{00}^{t_i-t_{i-1}} b^{\pi}_{t_i-1}((0,0))}{p_{00}^{t_i-t_{i-1}} b^{\pi}_{t_i-1}((0,0)) + n_1 \frac{p_{01}(p_{11}^{t_i-t_{i-1}} - p_{00}^{t_i-t_{i-1}})}{p_{11}-p_{00}} b^{\pi}_{t_{i-1}}((1,0))}, \\ b^{\pi}_{t_i}((1,0)) &= 1 - b^{\pi}_{t_i}((0,0)). \end{aligned}$$

   Now, let us compare two policies π and π', where $\pi_t = \pi'_t$ for $t \leq t_i$, $\pi_{t_{i+1}} = NT$, $\pi'_{t_{i+1}} = T_1$, $\pi_t = \pi'_t = NT$ for $t > t_i$ and $t \neq t_{i+1}$. $u^{\pi}_t(s)$ are the same as in the proof of Theorem 1 for every $s \in S$. When policy π' is used, we obtain the following:

$$\begin{aligned} u^{\pi'}_{t_{i+1}}((0,0)) &= -c_0 + u^{\pi}_{t_{i+1}}((0,0)), \\ u^{\pi'}_{t_{i+1}}((1,0)) &= -c_0 + (1-n_1) u^{\pi}_{t_{i+1}}((1,1)) + n_1 u^{\pi}_{t_{i+1}}((1,0)), \\ u^{\pi'}_{t_{i+1}}((2,1)) &= u^{\pi}_{t_{i+1}}((2,1)). \end{aligned}$$

   We then find the expected values of utility-to-go function for π as

$$\begin{aligned} \mathbb{E}[u^{\pi}_{t_{i+1}}] &= b^{\pi}_{t_i}((0,0)) \Big\{ \Big(p_{00}^{t_{i+1}-t_i} \Big) u^{\pi}_{t_{i+1}}((0,0)) \\ &\quad + \Big(p_{01} \sum_{j=0}^{t_{i+1}-t_i-1} p_{00}^{j} p_{11}^{t_{i+1}-t_i-j-1} \Big) u^{\pi}_{t_{i+1}}((1,0)) \\ &\quad + \Big(p_{01} p_{12} \sum_{j=0}^{t_{i+1}-t_i-2} p_{00}^{j} \sum_{k=0}^{t_{i+1}-t_i-j-2} p_{11}^{k} p_{22}^{t_{i+1}-t_i-j-k-2} \Big) u^{\pi}_{t_{i+1}}((2,1)) \Big\} \\ &\quad + b^{\pi}_{t_i}((1,0)) \Big\{ \Big(p_{11}^{t_{i+1}-t_i} \Big) u^{\pi}_{t_{i+1}}((1,0)) \\ &\quad + \Big(p_{12} \sum_{j=0}^{t_{i+1}-t_i-1} p_{11}^{j} p_{22}^{t_{i+1}-t_i-j-1} \Big) u^{\pi}_{t_{i+1}}((2,1)) \Big\}, \end{aligned}$$

and the utility-to-go function for π' can be calculated similarly. Then, we can find a condition which guarantees that $\mathbb{E}[u^{\pi'}_{t_{i+1}}] \geq \mathbb{E}[u^{\pi}_{t_{i+1}}]$ will hold.

Similarly, if the test result at time t_i indicates that he is healthy and we want to compare two policies π and π'', where $\pi_t = \pi''_t$ for $t \leq t_i$, $\pi_{t_{i+1}} = NT$, $\pi''_{t_{i+1}} = T_1T_2$, $\pi_t = \pi''_t = NT$ for $t > t_i$ and $t \neq t_{i+1}$. $u^{\pi}_{t_{i+1}}(s)$ are as in the proof of Theorem 1, when policy π'' is used, we obtain the following:

$$\begin{aligned}
u^{\pi''}_{t_{i+1}}((0,0)) &= -c_0 + u^{\pi}_{t_{i+1}}((0,0)), \\
u^{\pi''}_{t_{i+1}}((1,0)) &= -c_0 - c_1 + (1-n_1)\Big(v_1 u^{\pi}_{t_{i+1}}((0,0)) + (1-v_1)u^{\pi}_{t_{i+1}}((1,1))\Big) \\
&\quad + n_1 u^{\pi}_{t_{i+1}}((1,0)), \\
u^{\pi''}_{t_{i+1}}((2,1)) &= u^{\pi}_{t_{i+1}}((2,1)).
\end{aligned}$$

Then, we can find the utility-to-go functions and a condition that guarantees that $\mathbb{E}[u^{\pi''}_{t_{i+1}}] \geq \mathbb{E}[u^{\pi}_{t_{i+1}}]$ will hold.

2. If the individual is tested and found to be in disease stage 1 and the treatment is successful, then this case is the similar to the previous case, assuming that after the treatment, individual's health state can be determined accurately. We use $b^{\pi}_{t_i}((0,0)) = 1$, and $b^{\pi}_{t_i}((1,0)) = 0$ when finding the expected utility-to-go functions and otherwise proceed as above.

3. If the individual is tested and found to be in disease stage 1 and the treatment is not successful (or he is not treated) at time t_i, then this case is similar to the cases without false negatives when the person is known to be in disease stage 1 at time t_i.

CHAPTER IV

TESTING FOR HEPATITIS C VIRUS: IMPLICATIONS FOR POLICY

4.1 Introduction

Hepatitis C virus (HCV) is a blood-borne disease present in 3.9 million people in the US (48% are unaware) that can cause end stage liver disease [25]. It is the leading cause for liver transplants and the 10th leading cause of death in the US. It is generally asymptomatic for decades, and many people are unaware of the presence of the disease until end stage liver disease begins and treatment is no longer effective. HCV can be transmitted during the asymptomatic period. There is currently no vaccine, and treatments are somewhat effective if caught early enough. Once infected about 85% will go on to chronic HCV which leads to liver cirrhosis in over 20% of patients [79]. The risk of hepatocellular carcinoma is 17 times higher in patients with HCV [42]. 8,000 to 10,000 deaths per year in the U.S. are attributed to chronic liver disease, 40-60% of which are caused by HCV [21]. Prevalence in the overall population is 2% and as high as 40-60% in injection drug users, which is twice that for HIV. [28].

Several researchers have used models to assess the cost-effectiveness of treatment of HCV. The general consensus is that peginterferon therapy with ribavirin is cost-effective in treating individuals who are known to have HCV [57, 112, 138, 7]. There is far less consensus on the literature regarding testing for HCV. For example, Castelnuovo [14] finds that testing is cost-effective for injection drug users based on a decision analytic model; Gordon [54] also reports that testing can be cost-effective for high-risk groups. However, Plunkett [96] finds that screening of asymptomatic pregnant women for HCV infection is not cost-effective, even when the benefits to the child are

considered. In addition, Pereira [91] used a Markov simulation model to show that testing blood donors for HCV was cost-saving for the health care system. Finally, Singer [114] found that routine HCV testing was not cost-effective in asymptomatic, average-risk adults using a Monte Carlo Markov Simulation where the age of testing was assumed a priori, albeit assuming an older treatment than today's standard.

There is also little consensus among public health agencies on testing recommendations. For example, the US Public Health Service (USPHS) and Infectious Diseases Society of America (IDSA) recommend testing HIV positive persons for HCV. The US Preventive Services Task Force (USPSTF) recommends against routine testing for HCV in the general population and does not make a recommendation for at risk population. The Centers for Disease Control and Prevention (CDC), on the other hand, does recommend routine testing for population groups with a high risk for acquiring HCV such as drug users or commercial sex workers.

Awareness of HCV can effect the disease progression (by reducing alcohol consumption, if any) and potentially reduce secondary infections. Neither the literature on cost-effectiveness for screening of HCV nor the public health agencies make recommendations about timing of testing. The explicit consideration of timing is important, not only because it reveals the best time to test, but also because it could alter whether or not testing is cost-effective in the first place. The aim of this Chapter is to determine whether (and when) it is cost-effective to test and treat HCV, and to evaluate the best ages for testing.

4.2 Materials and Methods

We develop a Markov model of the natural history of Hepatitis C infection following those of Bennett [7], Singer [114], Stein [118], and Hornberger [57]. The states of the Markov chain are displayed in Figure 6, and the transition probabilities are listed in Table 6. Transition probabilities are annual so that each year of an individual's

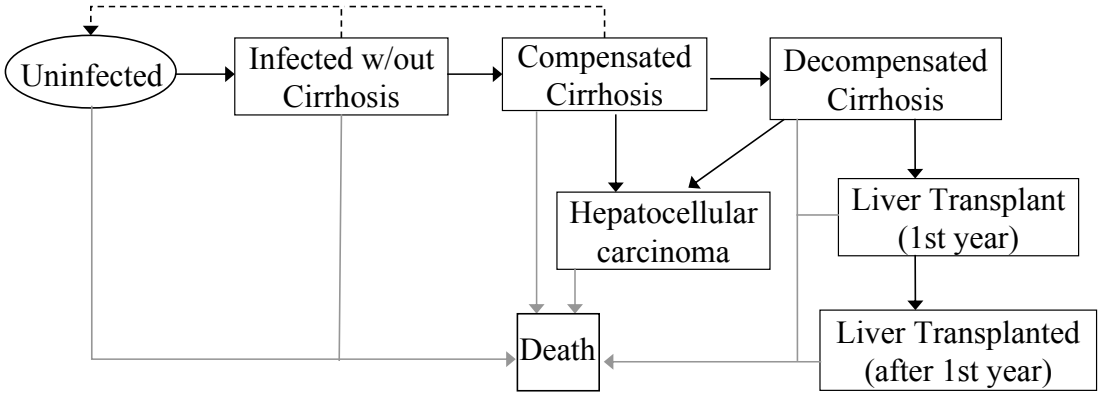


Figure 6: The states of HCV where solid lines represent natural history transitions and dashed lines represent transitions due to treatment success. Not shown in the diagram but included in the model are transitions to death from all states due to causes other than HCV.

life he or she will transition to another health state or remain in the same state. Each individual begins uninfected at age 13 and may or may not acquire HCV in the following years of life. Transition probabilities vary according to age, risk group, alcohol consumption, and awareness of HCV status. Each health state has an associated annual cost and utility measured in quality adjusted life years which are listed in Table 7. We discount the costs and QALY's by 3% [114] and compute the total discounted lifetime cost and QALYs of individuals with and without testing. All costs and QALYs are discounted to age 13 which represents the first period of the model.

We allow possibility of false positives and negatives. When a true positive is detected the individual will decrease his transmission to others by half, reduce alcohol consumption to less than 50 grams per day (if previously drinking greater than 50 grams per day), and begin treatment in the same year (if eligible as determined by the probability in Table 10).

We assume the treatment for all patients to be peginterferon plus ribavirin which is

the current most cost-effective and successful treatment for HCV. We include the possibility of individuals quitting treatment without completing it, as well as accounting for the percentage of people ineligible for treatment all together due to contraindications or other reasons. We account for the different success rates of treatments based on genotype, and include a probability of individuals being of each genotype. When treatment is not successful, and/or the individual drops out of treatment early, the individual continues through the natural history of HCV. When the treatment is successful, the individual is then assumed uninfected and still susceptible to acquiring HCV in the future.

Tables 7 - 10 lists the various assumptions and values used in the model. When there does not exist a consensus in the literature for a parameter value, we choose conservative estimates. For example, Chong [31] cites several different utility values for the health states of individuals with HCV used in previous papers; some derived from patient survey and others from expert elicitation. In our base case analysis we chose QALY values from Singer which are the most conservative, (i.e., higher QALY values for each state than other studies). We also choose a conservative estimate for the effect alcohol has on progression to cirrhosis (Freeman [51] versus Poynard [98]). Consequently, our results are not biased towards screening.

The model was written in Matlab 6.5 and run on a computing cluster with nodes that contain two 2.4GHz Xeon processors and 2GB RAM each.

4.3 Results

We considered up to 5 tests per lifetime in our analysis. Tests at all age combinations were considered with a minimum of two years between tests. We report the results of the model using additional discounted cost and additional discounted QALYs resulting from the testing policy.

Table 6: Parameter values for transition probabilities (reported in annual terms)

Base case transition probabilities	Value	References
Progression rate to Compensated Cirrhosis for heavy drinkers	0.0072	[51]
Progression rate to Compensated Cirrhosis without heavy drinking	0.0036	[98]
Progression rate to Decompensated Cirrhosis	0.0390	[7]
Progression rate to HCC from Cirrhosis or Decompensated Cirrhosis	0.0268	[36]
Rate of liver transplant from Decompensated Cirrhosis	0.0300	[7]
Death rate from Decompensated Cirrhosis	0.2180	[48]
Death rate from HCC	0.4270	[48], [7]
Death rate after liver transplant first year	0.1370	[50]
Death rate after liver transplant after first year	0.0520	[50]
Death rate from other causes		[23]

4.3.1 Overall Population

Our results indicate that it costs less than \$50,000/QALY to test (and possibly treat) members of the overall population for HCV during ages 20-51. These results show that the timing of HCV testing is important in determining its cost-effectiveness. It is never cost-effective to test only (without treatment) for the overall population. The test age that results in the highest additional QALYs is 33. This age balances the tradeoffs of testing later to decrease the likelihood of testing before infection occurs while testing early enough so that treatment is still effective and more secondary infections can be averted.

Figure 7 is a plot of the additional discounted QALYs gained versus the additional discounted cost for testing the overall population at various ages. Points below the \$50K/QALY line represent ages at which it is cost-effective to test. The results are shown for three scenarios: the base case (where anyone who drinks more than 50 grams of alcohol per day reduces consumption to less than 50 grams per day after a positive HCV test), under the assumption that no one reduces alcohol consumption

Table 7: QALY values for various health states

Health State	QALY Value	Reference
Uninfected	1.00	[114]
Infected without Cirrhosis	0.96	[114]
Compensated Cirrhosis	0.80	[114]
Decompensated Cirrhosis	0.56	[114]
Transplantation (1st year)	0.80	[114]
Transplantation (after 1st year)	0.95	[114]
Hepatocellular Carcinoma	0.25	[114]
Treatment	0.93	[7]

Table 8: Annual probability of HCV infection by type of risk group

Risk Group	**Annual Probability of Infection**	**References**
Overall population	0.0004	[25]
Injection drug users	0.014	[28]
Incarcerated individuals	0.0016	[28]
Commercial sex workers	0.0012	[80]
STD clinic attendees (non-IDU)	0.0008	[80]

following an HCV test, and when only patients with genotype 2 are treated for HCV. The ages corresponding to each point are listed only for the base case. The cost-effective age range for various populations, including those shown in Figure 7, are displayed in Table 14.

Figure 8 shows the results of one and two tests per lifetime for the overall population. Each dark blue point represents the additional cost and additional QALYs resulting from a test at the specified age. We can see that testing between ages 20 and 51 results in a cost/QALY smaller than \$50,000/QALY. Each light blue point represents the additional cost and additional QALYs resulting from some two test policy (compared to no testing at all) where the ages are not listed on the figure. Table 11 lists the policies in which the second test is incrementally cost-effective given the first test. This table can be used for an individual that has already had an HCV test

Table 9: Costs and discount value (where costs are in 2000 dollars)

Parameter	Value	References
Cost of screening test (ELISA)	$ 24.42	[118]
Cost of combination therapy (peginterferon + ribavirin)	$ 22,896	[40]
Discount factor for costs	3%	[68]
Discount factor for QALYs	3%	[114]
Health State		
Compensated Cirrhosis	$ 494 / year	[122]
Decompensated Cirrhosis	$ 25,691 / year	[122]
Transplantation (1st year)	$ 312,804 / year	[122]
Transplantation (after 1st year)	$ 30,121 / year	[122]
Hepatocellular Carcinoma	$ 16,748 / year	[122]

Table 10: Genotype, testing, and infection values

Factor	Value	Reference
Percent of population that is Genotype 1 (G1)	60%	[57]
Treatment success rate for G1	29%	[57]
Treatment success rate for non-Genotype 1 (G2)	62%	[57]
Probability of ELISA test false negative	0.014	[114]
Probability of ELISA test false positive	0.009	[114]
QALY change from infecting others	-1.1	calculated by model
Percent of population of heavy drinkers	4.90%	[78]

and is interested in whether or not a testing at another age would be incrementally cost-effective. However, it is important to note that Table 11 does not suggest that the policies listed in it are the best two test policies. When multi-test policies are considered as one single intervention, rather than multiple interventions at different ages, we can find the policies that produce the maximum QALY gained while still being cost-effective overall. These results are listed in Table 12.

Table 12 also shows the incremental cost-effectiveness of multiple tests. Since there are many policies that are cost-effective, we pick the single test policy that results in the highest QALYs and compare it to the two test policy with the highest QALYs (only policies that are cost-effective according to the $50,000/QALY threshold are

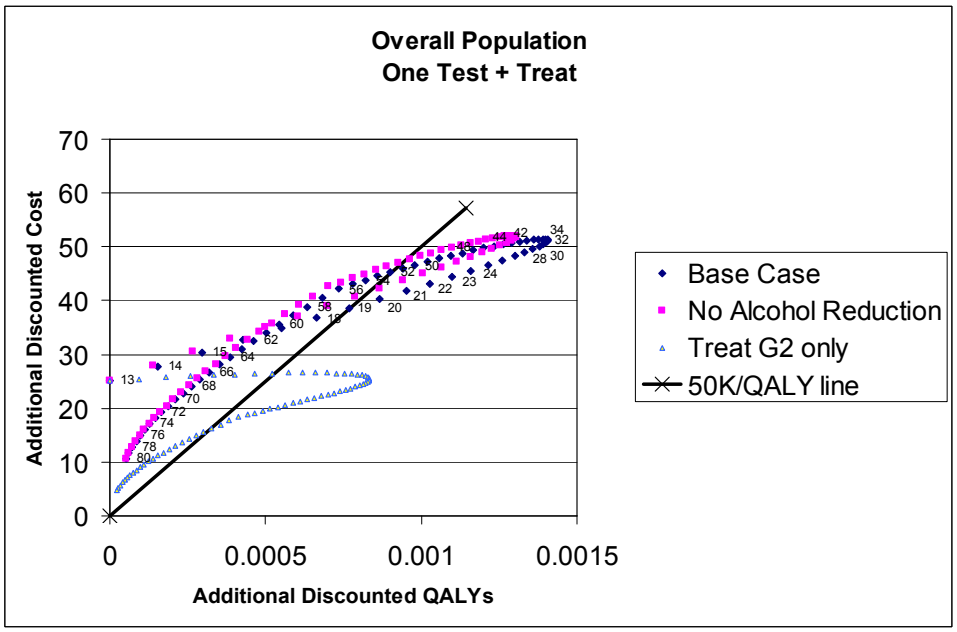


Figure 7: Cost versus QALY for one test policies of the overall population under various assumption.

considered). Table 12 displays the incremental cost-effectiveness of adding additional tests in this way for both the overall and the IDU populations.

4.3.2 At Risk Populations

Populations at risk for acquiring HCV have even wider age ranges for which single testing is cost-effective. Table 13 indicates the age ranges for testing only and testing with possible treatment various risk groups. For the riskiest group, injection drug users, the single test cost-effective age range is 15-73 for those that drink more than 50 grams of alcohol per day, and 16-54 for those who drink less than 50 grams of alcohol per day. The single test age yielding the highest QALYs is 32 and 35, respectively, yielding $1,950.29/QALY and $16,954.84/QALY, respectively. In fact, testing alone (without treatment) is itself cost savings for all ages greater than 15 for alcohol drinkers and from 16-50 for non-alcohol drinkers.

Figure 9 shows the results for testing IDU's that drink more than 50 grams/day of alcohol multiple times. All possible testing policies starting at one test per lifetime

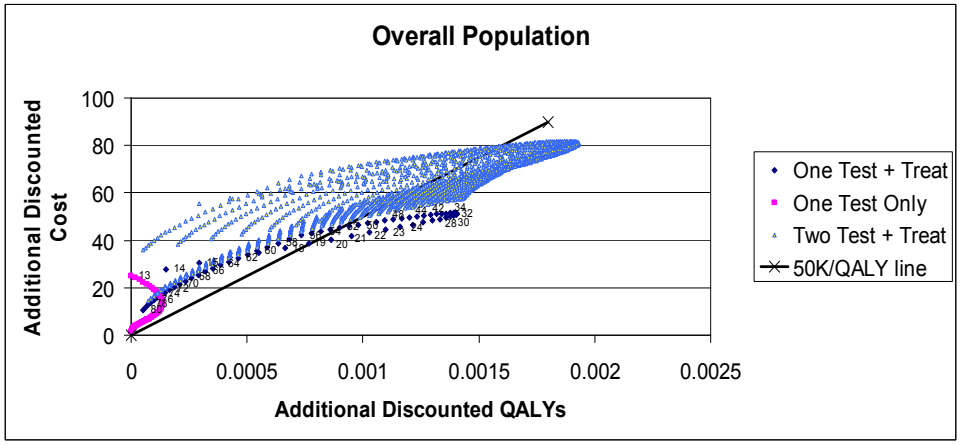


Figure 8: Cost versus QALY for one and two test policies.

up to five tests per lifetime are displayed. One can see diminishing returns for higher number of tests, but a reasonable increase in QALYs for a modest increase in cost. Since there are far too many policies to explicitly list, we indicate the testing ages of the policies that result in the highest QALYs for each of 1-5 tests per lifetime. We also indicate the testing ages of policies on the efficient frontier (i.e., policies that result in highest QALYs for least cost) at the points where the frontier changes from 1 test to 2 tests per lifetime, 2 to 3 tests per lifetime, etc. We also indicate the testing ages of one of the poorer policies that results in fewer QALYs gained at higher cost.

While the vast majority of policies are cost-effective, beginning the first test early and spacing out the tests during the lifetime is most efficient. Virtually all age combinations where the ages fall within the single test cost-effectiveness age range are also cost-effective. Since it may be difficult to obtain IDU's for testing, a reasonable policy would be to test an IDU whenever given the opportunity, even if they've had previous tests. Ideally, however, the tests should be reasonably spaced apart so that the policy lies closer to the efficient frontier. The resulting cost and QALYs depend on several factors including the spacing between tests, the number of tests, and the

Table 11: Incrementally cost-effective two test policies for testing the overall population for HCV

Two test policies where first and second tests are Incrementally cost-effective (ICE)			
first test age, second test age range	second test age with best ICE	ICE of first test	ICE of second test with best ICE
20, 28-49	37	\$46,631	\$42,448
21, 30-49	37	\$43,989	\$43,506
22, 31-48	38	\$42,001	\$44,585
23, 33-47	39	\$40,480	\$45,722
24, 35-47	40	\$39,306	\$46,915
25, 37-45	41	\$38,395	\$48,171
26, 39-44	41	\$37,693	\$49,466

age of the first test. However, one can see from the multi-test policies that result in the highest QALYs that policies in which the tests are somewhat centered around age 32 and fall within the single test cost-effective age range are quite efficient.

From Table 13 we can see that for the several at-risk populations the testing age resulting in highest QALYs is near 32 for those that drink more than 50 grams of alcohol per day, and near 35 for those who do not. The principle of centering the tests around this age while stilling falling within the single test cost-effective age range also holds and results in very cost-effective testing policies.

4.4 Sensitivity Analysis

Our estimate of incidence of HCV in the overall population does not include those that acquired HCV through a blood transfusion prior to 1992. However, we also ran the model with values for incidence that does include transfusion related HCV infection because many of those infections are found decades later. We found that the results change very little due to the fact that the number of people that have had transfusions prior to 1992 is very small compared to the total population. By including those with transfusions, however, the cost-effective single test age range changes to 19-52.

Table 12: Multiple tests per lifetime for testing the overall on IDU populations for HCV

For each number of tests, the policy with the highest QALY's is listed					
Overall Population					
	1 test	2 tests	3 tests	4 tests	5 tests
Test Ages	33	26,44	23,35,51	21,31,41,55	19,25,33,43,55
Cost-effectiveness ($/QALY)	$36,340	$41,958	$46,883	$51,464	$57,425
Incremental cost-effectiveness		$57,008	$82,516	$115,891	$202,482
IDU Alcohol Population					
Test Ages	32	26,41	23,33,47	21,29,37,49	19,25,31,39,49
Cost-effectiveness ($/QALY)	$1,950	$2,458	$2,859	$2,894	$2,822
Incremental cost-effectiveness					

We also found that reducing the incidence of the overall population by a factor of 2.5 (i.e., 0.016%) was just enough to cause testing (with possible treatment) to no longer be cost-effective at any age. When the probability of acquiring HCV reaches 0.0595%, testing only starts to become cost-effective (at age 41 only)

There is also some concern over the disutility of awareness of an HCV infection due to psychological harm. Our model indicates that only if the cost of this harm exceeds $4,850 is testing (and possibly treating) the overall population no longer cost-effective for any age. Moreover, if the cost of the HCV test increases by at least an additional $38 (or if the disutility of taking a test equates to $38), then testing (and possibly treating) the overall population is no longer cost-effective for any age.

Table 14 highlights the results of the model using other values for various parameters. The first column represents the parameter that is being changed from the base case. The second column represents the single test policy that results in the highest QALYs. The third column represents the single test+treat policy that results in the highest QALYs. Finally, the last column represents the cost/QALY of the policy

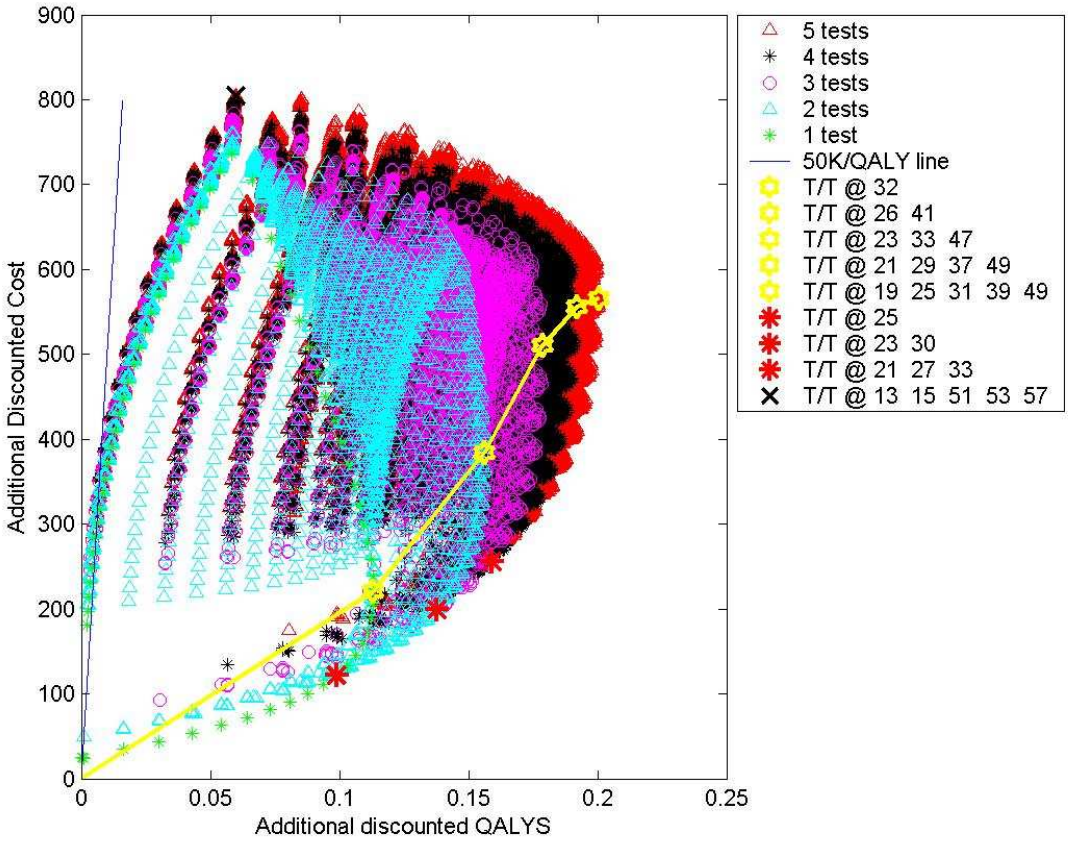


Figure 9: Cost versus QALY of up to 5 tests per lifetime for the IDU population .

in the previous column. The table shows results for various changes of the parameter values, including using more liberal estimates of the effect of alcohol and quality adjusted life years.

4.5 Discussion

Our results show that when determining the cost-effectiveness of HCV testing, considering timing and the effect of alcohol on disease progression are critical. If we ignore the effect of alcohol, and average the costs and QALYs of testing over all of the ages considered, then testing for HCV is no longer cost-effective, with or without treatment. However, when we explicitly consider the ages and the effects of alcohol, we see that testing the overall population for HCV is indeed cost-effective for many ages.

Our results differ from Singer that state that testing the overall population is not cost-effective, however Singer's analysis did not include the effect of alcohol and was conducted when there was no available data on peginterferon therapy. Consequently, interferon therapy (plus ribavirin) was the assumed treatment, which has a much higher rate of ineligibility (80%) [114]. Singer states that the model is sensitive to this value, and that when 50% of the population is eligible for treatment, then testing the overall population *is* cost-effective. Since our analysis assumes 50% ineligibility for peginterferon therapy (plus ribavirin) based on [60, 101], our results agree with Singer in this regard. Under Singer's assumptions on treatment (i.e., cost, treatment success rate, % ineligible), and without including the effect of alcohol, our model agrees with Singer's result that it is not cost-effective to test the overall population at any age for HCV. However, even under the treatment assumptions from Singer, when 4.9% of the population drinks greater than 50 grams of alcohol per day (and reduce this consumption to less than 50 grams per day following a positive HCV test) our model suggests that it is cost-effective to test the overall population between ages 30 - 40, while using the conservative estimates on the effect of alcohol from Freeman [51].

Our results are in agreement with the CDC recommendations for high risk groups (although these do not explicitly consider age, frequency or alcohol). Our results additionally indicate, however, that testing (and treating) the overall population is also cost-effective for several ages. This result disagrees with the U.S. Preventive Services Task Force (USPSTF) which recommends against testing the overall population and makes no recommendation for high risk groups. The USPSTF believes that the "psychological harm" of screening such as anxiety or the impact on partner relationships can be an important factor. In addition, they argue that the natural progression of HCV is unclear in that treatment is only effective for a subset of the population and it is difficult to determine who these individuals are a priori. We do explicitly model

the disutility to the individual of ineffective treatment, though we did not explicitly consider the potential psychological harm from the screening process since there are not good estimates of this cost for HCV. We considered this issue in sensitivity analysis and found the dollar value that would be necessary to make testing no longer cost-effective. It is noteworthy that the anxiety of a positive test would be greater for HIV for which universal screening is recommended. Additionally, for an infectious disease that can lead to death, averting secondary infections must also be a primary consideration.

Our results indicate that for high risk groups, more than one test per lifetime may be the best policy. We found that centering the testing ages around age 32 results in the highest QALYs gained when testing begins in early twenties and ends in late forties. Our sensitivity analysis shows that our results are robust under our assumptions.

New HCV testing technologies are being developed that could impact the recommendation regarding universal screening. OraSure Technologies Inc. (www.orasure.com) is commercializing a rapid HCV test that uses a saliva sample, which should bring down costs and be easier to administer.

Is is noteworthy to recall that the conclusion that it is cost-effective to test the overall population for HCV is made from a public health or societal perspective. That is, a particular insurance company or medical practice may have different incentives and may or may not find it in their best interest to test for HCV. Many of the benefits of an HCV test are realized several years, perhaps decades, later. Consequently, the insurance company that pays for an individual's HCV test may not be the same one that receives the benefits of the test (in averted health costs). Indeed, this unfortunate misalignment of incentives regarding health outcomes and costs is true for many medical interventions. Fortunately, other incentives exist for insurance companies to reimburse expenditures associated with interventions that are regarded

as cost-effective (from a societal perspective) by the medical community.

Limitations of our model include the fact that we do not include productivity losses from morbidity/mortality (but we do include death rates from other causes). Consequently, our results are a conservative estimate. Due to limited availability of data, we use the same drop out and ineligibility rates for all populations, however, we again use conservative estimates. Lastly, the age ranges for increased probability of infection for high risk groups were taken from HIV data since it is not available for HCV. Because the viruses are transmitted through similar means we feel this estimate is reasonable. Our sensitivity analysis shows that the model is not sensitive to these values.

Table 13: Results for testing various groups at risk for HCV. CSW = Commercial Sex Workers. CE = cost-effectiveness.

Population	**Sub-population /Parameter**	**1 test CE ages**	**1 test+treat CE ages**	**1 test+treat with highest QALY**	**$/QALY for single test w/ highest QALY**
Overall	drink > 50 grams alcohol / day	>15	15-71	31	$9,047.97
Overall	drink < 50 grams alcohol / day	empty	22-48	34	$40,996.62
Overall	4.9% drink > 50 grams alcohol / day	empty	20-51	33	$36,340.09
CSW	drink > 50 grams alcohol / day	>15	15-72	32	$6,702.86
CSW	drink < 50 grams alcohol / day	empty	17-52	35	$34,474.00
CSW	4.9% drink > 50 grams alcohol / day	empty	17-55	34	$30,148.82
Incarcerated	4.9% drink > 50 grams alcohol / day	19-66	15-55	33	$27,522.30
IDU	drink > 50 grams alcohol / day	15-80	15-73	32	$1,950.29
IDU	drink < 50 grams alcohol / day	16-50	16-54	35	$16,954.84

Table 14: Sensitivity analysis for testing the overall population for HCV. CE = cost-effective

Sub-population /Parameter	**1 test+treat CE ages**	**1 test+treat with highest QALY**	**$/QALY for single test w/ highest QALY**
No Alcohol Reduction	21-49	34	$39,804.28
Treat Genotype 2 Only	21-59	33	$30,480.54
Infect age start = 20	29-49	39	$42,984.19
Infect age end = 45	20-51	33	$36,597.72
Cost of treatment × 0.75	19-56	33	$31,455.21
Poynard [98] Alcohol progression value	17-58	33	$27,177.07
QALY values from Shiell [110]	18-59	34	$30,649.01
Progression to Cirrhosis × 0.75	20-51	33	$36,340.09
Progression to Decomp. Cirrhosis × 0.75	20-51	34	$37,709.55
Progression to HCC × 0.75	20-51	33	$37,051.31
Progression to Liver Transplant × 0.75	20-51	33	$36,443.38
Treatment failure rate × 0.75	17-61	33	$26,368.78
Probability of Infection × 0.4	empty	33	$50,899.45
Drop out rate × 2	22-46	33	$41,369.50

CHAPTER V

BIAS IN MARKOV MODELS OF DISEASES

5.1 Introduction

Markov models are commonly used to simulate diseases when evaluating various medical interventions. Modeling diseases allows us to consider long term consequences and other implications not practical via clinical trials. Examples include an analysis of cost-effectiveness of expanded HIV screening in the US[108, 84], an analysis of the optimal age of vaccination[6], the optimal timing of liver transplantation[2, 3], and an analysis of dynamic multi-drug therapies for HIV[133]. Markov models are also used to forecast and estimate future morbidity, mortality, prevalence, and costs of various diseases[138]. Markov models of diseases are often coupled with clinical trial data to determine the cost-effectiveness of medical interventions. Given the limited amount of resources available, Markov models of diseases can potentially be a relatively inexpensive, yet powerful tool to evaluate medical interventions.

However, due to limited disease data or model complexity, simplifying assumptions are made in many Markov disease models. In this Chapter, we analyze the effect of a common simplification, namely, that of modeling diseases that have nonlinear progression with Markov models that assume constant progression. The assumption of constant disease progression is used in many disease studies [4, 5, 35, 47, 51, 57, 64, 106, 114]. Studies often assume constant disease progression when there is insufficient patient data to characterize non-linearities or changes over time. Constant progression is also often assumed in order to reduce the model's complexity, particularly when deriving analytical results. In Chapter 3, for example, we used a *reduced* the state space model of Hepatitis C for analytical tractability. Consequently, the model did

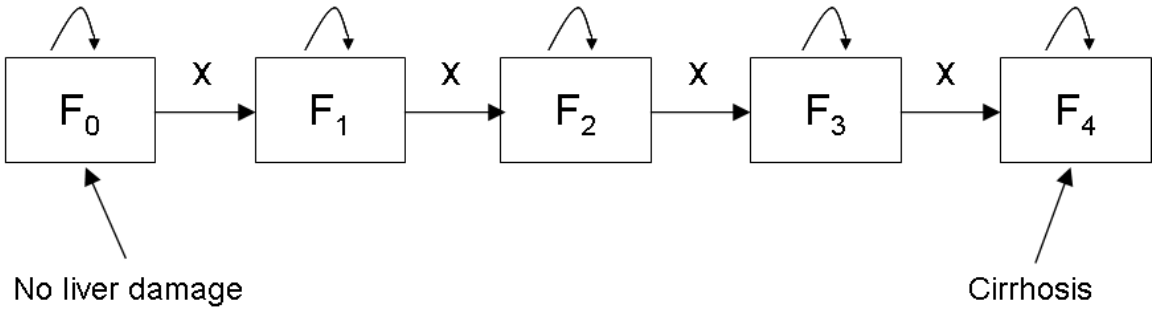


Figure 10: METAVIR standard for liver disease progression.

not allow for disease progression to change as a function of disease severity. The reduced state space was necessary to construct a dynamic policy of testing for the disease.

When constant disease progression is assumed, caution must be exercised in order to ensure that the disease progression is not overestimated or underestimated, which we call *bias*. This Chapter explores bias in Markov models of disease when constant disease progression is assumed. We consider diseases in which the progression (i.e., transition probabilities) depends on the severity of the disease (i.e., state) and diseases in which progression varies with time (i.e., age, time spent in a state). In both cases, we use Markov models and compare the use of state/time dependent transition probabilities with the use of transition probabilities that assume constant progression by aggregating data. We make such a comparison in order to determine if, and under what conditions, Markov models that assume linear progression underestimate or overestimate disease progression. We then use medical data to we assess the magnitude of the bias for Hepatitis C (HCV), Alzheimer's disease (AD), and lung cancer.

5.2 Markov Models of Various Diseases

For Markov disease progression models that have state dependent transition probabilities, we consider models of the type in Figure 10 and compare it to the case where the

probability transitions between states are not identical. Figure 10 is a standard disease progression model used for liver disease that uses the METAVIR standard for liver disease, which is the motivating disease for this analysis [57, 107, 114]. METAVIR is a scoring system specifically designed for HCV patients in which the scores F0-F4 represent different degrees of liver fibrosis. F0 represents no liver scarring and F4 represent cirrhosis or advanced liver scarring; the states in between represent intermediate levels of liver damage. Patients are determined to be in one of the five METAVIR states via a liver biopsy. Each state in Figure 10 corresponds to a score on the METAVIR system, which are ordered by increasing liver damage. The x values in the figure are determined by diving the time since infection by the number of states progressed during that time, and averaging over many patients. For example, if a patient is determined to be in state F3 (as determined by a liver biopsy) with an infection length of 30 years, the rate of liver disease progression is thus $3/30 = 0.1$, which is then assumed to be the constant rate of progression between all states. Consequently, the transition probabilities are identical by construction. This method of arriving at a single aggregate transition probability is referred to as the *indirect* method and is used in the majority of studies analyzing progression rates of HCV [97, 145, 146]. We will refer to the value computed using the indirect method as the indirect value.

When few patients in a data set have serial biopsies (i.e., biopsies at different points in time for the same patient), the modeler has little choice but to use the indirect method to estimate transition probabilities and implicitly assume constant disease progression. However, when patient data with serial biopsies is available, the modeler can potentially estimate the transition probabilities between METAVIR states so that the disease progression is not assumed to be constant. This Chapter provides insight into the value of having richer disease data (i.e., serial biopsies)

Studies involving liver disease progression typically assume a constant progression

of liver disease [51, 57, 114] as described above. Few studies allow for state dependent liver disease progression such as [7]. Studies have shown, however, that liver disease progression is not constant and varies significantly between METAVIR states [71, 98, 113, 145]. Matsumura [71] and Yi, et al, [145] computed estimates for the transition probabilities between the METAVIR states using patient data with serial biopsies. Both studies also compute the x value using the indirect method. No studies have analyzed, however, the consequences of *using* the single transition probability value that assumes constant progression versus using the transition probabilities that vary between states. Our study analyzes this question.

Many diseases are modeled by similar Markov chains to that of Figure 10, particulary those where the disease states represent scores on an ordinal rating system such as the METAVIR standard. These diseases range from Alzheimer's disease (AD) to Glaucoma to Amyotrophic Lateral Sclerosis (ALS, also referred to as Lou Gehrig's Disease). These diseases are progressive, sometimes degenerative, diseases that may or may not be fatal. The severity of AD, for example, is commonly ranked using the Mini-Mental State Examination (MMSE) which is based on a scale of 0-30. The Alzheimer's Disease Assessment Scale and the Blessed Information Memory Concentration are other common, ordered rating systems for Alzheimer's disease. In practice, the states are typically grouped to represent states such as "healthy," "mild," "moderate," and "severe."

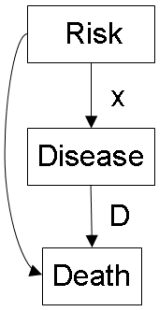
The majority of studies of AD assume a constant progression rate and use the same single transition value between states of different severity [5, 35, 47, 64, 106]. Studies that analyze the effect of different medical interventions on AD typically measure the degree to which the intervention decreases the constant progression rate. Jonsson [59], for example, determined that treatment with donepezil 5 mg corresponds to multiplying the transition probability to a state with a lower MMSE score by (1 - 0.4636) and by (1 - 0.4807) for treatment with donepezil 10 mg. Studies have shown,

however, that the progression is far from constant and is typically slower in less severe states [52, 119, 121]. Accordingly, some studies *do* use state dependent transition probabilities [120, 121]. Like in the case of HCV, studies have not considered the consequences of *using* the single transition probability value that assumes constant progression versus using the transition probabilities that vary between states.

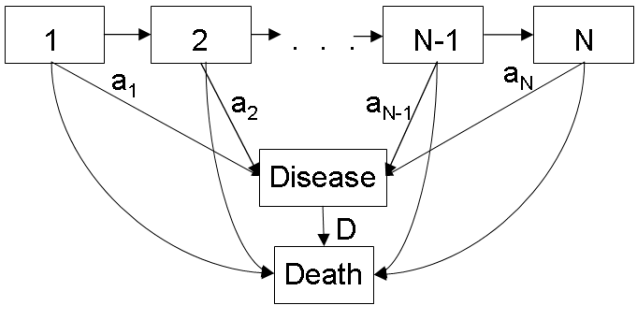
Another example of a disease that can be modeled by a Markov chain similar to Figure 10 is the progression of ALS, which is often scored on a scale of 1-5 (Mild, Moderate, Severe, Terminal, Death) [124]. Progression is often modeled as constant over time [4]. Similar to the previous examples, studies have shown, however, that ALS has nonlinear progression [124].

For Markov disease progression models that have time dependent transition probabilities, we consider models of the type in Figures 11(a) and 11(b). In this case, we can interpret states 1-N in Figure 11(b) as ages or as time spent in a risk state. The former is appropriate to model diseases that progress faster with age such as HCV or AD; the latter is appropriate to model diseases such as lung cancer whose risk of occurrence increases with the number of years of smoking. The model in Figure 11(b) allows for the transitions to vary with time since the $a_i's$ can each have different values, whereas the model in Figure 11(a) does not capture the time dependency and is equivalent to constraining the $a_i's$ to be identical in Figure 11(b). Scherrer [109] began a framework for analyzing bias in models with time dependent transition probabilities and compared the two models in the case of two risk states plus a disease state for two time periods where the disease state is an absorbing state.

Scherrer showed that the single transition probability of moving from the risk state to the disease state obtained by averaging the true transition probabilities of many risk states underestimates the likelihood of ending up in the disease state (when the probability of transitioning to the disease state increases with time spent in the risk state, and the entire population begins in the risk state). Scherrer also showed that



(a) constant transition probabilities

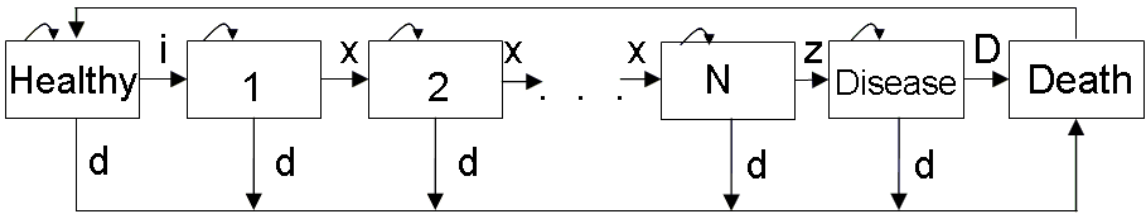


(b) time dependent transition probabilities

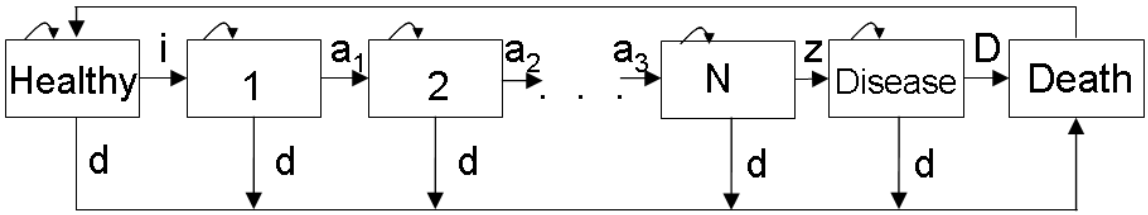
Figure 11: Markov models for a disease with time dependent transition probabilities.

the result also holds when the population is initially evenly split between the two risk states. We modify the Markov model in Scherrer (while preserving the characteristics that cause bias) so that we can solve for a stationary distribution for n risk states and arrive at conditions for bias that holds for the Markov model in steady state (i.e., independent of the initial population distribution).

Markov models with time dependent transition probabilities of the type in Figure 11(b) are appropriate for many other diseases where the transitions vary over time or by age, including HCV, AD, lung cancer, cardiovascular disease and diabetes. In the case of HCV, many studies utilize transition probabilities that vary by age



(a) transition probabilities are constant



(b) transition probabilities are state dependent

Figure 12: Markov model for a disease with state dependent transition probabilities as used in Monte Carlo simulations.

[57, 96, 123], whereas many others only use a single transition probability that applies for all ages [114, 118] thereby not capturing the large differences in progression by age. Our analysis studies the difference in the expected time to progress through the METAVIR states of the HCV model using the single transition probability versus using age dependent transition probabilities.

This model can also be used to capture incidence rates of diseases that vary with time. For example, incidence of AD increases with age [93]; in this case, each age of person's life can represent a Markov state and can have a different transition rate to the disease state. Similarly, since the probability of acquiring lung cancer increases with the number of years of smoking [127], the Markov states can represent the number of years of smoking, each with different transition rates to lung cancer.

5.3 Monte Carlo Simulations of Markov Models of Diseases

A common method of analyzing Markov chains of disease progression in the medical literature is to employ the use of Monte Carlo simulations [84, 108, 114]. In these

models, each individual patient's clinical course is followed from the time of entry into the model until death or the final state of interest is reached. A running tally is kept of the amount of time spent in each health state including the associated costs and quality adjusted life years (which measures the benefits gained). Upon the patient's death (or arrival in the final state), the next patient is introduced into the model. The process repeats over a large number of patients and the results are averaged over all patients. Figure 12(a) graphically represents the scenario where we require the transition probabilities for the disease progression to be identical in a Monte Carlo simulation, and we have generalized the number of states and introduced additional states before and after the disease progression states. In modeling HCV, for example, state "Healthy" represents uninfected, states "F0"-"F4" represent the METAVIR states, and the final state before death represents disease complications including HCC and decompensation. Figure 12(b) shows a disease progression with state dependent transition probabilities.

We are interested in comparing the model shown in Figure 12(a) with the model shown in Figure 12(b). When the final state is "Death" then the direct transitions to it (from states "Healthy - "N") represent death from causes other than the disease being modeled. It is worth noting that the final state need not be "Death." When the final state represents a diseased state (perhaps a non-fatal disease), the direct transitions to it represent the probability of acquiring the disease for reasons other than those captured by the model's states. In the case of time-dependent transition probabilities, we are interested in comparing the model in Figures 11(b) and 11(a) where we also assume that the "death" states feeds back into state 1 as it would in a Monte Carlo simulation. Similarly, "Death" need not be considered the final state.

The remainder of the chapter is organized as follows. We first derive at stationary distributions for models where the progression changes with state or with time. Then we use the stationary distribution to arrive at conditions for bias. We then discuss

factors that affect the degree of the bias. We apply the results to HCV, AD, and lung cancer using medical data. Finally, we discuss the implications of the results when Markov models of disease progression are used in studies of cost-benefit analysis, forecasting future prevalence, and estimating future disease burden.

5.4 Stationary Distributions

5.4.1 State dependent transition probabilities

In this section we solve for the stationary distributions for the various models we are considering. The stationary distribution of a Markov chain describes the steady state behavior of the Markov chain, which allows us to draw conclusions about Markov models in the long term. The stationary distribution of a Markov chain does not depend on the initial state of the chain. In the case of modeling diseases, this means that we do not need to assume an initial health state in our model; the results derived from the stationary distribution will hold regardless of the initial health state or initial distribution of health states (in the case of modeling populations). We will use the stationary distributions derived in this section in our analysis of model bias in Section 5.5

The Markov models in Figure 12 are irreducible aperiodic Markov chains where all of the states are positive recurrent (i.e., they are ergodic). If we add a direct path from "death" to state 1 in the models in Figure 11 (as would be the case in Monte Carlo simulations), then they are also ergodic. By Theorem 4.3.3 in Ross [104], there exists a unique stationary distribution π for each of those ergodic Markov chains.. Moreover, by Theorem 4.3.1 of Ross, the expected return time to state i (once the state is entered) is given by $\frac{1}{\pi_i}$. The death state is included for illustrative purposes and the time spent in it is always one time period. Keeping the death state in the model implies that $\frac{1}{\pi_{Death}} - 1$ represents the expected time from state "Healthy" to state "Death". It should be clear that solving the stationary distribution of the Markov

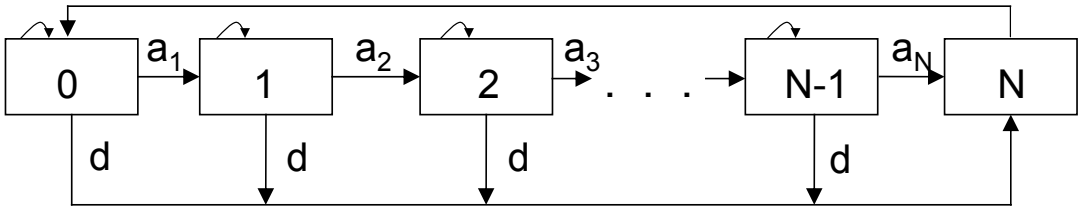


Figure 13: Generalized Markov model of disease progression with state dependent transition probabilities

chain in Figure 12(b) is equivalent to solving the stationary distribution of Figure 13, only differing in the labels given to the transition probabilities. Without loss of generality, we use the Markov chain in Figure 13 in our analysis for the remainder of this chapter. That is, the results throughout this chapter assume the comparison of Figure 13 to the corresponding single probability transition Markov chain (i.e., where $a_i = x \ \forall \ i$). We subsequently prove, however, that all of our results remain true in the more general case of comparing Figures 12(a) and 12(b), even when we add additional states before state 1 and after state N (as long as they do not directly communicate with any of the states $1 - N$ as shown in the Figures).

We begin with Figure 13 where $d = 0$ (i.e., there are no direct transitions to the final state). In this case the analysis is simple. Solving

$$\begin{bmatrix}\pi_0 & \pi_1 & \cdots & \pi_N\end{bmatrix} = \tag{16}$$
$$\begin{bmatrix}\pi_0 & \pi_1 & \cdots & \pi_N\end{bmatrix} \begin{bmatrix} 1-a_1 & a_1 & 0 & 0 & \cdots & 0 \\ 0 & 1-a_2 & a_2 & 0 & \cdots & 0 \\ 0 & 0 & 1-a_3 & a_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1-a_N & a_N \\ 1 & 0 & 0 & 0 & \cdots & 0 \end{bmatrix}$$

for π gives us the stationary distribution. It can easily be shown that

$$\pi_N = \frac{1}{1 + \sum_{1 \leq j \leq N} \frac{1}{a_j}}. \tag{17}$$

The proof of the solution to Equation 16 is a special case (where $d = 0$) of Proposition 1, which is proved in the Appendix. In the case that $a_i = x \ \forall \ i$, then $\pi_N = \frac{1}{1+\frac{N}{x}}$.

Next we consider Figure 13 where $d > 0$ (i.e., there are direct transitions to the final state). In this case the analysis is more complicated. Solving

$$\begin{bmatrix}\pi_0 & \pi_1 & \cdots & \pi_N\end{bmatrix} = \tag{18}$$
$$\begin{bmatrix}\pi_0 & \pi_1 & \cdots & \pi_N\end{bmatrix} \begin{bmatrix} 1-a_1-d & a_1 & 0 & 0 & \cdots & d \\ 0 & 1-a_2-d & a_2 & 0 & \cdots & d \\ 0 & 0 & 1-a_3-d & a_3 & \cdots & d \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1-a_N-d & a_N+d \\ 1 & 0 & 0 & 0 & \cdots & 0 \end{bmatrix}$$

for π gives us the stationary distribution.

Proposition 1. *The stationary distribution to the Markov chain in Figure 13 (i.e.,*

the solution to Equation 18) is

$$\pi_N = \frac{1}{1 + \sum_{1 \le j \le N} \frac{1}{a_j} \prod_{1 \le i \le j} (\frac{a_i}{a_i + d})}. \tag{19}$$

The proof of Proposition 1 is in the Appendix. Consequently, the expected time from state 0 to N is equal to

$$E_a = \sum_{1 \le j \le N} \frac{1}{a_j} \prod_{1 \le i \le j} (\frac{a_i}{a_i + d}). \tag{20}$$

It should be clear that if we consider the case where the $a_i = x \ \forall \ i$ as it is in the METAVIR standard and in Figure 13, then the expected time from state 0 to N is equal to

$$E_x = \sum_{1 \le j \le N} \frac{1}{x} \prod_{1 \le i \le j} (\frac{x}{x + d}). \tag{21}$$

5.4.2 Time dependent transition probabilities

Next we solve for the stationary distribution to the Markov chain in Figure 11(b). In this case, we must solve

$$\begin{bmatrix} \pi_1 & \pi_2 & \cdots & \pi_N & \pi_C & \pi_D \end{bmatrix} = \tag{22}$$

$$\begin{bmatrix} \pi_1 & \pi_2 & \cdots & \pi_N & \pi_C & \pi_D \end{bmatrix} \begin{bmatrix} 0 & 1 - a_1 - d & 0 & 0 & \cdots & a_1 & d_1 \\ 0 & 0 & 1 - a_2 - d & 0 & \cdots & a_2 & d_2 \\ 0 & 0 & 0 & 1 - a_3 - d & \cdots & a_3 & d_3 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \\ 0 & 0 & 0 & \cdots & 0 & a_N & d_N \\ 0 & 0 & 0 & \cdots & 0 & 1 - D & D \\ 1 & 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix}$$

for π.

Proposition 2. *The stationary distribution to the Markov chain in Figure 11(b) (i.e., the solution to Equation 22) is*

$$\pi_D = \frac{1}{1 + \sum_{1 \le j \le N} (1 + \frac{a_j}{D}) \prod_{1 \le i \le j-1} (1 - a_i - d_i)}. \quad (23)$$

The proof of Proposition 2 is in the Appendix. Consequently, the expected time from state 1 to Death is equal to

$$E_a = \sum_{1 \le j \le N} (1 + \frac{a_j}{D}) \prod_{1 \le i \le j-1} (1 - a_i - d_i). \quad (24)$$

Figure 11(a) is a special case (N = 3) of Figure 12(b). Consequently, we can use the corresponding development to arrive at the expected time from state 1 to Death as

$$E_x = \frac{x + d + D}{(x + d)(d + D)}. \quad (25)$$

It should also be clear that if we consider the case where the $a_i = x \; \forall \; i$ where the $d_i's$ may vary, then the expected time from state 1 to Death is equal to

$$E_x = \sum_{1 \le j \le N} (1 + \frac{x}{D}) \prod_{1 \le i \le j-1} (1 - x - d_i). \quad (26)$$

5.5 Analysis of Bias

In this section, we analyze bias in the models described above. In our analysis, we take the model with state/time dependent transition probabilities to be the *status quo*, and define bias as the change in the expected time it takes to reach the final state from the initial state when using constant transition probabilities (in place of the state/time dependent transition probabilities). The proofs of the theorems below are in the Appendix to this chapter.

5.5.1 Conditions for the existence of model bias

5.5.1.1 State dependent transition probabilities

For the special case where there are no direct transitions to the final state (i.e., $d = 0$) in the model in Figure 13, it follows from Equation 17 that the bias, $B_{a,x}$, resulting

from the use of a single probability transition, x, instead of the state dependent probability transitions a_i, is equal to

$$B_{a,x} = E_a - E_x = \frac{1}{1 + \sum_{1 \leq j \leq N} \frac{1}{a_j}} - \frac{1}{1 + \frac{N}{x}}.$$

The model with the single transition probabilities overestimates (underestimates) the disease progression, therefore, if and only if $x > (<) \frac{N}{\sum_{1 \leq j \leq N} \frac{1}{a_j}}$ which is the harmonic mean of the $a_i's$.

We turn our attention to the case when there are direct transitions to the final state (i.e., $d > 0$), which represent death from natural causes when the final state represents death (or an alternative means of obtaining disease when the final state represents a particular disease state). In this case, the bias condition is more complicated. We are interested in comparing the models in Figures 12(a) and 12(b) to analyze the consequences of using a single transition probability. With the framework introduced in the previous section, we can arrive at sufficient conditions for bias.

We start by subtracting Equation 21 from 20 to arrive at the expression for difference in the expected time from state 0 to N when using a single transition probability versus allowing the transition probabilities to vary by state, which gives us the following Theorem.

Theorem 3. *For a disease progression of the type in Figure 13, the bias resulting from the use of a single transition probability (i.e., $a_i = x \ \forall \ i$) versus allowing the transition probabilities to vary is*

$$B_{a,x} = \sum_{1 \leq j \leq N} \frac{1}{x} \prod_{1 \leq i \leq j} \left(\frac{x}{x+d}\right) - \sum_{1 \leq j \leq N} \frac{1}{a_j} \prod_{1 \leq i \leq j} \left(\frac{a_i}{a_i + d}\right) \tag{27}$$

where we measure bias by the difference in expected time from state 0 to state N.

Theorem 3 indicates the magnitude and direction of the bias. When x is calibrated such that the expression in 27 is zero, then there is no bias. That is, the expected time

from state 0 to state N is the same when using the state dependent probabilities or that particular value of x. Using any other value for x results in bias in the direction indicated by Theorem 3. Next, we describe sufficient conditions on the value of x for the presence of bias.

Theorem 4. *For a disease progression of the type in Figures 12(a) and 12(b), when x is less than or equal to the harmonic mean of the $a_i's$, and the $a_i's$ do not all equal x, then the model with aggregated transition probabilities (Figure 12(a)) strictly underestimates disease progression when compared to the model with state dependent transition probabilities (Figure 12(b)).*

The proof is in the Appendix. Theorem 4 states that the expected time from state 0 to N is larger when the harmonic mean of the state dependent transition probabilities is used as a constant transition probability than when the state dependent transition probabilities themselves are used. That is, model *bias* is strictly negative. The result holds for any value less or equal to than the harmonic mean as well. It follows from Theorem 4 that when a value smaller than the harmonic mean for x is used, as the value of x decreases the degree of underestimation (i.e., bias) increases. This is true since as x decreases, the time from state 0 to N increases, which is apparent by examining Figure 13. The importance of Theorem 4 is that it shows that when modeling a disease whose progression is not constant, using the harmonic mean of the state dependent transition probabilities (instead of the state dependent transition probabilities themselves) will cause the modeler to underestimate the progression of the disease. When using a value smaller than the harmonic mean, the modeler will increasingly underestimate the progression of the disease.

Corollary 1. *Consider a disease progression of the type in Figures 12(a) and 12(b) where the $a_i's$ are increasing. When x is computed using the indirect method, the model with aggregated transition probabilities (Figure 12(a)) strictly underestimates disease*

progression when compared to the model with state dependent transition probabilities (Figure 12(b)).

Corollary 1 is a special case of Theorem 4. It is important to note than when using the indirect method as described in section 5.2, the modeler can average the computed indirect value of many patients by taking the arithmetic mean of the indirect values, or more correctly, by taking the harmonic mean of the indirect values. The choice of which mean to use is left to the modeler, however, the harmonic mean is more appropriate when averaging rates. When averaging the expected times between states, the arithmetic mean is more appropriate. Indeed, the harmonic mean of the progression rates (indirect values) is equivalent to the arithmetic mean of the expected times between states. The proof of Corollary 1 in the Appendix discusses the use of the arithmetic and harmonic means further. In Corollary 1, we assume x is computed by taking the harmonic mean of the indirect values of many patients (or equivalently, the arithmetic mean of the expected times between states). Corollary 2 discusses the case where we use the arithmetic mean of the indirect values. The importance of Corollary 1 is that it suggests that studies that use the indirect method (the most common method of computing constant transition probabilities) due to lack of disease data are underestimating the disease progression for diseases such as HCV and AD that are believed to have increasing progression rates (i.e., the sick get sicker faster).

Note that the converses of Theorem 4 and Corollary 1 do not follow, which can be seen by a number of counter examples. We can, however, arrive at a result for model bias in the opposite direction (overestimation) using the development in the proof of Theorem 4.

Theorem 5. *For a disease progression of the type in Figures 12(a) and 12(b), when x is greater than or equal to the geometric (or arithmetic) mean of the $a_i's$, and the $a_i's$ do not all equal x, then the model with aggregated transition probabilities (Figure*

12(a)) strictly overestimates disease progression when compared to the model with state dependent transition probabilities (Figure 12(b)).

The proof is in the Appendix. Theorem 5 states that the expected time from state 0 to N is smaller when the geometric mean of the state dependent transition probabilities is used as a constant transition probability than when the state dependent transition probabilities themselves are used. That is, model *bias* is strictly positive. The result holds for any value greater than or equal to the geometric mean, including the arithmetic mean. It follows from Theorem 5 that when a value larger than the geometric mean for x is used (e.g., the arithmetic mean), as the value of x increases the degree of overestimation (i.e., bias) increases. This is true since as x increases, the time from state 0 to N decreases. The significance of Theorem 5 is that when modeling a disease whose progression is not constant, using the geometric mean of the state dependent transition probabilities (instead of the state dependent transition probabilities themselves) will cause the modeler to overestimate the progression of the disease. When using a value larger than the geometric mean, the modeler will increasingly overestimate the progression of the disease.

Corollary 2. *Consider a disease progression of the type in Figures 12(a) and 12(b) where the $a_i's$ are decreasing. When x equals the arithmetic mean of the computed indirect values for each patient, the model with aggregated transition probabilities (Figure 12(a)) strictly underestimates disease progression when compared to the model with state dependent transition probabilities (Figure 12(b)).*

Corollary 22 is a special case of Theorem 5. Contrary to Corollary 1, we assume x is computed by taking the *arithmetic* mean of the computed indirect values of the patients. The importance of Corollary 2 is that it suggests that studies that use the indirect method in this way are overestimating the disease progression for diseases that have decreasing progression.

5.5.1.2 *Time dependent transition probabilities*

Next, we consider Markov models with time dependent transition probabilities. Subtracting Equation 26 from 24 gives us the expression for the difference in the expected time from state 1 to state to Death when using a single transition probability versus allowing the transition probabilities to vary over time. Solving for x gives us a condition for bias, which we formulate as our next theorem.

Theorem 6. *A Markov model without time varying transition probabilities of the type in Figure 11(a) overestimates (underestimates) the expected time to death when compared to a model with time dependent transition probabilities (Figure 11(b)) when*

$$x > (<)\frac{E_a d - 1}{\gamma - E_a} \tag{28}$$

where E_a is defined above and $\gamma = \frac{1}{d+D}$

Theorem 6 indicates that when x equals the right hand side of inequality 28 the bias is zero. That is, the expected time from state 1 to death is the same when using the time dependent transition probabilities or that particular value of x. Using any other value for x will result in bias in the direction indicated by Theorem 6.

The sufficiency theorems that are true for the state dependent transition probabilities model are not true for the time dependent transition probabilities model. They can be shown to be false by simple numerical counter examples, which we show in Section 5.6

5.5.2 Factors that affect the degree of model bias

Thus far we have compared models where the transition probabilities are allowed to vary between states with models that use a single aggregate transition probability between several states. In this section, we consider intermediate levels of data aggregation. That is, instead of considering only the use of a single transition probability, we combine transition probabilities in groups as shown in Figure 14, where b_1

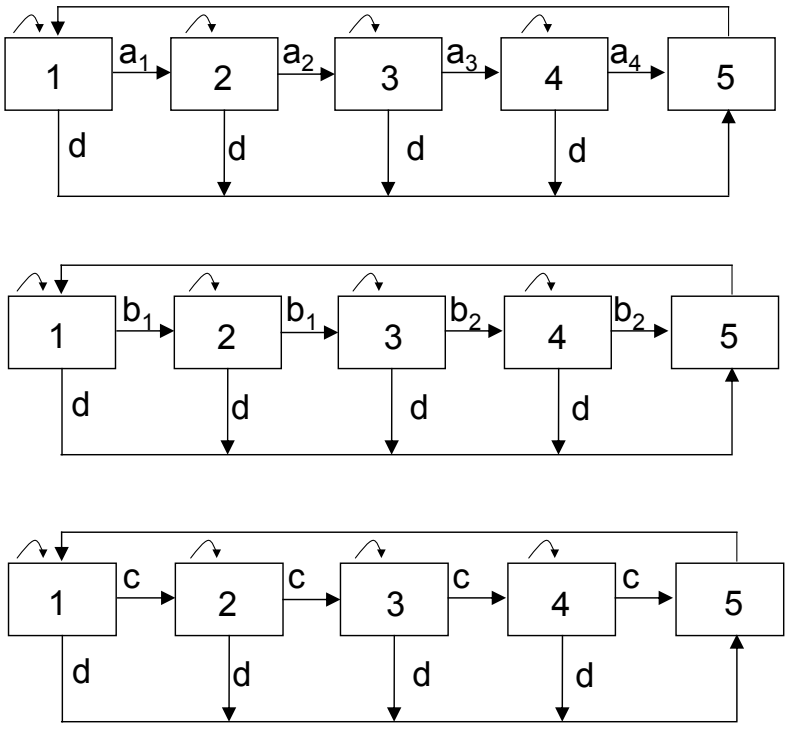


Figure 14: Markov models of disease with different degrees of aggregation

is formed by taking some mean of a_1 and a_2 only. By comparing the three models in Figure 14 we will be able to analyze the effect of the *degree* of data aggregation on model bias. The degree of data aggregation is increasing as we switch from using the $a_i's$ to using the $b_i's$ to using c as the transition probabilities for the Markov chain.

Theorem 7. *When using any value greater than or equal to the geometric mean (e.g., the arithmetic mean) of the transition probabilities in Figure 14, the models with aggregated transition probabilities increasingly overestimate disease progression as the degree of aggregation increases.*

It is worth recalling that the value computed using the indirect method (and averaging over many patients using the arithmetic mean) is greater than the arithmetic mean of the $a_i's$ for diseases that have decreasing progression. For this computed value, Theorem 7 applies for diseases that have decreasing progression.

Theorem 8. *When using any value less than or equal to the harmonic mean of the transition probabilities in Figure 14, the models with aggregated transition probabilities increasingly underestimate disease progression as the degree of aggregation increases.*

Recall that the value computed using the indirect method (and averaging over many patients using the harmonic mean) is less than the harmonic mean of the $a_i's$ for diseases that have decreasing progression. For this computed value, Theorem 7 applies for diseases that have decreasing progression. Theorems 7 and 8 state that bias increases as the degree of data aggregation increases, and that the increased bias is in the same direction as determined by Theorems 4 and 5. They are both proven in the Appendix to this Chapter.

5.6 Bias in Models of Hepatitis C, Alzheimer's Disease and Lung Cancer

In this section we use medical data to calculate the bias in models of three diseases introduced in Section 5.2. We start with models whose transition probabilities vary by state and use HCV and AD as examples. Disease progressions of HCV and AD were introduced in Section 5.2, and HCV with its corresponding disease model was extensively discussed in Chapter 4. We also use HCV, AD and lung cancer to show bias in models with time dependent transitions.

5.6.1 Bias in diseases with state dependent transition probabilities: Hepatitis C and Alzheimer's Disease

In this section we use HCV and AD as examples for bias in Markov models of diseases with state dependent transition probabilities when we aggregate the data and assume constant transition probabilities. In the case of HCV, the progression through the METAVIR states has been shown to vary by liver disease severity [71, 145]. In this case, model states represent METAVIR states where the transitions between states vary as shown in Table 15, where a_i is the transition probability of going from state i to

state $i+1$. We consider several different data sets for the same disease. The transition probability labeled "indirect" corresponds to the single transition probability obtained using the indirect method introduced in Section 5.2 where a constant linear disease progression is assumed. The indirect value reported in each case was obtained by the same corresponding study that computed the $a_i's$ in order to ensure that we do not introduce *ab extra* bias by using parameters derived from different data sets.

In the case of AD, the progression has also been shown to vary significantly by severity [119, 121]. The Markov model states in this case represent the grouped MMSE scores (see section 5.2 for an introduction) of an individual with AD, where the probability of transitioning between states varies according to Table 15.

There are very few studies that compute $a_i's$ for HCV, while several studies have computed them for AD. We chose studies that arrived at very different values for the $a_i's$ (due to studying different populations) so that we can test for bias under very different scenarios. For example, the $a_i's$ for Yi 1 pertain to 1138 chronic HCV patients in liver clinics who have disproportionately faster progression than those in Yi 2, which is made up of previously healthy women infected by exposure to contaminated anti-D immune globulin who were then screened for HCV. The data from Matsumura was obtained from Japanese patients infected with HCV that have chronic liver disease. Similarly, while the Stern and Suh data sets are on different scales, they also report very different progressions based on the populations considered.

Table 15: Parameter values for transition probabilities

	HCV			**AD**	
	Yi 1	**Yi 2**	**Matsumura**	**Stern**	**Suh**
$\boldsymbol{a_1}$	0.169	0.042	0.049	0.176	0.097
$\boldsymbol{a_2}$	0.118	0.045	0.217	0.588	0.070
$\boldsymbol{a_3}$	0.225	0.097	0.556	0.765	0.053
$\boldsymbol{a_4}$	0.207	0.070	0.294	0.412	
indirect	0.150	0.045	0.120	0.562	0.077
Source	Yi [145]	Yi [145]	Matsumura [71]	Stern [119]	Suh [121]

Table 16 displays the results for comparing the use of a single transition probability versus using the $a_i's$ for HCV and AD. The expected time to death is displayed using the indirect value as well as the harmonic (harm), geometric (geo) and arithmetic (arith) means, respectively. We use the death rates from the CDC [23]. For each data set, we compute the percent change ($\%\triangle$) relative to using the $a_i's$. Note that the indirect method results in greater bias in all cases but one (the arithmetic mean using data from Matsumura for HCV). In all cases, the harmonic mean results in less bias than any other single transition probability considered.

The arithmetic mean consistently results in the largest bias among the three means. We also compare the use of the $a_i's$ to the use of two probability transition values as in Figure 14 where we use the three different means. Using two probability transition values represents a lower degree of data aggregation than using only one. The direction of the bias remains constant, while the magnitude decreases as predicted by Theorems 7 and 8. In most cases, the bias decreases significantly between using a single transition probability versus two. It is also interesting that for the data sets where the $a_i's$ vary a great deal, the bias is much larger.

5.6.2 Bias in diseases with time/age dependent transition probabilities: Hepatitis C and Alzheimer's Disease and Lung Cancer

In this section we use HCV, AD and lung cancer as examples for bias in Markov models of diseases with transition probabilities that vary by time/age. In the case of HCV, the progression through the METAVIR states have been shown to vary by age [57, 96, 123]. The Markov model states in this case represent ages of an individual (while infected with HCV), where the transition to the disease state (cirrhosis) vary by age as shown in Table 17. In the case of AD, the risk of acquiring AD has been shown to increase with age [93]. The Markov model states in this case represent the ages of an individual without AD, where the probability of transitioning to the disease state (i.e., acquiring AD) varies by age according to Table 18. In the case of lung

Table 16: Results for HCV and AD with state dependent transition probabilities for multiple degrees of data aggregation

		1 transition probability				2 transition probabilities		
HCV								
	$a_i's$	**indirect**	**harm**	**geo**	**arith**	**harm**	**geo**	**arith**
Yi 1	18.57	20.47	18.61	18.13	17.71	18.58	18.44	18.29
%△		10.22%	0.23%	-2.36%	-4.66%	0.08%	-0.72%	-1.51%
Yi 2	39.16	44.06	39.38	38.21	36.96	39.18	39.07	38.95
%△		12.51%	0.57%	-2.42%	-5.62%	0.05%	-0.24%	-0.53%
Mat	21.72	24.27	22.54	15.93	12.18	22.16	18.80	15.90
%△		11.71%	3.73%	-26.69%	-43.91%	1.99%	-13.46%	-26.83%
AD								
Stern	9.63	6.43	9.70	8.34	7.37	9.68	8.62	7.69
%△		-33.25%	0.69%	-13.46%	-23.44%	0.49%	-10.49%	-20.14%
Suh	28.19	26.37	28.29	27.74	27.19	28.21	28.07	27.93
%△		-6.47%	0.34%	-1.61%	-3.56%	0.08%	-0.43%	-0.93%

cancer, we use the model in Figure 11(b) where each state represents the number of years an individual has smoked cigarettes. In this case we use the function

$$p(n) = 1.845 \times 10^{-10} n^{4.5} \tag{29}$$

as the probability of transitioning to lung cancer after smoking for n years, which is obtained from Scherrer's study [109]. The data used by Scherrer to arrive at this function was originally based on Doll and Peto [41] for individuals who smoke 20 cigarettes per day beginning at age 16.

Table 17: Liver Fibrosis estimates for HCV model. Source [57]

	Annual Fibrosis Progression by age (a_i)
≤30	0.0125
31-40	0.0225
41-50	0.0135
51-60	0.03125
61-70	0.05525
70+	0.07525

Table 18: Incidence of AD. Source [93]

	AD Incidence by age (a_i)
60-64	0.06
65-69	0.19
70-74	0.51
75-79	1.17
80-84	2.31
85-89	3.86
90-94	5.49
95+	6.68

Table 19 displays the results for comparing the use of a single transition probability versus using the $a_i's$ for HCV and AD. For HCV, we assume the individual acquires HCV at age 20 and we compute the expected time until death from age 20 onwards using the death rates from the CDC from Chapter 4. For AD, we consider the incidence starting at age 60 (since the incidence before age 60 is very small) and compute the expected time to death from age 60 onwards using the same death rates. We compare the use of the harmonic, geometric and arithmetic means. Note that the results serve as counterexamples to show that Theorems 4 and 5 do not hold for models with time dependent transition probabilities as they do for models with state dependent transition probabilities. The harmonic mean results in smaller bias for both cases, however this is not necessarily always the case.

Table 20 displays the results for comparing the use of a single transition probability versus using the $a_i's$ for lung cancer. We assume the individual begins smoking 20 cigarettes per day starting at age 15 and use same death rates from the CDC as in the previous table. We compare the use of the harmonic, geometric and arithmetic means. Note that the harmonic mean does not always result in smaller bias. We also show the results for different levels of data aggregation. We compare the use of the $a_i's$ to the use of two, three and four probability transition values in the way described in Figure 14 where we use the three different means. Note that the direction of the

Table 19: Results for HCV and AD with time dependent transition probabilities

	HCV, acquired at age 20			
	$a_i's$	**arithmetic**	**geometric**	**harmonic**
Expected time to death from age 20	33.91	25.16	28.89	32.32
%△		-25.79%	-14.82%	-4.70%
	AD incidence starting at age 60			
Expected time to death from age 60	20.75	16.52	19.47	21.35
%△		-20.35%	-6.14%	2.89%

bias does not always remain constant. However, the bias does decrease significantly when decreasing the degree of data aggregation. The results in Tables 19 and 20 are not sensitive to the death rates since they are driven by the $a_i's$.

Table 20: Results for Lung Cancer with time dependent transition probabilities for multiple degrees of data aggregation (smoking 20 cigarettes/day since 15 years old)

		1 transition probability			**2 transition probabilities**		
	$a_i's$	**arith**	**geo**	**harm**	**arith**	**geo**	**harm**
Expected time to death from age 15	59.98	45.64	61.47	62.97	57.67	61.46	59.87
%△		-23.92%	2.47%	4.97%	-3.86%	2.46%	-0.19%
		3 transition probabilities			**4 transition probabilities**		
		arith	**geo**	**harm**	**arith**	**geo**	**harm**
Expected time to death from age 15		59.30	59.74	60.12	59.82	59.92	60.02
%△		-1.13%	-0.40%	0.23%	-0.28%	-0.10%	0.07%

5.7 Discussion

The results in this Chapter suggest that efforts should be made to obtain/use disease data that allow for estimating transition probabilities that are state and/or time dependent in order to minimize model bias. When constant transition probabilities must be used due to limited data, or for analytical tractability, caution must be exercised when choosing the value for the transition probability in order to reduce the affect of bias. The Theorems in Section 5.5 give us simple sufficient conditions to test for the presence of bias in two different Markov models of diseases. The disease models can include state dependent transition probabilities or time/age dependent probabilities. For the case of state dependent transition probabilities, we also show analytically that the bias increases as the degree of data aggregation increases, and that the increased bias is in the same direction as determined by Theorems 4 and 5.

The results in Section 5.6 show that the bias can be significant in many cases when models use constant transition probabilities. Using the indirect method, which is a common practice in each of the diseases discussed, resulted in larger bias than any of the three means considered in all cases except for one. Additionally, the bias increases dramatically when the degree of data aggregation increases.

The case where insufficient disease data causes the modeler to use the indirect method of obtaining a constant transition probability is of particular interest since it is such a common occurrence. We showed that the transition probability derived using the indirect method is less than harmonic mean of the state dependent transition probabilities when they are increasing. This result is important because it means that the use of the indirect method for diseases with increasing progression rates leads to an underestimation of disease progression. We also showed that when the arithmetic mean is used to average the indirect values of many patients, the disease progression is overestimated for diseases that have decreasing progression rates.

It is noteworthy that many studies that produce estimates for the state dependent

transition probabilities of diseases with increasing progression rates don't always arrive at strictly increasing transition probabilities. Results often include progression rates that increases until the final states where progression then slows down, such as those used in section 5.6. In light of the progressive nature of the diseases studied, the authors typically attribute the observed slower progression in the final states due to ceiling effects of the scoring system (e.g. MMSE), and not to the nature of the disease itself, such as in Park, et al, [85] and Aguero-Torres et al, [1]. Nevertheless, we used the transition rates as reported in the literature and found the bias to be significant when using the indirect method.

The implication of our results is that many of the studies that assume constant transition probabilities are arriving at conclusions based on biased models, bias potentially as large as the examples in Section 5.6. Analyses that use Markov models of disease progression that ignore the state/time dependency of the probability transitions to determine the cost-effective of medical interventions such as pharmaceutical drugs, therapies or screening programs may be significantly biased towards or away from cost-effectiveness depending on the type of aggregation used. A cost-effectiveness study of HCV, for example, that assumes a constant disease progression and uses the harmonic mean (or any smaller value) for the disease progression could potentially determine a cost-effective medical intervention to *not* be cost-effective because the disease progression was underestimated.

Similarly, when estimating future prevalence, if a disease model overestimates (underestimates) the progression to death, *ceteris paribus*, then future prevalence is underestimated (overestimated). For many diseases, the health states become costlier and associated with lower quality of life as the disease progresses, as is the case with HCV, AD and lung cancer. In these cases, it is clear that if we calibrate the models such that they have the same expected time from the first state to the final state, when the disease progression increases (decreases) in time/state, the Markov models

that use a single probability transition overestimate (underestimate) the time spent in the later states. This can be important since the later states can be significantly costlier than the earlier states in the disease progression causing the disease burden forecast to be biased.

Finally, we close with a discussion of an important aspect of Markov models of disease progression that causes bias in addition to the bias caused by the issues studied herein. The additional bias introduced into Markov models of disease progression is due to the fact that transition probability estimates are typically inherently underestimated. This is true because when a patient is determined to be in a particular health state at some point in time (by a liver biopsy, for example), there is no way of knowing how long the patient was in that state. Consequently, the estimate for the rate of progression is commonly made assuming the patient entered that diagnosed state in the time period of the diagnosis. For this reason, estimates of the disease progression are typically lower bounds. As a result, the bias introduced by the use of a single transition probability (instead of state/time dependent probabilities) can be either increased or decreased by this measurement effect. In the case that the use of a single transition probability causes an underestimation of disease progression, the bias due to measurement will cause the disease progression to be underestimated even more so than the results in Section 5.6 suggest.

5.8 Appendix

For the proofs in this appendix, it is helpful to first introduce some mathematical tools that we will use in our analysis of model bias. We start by introducing elementary symmetric polynomials, which are special cases of symmetric polynomials. The elementary symmetric polynomial of degree m in n variables (where $m \leq n$) is the sum of all distinct products of degree m of the n variables. In other words, we form all m-tuples of the n variables and add them up.

Definition 1. *The elementary symmetric polynomial of degree m in n variables is defined as*

$$e_{n,m}(x_1, x_2, \cdots, x_n) = \sum_{1 \le j_1 \le j_2 \cdots \le j_m \le n} x_{j_1} x_{j_2} ... x_{j_n}.$$

For example, when $n = 3$, the elementary symmetric polynomials are

$$\begin{aligned} e_{3,0}(x_1, x_2, x_3) &= 1 \\ e_{3,1}(x_1, x_2, x_3) &= x_1 + x_2 + x_3 \\ e_{3,2}(x_1, x_2, x_3) &= x_1x_2 + x_1x_3 + x_2x_3 \\ e_{3,3}(x_1, x_2, x_3) &= x_1x_2x_3. \end{aligned}$$

Definition 2. *An alternative definition for $e_{n,m}(x_1, x_2, \cdots x_n)$ is that of the coefficient of x^m in the expansion of $\prod_{1 \le i \le n} (x_i + y)$.*

It should be clear that the $\frac{e_{n,m}}{e_{n,n}} = \overline{e}_{n,n-m}$ where $\overline{e}_{n,n-m}$ are the elementary symmetric polynomials of $\frac{1}{x_1}, \frac{1}{x_2}, \cdots \frac{1}{x_n}$. Next, we use the elementary symmetric polynomials in the following definition.

Definition 3. *We define*

$$S_{n,m} := \frac{e_{n,m}}{\binom{n}{m}}.$$

$S_{n,m}$ is the elementary symmetric polynomial of degree m in n variables divided by the number of terms in this polynomial. We use $S_{n,m}$ extensively in the proofs below to simplify notation. Additionally, we use a relationship between the $S_{n,i}$ called MacLaurin's inequality.

MacLaurin's inequality states that

$$S_{n,1} \ge (S_{n,1})^{1/2} \ge (S_{n,2})^{1/3} \ge \cdots \ge (S_{n,n})^{1/n}$$

with equality if and only if all of the x_i values are equal. We use MacLaurin's inequality in the proofs for Theorems 4 and 5.

Lastly, we introduce the following Lemma, which we use to reformulate the expression for the stationary distribution of Markov models with state dependent transition probabilities into a form that is easier to manipulate.

Lemma 2.

$$\sum_{1\leq j\leq n}\frac{1}{a_j}\prod_{1\leq i\leq j}(\frac{a_i}{a_i+d})=\frac{(\prod_{1\leq i\leq n}(a_i+d)-\prod_{1\leq i\leq n}a_i)\frac{1}{d}}{\prod_{1\leq i\leq n}(a_i+d)} \tag{30}$$

which we prove below.

Proof of Proposition 1

Multiplying the right hand side of Equation 18 gives us

$$\begin{aligned}
\pi_1 &= \pi_1(1-a_1-d)+\pi_N \\
\pi_2 &= \pi_1 a_1+\pi_2(1-a_2-d) \\
\pi_3 &= \pi_2 a_2+\pi_3(1-a_3-d) \\
\vdots &\quad \vdots \\
\pi_{N-1} &= \pi_{N-2}a_{N-2}+\pi_{N-1}(1-a_{N-1}-d) \\
\pi_N &= \pi_{N-1}a_{N-1}+d(\pi_1+\pi_2+\cdots+\pi_{N-1})
\end{aligned} \tag{31}$$

which can be written as

$$\begin{aligned}
\pi_1 a_1 &= \pi_2(a_2+d) \\
\pi_2 a_2 &= \pi_3(a_3+d) \\
\pi_3 a_3 &= \pi_4(a_4+d) \\
\vdots &\quad \vdots \\
\pi_{N-1}a_{N-1} &= \pi_N(1+d)-d \\
\pi_N &= \pi_1(a_1+d).
\end{aligned} \tag{32}$$

We know that for a stationary distribution

$$\pi_1 + \pi_2 + \pi_3 + \cdots + \pi_N = 1 \tag{33}$$

must hold. Rewriting Equation 33 in terms of π_N using the identities in Equations 32 we get

$$\begin{aligned} &\frac{\pi_N}{a_1 + d} + \frac{\pi_N a_1}{(a_1 + d)(a_2 + d)} + \frac{\pi_N a_1 a_2}{(a_1 + d)(a_2 + d)(a_3 + d)} \\ &+ \cdots + \frac{\pi_N a_1 a_2 \cdots a_n}{(a_1 + d)(a_2 + d)(a_3 + d) \cdots (a_{N-1})} + \pi_N = 1 \end{aligned} \tag{34}$$

which can be rewritten as Equation 19. □

Proof of Proposition 2

Multiplying the right hand side of Equation 22 gives us

$$\begin{aligned} \pi_1 &= \pi_D \\ \pi_2 &= \pi_1(1 - a_1 - d) = \pi_D(1 - a_1 - d) \\ \pi_3 &= \pi_2(1 - a_2 - d) = \pi_D(1 - a_1 - d)(1 - a_2 - d) \\ \vdots &\quad \vdots \\ \pi_N &= \pi_{N-1}(1 - a_{N-1} - d) = \pi_D(1 - a_1 - d)(1 - a_2 - d) \cdots (1 - a_{N-1} - d) \\ \pi_C &= \pi_1 a_1 + \pi_2 a_2 + \cdots + \pi_N a_N + \pi_C(1 - D) \\ \pi_D &= \pi_1 d_1 + \pi_2 d_2 + \cdots + \pi_N d_N + \pi_C(D). \end{aligned}$$

Since we know that Equation 33 must hold for a stationary distribution, rewriting Equation 33 in terms of π_D using the above equations gives us the desired expression for π_D and we are done. □

Proof of Lemma 2

Let us define the following string of identities

$$\begin{aligned}
\prod_{1\le i\le N}(a_i+d) &= d\prod_{2\le i\le N}(a_i+d)+a_1\prod_{2\le i\le N}(a_i+d)\\
\prod_{2\le i\le N}(a_i+d) &= d\prod_{3\le i\le N}(a_i+d)+a_2\prod_{3\le i\le N}(a_i+d)\\
\prod_{3\le i\le N}(a_i+d) &= d\prod_{4\le i\le N}(a_i+d)+a_3\prod_{4\le i\le N}(a_i+d)\\
&\vdots \quad\quad \vdots\\
\prod_{N-2\le i\le N}(a_i+d) &= d\prod_{N-1\le i\le N}(a_i+d)+a_{n-2}\prod_{N-1\le i\le N}(a_i+d)\\
\prod_{N-1\le i\le N}(a_i+d) &= d\prod_{N\le i\le N}(a_i+d)+a_{n-1}\prod_{N\le i\le N}(a_i+d).
\end{aligned}$$

Substituting from the bottom up recursively we have

$$\begin{aligned}
\prod_{1\le i\le N}(a_i+d) &= d\prod_{2\le i\le N}(a_i+d)+da_1\prod_{3\le i\le N}(a_i+d)+da_1a_2\prod_{4\le i\le N}(a_i+d)\\
&+ \cdots+da_1a_2\cdots a_{N-2}(a_N+d)+da_1a_2\cdots a_{N-1}. \qquad (35)
\end{aligned}$$

Now, we can rewrite the left hand side of Lemma 2 with a common denominator to be

$$\begin{aligned}
LHS &= \sum_{1\le j\le N}\frac{1}{a_j}\prod_{1\le i\le j}\left(\frac{a_i}{a_i+d}\right)\\
&= \frac{\prod_{2\le i\le N}(a_i+d)+a_1\prod_{3\le i\le N}(a_i+d)+a_1a_2\prod_{4\le i\le N}(a_i+d)+\cdots+a_1a_2\cdots a_{N-1}}{\prod_{1\le i\le N}(a_i+d)}.
\end{aligned}$$

Define the numerator of LHS to be

$$\gamma=\prod_{2\le i\le N}(a_i+d)+a_1\prod_{3\le i\le N}(a_i+d)+a_1a_2\prod_{4\le i\le N}(a_i+d)+\cdots+a_1a_2\cdots a_{N-1}.$$

From Equation 35 we have

$$\prod_{1\le i\le N}(a_i+d)=\gamma\times d+\prod_{1\le i\le N}(a_i)$$

and we are done. □

Proof of Theorem 4

We start by analyzing the Markov disease model in Figure 13 and comparing it to the case where $a_i = x \ \forall \ i$. Then we extend the result to be true in the more general case of Figures 12(a) and 12(b). When x is less than or equal to the harmonic mean of the $a's$, we write this mathematically as

$$x \leq \frac{N}{\sum_{i=1}^{N} \frac{1}{a_i}}.$$

From Equation 21, for the disease progression in Figure 13, we know that the expected time from state 0 to state N is

$$E_a = \sum_{1 \leq j \leq N} \frac{1}{a_j} \prod_{1 \leq i \leq j} (\frac{a_i}{a_i + d}).$$

By Lemma 2, we can rewrite this as

$$E_a = \frac{(\prod_{1 \leq i \leq N} (a_i + d) - \prod_{1 \leq i \leq N} a_i)\frac{1}{d}}{\prod_{1 \leq i \leq N} (a_i + d)}.$$

Using Definition 2 on the numerator and the denominator we can rewrite the above expression as

$$E_a = \frac{(d^N + e_{N,1}d^{N-1} + e_{N,2}d^{N-2} + \cdots + e_{N,N-1}d + e_{N,N} - \prod_{1 \leq i \leq N} a_i)\frac{1}{d}}{d^N + e_{N,1}d^{N-1} + e_{N,2}d^{N-2} + \cdots + e_{N,N-1}d + e_{N,N}}. \quad (36)$$

Since $\prod_{1 \leq i \leq N} a_i = e_{N,N}$, we can rewrite Equation 36 as

$$E_a = \frac{d^{N-1} + e_{N,1}d^{N-2} + e_{N,2}d^{N-3} + \cdots + e_{N,N-1}d}{d^N + e_{N,1}d^{N-1} + e_{N,2}d^{N-2} + \cdots + e_{N,N-1}d + e_{N,N}}.$$

Finally, dividing both numerator and denominator by $e_{N,N}$ we get

$$E_a = \frac{\overline{e}_{N,N}d^{N-1} + \overline{e}_{N,N-1}d^{N-2} + \overline{e}_{N,N-2}d^{N-3} + \cdots + \overline{e}_{N,1}}{\overline{e}_{N,N}d^N + \overline{e}_{N,N-1}d^{N-1} + \overline{e}_{N,N-2}d^{N-2} + \cdots + \overline{e}_{N,1}d + 1} \quad (37)$$

where $\overline{e}_{i,j}$ are the elementary symmetric polynomials of $\frac{1}{a_1}$, $\frac{1}{a_2}$,$\cdots$ $\frac{1}{a_N}$.

Note that since the denominator can be written by multiplying the numerator by d and adding 1, Equation 37 can be written as

$$E_a = \frac{A}{Ad+1}$$

where

$$A = \overline{e}_{N,N}d^{n-1} + \overline{e}_{N,N-1}d^{N-2} + \overline{e}_{N,N-2}d^{N-3} + \cdots + \overline{e}_{N,1}.$$

We can arrive at a similar expression for the expected time from state 0 to state N in the model with aggregated transition probabilities where we have $x's$ instead of $a's$. Using the same procedure as above we get that the expected time from state 0 to state N (using $x's$) is

$$E_x = \frac{\frac{1}{x^N}d^{N-1} + \binom{N}{1}\frac{1}{x^{N-1}}d^{N-2} + \binom{N}{2}\frac{1}{x^{N-2}}d^{N-3} + \cdots + \frac{N}{x}}{\frac{1}{x^N}d^N + \binom{N}{1}\frac{1}{x^{N-1}}d^{N-1} + \binom{N}{2}\frac{1}{x^{N-2}}d^{N-2} + \cdots + \frac{N}{x} + 1}. \tag{38}$$

Similarly, since the denominator can be written by multiplying the numerator by d and adding 1, Equation 38 can be written as

$$E_x = \frac{B}{Bd+1}$$

where

$$B = \frac{1}{x^N}d^{N-1} + \binom{N}{1}\frac{1}{x^{N-1}}d^{N-2} + \binom{N}{2}\frac{1}{x^{N-2}}d^{N-3} + \cdots + \frac{N}{x}.$$

We would like to compare the expected time from state 0 to state N using $a's$ with the expected time using $x's$. If we can show that

$$\frac{B}{Bd+1} > \frac{A}{Ad+1} \tag{39}$$

then it follows that the model in which $x's$ are used underestimates the disease progression since the expected time from state 0 to state N is longer. Since the inequality in 39 is always true whenever $B > A$, we need only to show that

$$\underbrace{\overline{e}_{N,N}}_{a}d^N + \underbrace{\overline{e}_{N,N-1}}_{b}d^{N-1} + \underbrace{\overline{e}_{N,N-2}}_{c}d^{N-2} + \cdots + \underbrace{\overline{e}_{N,1}}_{d}d <$$
$$\underbrace{\frac{1}{x^N}}_{a'}d^N + \underbrace{\binom{N}{1}\frac{1}{x^{N-1}}}_{b'}d^{N-1} + \underbrace{\binom{N}{2}\frac{1}{x^{N-2}}}_{c'}d^{N-2} + \cdots + \underbrace{\frac{N}{x}}_{d'}d. \tag{40}$$

The expression labeled d is less than or equal to the expression labeled d' since

$$\overline{e}_{N,1} = \sum_{1\leq i\leq N} \frac{1}{a_i}$$

by hypothesis, and $x \leq \frac{N}{\sum_{1\leq i\leq N} \frac{1}{a_i}}$ implies $\frac{N}{x} \geq \sum_{1\leq i\leq N} \frac{1}{a_i}$.

Next we must show that the terms labeled a, b, c, $\cdots$ are strictly less than the terms labeled a', b', c', $\cdots$ (whenever the a_i are not equal), which is true when

$$\begin{aligned} \frac{\binom{N}{k}}{x^{N-k}} &> \overline{e}_{N,N-k} \quad \forall\ k \\ \frac{\binom{N}{m}}{x^m} &> \overline{e}_{N,m} = \binom{N}{m} S_{N,m} \quad \forall\ m \\ \frac{1}{x^m} &> S_{N,m} \quad \forall\ m. \end{aligned} \tag{41}$$

But since $\frac{1}{x} \geq \frac{\sum_{1\leq i\leq N} \frac{1}{a_i}}{N} = S_{N,1}$ and by Maclaurin's inequality $S_{N,1} > (S_{N,m})^{\frac{1}{m}}$ for all m whenever the a_i are not all equal, then inequality 41 is always true. The proof that the result remains true in the general case of the model in Figure 12(b) is below. □

Proof of Corollary 1

Given Theorem 4, we need only to show that when the $a_i's$ are increasing, the value computed using the indirect method is less than the harmonic mean of the $a_i's$. The indirect method is computed without taking into account death from other causes (i.e., when $d = 0$). It is a well known property of the harmonic mean that it equals the arithmetic mean of the reciprocals. Indeed, the harmonic mean is most useful when averaging rates. Since a_i represents the transition probability from state i to $i+1$, when $d = 0$ the expected time from state i to $i+1$ is $\frac{1}{a_i} := t_i$. Consequently, the harmonic mean of the $a_i's$ is equal to the arithmetic mean of the $t_i's$. When the $a_i's$ are increasing (i.e., $t_i's$ are decreasing), we can see that the value obtained using

the indirect method,

$$x_{indirect} = \frac{1}{\frac{1}{N}(t_1 + \frac{t_1+t_2}{2} + \cdots + \frac{t_1+t_2 \cdots t_N}{N})} \tag{42}$$

is less than the value using the arithmetic mean of the $t_i's$,

$$x_{arithmetic} = \frac{1}{\frac{1}{N}(t_1 + t_2 + \cdots + t_N)} = \frac{N}{(1/a_1 + 1/a_2 + \cdots 1/a_N)} \tag{43}$$

which is equal to the harmonic mean of the $a_i's$, and we are done. □

Proof of Theorem 5

As in the proof of the previous theorem, we start by analyzing the Markov disease model in Figure 13 and compare it to the case where $a_i = x \ \forall \ i$. We start with Equations 37 and 38 from the proof of Theorem 4. In this case, however, we want to show that $E_x < E_a$ or equivalently that

$$\frac{B}{Bd+1} < \frac{A}{Ad+1} \tag{44}$$

which is true when $B < A$. In this case, we must show that the expression on the left hand side of Equation 40 is greater than the expression on right hand side of Equation 40. By hypothesis, a is greater than or equal to the expression labeled a' since

$$\overline{e}_{N,N} \ = \ (\prod_{1 \le i \le N} \frac{1}{a_i})^{\frac{1}{N}}. \tag{45}$$

which is the reciprocal of the geometric mean of the $a_i's$. It remains to show that the terms labeled b, c, $\cdots$, d are strictly greater than the terms labeled b', c', $\cdots$, d' in inequality 40 (whenever the a_i are not equal), which is true when

$$\begin{aligned} \frac{\binom{N}{k}}{x^{N-k}} &< \overline{e}_{N,N-k} \quad \forall \ k \\ \frac{\binom{N}{m}}{x^m} &< \overline{e}_{N,m} = \binom{N}{m} S_{N,m} \quad \forall \ m \\ \frac{1}{x^m} &< S_{n,m} \quad \forall \ m. \end{aligned} \tag{46}$$

But since $\frac{1}{x} \leq (\prod_{1 \leq i \leq N} \frac{1}{a_i})^{\frac{1}{N}} = (S_{N,N})^{\frac{1}{N}}$ and by Maclaurin's inequality $(S_{N,N})^{\frac{1}{N}} < (S_{N,m})^{\frac{1}{m}}$ for all m whenever the a_i are not all equal, then inequality 46 is always true. Since the geometric mean is always less than or equal to the arithmetic mean, the result holds for the arithmetic mean as well. The proof that the result remains true in the general case of the model in Figure 12(b) is below. □

Proof of Corollary 2

Given Theorem 5, we need only to show that when the $a_i's$ are decreasing, the arithmetic mean of the indirect values of the patients is greater than the arithmetic mean of the $a_i's$. When the $a_i's$ are decreasing, the $t_i's$ are increasing, and we can see that the arithmetic mean of the indirect values,

$$x_{indirect} = \frac{1}{N}(\frac{1}{t_1} + \frac{2}{t_1 + t_2} + \cdots \frac{N}{t_1 + t_2 + \cdots t_N}) \tag{47}$$

is greater than

$$x_{arithmetic} = \frac{1}{N}(\frac{1}{t_1} + \frac{1}{t_2} + \cdots + \frac{1}{t_N}) = \frac{1}{N}(a_1 + a_2 + \cdots + a_N), \tag{48}$$

which is equal to the arithmetic mean of the $a_i's$, and we are done. □

Proof that Theorems 4 and 5 are true in the general case of the model in Figure 12(b).

We have that

$$\sum_{1 \leq j \leq N} \frac{1}{a_j} \prod_{1 \leq i \leq j} (\frac{a_i}{a_i + d}) > \sum_{1 \leq j \leq N} \frac{1}{x} \prod_{1 \leq i \leq j} (\frac{x}{x + d})$$

which expands to

$$\begin{aligned} & \frac{1}{a_1 + d} + \frac{a_1}{(a_1 + d)(a_2 + d)} + \frac{a_1 a_2}{(a_1 + d)(a_2 + d)(a_3 + d)} \\ & + \cdots + \frac{a_1 a_2 \cdots a_n}{(a_1 + d)(a_2 + d)(a_3 + d) \cdots (a_{N-1})} \\ > \quad & \frac{1}{x + d} + \frac{x}{(x + d)^2} + \frac{x^2}{(x + d)^3} \\ & + \cdots + \frac{x^N}{(x + d)^{N-1}}. \end{aligned} \tag{49}$$

We want to show that the inequality remains true when we add the transition probabilities i, z, and D to the stationary distribution according to the Markov model in Figure 12(b). By updating Equation 34 (where we derived the stationary distribution) to include i, z, and D, and dividing by π_N, it should be clear that we want to show

$$\begin{aligned} & \frac{1}{i+d} + \frac{i}{(i+d)(a_1+d)} + \frac{ia_1}{(i+d)(a_1+d)(a_2+d)} \\ & + \cdots + \frac{ia_1 \cdots a_n zD}{(i+d)(a_1+d)(a_2+d)\cdots(a_{N-1})(z+d)(D+d)} \\ > \quad & \frac{1}{i+d} + \frac{i}{(i+d)(x+d)} + \frac{ix}{(i+d)(x+d)^2} \\ & + \cdots + \frac{ix^N zD}{(i+d)(x+d)^{N-1}(z+d)(D+d)} \end{aligned}$$

which is clearly true whenever inequality 49 is true. The result is also true for the reverse inequality. □

Proof of Theorem 7

To show that bias increases as the degree of aggregation increases, we must prove that

$$E_a > E_b > E_x$$

when the $a_i's$ and $b_i's$ are not all equal to x, and $b_1 = \sqrt{a_1 a_2}$ and $b_2 = \sqrt{a_3 a_4}$. $E_a > E_x$ and $E_b > E_x$ are true by Theorem 5 from the previous section. It remains to show that $E_a - E_b$ is strictly positive, which expands to

$$E_a - E_b = \tag{50}$$

$$\frac{a_1 a_2 a_3 a_4}{(a_1+d)(a_2+d)(a_3+d)(a_4+d)(\sqrt{a_1 a_2}+d)^2(\sqrt{a_3 a_4}+d)^2} \tag{51}$$

$$(-2\sqrt{a_1 a_2}d^2 - 2d^2\sqrt{a_3 a_4} + a_2 d a_4 + a_1 a_3 a_4 + a_1 d a_3 + a_1 d a_4 + d^2 a_3 + a_2 a_3 a_4 \tag{52}$$

$$+a_2 d^2 + a_2 a_3 d + a_1 a_2 a_4 + a_1 a_2 a_3 + a_1 d^2 + d^2 a_4 - 2a_3 a_4\sqrt{a_1 a_2} - 4\sqrt{a_1 a_2}\sqrt{a_3 a_4}d$$

$$-2a_1 a_2\sqrt{a_3 a_4})/((a_1+d)(a_2+d)(a_3+d)(a_4+d)(\sqrt{a_1 a_2}+d)^2(\sqrt{a_3 a_4}+d)^2).$$

Since the fraction in line 51 is always positive, it suffices to show that lines 52 -

53 are strictly positive. These lines can be rewritten as a polynomial in d as

$$\begin{aligned} d^2 \quad & (-2\sqrt{a_1a_2} + a_1 + a_2 + a_3 + a_4 - 2\sqrt{a_3a_4}) \\ +d \quad & (a_1a_3 + a_1a_4 + a_2a_4 + a_2a_3 - 4\sqrt{a_1a_2}\sqrt{a_3a_4}) \\ + \quad & (-2a_1a_2\sqrt{a_3a_4} - 2a_3a_4\sqrt{a_1a_2} + a_1a_2a_4 + a_1a_3a_4 + a_2a_3a_4 + a_1a_2a_3). \end{aligned}$$

Next we show that each coefficient of the powers of d is strictly positive. Rewriting the above expression, we can see that this is true if and only if

$$\begin{aligned} \frac{a_1 + a_2}{2} + \frac{a_3 + a_4}{2} &> \sqrt{a_1a_2} + \sqrt{a_3a_4} \\ \frac{a_1 + a_2}{2}\frac{a_3 + a_4}{2} &> \sqrt{a_1a_2}\sqrt{a_3a_4} \\ a_1a_2\frac{a_3 + a_4}{2} + a_3a_4\frac{a_1 + a_2}{2} &> a_1a_2\sqrt{a_3a_4} + a_3a_4\sqrt{a_1a_2}. \end{aligned}$$

The above inequalities are true by repeatedly applying the inequality of the arithmetic and geometric mean, which states

$$\frac{a_1 + a_2}{2} > \sqrt{a_1a_2}$$

whenever $a_1 \neq a_2$ (which is a special case of the Maclaurin inequality where n=2). It should be clear that the result holds for any mean greater than or equal to geometric mean. In the same way that Theorems 4 and 5 were shown to still be valid in the more general setting where we include additional states before and after the disease progression, this theorem is also true in the more general setting. The proof is omitted since it follows the same steps. □

Proof of Theorem 8

We omit the proof since it very closely follows that of Theorem 7, where instead of using the inequality of the arithmetic and geometric means, we use the identity

$$a_1^2 + a_2^2 > 2a_1a_2$$

whenever $a_1 \neq a_2$ and $a_1, a_2 > 0$ which can be seen to be true since $(a_1 + a_2)^2 > 0$. It should be clear that the result holds for any mean less than or equal to the harmonic mean. In the same way that Theorems 4 and 5 were shown to still be valid in the more general setting where we include additional states before and after the disease progression, this theorem is also true in the more general setting. The proof is omitted since it follows the same steps. □

CHAPTER VI

CONCLUSIONS AND FUTURE RESEARCH DIRECTIONS

This thesis introduced several issues regarding the study of diseases. In Chapter 2, we studied a specific question in HIV policy regarding the proposed closure of bathhouses as a control measure to reduce the spread of HIV among MSM. To answer the question, we developed a Bernoulli process transmission model of a heterogeneous population with multiple risk groups. We included the effect of co-infection with other diseases, such as Syphilis, which increase the probability of transmission when present. We showed that the HIV attack rate is concave as a function of the proportion of the bathhouse patrons' contacts with other bathhouse patrons. We used this fact to draw conclusions on the effect of closing bathhouses under certain assumptions.

We populated the model of HIV transmission with data from a survey of four major cities in the US and found that the impact on HIV incidence from the disproportionate mixing of the population due to the presence of bathhouses is small compared to the impact from changes in some key parameter values, such as condom usage. The effect that closing bathhouses will have on these parameter values is not clear; however, the result suggests that alternative interventions targeted at individuals in bathhouse venues could have greater effects on the spread of HIV than closing bathhouses.

In Chapters 3 and 4, we built a mathematical model to examine the timing of testing and treatment for diseases, particularly Hepatitis C. We studied the problem with a dual approach. We developed Markov Decision Process model that arrives at a dynamic testing policy for a disease model with a simplified state space (Chapter 3) and one in which we examine all possible testing policies for the full disease state

space model via simulation (Chapter 4). The model allows for the awareness of a disease to change behavior including the consumption of alcohol and transmission to others.

We use medical data of Hepatitis C in both Chapters 3 and 4, and find that the current policy recommendations of testing for Hepatitis C are too restrictive, and that it is indeed cost-effective to test the overall population. We also demonstrate the importance of including behavior changes in the model by comparing the results to previous studies.

The topic in Chapter 3 has potential to motivate future research. In the future, we plan to expand the dynamic model of testing for diseases to include a larger state space. The analytical expressions become more cumbersome, but closed form expressions can still be obtained. With an expanded model we will more accurately represent diseases, and be able to model more complicated diseases. We intend to apply both the dynamic and simulation models to other diseases where there is no consensus regarding screening policies. Possible diseases include cancer and other STD's.

The Markov model used in the study of Hepatitis C in Chapters 3 and 4 motivated the topic in Chapter 5 where we examine bias in Markov models of diseases, including the one studied in Chapters 3 and 4. We consider two common types of diseases and the associated Markov models commonly used to model them: ones in which the disease progression changes by severity of the disease, and ones in which the progression of the disease changes in time or by age. We find sufficient conditions for bias to exist in models with aggregated transition probabilities when compared to models with state/time dependent transition probabilities. We also find that when aggregating data to compute transition probabilities, the bias increases with the degree of data aggregation.

We examine the bias in Markov models of Hepatitis C, Alzheimer's disease and

lung cancer using medical data and find that the bias is significant depending on the method used to aggregate the data. The common method of acquiring a constant probability transition, the indirect method, typically causes bias greater than that of using the harmonic or geometric means. The key implication is that by not incorporating state/time dependent transition probabilities, studies that use Markov models of diseases may be significantly overestimating or underestimating disease progression, depending on the type of data aggregation used. This could potentially result in incorrect recommendations in cost-effectiveness studies and incorrect future prevalence and disease burden forecasts.

The topic in Chapter 5 also has potential for future study. We plan to generalize the state space so that the disease progression may lead to two different final states before death. This will be useful since, in the case of Hepatitis C for example, liver disease can lead to decompensated cirrhosis and/or hepatocellular carcinoma, both of which are fatal. We also intend to consider bias in other types of Markov disease models, including ones with more complicated state spaces, or perhaps when the transition probabilities depend on time and state.

REFERENCES

[1] Aguero-Torres, H., Fratiglioni, L., and Winblad, B., "Natural history of alzheimer's disease and other dementias: Review of the literature in the light of the findings from the kungsholmen project," *International Journal of Geriatric Psychiatry*, vol. 13, no. 11, pp. 755–766, 1998.

[2] Alagoz, O., Maillart, L. M., Schaefer, A. J., and Roberts, M. S., "The optimal timing of living-donor liver transplantation," *Management Science*, vol. 50, no. 10, pp. 1420–1430, 2004.

[3] Alagoz, O., Maillart, L. M., Schaefer, A. J., and Roberts, M. S., "Determining the acceptance of cadaveric livers using an implicit model of the waiting list," *To appear in Operations Research*, vol. 50, no. 10, pp. 1420–1430, 2006.

[4] Armon, C., Graves, M. C., Moses, D., Forte, D. K., Sepulveda, L., Darby, S. M., and Smith, R. A., "Linear estimates of disease progression predict survival in patients with amyotrophic lateral sclerosis," *Muscle and Nerve*, vol. 23, no. 6, pp. 874–882, 2000.

[5] Ballard, C., O'Brien, J., Morris, C. M., Barber, R., Swann, A., Neill, D., and McKeith, I., "The progression of cognitive impairment in dementia with lewy bodies, vascular dementia and alzheimer's disease," *International Journal of Geriatric Psychiatry*, vol. 16, no. 5, pp. 499–503, 2001.

[6] Barzilay, E. J., O'Brien, K. L., Kwok, Y. S., Hoekstra, R. M., Zell, E. R., Reid, R., Santosham, M., Whitney, C. G., and Feikin, D. R., "Could a single dose of pneurnococcal conjugate vaccine in children be effective? modeling the optimal age of vaccination," *Vaccine*, vol. 24, no. 7, pp. 904–913, 2006.

[7] Bennett, W. G., Inoue, Y., Beck, J. R., Wong, J. B., Pauker, S. G., and Davis, G. L., "Estimates of the cost-effectiveness of a single course of interferon-alpha 2b in patients with histologically mild chronic hepatitis C," *Annals Of Internal Medicine*, vol. 127, no. 10, pp. 918–920, 1997.

[8] Berube, A., "The history of gay bathhouses," *J Homosex*, vol. 44, no. 3/4, pp. 33–53, 2003.

[9] Binson, D., Blea, L., Cotten, P., Kant, J., and Woods, W., "Building an hiv/sti prevention program in a gay bathhouse: a case study," *AIDS Educ Prev*, vol. 17, no. 4, pp. 386–399, 2005.

[10] Binson, D., Woods, W., Pollack, L., Paul, J., Stall, R., and Catania, J., "Differential hiv risk in bathhouses and public cruising areas," *Am J Public Health*, vol. 91, no. 9, pp. 1482–1486, 2001.

[11] Brenner, H., Chang-Claude, J., Seiler, C. M., Sturmer, T., and Hoffmeister, M., "Does a negative screening colonoscopy ever need to be repeated?," *Gut*, vol. 55, no. 8, pp. 1145–1150, 2006.

[12] Bull, S., McFarlane, M., Lloyd, L., and RietMeijer, C., "The process of seeking sex partners online and implications for std/hiv prevention," *AIDS Care*, vol. 16, pp. 1012–20, 2004.

[13] Burris, S., "Legal aspects of regulating bathhouses: cases from 1984 to 1995," *J Homosex*, vol. 44, no. 3, pp. 131–151, 2003.

[14] Castelnuovo, E., Thompson-Coon, J., Pitt, M., Cramp, M., Siebert, U., Price, A., and Stein, K., "The cost-effectiveness of testing for hepatitis c in former injecting drug users," *Health Technology Assessment*, vol. 10, no. 32, pp. 1–93, 2006.

[15] Catania, J., Osmond, D., Stall, R., Pollack, L., Paul, J., Blower, S., Binson, D., Canchola, J., Mills, T., Fisher, L., Choi, K., Porco, T., Turner, C., Blair, J., Henne, J., Bye, L., and Coates, T., "The continuing hiv epidemic among men who have sex with men," *Am J Public Health*, vol. 91, no. 6, pp. 907–914, 2001.

[16] Centers for Disease Control and Prevention, "1999 usphs/idsa guidelines for the prevention of opportunistic infections in persons infected with human immunodeficiency virus," *MMWR*, vol. 48, no. RR10, pp. 1–59, 1999.

[17] Centers for Disease Control and Prevention, "Increases in unsafe sex and rectal gonorrhea among men who have sex with men–san francisco, 1994-1997," *MMWR*, vol. 48, no. 3, pp. 45–48, 1999.

[18] Centers for Disease Control and Prevention, "Hiv/aids surveillance report 2003," vol. 15, pp. 1–46, 2004.

[19] Centers for Disease Control and Prevention, "Recommendations for prevention and control of hepatitis c virus (hcv) infection and hcv-related chronic disease," *MMWR*, vol. 47, no. RR19, pp. 1–39, 2004.

[20] Centers for Disease Control and Prevention, "Trends in primary and secondary syphilis and hiv infections in men who have sex wtih men – san francisco and los angeles, california, 1998–2002," *MMWR*, vol. 53, no. 26, pp. 575–578, 2004.

[21] Centers for Disease Control and Prevention, "Disease burden from hepatitis A, B, and C in the United States," 2005.

http://www.cdc.gov/ncidod/diseases/hepatitis/resource/PDFs/disease_burden.pdf, Accessed December 2006.

[22] Centers for Disease Control and Prevention, "Hiv prevalence, unrecognized infection, and hiv testing among men who have sex with men - five us cities, june 2004-april 2005," *MMWR*, vol. 54, pp. 597–601, 2005.

[23] Centers for Disease Control and Prevention, "Deaths: Final data for 2003," *National Vital Statistics Reports*, vol. 54, no. 13, 2006. http://www.cdc.gov/nchs/data/nvsr/nvsr54/nvsr54_13.pdf, Accessed December 2006.

[24] Centers for Disease Control and Prevention, "Epidemiology of hiv/aids-united states, 1981-2005," *MMWR*, vol. 55, pp. 589–592, 2006.

[25] Centers for Disease Control and Prevention, "Hepatitis surveillance Report no. 61," 2006. http://www.cdc.gov/hepatitis, Accessed November 2006.

[26] Centers for Disease Control and Prevention, "Revised recommendations for hiv testing of adults, adolescents, and pregnant women in health-care settings," *MMWR*, vol. 55, no. RR-14, pp. 1–17, 2006.

[27] Centers for Disease Control and Prevention, "Twenty-five years of hiv/aids united states, 1981-2006," *MMWR*, vol. 55, pp. 585–589, 2006.

[28] Centers for Disease Control and Prevention, "Viral hepatitis C," 2006. http://www.cdc.gov/ncidod/diseases/hepatitis/c/plan/HCV_infection.htm, Accessed December 2006.

[29] Chapman, R., Berger, M., Weinstein, M., Weeks, J., Goldie, S., and Neumann, P., "When does quality-adjusting life-years matter in cost-effectiveness analysis?," *Health Economics*, vol. 13, no. 5, pp. 429–436, 2004.

[30] Chen, S., Gibson, S., Katz, M., Klausner, J., Dilley, J., Schwarcz, S., Kellogg, T., and McFarland, W., "Continuing increases in sexual risk behavior and sexually transmitted diseases among men who have sex with men, san francisco, 1999-2001," *Am J Public Health*, vol. 92, no. 9, pp. 1387–1388, 2002.

[31] Chong, C., Gulamhussein, A., Heathcote, E., Lilly, L., Sherman, M., Naglie, G., and Krahn, M., "Health-state utilities and quality of life in hepatitic C patients," *The American Journal of Gastroenterology*, vol. 98, no. 3, pp. 630–638, 2003.

[32] Ciesielski, C. and Tabidze, I., "Repeat syphilis infections in men who have sex with men, 2000-2005, chicago, il [d7e]," May 8 - 11 2006. Presented at: National STD Prevention Conference; 2006; Jacksonville, FL.

[33] Colfax, G. and Guzman, R., "Club drugs and hiv infection: a review," *Clin Infect Dis*, vol. 42, no. 10, pp. 1463–1469, 2006.

[34] Colfax, G., Mansergh, G., Guzman, R., Vittinghoff, E., Marks, G., Rader, M., and Buchbinder, S., "Drug use and sexual risk behavior among gay and bisexual men who attend circuit parties: a venue based comparison," *J Acquir Immune Defic Syndr*, vol. 28, no. 4, pp. 373–379, 2001.

[35] Coreybloom, J., Galasko, D., Hofstetter, C. R., Jackson, J. E., and Thal, L. J., "Clinical-features distinguishing large cohorts with possible ad, probable ad, and mixed dementia," *Journal of the American Geriatrics Society*, vol. 41, no. 1, pp. 31–37, 1993.

[36] Degos, F., Christidis, C., Ganne-Carrie, N., Farmachidi, J. P., Degott, C., Guettier, C., Trinchet, J. C., Beaugrand, M., and Chevret, S., "Hepatitis C virus related cirrhosis: time to occurrence of hepatocellular carcinoma and death," *Gut*, vol. 47, no. 1, pp. 131–136, 2000.

[37] DeGruttola, V., Seage III, G., Mayer, K., and Horsburgh Jr, C., "Infectiousness of hiv between male homosexual partners," *J Clin Epidemiol*, vol. 42, no. 9, pp. 849–856, 1989.

[38] Diehl, K., Ivy, J., Maillart, L., and Ransom, S., "Assessing dynamic mammography schedules," 2006.

[39] Disman, C., "The san francisco bathhouse battles of 1984: civil liberties, aids risk, and shifts in health policy," *J Homosex*, vol. 44, no. 3, pp. 71–129, 2003.

[40] DMD America, "Analy$ource online," 2006. http://www.analysource.com/analy$ource.pdf, Accessed December 2006.

[41] Doll, R. and Peto, R., "Cigarette-smoking and bronchial-carcinoma - dose and time relationships among regular smokers and lifelong non-smokers," *Journal of Epidemiology and Community Health*, vol. 32, no. 4, pp. 303–313, 1978.

[42] Donato, F., Boffetta, P., and Puoti, M., "A meta-analysis of epidemiological studies on the combined effect of hepatitis B and C virus infections in causing hepatocellular carcinoma," *International Journal Of Cancer*, vol. 75, no. 3, pp. 347–354, 1998.

[43] Eddy, D. M., Hasselblad, V., McGivney, W., and Hendee, W., "The value of mammography screening in women under age 50 years," *Jama-Journal Of The American Medical Association*, vol. 259, no. 10, pp. 1512–1519, 1988. Times Cited: 225 0098-7484.

[44] Ekstrand, M., Stall, R., Paul, J., Osmond, D., and Coates, T., "Gay men report high rates of unprotected anal sex with partners of unknown or discordant hiv status," *AIDS*, vol. 13, no. 12, pp. 1525–1533, 1999.

[45] Elwood, W. and Williams, M., "Sex, drugs, and situation: attitudes, drug use, and sexual risk behaviors among men who frequent bathhouses," *J Psychol Hum Sexuality*, vol. 10, no. 2, pp. 23–44, 1998.

[46] Farley, T., "Cruise control," *Washington Monthly*, vol. Nov, pp. 37–41, 2002.

[47] Farrer, L. A., Cupples, A., Vanduijn, C. M., Connorlacke, L., Kiely, D. K., and Growdon, J. H., "Rate of progression of alzheimers-disease is associated with genetic risk," *Archives of Neurology*, vol. 52, no. 9, pp. 918–923, 1995.

[48] Fattovich, G., Giustina, G., Degos, F., Diodati, G., Tremolada, F., Nevens, F., Almasio, P., Solinas, A., Brouwer, J. T., Thomas, H., Realdi, G., Corrocher, R., Schalm, S. W., Bhalla, A., Casarin, C., Bonetti, P., Basho, J., Fuschi, P., Tocco, A., and Mura, D., "Effectiveness of interferon alfa on incidence of hepatocellular carcinoma and decompensation in cirrhosis type C," *Journal Of Hepatology*, vol. 27, no. 1, pp. 201–205, 1997.

[49] Fenton, K. and Imrie, J., "Increasing rates of sexually transmitted diseases in homosexual men in western europe and the united states: why?," *Infect Dis Clin N Am*, vol. 19, pp. 311–331, 2006.

[50] Forman, L. M., Lewis, J. D., Berlin, J. A., Feldman, H. I., and Lucey, M. R., "The association between hepatitis C infection and survival after orthotopic liver transplantation," *Gastroenterology*, vol. 122, no. 4, pp. 889–896, 2002.

[51] Freeman, A., Dore, G., Law, M., Thorpe, M., Van Overbeck, J., Lloyd, A. R., Marinos, G., and Kaldor, J. M., "Estimating progression to cirrhosis in chronic hepatitis c virus infection," *Hepatology*, vol. 34, no. 4, pp. 809–816, 2001.

[52] Galasko, D., Coreybloom, J., and Thal, L. J., "Monitoring progression in alzheimers-disease," *Journal of the American Geriatrics Society*, vol. 39, no. 9, pp. 932–941, 1991.

[53] Goldie, S. J. and Kuntz, K. M., "A potential error in evaluating cancer screening: A comparison of 2 approaches for modeling underlying disease progression," *Medical Decision Making*, vol. 23, no. 3, pp. 232–241, 2003.

[54] Gordon, F., "Cost-effectiveness of screening patients for hepatitis c," *American Journal of Medicine*, vol. 107, no. 6B, pp. 36S–40S, 1999.

[55] GrosfeldNir, A., "A two-state partially observable markov decision process with uniformly distributed observations," *Operations Research*, vol. 44, no. 3, pp. 458–463, 1996. Times Cited: 4 0030-364X.

[56] Hauskrecht, M. and Fraser, H., "Planning treatment of ischemic heart disease with partially observable markov decision processes," *Artificial Intelligence In Medicine*, vol. 18, no. 3, pp. 221–244, 2000. Times Cited: 3 0933-3657.

[57] Hornberger, J., Torriani, F. J., Dieterich, D. T., Brau, N., Sulkowski, M. S., Torres, M. R., Green, J., and Patel, K., "Cost-effectiveness of peginterferon alfa-2a (40 kda) plus ribavirin in patients with HIV and hepatitis C virus co-infection," *Journal Of Clinical Virology*, vol. 36, no. 4, pp. 283–291, 2006.

[58] Houshyar, A., "What can be done to screen progressive diseases optimally," *IIE Transactions*, vol. 23, no. 1, pp. 40–50, 1991.

[59] Jonsson, L., Lindgren, P., Wimo, A., Jonsson, B., and Winblad, B., "The cost-effectiveness of donepezil therapy in swedish patients with alzheimer's disease: A markov model," *Clinical Therapeutics*, vol. 21, no. 7, pp. 1230–1240, 1999.

[60] Jowett, S. L., Agarwal, K., Smith, B. C., Craig, W., Hewett, M., Bassendine, D. R., Gilvarry, E., Burt, A. D., and Basendine, M. F., "Managing chronic hepatitis c acquired through intravenous drug use," *Qjm-Monthly Journal of the Association of Physicians*, vol. 94, no. 3, pp. 153–158, 2001.

[61] Kaplan, E. H. and Satten, G. A., "Repeat screening for hiv: When to test and why," *Journal Of Acquired Immune Deficiency Syndromes*, vol. 23, no. 4, pp. 339–345, 2000.

[62] Kim, W. R., Poterucha, J. J., Hermans, J. E., Therneau, T. M., Dickson, E. R., Evans, R. W., and Gross, J. B., "Cost-effectiveness of 6 and 12 months of interferon-alpha therapy for chronic hepatitis c," *Annals Of Internal Medicine*, vol. 127, no. 10, pp. 866–874, 1997.

[63] Kirch, R. L. A. and Klein, M., "Surveillance schedules for medical examinations," *Management Science*, vol. 20, no. 10, pp. 1403–1409, 1974.

[64] Kraemer, H. C., Tinklenberg, J., and Yesavage, J. A., "How far vs how fast in alzheimers-disease - the question revisited," *Archives of Neurology*, vol. 51, no. 3, pp. 275–279, 1994.

[65] Lee, H. and Pierskalla, W., "Mass screening models for contagious disease with no latent period," *Operations Research*, vol. 36, no. 6, pp. 917–928, 1988.

[66] Lee, S., Huang, H., and Zelen, M., "Early detection of disease and scheduling of screening examinations," *Statistical Methods In Medical Research*, vol. 13, no. 6, pp. 443–456, 2004. Times Cited: 0 0962-2802.

[67] Lee, S. J. and Zelen, M., "Scheduling periodic examinations for the early detection of disease: Applications to breast cancer," *Journal Of The American Statistical Association*, vol. 93, no. 444, pp. 1271–1281, 1998. Times Cited: 9 0162-1459.

[68] Lipscomb, J., Weinstein, M., and Torrence, G., "Time preference," in *Cost-effectiveness in Health and Medicine* (Gold, M., Siegel, J., Russell, L., and Weinstein, M. C., eds.), New York, NY: Oxford University Press, 1996.

[69] Manns, M. P., McHutchison, J. G., Gordon, S. C., Rustgi, V. K., Shiffman, M., Reindollar, R., Goodman, Z. D., Koury, K., Ling, M. H., and Albrecht, J. K., "Peginterferon alfa-2b plus ribavirin compared with interferon alfa-2b plus ribavirin for initial treatment of chronic hepatitis C: a randomised trial," *Lancet*, vol. 358, no. 9286, pp. 958–965, 2001.

[70] Mansergh, G., Colfax, G., Marks, G., Rader, M., Guzman, R., and Buchbinder, S., "The circuit party men's health survey: findings and implications for gay and bisexual men," *Am J Public Health*, vol. 91, no. 6, pp. 953–958, 2001.

[71] Matsumura, H., Moriyama, M., Goto, I., Tanaka, N., Okubo, H., and Arakawa, Y., "Natural course of progression of liver fibrosis in japanese patients with chronic liver disease type c - a study of 527 patients at one establishment," *Journal of Viral Hepatitis*, vol. 7, no. 4, pp. 268–275, 2000.

[72] McCoy, C. and Inciardi, J., *Sex, drugs, and the continuing spread of AIDS*. Los Angeles: Roxbury Publishing, 1995.

[73] McLean, C., Kaur, A., Kerndt, P., Lawrence, M., Bolan, G., Markowitz, L., Levine, W., and Finelli, I., "Case-control study of risk factors for acquiring infectious syphilis among men who have sex with men-los angeles, 2000 [abstract]," *Int J STD AIDS*, vol. 12, no. suppl 2, 2001.

[74] Michaelson, J. S., Halpern, E., and Kopans, D. B., "Breast cancer: Computer simulation method for estimating optimal intervals for screening," *Radiology*, vol. 212, no. 2, pp. 551–560, 1999.

[75] Milioni, A. Z. and Pliska, S. R., "Optimal inspection under semi-markovian deterioration - basic results," *Naval Research Logistics*, vol. 35, no. 5, pp. 373–392, 1988.

[76] Milioni, A. Z. and Pliska, S. R., "Optimal inspection under semi-markovian deterioration - the catastrophic case," *Naval Research Logistics*, vol. 35, no. 5, pp. 393–411, 1988.

[77] Monahan, G. E., "Optimal stopping in a partially observable markov process with costly information," *Operations Research*, vol. 28, no. 6, pp. 1319–1334, 1980. Times Cited: 9 0030-364X.

[78] National Center for Chronic Disease Prevention and Health Promotion, "Behavioral Risk Factor Surveillance System." http://www.cdc.gov/BRFSS/, Accessed May 2006.

[79] National Institutes of Health, "Management of hepatitis C," vol. 15, no. 3, pp. 1–41, 1997. http://consensus.nih.gov/1997/1997HepatitisC105html.htm, Accessed December 2006.

[80] National Network of STD-HIV Training Centers, "Viral hepatitis," 2005. http://depts.washington.edu/nnptc/core_training/clinical/PDF/Viraltis.pdf, Accessed December 2006.

[81] Newell, G., Mansell, P., Wilson, M., Lynch, H., Spitz, M., and Hersh, E., "Risk factor analysis among men referred for possible acquired immune deficiency syndrome," *Prev Med*, vol. 14, no. 1, pp. 81–91, 1985.

[82] Ozekici, S. and Papazyan, T., "Inspection policies and processes for deteriorating systems subject to catastrophic failure," *Naval Research Logistics*, vol. 35, no. 4, pp. 481–492, 1988. Times Cited: 3 0894-069X.

[83] Ozekici, S. and Pliska, S. R., "Optimal scheduling of inspections - a delayed markov model with false positives and negatives," *Operations Research*, vol. 39, no. 2, pp. 261–273, 1991.

[84] Paltiel, A. D., Weinstein, M. C., Kimmel, A. D., Seage, G. R., Losina, E., Zhang, H., Freedberg, K. A., and Walensky, R. P., "Expanded screening for hiv in the united states - an analysis of cost-effectiveness," *New England Journal of Medicine*, vol. 352, no. 6, pp. 586–595, 2005.

[85] Park, H. L., O'Connell, J. E., and Thomson, R. G., "A systematic review of cognitive decline in the general elderly population," *International Journal of Geriatric Psychiatry*, vol. 18, no. 12, pp. 1121–1134, 2003.

[86] Parmigiani, G., "On optimal screening ages," *Journal Of The American Statistical Association*, vol. 88, no. 422, pp. 622–628, 1993. Times Cited: 22 0162-1459.

[87] Parmigiani, G., "Optimal scheduling of fallible inspections," *Operations Research*, vol. 44, no. 2, pp. 360–367, 1996. Times Cited: 3 0030-364X.

[88] Parmigiani, G., "Timing medical examinations via intensity functions," *Biometrika*, vol. 84, no. 4, pp. 803–816, 1997. Times Cited: 7 0006-3444.

[89] Parmigiani, G. and Skates, S., "Estimating distribution of the age of onset of detectable asymptomatic cancer," *Mathematical And Computer Modelling*, vol. 33, no. 12-13, pp. 1347–1360, 2001. Times Cited: 4 0895-7177.

[90] Paz-Bailey, G., Meyers, A., Blank, S., Brown, J., Rubin, S., Braxton, J., Zaidi, A., Schafzin, J., Weigl, S., and Markowitz, L., "A case-control study of syphilis among men who have sex with men in new york city," *Sex Transm Dis*, vol. 31, no. 10, pp. 581–587, 2004.

[91] Pereira, A. and Sanz, C., "A model of the health and economic impact of posttransfusion hepatitis c: application to cost-effectiveness analysis of further expansion of hcv screening protocols," *Transfusion*, vol. 40, no. 10, pp. 1182–1191, 2000.

[92] Peterman, T., Heffelfinger, J., Swint, E., and Groseclose, S., "The changing epidemiology of syphilis," *Sex Transm Dis*, vol. 32, no. 10 suppl, pp. s4–s10, 2005.

[93] Petersen, R. C., Stevens, J. C., Ganguli, M., Tangalos, E. G., Cummings, J. L., and DeKosky, S. T., "Practice parameter: Early detection of dementia: Mild cognitive impairment (an evidence-based review) - report of the quality standards subcommittee of the american academy of neurology," *Neurology*, vol. 56, no. 9, pp. 1133–1142, 2001.

[94] Pinkerton, S. and Abramson, P., "Effectiveness of condoms in preventing hiv transmission," *Soc Sci Med*, vol. 44, no. 9, pp. 1303–1312, 1997.

[95] Pinkerton, S., Abramson, P., and Holtgrave, D., "The bernoulli-process model of hiv transmission," in *Handbook of economic evaluation of HIV prevention programs*, pp. 13–32, New York: Plenum Press, 1998.

[96] Plunkett, B. A. and Grobman, W. A., "Routine hepatitis c virus screening in pregnancy: A cost-effectiveness analysis," *American Journal Of Obstetrics And Gynecology*, vol. 192, no. 4, pp. 1153–1161, 2005.

[97] Poynard, T., Bedossa, P., and Opolon, P., "Natural history of liver fibrosis progression in patients with chronic hepatitis c," *Lancet*, vol. 349, no. 9055, pp. 825–832, 1997.

[98] Poynard, T., Ratziu, V., Charlotte, F., Goodman, Z., McHutchison, J., and Albrecht, J., "Rates and risk factors of liver fibrosis progression in patients with chronic hepatitis C," *Journal Of Hepatology*, vol. 34, no. 5, pp. 730–739, 2001.

[99] Rabin, J., "The aids epidemic and gay bathhouses: a constitutional analysis," *J Health Polit Policy Law*, vol. 10, no. 4, pp. 729–747, 1986.

[100] Rapatski, B., Suppe, F., and Yorke, J., "Hiv epidemics driven by late disease stage transmission," *J Acquir Immune Defic Syndr*, vol. 38, no. 3, pp. 241–253, 2005.

[101] Rifai, M., Moles, J., and Short, D., "Hepatitis c treatment eligibility and outcomes among patients with psychiatric illness," *PSYCHIATRIC SERVICES*, vol. 57, no. 4, 2006.

[102] Rosenberg, P. S., Biggar, R. J., and Goedert, J. J., "Declining age at HIV-infection in the United States," *New England Journal Of Medicine*, vol. 330, no. 11, pp. 789–790, 1994.

[103] Rosenquist, C. J. and Lindfors, K. K., "Screening mammography in women aged 40-49 years - analysis of cost-effectiveness," *Radiology*, vol. 191, no. 3, pp. 647–650, 1994.

[104] Ross, S., *Stochastic Processes*. Wiley, 1995.

[105] Rottingen, J., Cameron, D., and Garnett, G., "A systematic review of the epidemiologic interactions between classic sexually transmitted diseases and hiv," *Sex Transm Dis*, vol. 28, no. 10, pp. 579–597, 2001.

[106] Salmon, D. P., Thal, L. J., Butters, N., and Heindel, W. C., "Longitudinal evaluation of dementia of the alzheimer type - a comparison of 3 standardized mental status examinations," *Neurology*, vol. 40, no. 8, pp. 1225–1230, 1990.

[107] Salomon, J. A. and Goldie, S. J., "Cost-effectiveness of treatment for chronic hepatitis c infection - reply," *Jama-Journal Of The American Medical Association*, vol. 290, no. 15, pp. 1994–1994, 2003.

[108] Sanders, G. D., Bayoumi, A. M., Sundaram, V., Bilir, S. P., Neukermans, C. P., Rydzak, C. E., Douglass, L. R., Lazzeroni, L. C., Holodniy, M., and Owens, D. K., "Cost-effectiveness of screening for hiv in the era of highly active antiretroviral therapy," *New England Journal of Medicine*, vol. 352, no. 6, pp. 570–585, 2005.

[109] Scherrer, C., *The Allocation of Scarce Resources in Public Health*. PhD thesis, Georgia Institute of Technology, 2005.

[110] Shiell, A., Brown, S., and Farrell, G. C., "Hepatitis c: an economic evaluation of extended treatment with interferon," *Medical Journal of Australia*, vol. 171, no. 4, pp. 189–193, 1999.

[111] Shilts, R., *And the band played on: politics, people, and the AIDS epidemic*. New York: St. Martin's Press, 1987.

[112] Siebert, U., Sroczynski, G., Marckmann, G., and Hillemanns, P., "Cost-effectiveness of human papillomavirus testing in cervical cancer screening. a review of decision analyses," *Gynakologe*, vol. 36, no. 4, pp. 341–348, 2003.

[113] Silini, E. M., Cavallero, A., Dal Bello, B., Bruno, R., Barbarini, G., Belloni, G., Cividini, A., Tinelli, C., and Mondelli, M. U., "Modelling liver fibrosis progression in chronic hepatitis c: A study of 563 patients with sequential liver biopsies," *Journal of Hepatology*, vol. 44, pp. S36–S36, 2006. Silini, E. M. Cavallero, A. Dal Bello, B. Bruno, R. Barbarini, G. Belloni, G. Cividini, A. Tinelli, C. Mondelli, M. U. Suppl. 2.

[114] Singer, M. E. and Younossi, Z. M., "Cost effectiveness of screening for hepatitis c virus in asymptomatic, average-risk adults," *American Journal Of Medicine*, vol. 111, no. 8, pp. 614–621, 2001.

[115] Smallwood, E. and Sondik, E. J., "Optimal control of partially observable markov processes over a finite horizon," *Operations Research*, vol. 21, no. 5, pp. 1071–1088, 1973.

[116] Sonnenberg, A., Delco, F., and Inadomi, J. M., "Cost-effectiveness of colonoscopy in screening for colorectal cancer," *Annals Of Internal Medicine*, vol. 133, no. 8, pp. 573–584, 2000.

[117] Stall, R., Hays, R., Waldo, C., Ekstrand, M., and McFarland, W., "The gay '90s: a review of research in the 1990s on sexual behavior and hiv risk among men who have sex with men," *AIDS*, vol. 14, no. Suppl 3, pp. S101–114, 2000.

[118] Stein, K., Dalziel, K., Walker, A., Jenkins, B., Round, A., and Royle, P., "Screening for hepatitis c in genito-urinary medicine clinics: a cost utility analysis," *Journal of Hepatology*, vol. 39, no. 5, pp. 814–825, 2003.

[119] Stern, R. G., Mohs, R. C., Davidson, M., Schmeidler, J., Silverman, J., Kramerginsberg, E., Searcey, T., Bierer, L., and Davis, K. L., "A longitudinal-study of alzheimers-disease - measurement, rate, and predictors of cognitive deterioration," *American Journal of Psychiatry*, vol. 151, no. 3, pp. 390–396, 1994.

[120] Stewart, A., Phillips, R., and Dempsey, G., "Pharmacotherapy for people with alzheimer's disease: A markov-cycle evaluation of five years' therapy using donepezil," *International Journal of Geriatric Psychiatry*, vol. 13, no. 7, pp. 445–453, 1998.

[121] Suh, G. H., Ju, Y. S., Yeon, B. K., and Shah, A., "A longitudinal study of alzheimer's disease: rates of cognitive and functional decline," *International Journal of Geriatric Psychiatry*, vol. 19, no. 9, pp. 817–824, 2004.

[122] Sullivan, S. D., Jensen, M., Bernstein, D. E., Hassanein, T. I., Foster, G. R., Lee, S. S., Cheinquer, H., Craxi, A., Cooksley, G., Klaskala, W., Pettit, K., Patel, K. K., and Green, J., "Cost-effectiveness of combination peginterferon alpha-2a and ribavirin compared with

interferon alpha-2b and ribavirin in patients with chronic hepatitis C," *American Journal Of Gastroenterology*, vol. 99, no. 8, pp. 1490–1496, 2004.

[123] Tanaka, J., Kumada, H., Ikeda, K., Chayama, K., Mizui, M., Hino, K., Katayama, K., Kumagai, J., Komiya, Y., Miyakawa, Y., and Yoshizawa, H., "Natural histories of hepatitis c virus infection in men and women simulated by the markov model," *Journal of Medical Virology*, vol. 70, no. 3, pp. 378–386, 2003.

[124] Tavakoli, M. and Malek, M., "The cost utility analysis of riluzole for the treatment of amyotrophic lateral sclerosis in the uk," *Journal of the Neurological Sciences*, vol. 191, no. 1-2, pp. 95–102, 2001.

[125] Taylor, M., Hawkins, K., Gonzalez, A., Buchacz, K., Aynalem, G., Smith, L., Klausner, J., Holmberg, S., and Kerndt, P., "Use of the serologic testing algorithm for recent hiv seroconversion (starhs) to identify recently acquired hiv infections in men with early syphilis in los angeles county," *J Acquir Immune Defic Syndr*, vol. 38, no. 5, pp. 505–508, 2005.

[126] Thompson, J., "Understanding the aids epidemic: A modeler's odyssey'," in *Mathematical Modeling* (Wallenius, D. S. and T., eds.), pp. 41–69, New York: CRC Press, 1999.

[127] United States Department of Health and Human Services, "The health consequences of smoking: A report of the surgeon general.," 2004. Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health.

[128] U.S. Preventive Services Task Force, "Screening for hepatitis c: Recommendation statement," *Annals Of Internal Medicine*, vol. 140, no. 6, pp. 462–464, 2004.

[129] Van Beneden, C., O'Brien, K., Modesitt, S., Yusem, S., Rose, A., and Fleming, D., "Sexual behaviors in an urban bathhouse 15 years into the hiv epidemic," *J Acquir Immune Defic Syndr*, vol. 30, no. 5, pp. 522–526, 2002.

[130] Vijan, S., Hofer, T. P., and Hayward, R. A., "Cost-utility analysis of screening intervals for diabetic retinopathy in patients with type 2 diabetes mellitus," *Jama-Journal Of The American Medical Association*, vol. 283, no. 7, pp. 889–896, 2000.

[131] Vittinghoff, E., Douglas, J., Judson, F., McKiman, D., MacQueen, K., and Buchbinder, S., "Per-contact risk of human immunodeficiency virus transmission between male sexual partners," *Am J Epidemiol*, vol. 150, no. 3, pp. 306–311, 1999.

[132] Wawer, M., Gray, R., Sewankambo, N., Serwadda, D., Li, X., Laeyendecker, O., Kiwanuka, N., Kigozi, G., Kiddugavu, M., lutalo, T., Nalugoda, F., Wabwire-Mangen, F., Meehan, M., and Quinn, T., "Rates of hiv-1 transmission per coital act, by stage of hiv-1 infection, in rakai, uganda," *J Infect Dis*, vol. 191, pp. 1403–1409, 2005.

[133] Wein, L. M., Zenios, S. A., and Nowak, M. A., "Dynamic multidrug therapies for hiv: A control theoretic approach," *Journal of Theoretical Biology*, vol. 185, no. 1, pp. 15–29, 1997.

[134] Weingberg, M. and Williams, C., "Gay baths and the social organization of impersonal sex," *Soc Probl*, vol. 23, pp. 124–136, 1975.

[135] White, C. C., "Optimal inspection and repair of a production process subject to deterioration," *Journal Of The Operational Research Society*, vol. 29, no. 3, pp. 235–243, 1978.

[136] Wiley, T. E., McCarthy, M., Breidi, L., and Layden, T. J., "Impact of alcohol on the histological and clinical progression of hepatitis c infection," *Hepatology*, vol. 28, no. 3, pp. 805–809, 1998.

[137] Wohlfeiler, D., "Structural and environmental hiv prevention for gay and bisexual men," *AIDS*, vol. 14, no. Suppl 1, pp. S52–S56, 2000.

[138] Wong, J. B., McQuillan, G. M., McHutchison, J. G., and Poynard, T., "Estimating future hepatitis c morbidity, mortality, and costs in the united states," *American Journal of Public Health*, vol. 90, no. 10, pp. 1562–1569, 2000.

[139] Woods, W., Binson, D., Mayne, T., Gore, L., and Rebchook, G., "Hiv/sexually transmitted disease education and prevention in us bathhouses and sex club environments," *AIDS*, vol. 14, no. 5, pp. 625–626, 2000.

[140] Woods, W., Binson, D., Mayne, T., Gore, L., and Rebchook, G., "Facilities and hiv prevention in bathhouse and sex club environments," *J Sex Res*, vol. 38, no. 1, pp. 68–74, 2001.

[141] Woods, W., Binson, D., Pollack, L., Wohlfeiler, D., Stall, R., and Catania, J., "Public policy regulating private and public space in gay bathhouses," *J Acquir Immune Defic Syndr*, vol. 32, no. 4, pp. 417–423, 2003.

[142] Woods, W., Tracy, D., and Binson, D., "Number and distribution of gay bathhouses in the united states and canada," *J Homosex*, vol. 44, pp. 55–70, 2003.

[143] Xia, Q., Tholandi, M., Osmond, D., Pollack, L., Zhou, W., Ruiz, J., and Catania, J., "The effect of venue sampling on estimates of hiv prevalence and sexual risk behaviors in men who have sex with men," *Sex Transm Dis*, vol. 33, no. 7, p. in press, 2006.

[144] Yang, Y. J. and Klutke, G. A., "Improved inspection schemes for deteriorating equipment," *Probability In The Engineering And Informational Sciences*, vol. 14, no. 4, pp. 445–460, 2000.

[145] Yi, Q., Wang, P. P., and Krahn, M., "Improving the accuracy of long-term prognostic estimates in hepatitis c virus infection," *Journal of Viral Hepatitis*, vol. 11, no. 2, pp. 166–174, 2004.

[146] Zarski, J. P., Mc Hutchison, J., Bronowicki, J. P., Sturm, N., Garcia-Kennedy, R., Hodaj, E., Truta, B., Wright, T., and Gish, R., "Rate of natural disease progression in patients with chronic hepatitis c," *Journal of Hepatology*, vol. 38, no. 3, pp. 307–314, 2003.

[147] Zelen, M., "Optimal scheduling of examinations for the early detection of disease," *Biometrika*, vol. 80, no. 2, pp. 279–293, 1993.

VITA

Daniel Mello Faissol was born in Los Angeles, CA on April 6th, 1980. He attended the University of California at Los Angeles where he obtained a Bachelor's Degree Electrical Engineering with a minor in Public Policy in 2003. One year later, he joined the PhD program in Industrial and Systems Engineering at the Georgia Institute of Technology where he also earned a Master's Degree in Economics.