

©Copyright 2012

Aditya S. Khanna

Bio-Behavioral Models for HIV Prevention

Aditya S. Khanna

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2012

Reading Committee:

Steven M. Goodreau, Chair

Matthew R. Golden

Vladimir Minin

Program Authorized to Offer Degree:
Quantitative Ecology and Resource Management

University of Washington

Abstract

Bio-Behavioral Models for HIV Prevention

Aditya S. Khanna

Chair of the Supervisory Committee:
Associate Professor Steven M. Goodreau
Anthropology

Background: In this dissertation, I model HIV transmission in two different populations: circular migrants in South Africa, and men who have sex with men (MSM) in Southern California.

Objectives: One objective of the circular migrations (Chapter 2) project is to compare compartmental and network models in how they explain the dynamics of a system in which some individuals engage in circular migrations. Another goal is to understand the role such movement patterns play in HIV transmission, given the critical interaction between acute infection and rate of movement.

In the population of MSM in Southern California, I model post-diagnosis behavior change (PDBC). The objective in Chapter 3 is to quantify the population-level effects of individual-level PDBC. A second objective (Chapter 4) is to compare various strategies for early diagnosis in terms of how many new infections they produce over a 10-year period.

Methods: The compartmental models in Chapter 2 are developed using ordinary differential equations (ODE's) and the network models are derived from the exponential random graph models (ERGM's). Epidemiological studies on circular migrations in the KwaZulu-Natal province conducted over the past decade provide the necessary motivation.

The models on PDBC in Chapters 3 and 4 are based on ERGMs. The behavioral components of these models are parameterized primarily using the Acute Infection and Early Disease Research Program (AIEDRP). Several other studies are used to model the

various biological and demographic components of these models (details in Appendix F).

Results: In Chapter 2, I find that compartmental models do not show any impact of the frequency of migrations on HIV prevalence. Under suitable assumptions, network models show that “frequent” migrations (that take place at intervals shorter than the window of acute infection) produce a larger epidemic than infrequent migrations.

In Chapter 3, I find that if there were no PDBC, HIV prevalence in Southern California MSM would be higher by about a third. In Chapter 4, I find that individual testing strategies where people test based on the number of partners or time since last test work better than other strategies I consider.

Discussion: Chapter 2 is a largely theoretical study, parameterized by qualitative descriptions of data. While the question of the precise impact of circular migrations on HIV transmission is still an open one, and likely contingent upon specific behavioral components in particular populations where circular migrations are practiced. However, our work demonstrates that we need to network models to understand this connection. We also demonstrate the potential impact of acute infection and its interaction with frequency of migration, and the effect of delayed path-acceleration.

In Chapters 3 and 4, I focus on behavioral interventions to prevent HIV infections. I do not consider “treatment as prevention” type strategies; though such considerations are suitable for future work and may act in synergy with the behavioral interventions I consider here.

TABLE OF CONTENTS

	Page
List of Figures	iv
List of Tables	v
Glossary	vi
Chapter 1: Introduction	1
1.1 Overview	1
1.2 Recent HIV Statistics in Sub-Saharan Africa and North America	2
1.3 Acute HIV Infection	3
1.4 Sexual Behavior and HIV	4
1.5 Concurrency and HIV	5
1.6 HIV Modeling	6
Chapter 2: Circular Migrations and HIV Transmission: A Comparison of Com- partmental and Network Modeling	10
2.1 Introduction	11
2.2 Background	13
2.3 Methods	18
2.4 Other Analyses	30
2.5 Results	30
2.6 Discussion	31
Chapter 3: Modeling the Impact of Post-Diagnosis Behavior Change on HIV Preva- lence in Southern Californian Men who have Sex with Men (MSM)	44
3.1 Introduction	44
3.2 Background	45
3.3 Methods	46
3.4 Results	53
3.5 Discussion	54

Chapter 4:	An Examination of Diagnosis Strategies to Maximize the Preventive Potential of Post-Diagnosis Behavior Change in Southern Californian Men who have Sex with Men	63
4.1	Introduction	64
4.2	Background	65
4.3	Methods	66
4.4	Results	69
4.5	Discussion	71
Chapter 5:	Concluding Thoughts	78
5.1	Circular Migrations and HIV Transmission	78
5.2	Individual-Level Post Diagnosis Behavior Change and Population-Level Effects	79
5.3	Examination of Diagnosis Strategies	80
Appendix A:	R_0 Computations for Compartmental Models in Chapter 2	81
Appendix B:	The Role of Acute Infection in Chapter 2	88
B.1	Introduction	88
B.2	Methods	88
B.3	Results	90
B.4	Discussion	90
Appendix C:	Restricted Network Models with Stochastic Migrations in Chapter 2	94
C.1	Introduction	94
C.2	Methods	94
C.3	Results	94
C.4	Discussion	95
Appendix D:	Degree Distributions in Network Models in Chapter 2	97
D.1	Introduction	97
D.2	Methods	97
D.3	Results	97
D.4	Discussion	99
Appendix E:	Isolating the Impact of Delayed Path-Acceleration and Momentary Degree Distribution in the Restricted Network Models in Chapter 2 .	100
E.1	Introduction	100

E.2	Methods	100
E.3	Results	101
E.4	Discussion	101
Appendix F: Technical Appendix to Chapter 3: Model Formulation, Parameter		
	Estimation, Simulation, and Comprehensive Listing of Data Sources .	106
F.1	Model for Non-Main Partnership Network	106
F.2	Model for Main Partnership Network	110
F.3	Statistics to Model Counter-Factual with No PDBC	113
F.4	Other Biological and Demographic Processes	113
F.5	Itemized List for Data Sources	115
Bibliography		120

LIST OF FIGURES

Figure Number	Page
2.1 Concurrency Structure in Circular Migrations	38
2.2 Transfer Diagram for Circular Migration System	39
2.3 HIV Prevalence as a Function of Time with Compartmental Models	41
2.4 HIV Prevalence as a Function of Time with Network Models	43
3.1 HIV Prevalence in the Baseline “Testing Frequency” Models	59
3.2 HIV Prevalence in the “Level of Awareness” Models	60
3.3 Protetective Effect of PDBC	61
3.4 Mean Degrees of Diagnosed and Undiagnosed Men in Testing Frequency and Level of Awareness Scenarios	62
4.1 Graphical Comparison of Each Strategy in the Two Scenarios	77
B.1 HIV Prevalence in Compartmental Models with No Acute Infection	92
B.2 HIV Prevalence in Network Models with No Acute Infection	93
C.1 HIV prevalence in Restricted Network Models with Stochastic Migrations	96
D.1 Cumulative and Momentary Degree Distributions in Network Models	98
E.1 HIV Prevalence in Constrained Network Models	103
E.2 Cumulative and Momentary Degree Distributions in Network Models	104
E.3 HIV Prevalence in Constrained Network Models with a 100-week Migration	105

LIST OF TABLES

Table Number	Page
2.1 Demographic and Biological Parameters	38
2.2 Initial Conditions for Compartmental Model	40
2.3 Biological and Behavioral Parameters in the Contact-As-Act Model	41
2.4 Biological and Behavioral Parameters in the Contact-As-Partnership Model	42
3.1 Proportion Diagnosed Early	59
4.1 Number of new infections in the Testing Frequency Scenario	75
4.2 Proportion of Cases Prevented in Testing Frequency Scenario	75
4.3 Number of New Infections in Level of Awareness Scenario	76
4.4 Proportion of Cases Prevented in Level of Awareness Scenario	76
F.1 Race Mixing Matrix for the Non-Main Network	118
F.2 Degree Distributions for Main Partnerships for network of size 5000.	118
F.3 Race Mixing Matrix for the Main Network	118
F.4 Race Mixing Matrix for the Non-Main Network in the “No PDBC” Counter-Factual	119

GLOSSARY

AIEDRP: Acute Infection and Early Disease Research Program.

ACUTE HIV INFECTION: high infectiousness during the first 12 weeks of HIV infection.

BASIC NETWORK MODELS: network models as close as possible to compartmental models in Chapter 2.

CIRCULAR MIGRATIONS: periodic movement between two or more locations.

CONCURRENCY: temporal overlap in partnerships.

CONSTRAINED NETWORK MODELS: network models where there is no restriction on the location of multiple partners of migrant men but the maximum number of partners of migrant men is limited to 2.

CUMULATIVE DEGREE DISTRIBUTION: lifetime distribution of number of partnerships.

DEGREE DISTRIBUTION: Distribution of number of partnerships.

ERGM: exponential family random graph model.

HIV: human immunodeficiency virus.

IT STRATEGIES: individualized testing strategies where men test after a certain number of partners or a certain number of days since the last test.

LEVEL OF AWARENESS: scenario in Chapters 3 and 4 where we match data that only about 56% of HIV infected men are aware of their status.

MOMENTARY DEGREE DISTRIBUTION: distribution of number of partnership at a given point in time.

MSM: men who have sex with men.

NODE: graph-theoretic term that represents individuals in the context of this work.

ODE: ordinary differential equation.

PDBC: post-diagnosis behavior change.

RESTRICTED NETWORK MODELS: network models in Chapter 2 where multiple partners of migrant men are required to be in separate locations

TESTING FREQUENCY: scenario in Chapters 3 and 4 where we assume men test on average once every 351 days.

TIE: graph-theoretic term that represents sexual partnership (or an act of UAI) in the context of the current work (also known as “edge”).

UAI: unprotected anal intercourse.

ACKNOWLEDGMENTS

This dissertation is the culmination of several years of work. Many individuals have contributed to my progress in graduate school, and the ultimate completion of my dissertation. I would like to thank them here.

First of all, I thank my advisor Dr. Steven Goodreau, who in addition to being a wonderful scientist and teacher, is also a great human being. I first approached him in late 2006 with the idea that a social scientist would help me ask the right questions, and I will figure out the modeling methods. It turned out that I learnt more than I could ever imagine about both of these from Steve, and much more.

Many thanks are due to my committee: Dr. Matthew Golden, Dr. Lisa Manhart, and Dr. Vladimir Minin. Vladimir helped me develop confidence in taking on research problems that I find interesting, and having the confidence to solve them. Lisa helped me learn how to think like an epidemiologist. Matt also helped me to learn how to use models to gain epidemiological insight and think about the connection between the science we practice in our labs and making a real difference in the communities we serve.

I also would like to thank the QERM program for giving me an opportunity to attend graduate school in what I really wanted to do. Special thanks are due to Dr. Loveday Conquest and Joanne Besch for their constant, unwavering support. I am also deeply grateful to the Center for Quantitative Science, its director Dr. Vince Gallucci and coordinator Linda Hegrenes, for providing me with teaching assistantships that made graduate school possible. I also thank Dr. Pamina Gorbach at UCLA for several years of intellectual and financial support that made much of my dissertation possible.

The Network Modeling Group (NMG) at the University of Washington has been a wonderful intellectual home to me; here, I learnt so much about all the different pieces of the work I do. The support from the students, post-docs, and faculty has been amazing. Dr.

Martina Morris posed two big questions that formed significant components of Chapters 2 and 4. Dr. Mark Handcock taught me the rudiments of the statistical analysis of networks. Dr. Susan Cassels gave me the first set of papers I read on migrations, and (along with Dr. Morris) authored the first paper I read in graduate school. Dr. Abernethy gave me great ideas on how my work fits in the bigger scheme of things, and some useful pointers on scientific presentations.

The Center for Studies in Demography and Ecology (CSDE) provided me with the space and the facilities that were needed. I thank them for office space over the past year. I also am deeply indebted to Matt Weatherford for teaching me so much about UNIX, and constantly keeping an open door and a smile on his face. I also thank Scott Sipes, Deborah Harper and Debbie Smith who made it easy for me to navigate the various financial and bureaucratic components of graduate school.

I want to thank the Statistics department at the University of Washington for an excellent experience in growing as a statistician, particularly Profs. Galen Shorack, Michael Perlman, Paul Sampson and David Yanez (Biostatistics). Dr. Dobromir Dimitrov, Fred Hutchinson Cancer Research Center, has been a wonderful teacher of mathematical epidemiology. Dr. Mark Kot, Applied Mathematics, provided many helpful ideas on the migrations project, and was extremely giving of his time as I looked for ideas on what I do after graduate school. I thank my undergraduate thesis advisors, Dr. Kelly Kriebel and Dr. Gordon Williams, who supervised my first experience in ecological modeling.

I express my gratitude to Bulldog News, Cafe Racer, Chili's South Indian restaurant, College Inn, Fuel Coffee (Montlake), Korean Tofu House, in no particular order, for providing space to do this work and/or sustenance that kept me going.

I have also been blessed to have so many wonderful friends here in Seattle and previous locations I lived in. During the course of graduate school many individuals come and go through your life, but their collective presence is still with me. I can't name all of them individually, but many many thanks to all of you.

I have an amazing host family. Dr. Susanna Cunningham, and her family, Kathryn,

Brian and Ian have been more than supportive and loving during my stay in Seattle.

Finally I want to thank my mother and father, Rama and Subhash Khanna, for allowing me to pursue my dreams at every step of my life, and my very loving sister Deepti and niece Lakshita. I couldn't ask for a more amazing family and I am blessed to have them.

DEDICATION

to my dear Mummy and Papa,
Rama and Subhash Khanna

Chapter 1

INTRODUCTION

1.1 Overview

Mathematical modeling of human immunodeficiency virus (HIV) transmission is an extremely valuable method to answer important, unanswered questions about the epidemic. The computer serves as a “laboratory” where we examine many questions where empirical investigation is either infeasible or unethical [1]. Mathematical models are especially useful in HIV prevention because they give us a framework to evaluate and compare different scenarios. In this dissertation, I build mathematical models to study HIV transmission in two different settings: a hypothetical setting that involves circular migrants, and men who have sex with men (MSM) in the Southern California region of the United States.

This work is divided into three chapters. In Chapter 2, I examine the transmission dynamics of HIV in a population of migrants and non-migrants, where the former engage in circular migrations between their home-village and workplace. These migrant men have partners in both locations and sexual contact with partners in either location is contingent upon the time they spend in each location. There are two primary goals; the first is to examine the relationship between the frequency of migration and HIV prevalence. The second goal is to model this system using compartmental and network models, and compare the dynamics observed through the two frameworks.

The second and third chapters are modeling efforts in a much larger project to understand the preventive potential of post-diagnosis behavior change (PDBC) in MSM in Southern California. At the heart of this work are longitudinal data collected as part of the Acute Infection and Early Disease Research [2, 3] program. These data report detailed behavioral information on the behavior of recently diagnosed men, at the time of diagnosis and at 3-monthly interviews thereafter.

I use these data to parameterize network models of HIV to answer specific questions that

involve PDBC. In Chapter 3, I focus on modeling the population-level effects of individual-level PDBC. I present results that demonstrate the extent to which HIV prevalence among Southern Californian MSM could be higher – were there no PDBC. I present a second set of results that indicate the importance of longitudinal data in evaluating effectiveness of PDBC. In Chapter 4, I examine how PDBC can be used as a strategy to prevent new infections, given the current state of HIV testing, and testing behavior. I compare these situations with newer measures that may be implemented in the near future, and compare their effectiveness.

There are several themes that are in common to the three chapters in this work. I describe the significance of these themes here, and their relevance of each to the questions pursued in this dissertation.

1.2 Recent HIV Statistics in Sub-Saharan Africa and North America

An estimated 34 million people worldwide were living with HIV at the end of 2010 [4]. This estimate is about 17% greater than the estimate for 2001, and is reflective of continued high incidence as well as increased life expectancies of infected individuals, (at least) partly due to availability of better therapies [4]. Sub-Saharan Africa continues to experience the greatest burden of HIV; in 2010, despite a decline in the number of new infections, an estimated 68% of the people infected with HIV resided in this region, which is home to about 12% of the total world population. Sub-Saharan Africa also accounted for about 70% of the new HIV infections worldwide. South Africa, in particular, is home to an estimated 5.6 million people with HIV [4]. While an estimated 33% decline in HIV incidence in South Africa from 2001 to 2009 shows that the epidemic is “leveling off,” the overall HIV epidemic here (and in other high prevalence countries such as Lesotho, Mozambique and Swaziland) remains “unacceptably high” [4].

In North America, on the other hand, we have seen little change in the number of new infections since 2004 [4]. In 2011, the Centers for Disease Control and Prevention (CDC) estimated that there were more than 1.1 million people living with HIV in the United States (according to 2008 data), and approximately 50,000 new infections occurred every year (about 48,100 in 2009) [5]. Men who have sex with men (MSM) constitute more than

half of both, the number living with HIV, and the number of new infections in the United States each year [6]. The other high-risk groups in the U.S. are injection drug users and high-risk heterosexuals; MSM is the only category in which the number of new infections has been increasing annually [6]. MSM in San Diego County show similar trends of high prevalence; from 1985-2004 non-IDU MSM accounted for 79% of the HIV transmission events [7].

The Southern California region of the United States is experiencing a particularly large HIV epidemic. The cities of Los Angeles (LA) and San Diego (SD), and their respective counties, are major population centers of this region. Los Angeles County (LAC) is home to an estimated 29% of the population of California [8], but is estimated to contain 41% of new HIV cases, and 76% of all AIDS cases in the state [9]. Sexual contact in MSM is estimated to account for 71.4% of the cumulative HIV cases in California males between 1983 and 2005 [10]; and, of all persons living with HIV and AIDS in this region at the end of 2009, 88% were male [11].

HIV is a global phenomenon, but two groups the infection continues to remain highly concentrated in – demographically, behaviorally, and geographically – are heterosexuals in South Africa, and in North American MSM. This thesis uses mathematical models to uncover population-specific patterns in these groups and examine questions related to prevention.

1.3 Acute HIV Infection

An HIV infected individual is most infectious during the first 6 to 12 weeks of infection [12, 13, 14, 15, 16, 17, 18]. While it is accepted that early detection of HIV during this phase is critical for infection management [19], the precise estimates for number of onward transmission events directly attributable to acute infection are variable [20]. Following this phase, there is a long period of “chronic infection,” that can last up to 10 years, and a final rise in infectivity before death [15, 16, 17, 18]. The precise duration of these phases and the relative probabilities of onward transmission during these phases are debated, but the importance of acute infection in onward transmission is well established.

Acute infection is a critical theme in all three chapters of this work. In the context of

circular migrations, if migrant men migrate back and forth between two locations while still acutely infected, onward transmission probabilities are much higher and consequently the expected number of transmission events is greater. Equivalently, uninfected migrants who travel frequently may be more likely to contact infected partners during the acute phase. Thus the frequency of migration itself is critical since this frequency determines at what stage of infection infected migrants may contact their partners (or vice-versa).

In Chapters 3 and 4 we focus on PDBC in MSM. It is documented that MSM reduce the level of their risky sexual behavior upon diagnosis to reduce risk to uninfected partners [2, 21, 3]. Thus diagnosis during the acute phase has the potential to reduce risky behavior during the most infectious phase in the course of infection, and therefore has enormous public health consequences. Naturally, the methods to diagnose infection early, especially during acute phase, have become critical in prevention. Thus acute infection plays a key role in understanding the population-level effects of PDBC in Chapter 3 and the various strategies we examine in Chapter 4 are centered around diagnosing infected individuals during this phase.

1.4 Sexual Behavior and HIV

HIV is an infectious disease that spreads from person to person. The selection forces by which people select partners (“selective mixing”) and the timing and distribution of those partnerships (“partnership timing”) are key components of understanding the dynamics of the disease [22]. Examples of selective mixing include relative attributes of partners including race, age, level of sexual activity, geographic location (in the context of mobile populations) and other attributes [22]. The timing of partnerships defines the risk structure of the population, and the connectivity of the population with regards to disease transmission (more on this topic in Section 1.5 where we discuss concurrency). For a long time, epidemiological study and data collection focused on individual-level attributes and behavior as predictors, and individual-level risk as outcome. Egocentric data – that report individual level behavior and attributes, and relevant information on attributes and history of behaviors the individual practiced with their partners – have made it possible to examine the problem of transmission at the level of the partnership, and the partner’s partners.

Besides egocentric data, other sampling (eg respondent driven sampling) methods to arrive incorporate partnership-level information are available [23]. Methodological developments in the form of software packages have made it possible to reliably estimate network structure using such egocentric data [24] and thus incorporate partner (or dyad) level attributes and information in epidemiological analyses. Indeed, we are now starting to see papers that combine egocentric data analysis and network modeling methods [25, 26, 27].

Thus partnership-level data and methods to analyze such data have made it possible to incorporate sexual behavior (and changing trends in such behavior) in studying questions of prevention. Sexual behavior is a persistent theme throughout the current work. In the context of migrations, sexual behavior of migrant men, due to differing norms is different from non-migrant men [28, 29]. The set of partners that migrants contact is also dependent on their spatial location and combined with periodic movement between two locations that leads to a unique sexual network structure. We discuss the epidemiological effects of this unique transmission structure, and how best to model these phenomena, extensively in Chapter 2.

Since PDBC is the primary focus of the work here on MSM, behavior (and change in risky sexual behavior) is at the core of the questions we are interested in. Moreover, these models are parameterized using the Southern California Acute Infection and Early Disease Research Program (AIEDRP) which is a longitudinal study that reports egocentric data on recently infected MSM [2, 3].

1.5 Concurrency and HIV

The timing of partnerships is a key component of understanding the overall structure of sexual networks. Broadly, we differentiate between “sequentially monogamous” and “concurrent” partnerships [22]. In the mid-1990s, a set of modeling papers established that populations with the same average number of partnerships but where one population engaged in sequentially monogamous partnerships, and the other where people engaged in concurrent partnerships experienced very different epidemics [30, 31, 32]. Subsequent papers that combine both modeling and data analysis [25, 26, 27] find two distinct mechanisms by which concurrency amplifies HIV spread. First, it doubles the number of “reachable paths” for

the disease (“path-doubling”); someone practicing sequential monogamy can “bridge” infection from their first partner to their second, but not vice versa, while someone practicing concurrency can do either [33]. Second, it speeds up the time by which such transmission chains can occur. That is, for someone practicing serial monogamy, if they are infected by their first partner they must wait until the end of that relationship and the formation of the next one before they can pass the virus onward; in the case of relational concurrency, that waiting period may be eliminated (“path-acceleration”). This latter effect is made especially important by the existence of acute HIV infection.

One criticism of the recent data-driven work on concurrency is that this work assumes equal amounts of sex with all partners (for people in multiple partnerships). In other words, the criticism is that this literature does not consider “coital dilution” [34]. The inherent activity and inactivity of partnerships in the context of circular migrations helps us consider this concern because disease transmission is only possible in partnerships where both partners are in the same location (“active” partnerships). The inactive partnerships are not able to transmit disease until the migrant partners return and these partnerships are re-activated. Thus Chapter 2 offers a special case to build a model that explores the question of concurrency as a driver in the HIV epidemic, when coital dilution is accounted for.

We dichotomize the partnerships of MSM in Chapters 3 and 4 as either “main” (those that involve significant emotional involvement) or “non-main.” The theme of concurrency is relevant because we model the distribution of the number of main partnerships, and the temporal overlap of one or more main and non-main partnerships at any given time.

1.6 HIV Modeling

As we noted earlier, mathematical models and computational methods serve as a scientist’s “laboratory” [1]. The value of the tools that models provide researchers has long been recognized. The earliest of these models adapted the system of ordinary differential equations (ODE’s) within the context of compartmental models (where individuals are aggregated based on some common characteristic, for example, infection status) and formalized their treatment [35]. Other modeling methods combined the compartmental framework with

stochastic dynamics of disease transmission, as in the Reed-Frost models (which themselves have a rich history) [36].

During the 1980's, compartmental models with ODE's were adapted to study dynamics of HIV transmission [37, 38, 39]. These models have greatly improved our understanding of various questions of HIV transmission [1] and continue to be used extensively to answer more recent questions in the field (such as the impact of anti-retroviral therapy) [40].

As we discussed earlier, formulation of transmission models needs explicit accounting of partner selection forces, and the timing and distribution of these partnerships. To represent this temporal overlap in partnership structure, and the distribution of these partnerships at any given time, we need a framework that allows us to model the actors and their ties individually. This need necessitated the development of network models that allow explicit representation of actors, their attributes, and the partnerships they form and dissolve. Indeed, the focus of an entire chapter in this dissertation (Chapter 2) is on comparing compartmental and network models.

Some network modelers have hypothesized a “preferential attachment” [41] process as underlying the formation of ties in networks [42]. This preferential attachment process, where a few nodes have many ties is represented using the power law distribution [42]. This power law has been used to model the distribution of ties in different contexts [41] and adapted to models of human sexual networks [42, 43].

One consequence of this power law distribution of partnerships is that the distribution of the number of contacts has infinite variance [44]. Jones and Handcock [44] explored the consequences of fitting the limiting distribution of this power law structure, the Yule distribution, to three different data-sets: Rakai-district (Uganda), Sweden, and USA [44]. They estimated the scaling parameter of the Yule distribution and found that in two of the three cases their estimates lie in the range where this idealized degree distribution has finite variance, (whereas the power law distribution predicts that the distribution of the number of partnerships has infinite variance) [44]. The estimate for American men is in the region where the degree distribution is infinite, but even in that case other distributions with finite variance fit the data better than the idealized Yule distribution [44, 45].

One reason the power law distribution fails is because the power law and the underlying

assumption of “preferential attachment” do not explain how human beings form partnerships [45]. The power law is a possible, even probable, mechanism for systems such as the size of cities (where highly populated cities attract more people) [46] or the citation of journal articles [46, 47]. Selection of sexual partners is a fundamentally different process, where people select their partners based on a number of criteria, including measurable ones such as age, race, class, and sexual-orientation [22]. This process of selective-mixing gives rise to complex networks, and we need to account for this complexity to develop realistic networks of transmission of HIV and other STIs. Moreover, formation of new partnerships by individuals is dependent upon whether they are currently in one (or more) partnerships. Thus there is inherent dependence in the structure of sexual partnerships [48]. This dependence poses additional challenges in the development of appropriate models.

Thus, while scale-free networks provide a succinct, parsimonious representation of networks, their structure is inappropriate to explain formation of sexual partnerships. Network models for HIV transmission need to be *grounded in data and behavioral and social theory*, and use tools that represent the measures implied by those data and that theory. Over the past decade, socio-behavioral scientists, statisticians, and mathematical modelers have collaborated to develop the framework of exponential random graph models (ERGMs, also known as p^* [49]). These models can represent sexual networks both in terms of mixing structures and dependence in ties [50].

Realistic models of human sexual networks also need to account for relational durations, formation and dissolution of partnerships, and the vital demographic processes of birth and death. With recent development in methodology, ERGM’s can now account for all of these considerations in a statistically principled fashion [51]. Software to implement such models is also available for use in epidemiology and other applications [24] and these models are now used to model HIV transmission [25, 26, 27]. We thus have a flexible framework that can incorporate realistic mixing structures to build models of transmission and thus answer questions that have proved intractable to date.

This is a “modeling dissertation” and all chapters in this work are based on gaining epidemiological insight using mathematical models. The chapter on circular migrations is a comparison of compartmental and network modeling to examine the impact of the unique

nature of temporal overlap on overall HIV transmission. The two chapters on PDBC in MSM use network models for their respective goals.

Chapter 2

**CIRCULAR MIGRATIONS AND HIV TRANSMISSION: A
COMPARISON OF COMPARTMENTAL AND NETWORK
MODELING***Coauthorship*

This is joint work with Dr. Dobromir T. Dimitrov (Fred Hutchinson Cancer Reserach Center) and Dr. Steven M. Goodreau (University of Washington). Dr. Dimitrov cross-checked the results from the compartmental models by coding them in Matlab (my code for these models is in R) and mentored me through the process of examining the differences between the two versions of compartmental models we have (especially the R_0 computations). Dr. Goodreau provided the basic idea of the project, and several key insights through the course of this work. Dr. Martina Morris (University of Washington) provided the initial idea of comparing comparmental and network models, though she is not a coauthor on this chapter.

Abstract

Background: Circular migrations – the periodic movement of individuals between two (or more) places [28], often between home and labor sites – is a relatively common practice in some countries of sub-Saharan Africa. The relationships between circular migration and HIV transmission are complex [52, 53], with circular migration providing multiple mechanisms that have the potential to amplify HIV transmission in a population.

Objectives: There are two primary goals of this paper; the first is to build multiple versions of our model using the compartmental modeling framework, and a network-based approach – to see how the models differ in their predictions for the relationship between migration frequency and HIV prevalence. Our second goal is to understand the dynamics of the interaction between acute infection and the frequency of migration and the impact of this interaction on HIV transmission.

Methods: We build Susceptible-Infected (S-I) models with births and deaths using ordinary differential equations, and a separate set of network models that adapt the framework of exponential random graph models (ERGM’s). We use biological data from [15] and [16] to parameterize the transmission components of these models. We test the relationship between migration frequency and HIV transmission at frequencies of 3 weeks and 30 weeks.

Results: We find that compartmental models show that there is no relationship between migration frequency and HIV prevalence. Under suitable assumptions, network models show that migrations every 3 weeks produced a bigger epidemic and statistically significant number of new infections (at the 95% level) than migrations at 30 week intervals. We also find that acute infection and its interaction with migration frequency are critical in sustaining these significant differences and demonstrate the contribution of delayed “path-acceleration” to this finding. We discuss the epidemiological implications of these results and the qualitative differences in the two modeling frameworks.

2.1 Introduction

Circular migrations – the periodic movement of individuals between two (or more) places [28], often between home and labor sites – is a relatively common practice in some countries in sub-Saharan Africa (for example South Africa [52, 54]) and the relationship between migrations and HIV has been the focus of much recent epidemiological work (South Africa [55, 56, 57], Kenya [58], Zimbabwe [59, 60], Tanzania [61]). The relationships between circular migration and HIV transmission are complex [52, 53]), with circular migration providing multiple mechanisms that have the potential to amplify HIV transmission in a population. Among these is the fact that some circular migrants have the opportunity to maintain ongoing sexual partnerships at their home location and other partnerships in their workplace [52].

A separate literature has used both data analysis [25, 26] and mathematical modeling [30, 31, 32, 62] to consider the impact of relational concurrency – the existence of partnerships that overlap in time – on HIV spread, relative to similar numbers of partnerships occurring strictly sequentially. This literature finds two distinct mechanisms by which concurrency amplifies HIV spread. First, it doubles the number of reachable paths for the disease (“path-

doubling”); someone practicing sequential monogamy can bridge infection from their first partner to their second, but not vice versa, while someone practicing concurrency can do either [33]. Second, it speeds up the time by which such transmission chains can occur. That is, for someone practicing serial monogamy, if they are infected by their first partner they must wait until the end of that relationship and the formation of the next one before they can pass the virus onward; in the case of relational concurrency, that waiting period may be eliminated (“path-acceleration”). This latter effect is made especially important by the existence of the HIV acute infection window – a period of 2-3 months immediately after infection when an individual is much more infectious to others than they are at subsequent times [15, 16]. Any phenomenon that increases the chances that people have an additional partner beyond the one from whom they obtained the infection during that period can be crucial in sustaining an epidemic.

In this paper, we develop a set of models to explore the relationships between the rates of circular migration for men with ongoing partnerships in each of two locations and HIV prevalence. The structure of concurrency in the context of circular migrations is unique because while migrants may have partners in multiple locations, the set of “active” partnerships is restricted to those in which the partners are in the same location as them (We define partnerships with both partners in the same location as “active”; if the migrant is in a different location from their partner the partnership is “inactive”). This unique form of concurrency allows us to build models that address some of the prior concerns about the modeling of concurrency (more details are in Section 2.2).

We do not suggest that this inherent uniqueness in the structure of concurrency is the only means (or even the chief means) by which circular migration affects HIV transmission, but we limit our exploration to this form for the purposes of this work. At this stage, then, the work is largely conceptual and theoretical; parameterizing our model using more empirical data constitutes a next step. Our hypothesis is that all else being equal, migrations at intervals shorter than the window of acute infection (“frequent” migrations) will result in higher HIV transmission than migrations at intervals longer than the length of acute infection (“infrequent” migrations). The magnitude of the difference in transmission between frequent and infrequent migrations will help us disentangle the potential roles of

path doubling and path acceleration in amplifying HIV spread. Adaptations of these models are being developed to answer more contemporary questions in the epidemiology of HIV [40, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73].

The main modeling approach used in the HIV literature is compartmental, or ordinary differential equation (ODE), modeling [74]. Compartmental models have a rich history in what they have helped us understand about HIV transmission and continue to be the “work-horse” of model development [1]. Although this approach has many strengths, its ability to represent the full impacts of relational interactions is limited, given its fundamental approach of aggregating individuals and their relationships. An alternate approach is to use network methods that allow explicit representation of actors, and their attributes and partnerships. We hypothesize that given the novel nature of the underlying partnership structure (given potentially concurrent partnerships of migrant men in two separate locations) such explicit accounting of actors and their partnerships is necessary to model the effect of circular migrations on HIV transmission.

There are thus two primary goals of this paper; the first is to build multiple versions of our model – using the compartmental modeling framework, and a network-based approach – to see how the models differ in their predictions for the relationship between migration frequency and HIV prevalence. Our second goal is to understand the dynamics of the interaction between acute infection and the frequency of migration and the impact of this interaction on HIV transmission.

2.2 Background

2.2.1 Circular Migrations and HIV in South Africa

Of the 33 million HIV-infected people in 2009, South Africa alone had 5.6 million [75]. The prevalence of HIV in South Africa increased from 0.76% in 1990 to 26.5% in 2002 [52], and was reported at about 17.8% in 2009 [76]. South Africa has had a history of circular migrations, and migration in general has been an important component of the economic and political structure for the past century [54, 77, 78]. As early as 1886, there was demand for migrant labor in the gold mines of Transvaal [77]. Most of the gold was at great depths, and

labor, especially low-cost labor, was urgently needed. These mines in South Africa started bringing in workers from neighboring countries. To keep costs of housing low, and because black people were not allowed to live permanently on these lands anyway, single-sex hostels were established [77].

However, patterns of migration have changed over the past three decades [79], especially after apartheid was lifted in 1994 [54]. Currently most labor migration is for a longer period of time, and workers are generally allowed to visit home at shorter intervals. Approximately 90% of black laborers in the mines are migrants, with restricted opportunities to visit family, and a low likelihood of visits by their partners [54]. The Hlabisa district of the Kwa-Zulu Natal province is a major migrant sending location; the coastal mining town of Richards Bay and the inland town of Carletonville are two locations where men migrate to. Richards Bay is only about 100 km from Hlabisa, whereas Carletonville, at a distance of 700 km, is farther away [57]. Consequently, migrants from Richards Bay are able to return home more frequently than migrants from Carletonville, and the frequency of migrant travel therefore becomes an important factor in transmission dynamics of HIV [57].

Thus migrant labor continues to constitute an essential component of the South African economic structure.

Migrations and HIV Transmission

The linkage between human population mobility and infectious disease dispersal is well established in epidemiological literature [80]. Oscillatory movements lead to weakening of the family structure and the associated network of social and psychological support, leaving people more willing to engage in high-risk sexual behavior [77]. However, the precise mechanism underlying this connection between HIV and mobility is not completely understood [52]. The general consensus seems to be that migrants are more likely to engage in such high-risk behaviour due to leaving the normative lifestyle of home [28, 29] in the following ways [80]: (i) migrants may change partners more frequently, (ii) migrants may form partnerships with individuals who are at a high risk of infection, such as commercial sex workers, (iii) migrants may have a higher number of relationships that overlap in time. One

study from the Limpopo province in South Africa has also found migrants to have a very low perceived risk [81]. It was hoped that high-risk sexual behavior amongst migrant men might reduce With improvements in roads and the travel infrastructure; however, research from the Limpopo province has found no evidence to support this hypothesis [81].

Of the nine provinces in South Africa, Kwa-Zulu Natal has had the highest HIV prevalence (39.5% in 2009) [82]. In this province, migration is an historically important part of the local economy [52]. Moreover, several studies have found migrants to have a higher rate of HIV prevalence than non-migrants [52, 54, 55, 83, 84]. Higher prevalence rates among migrants have been observed in other African countries, notably Uganda, Zimbabwe, Senegal [52, 54, 55], and Kenya [58]. However, other studies in Zimbabwe [59, 60] have found no link between migration/mobility and HIV. Another study in Tanzania [61] found no significant differences in the sexual behavior of migrant and non-migrant men [53]. Migration is not a homogeneous process; when studying migration and its epidemiological implications, it is imperative to clearly define the pattern of mobility, and the populations at risk [53].

An important study of the differences in HIV prevalence among migrants and non-migrants in South Africa was conducted in the Kwa-Zulu Natal province of South Africa [56, 57]. In the Hlabisa district of the province, 62% of adult men migrate for employment, mostly to the towns of Carletonville and Richards Bay. The two places represent different patterns of migration, with the former being farther away from the rural area of Hlabisa, and allowing the migrants relatively fewer trips home. Richards Bay, on the other hand, is much closer, allowing migrants to return home more often [55, 57, 83, 84]. It was found that migrant men in Carletonville were 3.2 times as likely to be infected as their rural partners. In Richards Bay migrant laborers were as likely to be infected as their rural partners, but migrant couples were twice as likely to be infected as non-migrant couples, and in one-third of cases, the woman was the infected partner [83]. This finding is critical to our understanding of the impacts of circular migration on HIV transmission, because it established that it is not just the migrant laborer who is acting as a transmission agent between two otherwise disconnected areas, but that the non-migrating women are often the infected partner [56].

2.2.2 *Circular Migrations and Concurrency*

Concurrent sexual partnerships are those that temporally overlap in time. Through simulation studies, Morris and Kretzschmar [30, 31, 32] showed the impact of concurrent partnerships, i.e. relationships that overlap in time, on the spread of HIV. The recent development of tools [24] to incorporate relational concurrency in mathematical models in more systematic and generalizable ways has led to data-based work on the examination of the effects of concurrency on HIV transmission [25, 26].

Circular migrations are unique because of the underlying nature of temporal overlap in partnerships. While migrant men may have partners in both the home village and the workplace, only one of these partnerships (or set of partnerships) can be “active” – i.e. these partnerships involve ongoing coital acts – at a given time. Migrants often return to the same partners in the home locations (due to the presence of their families there) [79] and may practice risky sexual behavior in the work location [79]. Thus the location of a migrant decides which partnerships (or set of partnerships) are active.

While this structure appears similar to “serial monogamy” [85], it is different from a monogamous structure in one key respect, i.e. the migrant men return to the same partner(s) in at least one (or both) locations. The temporal overlap resulting from this periodic activation and deactivation result in a unique form of concurrency.

In Figure 2.1, we consider a migrant man who travels back and forth between two locations A and B, and has one partner in each location. When he is in location A, his partnership in that location is active. Similarly, when he is in location B, only his partnership in location B is active. The non-migrant man, on the other hand, stays in one location only, and the partnership stays active for the entire duration of the partnership. Thus while the migrant man has twice as many partners as the non-migrant man (path-doubling), both men have the same amount of sex.

Figure 2.1 shows that the coital acts of the migrant man are distributed over two partnerships in separate locations, and coital acts are contingent upon location by design. One criticism of the recent work on concurrency is that these models assume daily sex with every partner and thus do not account for “coital dilution” in the presence of multiple overlapping

partnerships [34]. The inherent activation and deactivation of partnerships in the context of circular migrations helps us consider the criticism of modeling coital dilution because disease transmission is only possible in active partnerships. The inactive partnerships are not able to transmit disease until they are reactivated. Thus the current work offers a special case to build a model that explores the question of concurrency as a driver in the HIV epidemic, when coital dilution is accounted for. We do not suggest that this inherent uniqueness in the structure of concurrency is the only means (or even the chief means) by which circular migration affects HIV transmission, but we limit our exploration to this form for the purposes of this work.

2.2.3 Circular Migrations and Acute HIV Infection

Based on data presented by Wawer et al. [15], Hollingsworth et al. have shown that the infectivity of an individual after sero-conversion is highest for a period of about 2.9 months after infection [16]. This period of acute infection is followed by a long period of stable chronic infectivity and a final, late-stage rise before death [15, 16]. Since HIV has a long latency period, acute infection plays a key role in the context of circular migrations. Infected migrant individuals may migrate between two locations without knowledge of their infection status while they are acutely infected, and thus put their partners (and their partners' partners) at risk of infection. Conversely, uninfected migrant men who are frequently traveling between their village and the workplace may be at a greater risk to contract infection from an acutely infected partner, and subsequently put their partners (and their partners' partners) in the other location at greater risk. If, on the other hand, travel occurs at intervals longer than the period of acute infection, then the likelihood of onward transmission due to acute infection may be greatly reduced.

The aim of this project is to make the relationship between the interaction of circular migrations, acute HIV infection, concurrent sexual partnerships and consequent HIV transmission more precise. To do so we use both the classical framework of compartmental models (formulated using ordinary differential equations or ODE's) and network-based models that have recently been developed [24]. We use these tools to study the impact of

circular migration frequency on HIV prevalence and compare the features and predictions of the two frameworks.

2.3 Methods

We begin by describing the features common to all of our models, and then proceed to describe the particular features of compartmental models in Section 2.3.1 and network models in 2.3.2. We classify individuals (also called actors) in the population based upon their infection status, migrating status, sex, and current location. The population is equally divided between two locations: a rural “village” and an urban “workplace.” We assume that some, but not all, men, migrate; we do not consider female migration in this model. As time evolves, the migrant men in the workplace and village change their locations. Thus the population consists of six classes when considering migration status (whether someone is a migrant), sex, and current location. The epidemiological papers [56, 57] we consider to parameterize our models only present data on migrations by men; therefore, these modeling assumptions we make are to be consistent with this underlying epidemiological literature from KwaZulu-Natal.

There are four states of infection: susceptibility, acute infection, chronic infection, and late-stage infection. We denote these states by S , A , C and L respectively. Thus we describe a state by the notation S_{XYZ} , A_{XYZ} , C_{XYZ} or L_{XYZ} . For the subscript, X denotes the migration status of the individual (migrant M or non-migrant N); Y denotes the sex (male M or female F); and Z denotes the location (urban workplace U or rural village R). Thus, for example, S_{MMU} represents a susceptible migrant male in the urban area. Since females do not migrate in the model, the first subscript N for females is redundant, but we include it for symmetry.

Our four infectious states and six different classifications yield a total of 24 state variables. Transitions are of four types: arrivals (due to fertility), departures (natural and disease-related mortality), migrations (at periodic intervals), and infections (progression from the susceptible to acute, chronic, and late-stage infection). Infections occur when members of one of the susceptible classes contract HIV upon contact with an infected partner in the same location.

We assume constant and uniform rates of migration (for migrating men only), and for progression through disease stages, background mortality, and AIDS-specific mortality (for those in late-stage infection only). Introductions due to fertility in the population are also assumed to be constant.

Figure 2.2 shows a schematic for the system.

All simulations are run in the **R** programming language: for compartmental models we use the `odesolve` package [86] (with examples as shown in [87]) and for network models we utilize the `statnet` package [24].

2.3.1 Compartmental Models

In compartmental epidemic models of HIV/AIDS, a contact is typically defined in one of two ways: as the initiation of a partnership or as an individual sex act. The “contact-as-partnership” approach aggregates the coital acts over the course of a partnership. The “contact-as-act” approach, defines a contact as a sexual episode. In the contact-as-partnership approach, contact rates are relatively low. In contrast, the probability of transmission given contact is relatively high, since it reflects the probability of transmission over the entire course of the relationship. The contact-as-act approach has low transmission probability per contact but a high partner-change rate. We denote this contact rate by t instead of the more common c to avoid confusion with the chronic stage of infection (for which we use C).

In this paper, we explore both interpretations of contact using compartmental models.

The probability that a susceptible individual becomes infected at any time point depends upon [38, 88]:

1. the average number of contacts;
2. the probability that their partner is in one of the three infected states. We assume that all individuals in the population have the same amount of sex in both interpretations of contacts. In the contact as partnership model, migrant men have twice as many partners per unit time as non-migrant men, but half as much sex. We further assume that conditional on these rates, mixing is random with regard to serostatus within

location. Thus, the probabilities that a susceptible individual chooses a partner in each of the three stages is simply: $\frac{A_{XYZ}}{N_{XYZ}}$ for acutely infected individuals, $\frac{C_{XYZ}}{N_{XYZ}}$ for chronically infected individuals, and $\frac{L_{XYZ}}{N_{XYZ}}$ for infected individuals in the late-stage where N_{XYZ} is the total number of individuals of group XYZ ;

3. the probability of acquiring infection per coital act from an infected partner over the course of the relationship (or in a single act depending upon the interpretation of contact being used), described by $\beta_{..}$. The first subscript for $\beta_{..}$ describes the stage of infection of the positive partner at the initiation of the partnership – Acute (A), Chronic (C) or Late-Stage (L) – and the second subscript describes whether the partnership involves a migrant (M) or non-migrant (N) male. The latter is relevant because it affects the number of coital acts that a man will have with his partner during the relationship, given that a migrant man is only present for part of the time during the course of the partnership (more details below).

Note that the two interpretations of contact do not change the structure of differential equations in the model; it is the probabilities of transmission over the course of a “contact” (β) and the contact rate t that are different. The β parameter comes from published data [15, 16]. The rate parameter t is selected based on qualitative descriptions of data in [56, 57].

Tables 2.1 and 2.2 show the demographic parameters and initial values of the state variables that remain constant for either interpretation. The system of equations (described below in Section 2.3.1) is simulated over approximately 100 years with the parameter values in Table 2.2. Each time step represents 1 week. Tables 2.4 and 2.3 provide values for the biological and behavioral parameters that are different for the contact-as-partnership and contact-as-act models respectively. These values are selected so the number of coital acts is the same across the two models, and in the mean per-act transmission probability.

Differential Equations

To assemble these pieces, we begin by considering susceptible migrating urban males. We describe the rate of change in this population (with one week taken as one time unit) as

$$\begin{aligned} \frac{dS_{MMU}}{dt} = & \nu/8 - S_{MMU}t_{MMU} \frac{A_{NFU}}{N_{NFU}} \beta_{A,M} - S_{MMU}t_{MMU} \frac{C_{NFU}}{N_{NFU}} \beta_{C,M} - \\ & S_{MMU}t_{MMU} \frac{L_{NFU}}{N_{NFU}} \beta_{L,M} - \delta S_{MMU} + \delta S_{MMR} - \mu S_{MMU}, \end{aligned} \quad (2.1)$$

where δ is the rate of migration between the urban and rural area. The parameter ν is the number per unit time that enter the entire population. Since urban migrant men account for one-eighth of the total population, the number of new arrivals per unit time in this group is $\nu/8$. The parameter μ denotes the rate of natural mortality.

The change in population of acutely infected migrating urban males is

$$\begin{aligned} \frac{dA_{MMU}}{dt} = & S_{MMU}t_{MMU} \frac{A_{NFU}}{N_{NFU}} \beta_{A,M} + S_{MMU}t_{MMU} \frac{C_{NFU}}{N_{NFU}} \beta_{C,M} + \\ & S_{MMU}t_{MMU} \frac{L_{NFU}}{N_{NFU}} \beta_{L,M} - \delta A_{MMU} + \delta A_{MMR} - \\ & \gamma_1 A_{MMU} - \mu A_{MMU}, \end{aligned} \quad (2.2)$$

where γ_1 is the rate at which an acutely infected individual becomes chronically infected.

The change in the population of chronically infected migrating urban males is

$$\frac{dC_{MMU}}{dt} = \gamma_1 A_{MMU} - \gamma_2 C_{MMU} - \delta C_{MMU} + \delta C_{MMR} - \mu C_{MMU}, \quad (2.3)$$

where γ_2 is the rate of chronic to late-stage transition. The change in population of late-stage migrating urban males is

$$\frac{dL_{MMU}}{dt} = \gamma_2 C_{MMU} - \delta L_{MMU} + \delta L_{MMR} - (\mu + \mu_d) L_{MMU}. \quad (2.4)$$

where μ_d is the rate of mortality due to the disease.

Now we consider the females in the urban area. An urban female can become infected either by an infected male in one of the three infectious states, who is either a migrant or a non-migrant. Therefore we have

$$\begin{aligned} \frac{dS_{NFU}}{dt} = & \nu/4 - S_{NFU}t_{NFU} \frac{A_{MMU}}{N_{MMU}} \beta_{A,M} - S_{NFU}t_{NFU} \frac{C_{MMU}}{N_{MMU}} \beta_{C,M} - \\ & S_{NFU}t_{NFU} \frac{L_{MMU}}{N_{MMU}} \beta_{L,M} - S_{NFU}t_{NFU} \frac{A_{NMU}}{N_{NMU}} \beta_{A,N} - \\ & S_{NFU}t_{NFU} \frac{C_{NMU}}{N_{NMU}} \beta_{C,N} - S_{NFU}t_{NFU} \frac{L_{NMU}}{N_{NMU}} \beta_{L,N} \\ & - \mu S_{NFU}, \end{aligned} \quad (2.5)$$

$$\begin{aligned}
\frac{dA_{NFU}}{dt} = & S_{NFU}t_{NFU,M} \frac{A_{MMU}}{N_{MMU}} \beta_{A,M} + S_{NFU}t_{NFU,M} \frac{C_{MMU}}{N_{MMU}} \beta_{C,M} + \\
& S_{NFU}t_{NFU,M} \frac{L_{MMU}}{N_{NMU}} \beta_{L,N} + S_{NFU}t_{NFU,NM} \frac{A_{NMU}}{N_{MMU}} \beta_{A,N} + \\
& S_{NFU}t_{NFU,NM} \frac{C_{NMU}}{N_{NMU}} \beta_{C,N} + S_{NFU}t_{NFU,NM} \frac{L_{NMU}}{N_{MMU}} \beta_{L,N} - \\
& \gamma_1 A_{NFU} - \mu A_{NFU},
\end{aligned} \tag{2.6}$$

$$\frac{dC_{NFU}}{dt} = \gamma_1 A_{NFU} - \gamma_2 C_{NFU} - \mu C_{NFU}, \tag{2.7}$$

and

$$\frac{dL_{NFU}}{dt} = \gamma_2 C_{NFU} - (\mu + \mu_d) L_{NFU}, \tag{2.8}$$

to describe the various interactions of females in the urban area.

The interactions of the non-migrant males are similar to those of the migrant males defined in equations (2.1) to (2.4), without the migration term. Thus,

$$\begin{aligned}
\frac{dS_{NMU}}{dt} = & \nu - S_{NMU}t_{NMU} \frac{A_{NFU}}{N_{NFU}} \beta_{A,N} - S_{NMU}t_{NMU} \frac{C_{NFU}}{N_{NFU}} \beta_{C,N} - \\
& S_{NMU}t_{NMU} \frac{L_{NFU}}{N_{NFU}} \beta_{L,N} - \mu S_{NMU},
\end{aligned} \tag{2.9}$$

$$\begin{aligned}
\frac{dA_{NMU}}{dt} = & S_{NMU}t_{NMU} \frac{A_{NFU}}{N_{NFU}} \beta_{A,N} + S_{NMU}t_{NMU} \frac{C_{NFU}}{N_{NFU}} \beta_{C,N} + S_{NMU}t_{NMU} \frac{L_{NFU}}{N_{NFU}} \beta_{L,N} - \\
& \mu A_{NMU} - \gamma_1 A_{NMU} - \mu A_{NMU},
\end{aligned} \tag{2.10}$$

$$\frac{dC_{NMU}}{dt} = \gamma_1 A_{NMU} - \gamma_2 C_{NMU} - \mu C_{NMU} \tag{2.11}$$

$$\frac{dL_{NMU}}{dt} = \gamma_2 C_{NMU} - (\mu + \mu_d) L_{NMU} \tag{2.12}$$

describe the changes in population for the four infection states of urban non-migrant males.

The basic structure of equations (2.1) to (2.12) is the same for the rural area with the location sub-script U replaced by R (for rural). Thus, we have a system of 24 interacting equations.

Infection Model

Wawer et al. (2005) [15] provide estimates of the per-act transmission probability during different stages of infection, and Hollingsworth et al. [16] provide estimates of the relative transmission probabilities of these various stages. We combine these two measures by taking the per-act chronic transmission probability as 0.0007 from [15] and the acute and late-stage transmission probabilities relative to the probability during chronic stage from [16] (since the duration of the various stages provided by Wawer et al [15] are debated [16, 17, 18]) and Hollingsworth et al. do not provide per act probabilities. The acute and late stage probabilities are then 26 and 7 times the chronic stage probabilities [16]. The parameter γ_1 is taken as $1/12$, so that the mean duration of acute infection is 12 weeks, which is in accordance with Hollingsworth (2008) [16]. Due to the structure of this ODE model, therefore, the proportion of acutely infected individuals that flow in to the chronic phase per unit time has an exponential distribution with a mean of 12 weeks. The parameter δ is varied to investigate the impact of varying migration rates on the dynamics of disease transmission. We consider the mean duration of the chronic stage to be 500 weeks (approximately 9.6 years) hence γ_2 is $1/500$. The parameter μ_d is $1/40$ fixing the mean duration of the late-stage at 40 weeks [16].

In the interpretation of contact as partnership, we take a hypothetical value of 100 weeks as the average duration of partnerships. The per-act transmission probabilities are converted to the partnership-level and are given in Table 2.4. To convert these probabilities, the constants $\beta_{A,M}$ and $\beta_{A,N}$ represent the probabilities that infection is transmitted during the course of a partnership when the partnerships involve acutely infected migrant and non-migrant men respectively. If p_i is the probability of transmission per coital act during the stage i of infection, then

$$\beta_{i,M} = 1 - (1 - p_i)^{d \times n/2} \quad (2.13)$$

and

$$\beta_{i,N} = 1 - (1 - p_i)^{d \times n} \quad (2.14)$$

where d is the average duration of the partnerships and n is the number of coital acts per time unit (one week in this case). The parameterization is such that while migrant men have twice as many partners per unit time as non-migrant men, the amount of sex men in both groups have is the same. That is, coital dilution is complete, since overall coital frequency for a man does not depend on his number of partners.

In the contact-as-act approach, β simply represents the per-act transmission probabilities as shown in Table 2.3.

Disease is introduced in the population by infecting fifty urban and fifty rural women infected at the start of the simulation. We assume that the proportion of infectives in each infection stage is equal to the length of that stage relative to the total time of infection.

The initial values for S_{NMU} , S_{NFU} and S_{MMU} are 1250, 2450 and 1250 respectively. We investigate migration frequencies of 3 and 30 weeks. We select these particular migration frequencies for two reasons; the first is to be consistent with the pattern of migration reported in the literature – that migrants in work-locations closer to home tend to return home *at least* once a month and those working farther away tend to return *at most* 3 or 4 times a year [52]. Secondly, given the relevance of acute infection to this project, our goal is to compare prevalence curves obtained at rates of migration that are shorter and longer than the acute phase of HIV with each other.

We also investigate a number of other migration frequencies, shorter (1, 4, 7 and 10 weeks) and longer (20, 40, 60, 80 and 100 weeks) than the acute phase. Collectively, we refer to migrations at intervals shorter than the length of acute HIV infection as “frequent migrations.” Conversely, migrations at intervals longer than the length of acute infection are “infrequent migrations.”

To examine the mathematical properties of epidemic potential in the two interpretations of contact, we present reproduction numbers R_0 for both interpretations. We perform these computations following an algorithm in [89]; details are in Appendix A.

Conservation of Sexual Acts

The parameter values in the contact-as-partnership case Tables 2.4 are chosen to reflect some basic characteristics of human sexual behaviour. The average number of partners per week for migrant men is greater than that for non-migrant men, though the average number of coital acts for migrant and non-migrant men is the same. This assumption is in accordance with the observation that migrant men are exposed to less stringent social norms, and generally have more sexual partners [52, 54], but because migrant men do not have more sex acts than non-migrant men we are also able to incorporate the effects of coital dilution. With the given parameter values we then have

$$\begin{aligned} t_{MM}N_{MM} &= t_{NF,M}N_{NF} \\ t_{NM}N_{NMU} &= t_{NF,NM}N_{NF} \end{aligned}$$

where $t_{NF,M}$ is the contact rate between women and migrant men and $t_{NF,NM}$ is the contact rate between women and non-migrant men. These relationships imply conservation of sexual contacts i.e. the total contacts of migrant and non-migrant males must equal the contacts of their female partners in either location. Thus migrant men in the contact-as-partnership version have twice as many partners, but since only half of those partnerships are active at any given time, the number of sexual acts for migrant and non-migrant men are equal.

Fertility and Mortality

We consider two types of mortality: natural (μ) and mortality due to the disease (μ_d). Thus $1/\mu$ is the average sexual lifespan of an uninfected individual. We assume that an uninfected individual will remain sexually active from the age of 15 years to the age of 60 years; thus setting $\mu = 1/(45 \times 52)$ sets the sexual lifespan of an individual to (45×52) weeks, or 45 years. The average lifespan of an infected individual, is 552 weeks, comprising acute, chronic and late stages that last on average for 12, 500 and 40 weeks respectively [16].

To solve for the number of arrivals in the population per unit time, we set equation (2.1)

equal to 0 in the disease free state. Then

$$\frac{dS_{MMU}}{dt} = \nu/8 - \mu S_{MMU} = 0$$

implying

$$\nu/8 = S_{MMU}\mu = \frac{625}{45 \times 52}. \quad (2.15)$$

Therefore, there are $625/(45 \times 52)$ individuals that enter the sub-population of susceptible migrant men in the urban location per week. The other sub-populations of men have the same number of arrivals per week, and each of the two sub-populations of women have $1250/(45 \times 52)$ individuals that arrive per unit time.

Equivalence of the two interpretations of contact

We base our parameterization on a total of 3000 partnerships in the population. Of these, 2000 are accounted for by migrant men, and 1000 are accounted for by non-migrant men. Since there are 1250 total non-migrant men in the population, the active “mean degree” (average number of active partnerships per person) of non-migrant men is $1000/1250=0.8$. On average, only half of the partnerships of migrant men are active at any given time – therefore the 1250 migrant men in our population account for 1000 active partnerships. The active mean degree for migrant men, therefore, is $1000/1250=0.8$. Since the active mean degrees for both migrant and non-migrant men is 0.8 – the active mean degree for the population is 0.8. Thus there are 2000 active partnerships in the population on average.

Since we assume 3 sexual acts per week over all time-steps, we have an average of $0.8 \times 3 = 2.4$ sexual acts that occur over active partnerships for both migrant and non-migrant men, and women. Our “contact-as-act” model is parameterized with 2.4 sexual acts per week to keep the two scenarios identical with respect to this key feature.

2.3.2 Stochastic Network-Models

Partnership and Demographic Structures

We now create alternate models of this system that explicitly account for the individuals in the population (instead of aggregated counts), their attributes (sex, location, migration

status, infection status, time of infection), relational timing, and temporal overlap in partnerships. These models are based on graphs, and we adopt graph-theoretic concepts and terminology. Consequently, individuals are “actors”, and partnerships are “ties.” We account for all men and women in the population, and the set of their ties. This undirected graph, at any given time, is what we call a “cross-sectional network,” or simply, a “network,” for brevity.

The cross-sectional network is then evolved over time to model the process of formation of new partnerships, dissolution of existing partnerships, and arrivals and departures from the population due to fertility and mortality. This process yields a temporal series of networks over discrete time steps (henceforth called a “network series”). We then simulate HIV transmission on these temporal networks.

An important point to note is that if individuals A and B are in a partnership, they do not necessarily have to be in the same location. For example, A might be a migrant man, married to, or in a long-term relationship with B . While B is fixed in one of the two locations throughout the simulation, A is changing his location periodically. To model disease transmission, however, we need to evaluate if A and B are in the same location; as defined earlier, if they are in the same location, their partnership is “active;” otherwise, this partnership is “inactive.”

We use the framework of separable-temporal exponential random graph models (ERGMs) of Krivitsky et al. [51] to model the partnership structure in the population. This method requires separate models that govern the process of formation and dissolution models.

The first model (described below) is set up to be as similar to our compartmental model as possible. We then describe additional models that utilize the possibilities of network models (described below in Section 2.3.2).

Basic Network Models

The formation model is

$$\text{logit}(p(y_{ij,t=1}|y_{ij,t=0} = 0)) = \theta\delta(e) + \sum_{p_1} \sum_{p_2} \theta_{m,p_1,p_2} \delta(mp_1, p_2) \quad (2.16)$$

where y_{ij} is an edge between the pair of persons i and j ; $y_{ij} = 1$ represents the existence of a partnership and e is the total number of partnerships in the network. The term θ_{m,p_1,p_2} is the coefficient associated with mixing between partners p_1 and p_2 . The terms p_1 and p_2 represent either migrant or non-migrant men, and women. There are several structural zeros: any partnership where p_1 and p_2 are the same sex is a structural zero, and if p_1 is a non-migrant man in one location, then the number of his partnerships with women in the other location will be zero (and vice-versa). The δ functions are “change statistics,” defined as changes in the values of these statistics associated with a toggle of a dyad (i.e. switching the value of any dyad) [51]. Migrant men form partnerships in both locations, and therefore these terms are not structural zeros.

Migrant men spend a given amount of time (on average) in one location. The inverse of this quantity gives the weekly probability of migrating to the other location in any particular week (time steps are in weekly units). Migration to the other location is thus a binary event. This model is qualitatively identical to the deterministic model described above, with respect to all key features of the process: (i) the structure of partnerships, (ii) the process of migrations, and (iii) the infection model.

The dissolution model is much simpler, where dissolution of each tie is a Bernoulli process, with the probability of dissolution determined by the average duration of partnerships. This dissolution model is

$$\text{logit}(p_{y_{ij,t=1}|y_{ij,t=0}=1}) = \theta_{\text{diss}} \quad (2.17)$$

where θ_{diss} is the log-odds of the probability that a partnership persists at a given time-step.

Restricted Network Models

A key feature of the circular migration system is that a migrant man who has two ongoing partnerships should be more likely to have one ongoing partnership in the village and one in the workplace than both in the same place. Our network model allows us to introduce this pattern in ways that ODE models do not. We refer to this as the “restricted model” because the partnership structure of migrant men with multiple partners is restricted to exclude multiple partners in the same location.

We restrict the partnership structure by adding additional terms to the model in (2.16). These terms restrict partnerships between a migrant man and multiple women in the same location. The consequences of this constraint are that no migrant man has more than two partners in total, and for migrant men with multiple partners one partner is in the urban and the other is in the rural location.

This restricted partnership model is

$$\begin{aligned} \text{logit}(p(y_{ij,t=1}|y_{ij,t=0} = 0)) &= \theta\delta(e) + \sum_{p1} \sum_{p2} \theta_{m,p1,p2} \delta(mp1,p2) + \\ &\theta_{m_1,\{m_{21},m_{22}\}}^{\text{urban}} \delta^{\text{urban}} m(m_1, \{m_{21}, m_{22}\}) + \\ &\theta_{m_1,\{m_{31},m_{32}\}}^{\text{rural}} \delta^{\text{rural}} m(m_1, \{m_{31}, m_{32}\}) \end{aligned} \quad (2.18)$$

where m_1 is a migrant-man, m_{21}, m_{22} are two distinct rural women and m_{31}, m_{32} are two distinct urban women. We thus restrict concurrent partnerships between a migrant man and multiple women in the same region by setting the statistics defining these δ terms as structural zeros. (The technical term for this parameter is a 2-star [90].)

The dissolution model here is identical to the dissolution model in the basic network model in (2.17).

We consider 5000 nodes at the beginning with 3000 total ties – the same number we chose in the analysis with compartmental models. Migrant men account for twice as many ties as non-migrant men, but both groups have equal numbers of active partnerships. We have an average of 2.4 sex acts per week over active partnerships. The average duration of partnerships is 100 weeks.

We also model migration as a step function where migrant men switch locations at fixed intervals rather than as probabilistic events (where the probability of migrations defined as the reciprocal of the specified time to be spent in each location) since this assumption might be more consistent with the process. The reason we investigate probabilistic migrations earlier is because that is how they are defined in the compartmental framework and we wanted identical models at the start of the simulation in both frameworks.

Infection Model

Disease is then simulated over the “network series” which consists of networks simulated over weekly time steps. The process of epidemic simulation is identical to the process for compartmental models. There are 50 infected females in the rural and urban areas each. Their time since infection is randomly selected between 0 and 552 weeks, the total life-span of an infected individual. Acute, chronic and late-infection last 12, 500 and 40 weeks respectively, and the probabilities for onward transmission per act are identical to those selected for the compartmental models.

As with the compartmental models, we vary the average time spent at each location between migrations and plot prevalence curves for the entire population. We select migration intervals of 3 weeks and 30 weeks to be consistent with the compartmental model. To account for the stochastic nature of these models, we conduct 10 repetitions of each experiment, in both the basic and the restricted network models.

2.4 Other Analyses

We perform other analyses to

(i) evaluate the impact of acute infection on HIV transmission in Appendix B, (ii) examine the restricted network structure but with stochastic migrations (instead of clockwork) migrations in Appendix C, (iii) to evaluate the reasons for differences in the two models, compare cumulative and momentary degree distributions in basic and restricted network models (Appendix D), and, (iv) tease apart the effects of delayed path acceleration and momentary degree distributions on the observed differences in HIV prevalence in the basic and restricted models in Appendix E.

2.5 Results

2.5.1 Compartmental Models

In Figure 2.3, we see that when contact is defined as a partnership, a small epidemic is produced (equilibrium prevalence: 1.6%) for any migration interval. When contact is defined as a sexual episode, we see that a large epidemic is quickly produced (equilibrium

prevalence: 36.7%), but the prevalence trajectories for different migration rates are identical.

We also find that the reproduction number R_0 in the “contact-as-partnership” and “contact-as-act” cases is 1.19 and 1.58 respectively.

2.5.2 Network Models

In Figure 2.4, the basic network models show that there is no statistical difference between frequent and infrequent migrations in the basic network models. Ten experimental repetitions showed a mean prevalence of 26.9% (95% CI: 25.7%, 28.2%) after 5000 weeks at an average migration frequency of 3 weeks. The average migration frequency of 30 weeks produced a mean prevalence of 25.2% (95% CI: 24.0%, 26.3%) after 6000 weeks. Also with migrations at an average interval of 3 weeks produced 6793 (95% CI: 6684, 6902) new infections and the migration frequency of 30 weeks produced 6680 (95% CI: 6589, 6771) new infections over 5000 weeks (both averaged over 10 experimental repetitions).

When the partnership structure for migrant men is restricted and migrations occur as deterministic step-functions, migrations every 3 weeks produced a prevalence of 5.7% (95% CI: 5.3%, 6.1%) and migrations every 30 weeks showed a final prevalence of 4.2% (95% CI: 3.9%, 4.5%) after 6000 weeks. In the former case, we saw 3735 (95% CI: 3613, 3857) new infections over 6000 weeks and 3132 (95% CI: 3019, 3245) new infections in the latter scenario. Restricted network models with stochastic migrations (Appendix C) show the same results qualitatively.

2.6 Discussion

The compartmental models show that the migrations do not impact the rate at which HIV transmits through the population. Network models show the same result when partnership structure of migrant men is not restricted i.e. multiple partners of migrant men can be in either location. When partnerships of migrant men are restricted so that their partners are required to be in both locations, and migrations occur as a step function, we see that migrations that occur every 30 weeks show a statistically significant lower prevalence than migrations that occur every 3 weeks. We also noticed that a (statistically significant) greater number of new infections is produced when the migration interval is 3 weeks than when it

is 30 weeks in the restricted network models. These significant differences are not seen in the basic network models that are as close to the compartmental models as possible.

The key qualitative difference in the migration frequencies we experiment with is that while one is shorter than the length of acute HIV infection, the other is longer. We investigated a case where the average frequency of migration is equal to the length of acute HIV infection (12 weeks) and found the prevalence trajectory to be closely aligned with the trajectory at a migration frequency of 3 weeks.

To understand why compartmental and network models show different impacts of migrations on HIV transmission, it is important to note that the chief distinction between the two frameworks is that compartmental models do not capture the relational dependence in the partnership structure of migrant men – especially in terms of distributing their partners in both locations. Epidemiologically, the impact of the rate of circular migrations depends on the rate at which migrant men contact partners in *two separate locations*. Compartmental models, on the other hand, model flows between compartments that occur at a fixed rate. As a consequence, the time spent in any one state has an exponential distribution. Compartmental models also do not distinguish the partnerships of the actors, because they model aggregated counts of individuals, and their flow between the various infection states (and locations).

Network models are easily adapted to capture the unique partnership structure of migrant men (with regard to multiple partnerships). Hence we notice that when the network models are set up identical to the compartmental models they do not show any impact of the frequency of migration on HIV prevalence. However, once the network models are set up to capture the fact that partners of migrant men are distributed between two locations, we notice different epidemic trajectories when migrations occur at intervals shorter and longer than the length of acute HIV infection.

Another key result that emerges from this work is the difference in epidemic behaviors with the two definitions of contact in the compartmental framework. As we saw, compartmental models allow for two separate definitions of sexual “contact”; either these contacts are single sexual-acts, or they are partnerships that last over a period of time. When contacts are defined as sexual acts, a large epidemic is quickly produced, whereas the

contact-as-partnership models produce a negligible epidemic.

The contact-as-act models assume proportional mixing between the sub-populations that make sexual contact with each other. This assumption implies that each coital act occurs with a different individual, which is clearly unrealistic for our system in which a significant number of sexual partnerships would be expected to be long term. With the alternate definition of contact as partnership, we evaluate discordant relationships at the start, and simulate disease transmission through these partnerships. We neglect the fact that during the course of a partnership, say between two susceptible individuals, one might seroconvert and thus put their partner at risk as well [88].

Thus, in the contact-as-act models, the partner change rate is high, but the probability of disease transmission is low (in any stage of infection). In the contact-as-partnership case, the partner change rate is low, but the disease transmission probability is averaged over the course of the partnership, and therefore high. As we have discussed, when contact is defined as partnership, sero-discordance of infection is evaluated at the start of the partnership, and averaged over the duration of the partnership. The higher partner-change rate in the contact-as-act case has a greater number of expected coital acts between sero-discordant partners, and therefore, we see a higher equilibrium prevalence. Mathematically, we verified this difference by estimating the reproduction number R_0 . We found that R_0 estimates are higher in the contact-as-act case than in the contact-as-partnership case (more details are in Appendix A). We do not present R_0 estimates for the network models because this is an open area of research.

Between the two interpretations of contacts, there is no middle ground, and both assumptions are unrealistic in their own ways. A stochastic network model allows us to avoid both the potential pitfalls mentioned above, because we explicitly model the mean duration of partnerships and evaluate sero-discordance in these partnerships at every time step, rather than just the beginning of the simulation.

Compartmental models that rely on ordinary differential equations also implicitly assume exponential waiting times for transfers between states. For us, this means that waiting times for transfers between infection states, and for migrant men, times for transfers (migrations) between locations, are exponentially distributed. This assumption may be more realistic

for transfers between infection states than for migrations between locations (hence, in the network models, we assume clockwork migrations in the “restricted” version but experiment with stochastic migrations in Appendix C). We may be able to investigate alternate modeling frameworks in the compartmental setting. Continuous time Markov chains are one possible approach that could retain the aggregation of actors in compartments, but allow us to explore alternative models for transfers between states. It is not immediately clear if such an approach will change our results qualitatively, but they may be worth exploring.

In the context of network models, since the final prevalences in restricted models are low, we built behavioral heterogeneity in the behavior of migrant men. We took the average duration of partnerships of migrant men in the urban area to be one-fourth of the average duration of their partnerships in the rural area. We observed the same qualitative results in this case – migrations at 3-week intervals produced a higher prevalence than migrations at 30-week intervals (when the partnership structure of migrant men is restricted so that their multiple partners are distributed in the two locations) but the prevalence itself did not change very much.

We also notice that the basic network models show a higher equilibrium prevalence than models where the partnership structure of migrant men is restricted. There are three possible reasons for this difference. The first reason is that when the partnership structure of migrant men is restricted their multiple partners are in separate locations, and therefore they have to wait to contact the partner in the other location until the time to migration has passed (a consequence of delayed path acceleration). Thus, the frequency of migration affects onward transmission of HIV, and this interpretation holds in light of our observation that migrations at intervals shorter than the length of acute infection produce a bigger epidemic than migrations at intervals longer than acute infection. Second, one consequence of how we modeled the restricted partnership of migrant men is that migrant men can have at most two concurrent partners, whereas in the basic models the number of partners of migrant men is not restricted. Third, the momentary degree distribution in the basic model has more isolates and more individuals with 3 or more partners than the restricted models (details are in Appendix D). Therefore, the higher equilibrium prevalence in the restricted models might be due to the higher number of individuals with multiple partnerships at a

given time.

We investigated whether the differences in the basic and restricted network models are due to delayed path acceleration being explicitly modeled only in the restricted network models, or differences in momentary degree distributions between the basic and restricted models (Appendix E). Here, we develop models “in between” the basic and restricted models. These “constrained models” have no restriction on the location of the partners of migrant men, but the number of partners at any unit time of migrant men is *constrained* to a maximum of 2. Thus, these models help isolate the effect of delayed path acceleration (that comes about by requiring migrants’ partners to be in both locations) from the effect of limiting the momentary number of partners of migrant men to 2.

The constrained models do not show the significant difference in epidemic trajectories or the number of new infections between 3 and 30-week migrations. Thus the significant differences we see between the frequent and infrequent migrations are a result of delayed path acceleration that slows down the rate at which migrant men contact their partners in separate locations, when migrations are less frequent.

Another feature we built in to the restricted models (as examined in Section 2.3.2) is the clockwork nature of the migration process. We felt this migration structure was a better representation of reality. However, this assumption implied that the restricted models departed from the basic models in two ways: (i) the partnership structure of migrant men, (ii) the model of the migration process. We examined a version of the restricted model (Appendix C) with stochastic migrations. Results from these models were qualitatively similar to results from restricted models with clockwork migrations; 3-week migrations produced a significantly larger epidemic and greater number of new infections through the course of the simulation. Thus the differences that we saw between the basic and restricted models in Section 2.3.2 are attributable to the restricted nature of the partnerships of migrant men, rather than the model behind the migration process itself.

The *interaction* between acute infection and migration frequency is a major focus of this work. In Appendix B, we demonstrate that without acute infection the significant differences in epidemic trajectories and number of new infections between 3 and 30-week migrations in the restricted network models disappear. We also see a smaller epidemic

without acute infection in the various models we investigate. Thus, acute infection, and its interaction with migration frequency, play a major role in the transmission dynamics of HIV in a circular migration system.

A natural question to ask is how would transmission dynamics change if there were no migrations in a population identical to the one we analyze here? In other words, how does *any* population mobility (as opposed to the rates of migration) impact the transmission of HIV? As expected, compartmental models show no difference in epidemic trajectories in a population with and without migrations. Modeling an identical population without migrations using network models is a more difficult problem conceptually. In the network models, we model individuals and partnerships explicitly, and as a consequence migrant men have particular partners in two locations. If we simply set the migration parameter equal to zero, then the migrant men will never contact half their partners, in which case those individuals are not really partners, and hence that analysis is meaningless. However, if we assume all contacts of “migrant” men are in one location, then we are in effect modeling two isolated regions with concentrated partnership structures. In this case, we therefore obtain a large epidemic quickly. Modeling the epidemiological consequences of any population mobility relative to no population mobility is therefore an open problem.

In the work by Lurie et al. [57], it is documented that the mining town of Richards Bay is much closer (about 100 km) to the rural migrant-sending district of Hlabisa than Carletonville, which is approximately 700 km away. Therefore, migrants from Richards Bay are able to return home much more frequently than those from Carletonville. Our restricted network models suggest that migrations at 3-week intervals produce a significantly larger epidemic than migrations at 30-week intervals. This finding implies that the behavioral impact of migrants who return home more frequently might be more important in the epidemic than migrants who return less frequently. But our finding is contingent upon the assumption that the behavior of migrants in urban towns (both close to and far away from home) is identical in terms of the number of sexual partners and partner-change rate. In reality, it is possible that the behavior of migrants that are farther away from home is much different from the behavior of migrants that are close to home. Conversations with experts in the subject have indicated that such differences are poorly understood. Therefore, a

useful extension of this project would be to examine different behavior patterns in migrants that are close to and far away from home.

In summary, the nature of the impact of circular migrations is still an open question, likely contingent upon specific behavior patterns in the context of home and work locations in the specific population of interest. However, from the current study, we learn that we need network models to even begin to examine the impact of circular migrations on HIV transmission.

This study is largely theoretical. Our behavioral assumptions relied on qualitative descriptions of data, primarily as presented in [56] and [57]. Next steps in this research project are to parameterize these network models to data from different populations where circular migrations are seen, to include more realistic migration structures that exist in those populations, and to parameterize the behavioral structures of these models to epidemiological data. We discuss possibilities of projects that might enable us to incorporate data in Chapter 5.

Figure 2.1: Concurrency structures in persistent partnerships of migrant (top) and non-migrant (bottom) men.

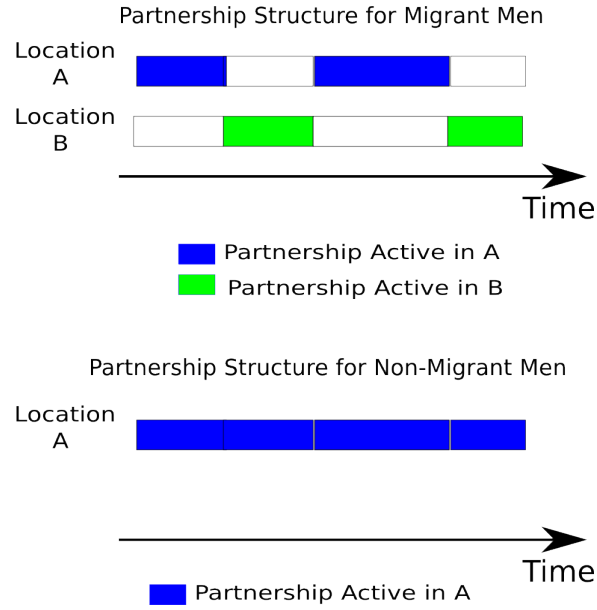


Table 2.1: Demographic and biological parameters: These parameters are identical in the contact-as-act and the contact-as-partnership cases.

Parameter	Notation	Value
Biological Parameters	Natural Mortality Rate	μ $1/(45 \times 52)$
	AIDS-related Mortality Rate	μ_d $1/40$
	Fertility Rate	ν $1250/(45 \times 52)$
Biological Parameters	Acute-Stage Duration	$1/\gamma$ 12 weeks
	Chronic-Stage Duration	$1/\eta$ 500 weeks
	Late-Stage Duration	$1/\mu_d$ 40 weeks

Figure 2.2: Transfer diagram for migration system. White and blue blocks show non-migrant men in each region. Pink blocks show non-migrant women. Infection is transmitted between infected men and susceptible women and infected women and susceptible men. Location A and Location B represent the rural home-town and the urban workplace respectively.

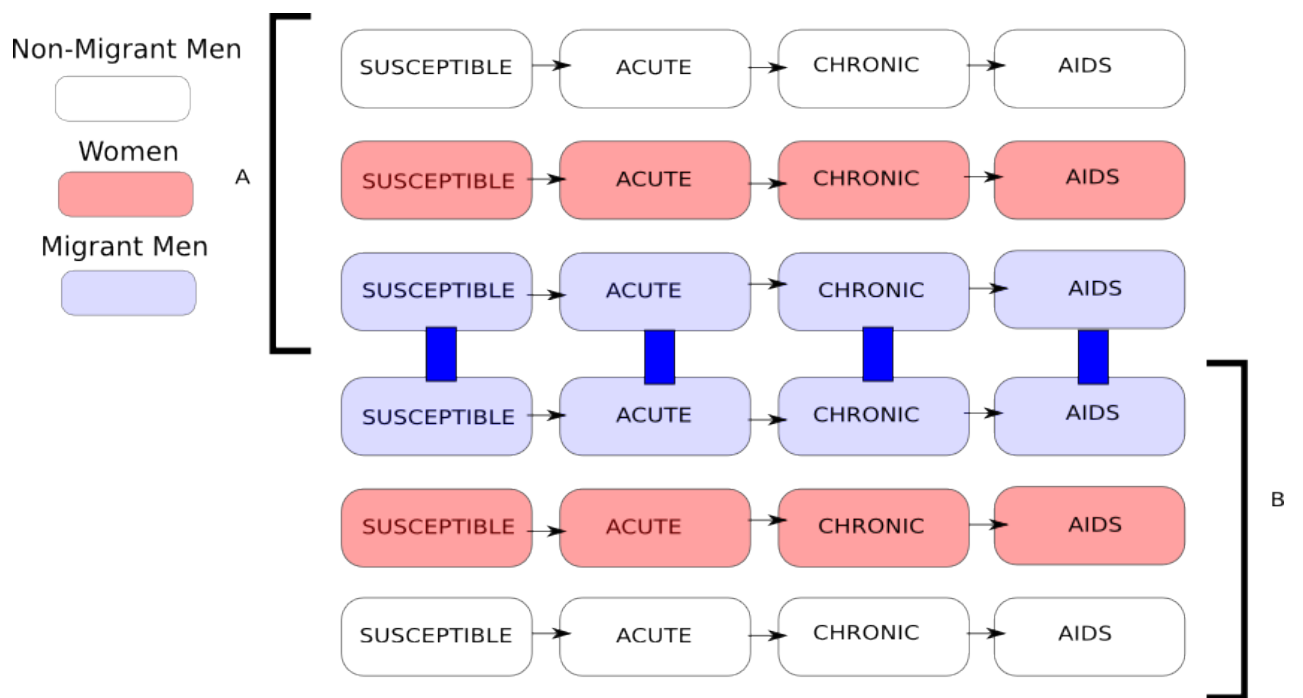


Table 2.2: Initial values for state variables. These initial values are the same in both contact-as-act and contact-as-partnership approaches. Total population size is 5000, including 2500 men and 2500 women. Half the men are migrant, and half the men are non-migrant. The population is equally divided between the urban and the rural areas.

State Variable	Value
Susceptible Migrant Males	1250
Acutely Infected Migrant Males	0
Chronically Infected Migrant Males	0
Late-Stage Infected Migrant Males	0
Susceptible Females	2450
Acutely Infected Females	$12/552 \times 50$
Chronically Infected Females	$500/552 \times 50$
Late-Stage Infected Females	$40/552 \times 50$
Susceptible Non-Migrant Males	1250
Acutely Infected Non-Migrant Males	0
Chronically Infected Non-Migrant Males	0
Late-Stage Infected Non-Migrant Males	0

Table 2.3: Biological and behavioral parameters in the contact-as-act model. Biological parameters $\beta_A, \beta_C, \beta_L$ are per-act transmission probabilities, subscript i represents one of the three infection-states, and β_i shows the computation of per-partnership transmission probabilities for the particular infection-state. Behavioral parameters t_{\dots} are set to 2.4 to have a mean number of 2.4 sexual acts per week in the population (details in Section 2.3.1). The subscript \cdot represents rural (R) or urban (U) regions.

Biological Parameters	β_C	0.0007
	β_A	0.0007×26
	β_L	0.0007×7
Behavioral Parameters	t_{MM}	2.4
	t_{NM}	2.4
	$t_{NF,M}$	2.4
	$t_{NF,NM}$	2.4

Figure 2.3: Comparison of HIV prevalence as a function of time using compartmental models. The top figure shows the result for sexual-contact defined as a long term partnership, and the bottom figure shows the result for sexual-contact defined as a sexual episode. We compare migrations at average intervals of 3 weeks and 30 weeks. In addition, we consider a number of intervals shorter (1, 4, 7, and 10 weeks) and longer than the acute phase (20, 30, 40, 60, 80 and 100 weeks).

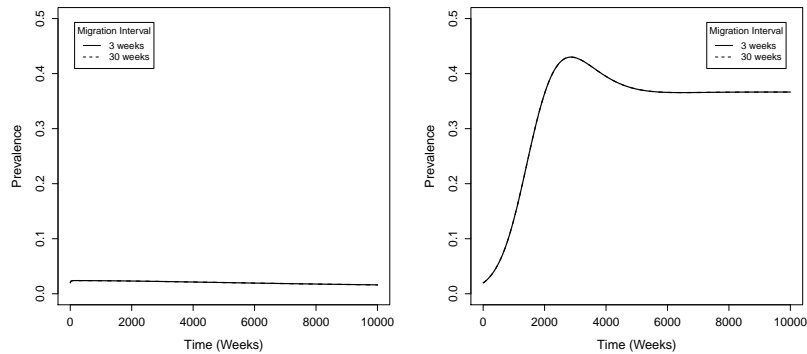
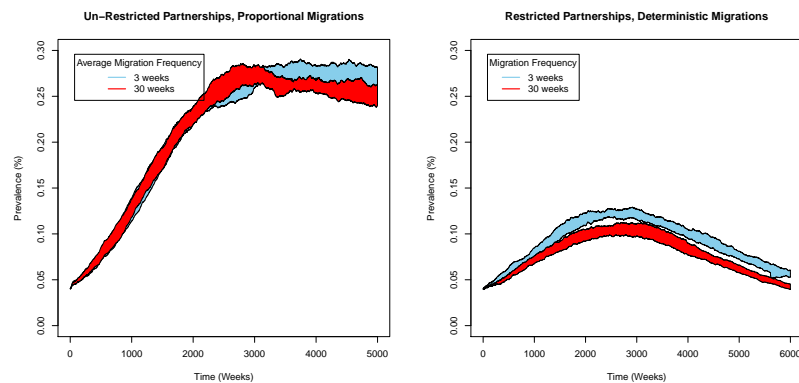


Table 2.4: Biological and behavioral parameters in the contact-as-partnership model. Biological parameters p_A , p_C , p_L are per-act transmission probabilities, subscript i represents one of the three infection-states, and β_i shows the computation of per-partnership transmission probabilities for the three infection states. Behavioral parameters $t_{MM.}$, $t_{NM.}$, show number of sexual partners per unit time for migrant and non-migrant men respectively. Behavioral parameters $t_{NF.,M}$, $t_{NF.,NM}$ show the number migrant and non-migrant partners per unit time of women respectively. The subscript . represents rural (R) or Urban (U) regions, and is identical in either case. The parameter d is the average duration of partnerships, taken as 100 weeks. The parameter n is the average number of coital acts per week, taken as 3. However, the number of coital acts that actually occur is 2.4, on account of a certain number of partnerships being inactive at any time step (details in Section 2.3.1).

	p_C	0.0007
	p_A	0.0007×26
Biological Parameters	p_L	0.0007×7
	β_i (for non-migrants)	$1 - (1 - p_i)^{d \times c}$
	β_i (for migrants)	$1 - (1 - p_i)^{d \times c/2}$
	n	3
	$t_{MM.}$	$1000/1250 \times 2$
Behavioral Parameters	$t_{NM.}$	$1000/1250$
	$t_{NF.,M}$	$\frac{t_{MM.} N_{MM.}}{N_{NF.}}$
	$t_{NF.,NM}$	$\frac{t_{NM.} N_{NM.}}{N_{NF.}}$

Figure 2.4: Comparison of HIV prevalence as a function of time using network models. The top figure shows the result for a network-model identical to the compartmental model. The bottom figure shows the result for when the migrants' partnership structure is restricted and consequently multiple partners are in both locations and migrations occur as deterministic step-functions. We consider 10 repetitions each of migration-intervals of 3 and 30 weeks. The colored regions correspond to 95% confidence intervals about the mean.



Chapter 3

MODELING THE IMPACT OF POST-DIAGNOSIS BEHAVIOR CHANGE ON HIV PREVALENCE IN SOUTHERN CALIFORNIAN MEN WHO HAVE SEX WITH MEN (MSM)*Coauthorship*

This is joint work with Dr. Steven M. Goodreau (University of Washington), Dr. Pamina M. Gorbach (University of California Los Angeles), Dr. Eric Daar (University of California Los Angeles), Dr. Susan J. Little (University of California San Diego).

Dr. Goodreau coded the initial versions of network models for HIV transmission in men who have sex with men and much guidance and mentoring along the way; Dr. Gorbach provided the main motivation for modeling the impact of post-diagnosis behavior change and the funding to do this work; Dr. Daar and Dr. Little led the Acute Infection and Early Disease Research Program.

3.1 Introduction

Men who have sex with men (MSM) form one of the highest risk groups for HIV in the United States [10, 91], with roughly half of new infections occurring in this population [92]. The CDC estimated HIV prevalence in the MSM community was 19% in 2008 [92].

Recent longitudinal studies have found that men who have sex with men (MSM) reduce risky sexual activity [2, 3, 21] upon receiving a positive HIV diagnosis. It seems likely that such post-diagnosis behavior change (PDBC) occurs on account of a desire in the MSM community to protect one's partners [93]. Such "community-initiated" [93] strategies may have much potential for prevention of new infections. However, a review of cross-sectional studies found that positively diagnosed MSM practice a high level of risky sexual activity [94]. Another cross-sectional study found an increase in condom use and/or abstinence among MSM, but that a high proportion of those who reported anal sex still reported no condom use [95]. Given the cross-sectional nature of the studies in this review, it is difficult

to assess the change in level of risky sexual activity upon diagnosis.

The timing, extent and heterogeneity in reduction of risky sex on account of behavior change are questions that are being studied; however, even short-term reductions among recently HIV infected individuals can be highly effective in reducing onward transmission events especially since recently diagnosed individuals may be in (or not far removed from) the stage of acute infection [96]. If behavior change to reduce risk of onward transmission occurs when individuals are most infectious, its preventive potential may be maximized. However, other papers in the literature question the effectiveness of risk-reduction approaches that men undertake and emphasize the “urgent” importance of public health messages around the effectiveness of such strategies [97].

In this paper, we use mathematical models to demonstrate the impact of PDBC on overall HIV prevalence in Southern Californian MSM when acute HIV infection and behavior change of MSM are taken into account. We also demonstrate the population-level effects of a temporal phenomenon such as PDBC cannot be estimated using cross-sectional data; we need longitudinal data to uncover patterns in individual-level PDBC that document sexual behavior before and after diagnosis. We end with a discussion on using PDBC as a preventive tool. We thus employ novel modeling tools in this paper to address questions of immediate public health relevance.

3.2 Background

Southern California is a populous and racially diverse area of the United States. The cities of Los Angeles (LA) and San Diego (SD), and their respective counties, are major population centers of this region. The HIV epidemic here is large – Los Angeles County (LAC) is home to 29% of the population of California [8], but is estimated to contain 41% of new HIV cases, and 76% of all AIDS cases in the state [9].

Sexual contact in MSM is estimated to account for 71.4% of the cumulative HIV cases in California males between 1983 and 2005 [10]. Of the 40,000 persons living with HIV and AIDS (PLWHA) in LA County at the end of 2009, 88% were male [11] and 72% of PLWHA were MSM [11]. MSM in San Diego County show similar trends of high prevalence; from 1985-2004 non-IDU MSM accounted for 79% of the HIV transmission events [7].

MSM have been shown to modify sexual behavior in the time following diagnosis, presumably with the goal of reducing risk of transmission to partners who are HIV-negative [21]. These modifications include reducing the number of sexual partners, especially casual ones [2, 21], reducing unprotected anal sex within partnerships [2, 21], and choosing partners with the same HIV status (serosorting) [2, 21]. Since transmission of HIV through unprotected anal intercourse (UAI) is more probable when the positive partner is insertive, rather than receptive [98], modifying sexual role in partnerships (sero-positioning) is another behavioral strategy MSM adopt to reduce transmission events [2, 21].

The use of methamphetamines (“meth”) has been relatively common among MSM since at least the 1990’s [99], and is especially common on the west coast [100]. Meth use is associated with behavioral disinhibition, multiple sexual partners, low rates of condom use, prolonged sexual activity, encounters with casual/anonymous partners [101] and low rates of adherence to HIV treatment regimens [102]. Thus MSM who use meth tend to have higher rates of risky sexual behavior than non-users, and meth use is an important covariate to consider in estimating population-level impacts of PDBC in this population.

MSM engage in multiple types of sexual contacts, ranging from stable main partnerships to casual, one-time contacts [2, 103]. Levels and patterns of PDBC appear to vary by partnership type [3], as may be expected, since the desire to protect one’s partner would reasonably vary with levels of emotional intimacy.

In this work, we demonstrate the impact of PDBC on overall HIV prevalence in the population when the natural history of HIV infection and behavior of MSM are taken in to account. We explicitly model both main partnerships and casual contacts, and disaggregate behavioral patterns between meth users and non-users.

3.3 Methods

3.3.1 Overview

We create a dynamic, stochastic network simulation based in the exponential random graph modeling (ERGM) framework, parameterized using biological, behavioral and demographic data. The model structure closely follows that developed for the Prevention Umbrella for

MSM in the Americas (PUMA) project, as do many of the biological parameters [27]. Our model incorporates births, deaths, aging, treatment, circumcision-status, testing behavior, treatment, meth use, diagnosis status and post-diagnosis behavior change, partnership types, temporal overlap in partnerships, viral load, and sexual role. Since treatment is an important component of HIV epidemiology in United States MSM, we adopted a basic model of treatment that accounts for race differentials as measured by the waiting time between onset of infection and commencement of treatment in accordance with [27]. Extensive detail on each component is included in the Appendix F; summary data are below.

The behavioral data are primarily from the Southern California Acute Infection and Early Disease Research Program (AIEDRP) study. Complete details on the AIEDRP study are elsewhere [2, 3]. In summary, newly diagnosed HIV positive men completed AIEDRP questionnaires at baseline. Follow-up information was collected by offering interviews every 3 months after initial enrollment. 193 respondents completed at least one follow-up survey. At baseline, respondents provided detailed information on their three most recent partners, and at follow-up, on the most recent partner, yielding a total of 1011 partner-reports. For simplicity, we dichotomize these partnerships as “main” and “non-main”; the former category includes partners reported as main, the latter category includes all other partners.

We use these data to create network models constructed using exponential random graph model framework (ERGMs, also called p^* in the literature). In their most basic form, ERGMs allow us to simulate cross-sectional networks (within stochastic variation) of a number of network features. These features are described in the form of statistics that may include various forms of dependence in the ties in the network. This capability is a major strength of the applicability of the ERGM framework to modeling the sexual partnerships since it allows us to capture temporal overlap in partnerships that classical compartmental methods for modeling HIV transmission cannot [25]. As in classical compartmental models, we are also able to account for various selection forces such as assortative mixing based on age, race and knowledge of HIV status. Separable Temporal ERGMs (known as STERGMs) [51] are an extension that allow us to model the network structure in a stochastically evolving dynamic network in a statistically principled fashion.

The ERGM framework is implemented using the **statnet** suite of packages [24] pro-

grammed in the **R** programming language.

3.3.2 Non Main Networks: Specification and Estimation

We model unprotected anal intercourse (UAI) contacts in non-main partnerships in our population as a series of cross-sectional networks. This network contains no information on duration of partnerships and consists only of statistics estimated from egocentric data (with detailed reports on the last three partners) that describe the network structure.

We therefore model non-main partnerships using cross-sectional ERGMs.

The specific features of the network structure that we capture are the mean number of partnerships per person, levels of sexual activity, assortative mixing by age and race, and differential levels of activity of meth users and non-users. In this network, PDBC is modeled by considering two network features: a reduction in total UAI with non-main partners of diagnosed individuals and selective mixing by diagnosis status so that the number of ties between individuals of the same diagnosis status is greater than would be expected by chance.

Data for all of these features (except race mixing and meth use) are obtained from the AIEDRP study. AIEDRP data on race mixing are very sparse and therefore we use a study of MSM in San Francisco [104]. We are not assuming that the race composition of San Francisco and LAC are the same. The assumption is that the patterns of mixing between the different racial groups in the two places is the same, i.e. the proportion of non-main ties that exist between any particular pairing of races in the two regions is the same.

The proportion of meth users in the AIEDRP sample is extremely high presumably because this is a sample of recent sero-converters. To use estimates more reflective of meth use in the population we are interested in (MSM in Southern California), we use data from a separate study that reports 11% of MSM recently diagnosed with AIDS in Los Angeles County have used meth in the past 12 months [105]. Our population of 5000 men therefore has 550 meth users. These race composition of the set of these 550 meth users is identical to the the composition in AIEDRP.

3.3.3 Main Network: Specification and Estimation

The main network incorporates duration information in addition to information on network statistics measurable from egocentric cross-sectional data. Therefore, we use the STERGM class of models [51]. This approach preserves the following features (stochastically): the proportion of individuals who report 0, 1, or 2 ongoing main partnerships at a given time; the mean difference in the square roots of the ages of main partners; race mixing (as described above for non-main networks). As in the case of non-main networks, all statistics except race mixing are from AIEDRP, and race mixing is from the same study of MSM in San Francisco [104].

We model transmission of infection in the main network through UAI acts that occur on a given day (more details below). PDBC in the main network is modeled by considering a reduction in the daily probability of UAI in main partnerships that are discordant by diagnosis status. The daily probability of UAI in non-main partnerships is obtained through AIEDRP, but the daily probability of UAI in partnerships discordant by diagnosis status is obtained from [27] (because AIEDRP data report proportion of unprotected sex acts and do not explicitly differentiate between oral and anal-sex acts).

Components of Simulation

We create an initial population of 5000 MSM using a simulated-annealing algorithm that approximates the statistics we estimated from our ego-centric data. We randomly infect 19% of our population in accordance with recent estimates of HIV prevalence in United States MSM [92]. We then simulate our model forward in daily time-steps, with each of the following steps:

1. Arrivals: Men enter our population at age 18, when they are untested for HIV, and are HIV-negative. We set the number of arrivals to have a population that grows slightly given a stable HIV epidemic.
2. Deaths/Departures: Non-HIV deaths follow US age-specific mortality rates derived from CDC life-tables [27]. AIDS-related deaths occur as a function of rising viral loads, which are themselves determined by time since infection, treatment status, and

suppression (see below). Individuals leave the population of interest at 65 years of age.

3. UAI in main partnerships: UAI events occur with a given probability on a particular day in accordance with parameters used in PUMA since these parameters are not directly estimable from AIEDRP. This daily probability is based on the relative sero-statuses of the two partners. In the non-main network, partnerships are UAI contacts on a particular day.
4. Transmission in main and non-main partnerships: Transmission of HIV occurs probabilistically in partnerships, where UAI events occur on a particular day. This probability is a function of the viral load of the infected partner and the pre-exposure prophylaxis (PREP) status and circumcision status of the uninfected partner. We consider transmission only due to UAI, and ignore transmission due to other events such as oral-intercourse, or needle-sharing.
5. Update network: We then update the network by considering changes in population-size due to births, deaths, and aging, formation and dissolution of main partnerships, formation of new non-main partnerships, and changes in other attributes of the nodes (viral-load, infectivity, treatment-status). Viral-load is a function of time since infection and the treatment-status. Treatment-status in turn depends upon the testing behavior of the individual. Details are in the appendix. We then repeat steps 4 and 5 over a 50-year period to allow the epidemic trajectory sufficient time to equilibrate.

Other biological and demographic processes that we model are explained in detail in Appendix F.

3.3.4 Modeling Post Diagnosis Behavior Change (PDBC)

To recap, we model the following mechanisms of PDBC:

1. Reduction in total number of non-main partnerships: We model a 25.6% reduction in mean-degree of non-main partnerships upon diagnosis in accordance with AIEDRP

data (as seen in Figure 1 of [3]) where the mean number of non-main partners is 8.37 in the three months prior to diagnosis, and 6.23 post-diagnosis, taking the mean over the four follow-up periods.

2. Selection of non-main partners by diagnosis status:

AIEDRP data classify the HIV status of partners as positive, negative or unknown. We re-classify the unknown partners as positive or negative according to the proportion of HIV positives in the population nationally.

Then, to account for selective mixing by diagnosis status, we model an approximate reduction of 40% (with stochastic variation) in the proportion of discordant non-main ties compared to the number expected if there was proportional mixing by diagnosis status. The precise extent of this reduction depends on the precise number that are infected at the start of any given simulation (drawn from a binomial distribution) and this number varies slightly due to the stochasticity of the process. A given proportion of these infecteds are then classified as diagnosed, which itself is a bernoulli process (and the precise number is again stochastically variable).

3. UAI in Main Partnerships: Within main partnerships, a reduction in daily probability of UAI from 0.156 to 0.109. As explained above, the former is obtained from AIEDRP data and the latter is from [27].

In this work, we do not model sero-positioning as a risk reduction strategy.

3.3.5 Counter-Factual Models

We explore counter-factual models that vary from the baseline models by modeling no reduction in total number of non-main partners upon diagnosis, proportional mixing by diagnosis status in non-main partnerships, and the same daily probability of UAI within negatively concordant and discordant (by diagnosis status) main partnerships.

Parameters from the AIEDRP follow-up data represent the behavior of diagnosed men. Therefore, our base scenario that models PDBC includes parameters from both baseline

and follow-up data-sets. The counter-factual that does not incorporate PDBC includes information only collected at baseline to capture the behavior of undiagnosed men.

3.3.6 Sensitivity Analysis

Our base model parameterized testing frequency using studies containing self-reports of the time since last test for negative men from a study of MSM in clinics in four major US Cities [106]. The study reports a median inter-test interval of 243 days [106]. If we assume a simple exponential waiting time distribution for the time till test, this statistic corresponds to median inter-test interval of 351 days. Our models indicate that at this level of testing, about 95% of infected MSM are aware of their status (at steady-state).

Other studies have reported different summary metrics on testing and awareness, e.g. the proportion of HIV positive men in a given setting who are aware of their status. For instance, based on NHBS data, the CDC reported that about 44% of infected MSM nationally are not aware of their status [92]. Our models indicate that to obtain this level of awareness of infection, the testing rate has to be extremely low – slightly less than an average of one test every ten years for all negative or undiagnosed men (though other ways of matching this level of infection awareness are possible – for details see Section 3.5).

These different summary metrics may generate very different pictures of testing and awareness, even when they come from the same study. Indeed, our dynamic models suggest that these two figures are incompatible with one another; it is not possible for testing frequencies to be that high and status awareness to be that low. Multiple explanations for this discrepancy exists (see the discussion). Since we cannot ascertain which is correct, we instead conducted a sensitivity analysis in which we varied the mean testing frequency, which in turn led to different proportions of the HIV+ population being aware of their status in the long run. We experimented with two main scenarios:

1. “Testing Frequency” (Baseline Models): In this setting we assume that men test once every 351 days on average in accordance with clinical data from [106].
2. “Level of Awareness”: Here we assume that the proportion of HIV-positive MSM who are aware of their status as reported by the CDC (56%) is correct [92]. We

experimented with scenarios iteratively in order to obtain the mean testing frequency that results in about 55-60% awareness of infection at equilibrium.

We repeated each experiment ten times. We then ran additional simulations in between these values, and varied the average testing rate between two times every ten years and ten times every years (at increments of two tests per ten years) to compare the difference in prevalence in counter-factuals that exclude and include PDBC, relative to the case that includes PDBC.

We also present proportion of infected individuals who are diagnosed early in each scenario.

3.4 Results

Figure 3.1 shows a comparison of prevalences in the baseline testing-frequency models with and without PDBC. We see that final equilibrium prevalence when PDBC is accounted for is 31.9% (averaged over 10 repetitions) and 41.7% when there is no PDBC. Thus, prevalences would be higher by close to a third (30.6%) without PDBC given our baseline models. With testing at this frequency, our model suggests that approximately 94.8% of HIV+ men would be aware of their status at the steady-state.

In contrast, under the level of awareness model (Figure 3.2), the effects of PDBC are much smaller: the mean equilibrium prevalence (over 10 repetitions of the experiment) with and without PDBC is 44.5% and 45.3% respectively. Recall that to reduce the level of awareness, we had to assume that the average testing frequency in this model is slightly less than once every 10 years (one test every 4000 days, or 10.9 years). The mean proportions of those infected who are aware of their infection-status in the cases with and without PDBC are 62.3% and 62.9% respectively.

Since these two models yield strongly different results, and have such different assumptions, we investigated various scenarios with average rates of testing in between the two extremes of the testing frequency and level of awareness cases. In Figure 3.3, we see that even at an average of two tests every ten years, the equilibrium prevalence without PDBC is 15.2% higher than when PDBC is accounted for.

Note that the prevalences we see in this paper are higher than national estimates for MSM, since our baseline model is parameterized by the behavior of men just before seroconversion. We discuss implications below.

We compare the mean number of non-main partners per person (mean degree) of diagnosed and undiagnosed (includes true negatives and undiagnosed positives) in Figure 3.4. We see that both “testing frequency” and “level of awareness” models show diagnosed individuals have greater number of non-main partners per individual than undiagnosed individuals.

In the testing frequency cases, we also found that the mean proportion of infected individuals who were diagnosed within the first 180 days (“early diagnosis”) is 31.6% and 30.5% with and without PDBC. In the level of awareness models, the mean proportion of individuals diagnosed early is 3.7% and 3.3% with and without PDBC. We present the proportion of infected individuals who are diagnosed early for the various intermediate cases shown in Figure 3.3 in Table 3.1.

3.5 Discussion

In this paper, we demonstrate population-level effects of individual-level change in behavior of recently diagnosed HIV positive men. Our baseline (“testing frequency”) models – that are based on testing behavior as reported in clinical data from four major metropolitan centers in the United States [106] – show that without PDBC, HIV prevalence in the MSM community would be higher by about 30.6% relative to current estimates. We call our second set of models the “level of awareness” models. This scenario matches CDC data that show approximately 56% of infected MSM are aware of their status [92], and are based on an average testing frequency of one test every 4000 days (or 10.9 years). These level-of-awareness models show that without PDBC, equilibrium prevalence would be higher by only about 1.8% relative to models that incorporate PDBC. An average rate of testing of one test every 4000 days is admittedly unrealistically low (more on this below). Our analyses indicate that even under very low rates of testing (twice every ten years on average) equilibrium prevalence without PDBC would be higher by about 15.2%.

We have seen the large difference in the proportion of infected who are aware of their

status in the testing frequency and level-of-awareness scenarios in Section 3.4. Another method to demonstrate this incompatibility is to assume that both of the defining conditions of the two scenarios (approximately yearly testing on average, and about 50% awareness of infection) are true. Then, if we consider HIV prevalence in MSM to be 20% [107], assume half of these infected individuals are unaware (in accordance with NHBS 2008 and the level-of-awareness scenario), and about 60% tested in the past year,¹ then we conclude that

$$100\% \times 1/5 \times 1/2/5 \approx 6\%$$

of MSM nationally tested negative at the last test and sero-converted in the past year (or some HIV positive MSM mis-reported results). However, it is established that about 1-2% of MSM become HIV positive in a given year [6]. Thus the testing frequency and level of awareness scenarios are incompatible with each other. On the other hand, if we assume only about 5% of positive MSM are unaware of their status – consistent with our modeling results in the testing frequency case – then we estimate about 1.2% of all MSM sero-converted in the past year, which is much closer to the accepted number.

As we saw earlier, the testing frequency data come from a clinical study in four major urban centers in the US and the level of awareness numbers come from 2008 NHBS data. These datasets use different sampling methods – the clinical studies report data on MSM attending HIV clinics, and the NHBS reports data on MSM frequenting venues popular in the community. Some of the incompatibility between these data might therefore be attributed to the sampling mechanisms; it is likely that men regularly attending an HIV clinic test more frequently than the average member of the population while younger MSM (sampled at popular venues) are more likely to be unaware of their infection status. However, it is unlikely that the extent of this incompatibility is entirely due to sampling.

It is believed that knowledge of HIV status is not as low as the NHBS data suggest (Patrick Sullivan, personal communication). Stigma against HIV and the resulting social-

¹ We arrive at the estimate that about 60% of negative or undiagnosed MSM tested in the past year by assuming that testing is a memoryless process with a constant daily probability p . Then under yearly testing (as per the testing frequency scenario and [106]), this daily testing probability p is $1/365$ and the proportion of men who tested within the past year is $1 - (1 - p^{365}) = 63.4\%$. We approximate this estimate as 60%.

desirability bias may be factors why men tend to under-report knowledge of positive HIV status. Moreover, venues where NHBS data are collected (bars, clubs) [92] tend to have an over-representation of young MSM, who are particularly likely to not be diagnosed. Thus the low infection awareness in the NHBS data, and in our level-of-awareness models is likely an artifact of the interplay of various factors, and less representative of the real infection awareness in MSM.

Analysis of other data, in particular NHBS data from San Francisco, New York and a clinical study in Massachusetts [108] have indicated testing levels close to our assumption of yearly testing (Steven Goodreau, personal communication). We saw that at any reasonable level of testing, even as low as an average of twice every 10 years per individual, a lack of PDBC results in an increase in equilibrium prevalence by about 15.2%. Thus HIV prevalence in MSM would be substantially higher were PDBC not a real phenomenon.

Given the importance of treatment in the HIV epidemiology of MSM in the United States, we adopted a basic model of treatment that accounts for race differentials as measured by the waiting time between onset of infection and commencement of treatment in accordance with [27]. However, the landscape of treatment is constantly evolving, and future work on the subject should account for changes in such treatment patterns. It is also important to remember that the focus here is on the effectiveness of post-diagnosis *behavior* change, and not on treatment as a prevention technique.

The prevalences we see in this paper are somewhat higher than the national estimates of 20-25% in MSM because the baseline component of these data come from a group of recently HIV-infected individuals. Since our sample consists of recent sero-converters, on average, the behavior of these men with regards to HIV acquisition is potentially more risky than a typical member of the MSM population. It is also likely that the follow-up data show more reduction in behaviors directly related to HIV acquisition than typically occur because these men are participants in the follow-up components of the study and are therefore likely more conscious of their behavior upon diagnosis than a typical member of the population. However, our work shows the extent to which PDBC may mitigate the epidemic – under the assumption that diagnosed individuals reduce their sexual activity with regards to the 3 mechanisms we model.

This study is not an attempt to model the historical trajectory of the epidemic. Therefore readers should not interpret Figures 3.1 and 3.2 as a prediction of HIV prevalence over the next 50 years. The purpose of this paper is to model HIV prevalence trajectories given current data on the sexual behavior of Southern California MSM and demonstrate how much higher prevalence could be if post-diagnosis behavior change was not present.

Figure 3.4 shows that even when we account for PDBC in the population, diagnosed individuals have more partners (per person) than undiagnosed individuals cross-sectionally. Longitudinal studies are needed to capture the true extent of PDBC; cross-sectional measures of levels of sexual activity among diagnosed and undiagnosed individuals are not suggestive of the extent to which PDBC exists in a population. Our level of awareness scenario is based on an average test every 4000 days. A number of possible techniques to match the report of 55-60% awareness are possible. One reasonable method to match this report might be to build in an attribute that decides the potential for the men in every race to mis-report their test result. Data on mis-reporting of HIV test results were presented at the Conference on Retroviruses and Opportunistic Infections (CROI) in Seattle, WA in March 2012, and further studies on this subject are in progress (Patrick Sullivan, personal communication). The models we have developed during the course of this work allow for easy extension to addition of an indicator function for individuals that mis-report their status (and this individual-level attribute can be easily made dependent on race too). The broader implication of mis-reporting of test results suggests the extent to which HIV stigma is present even today, particularly among Black MSM. HIV prevention efforts should consider the prevalence of such stigma when reaching out to communities that are especially at risk.

Our model did not include any rebound back to pre-diagnosis behavior for men at any point. Gorbach et al. [3] found evidence for a rebound in terms of recent UAI with partners whose HIV status is unknown. However, to parameterize our model, we considered a different set of behavioral metrics – specifically the likely number of truly HIV-negative men that a positively diagnosed man will have unprotected sex with per unit time. This metric combines both the number of non-main partners men report, and then the proportions of those men who are of each status. In this combined metric specifically, we did not see evi-

dence for rebound. However, rebound in risky sexual activity after a certain amount period of time since diagnosis has passed continues to be an area of research, and no consensus exists yet about the forms such behavior may take (i.e. the specific mechanisms of sexual behavior that it may be manifested in) or the timing of such behavior.

The models in this work are based on recent developments in the exponential random graph framework and are implemented in the “statnet” package [24] in the R-programming language. A major strength of this framework is that it allows us to account for temporal overlap in relationships and various selection forces to describe the partnership structure of MSM. Transmission of infections is then simulated as a diffusion process on these partnership-networks.

As we have seen, this work demonstrates that at any reasonable rate of testing, PDBC has a significant protective effect at the population level. This result helps drive home the fact that MSM engage in altruistic behavior that protects their HIV-negative partners, and the community at large.

In Chapter 4, we model the effect of various testing strategies on early diagnosis. Early diagnosis is a combination of two separate mechanisms: instituting tests that are capable of detecting HIV earlier than the typical HIV tests in use at present, and increasing the average frequency of testing in MSM, and the interaction between these mechanisms. This comparison of diagnosis strategies is centered around maximizing the potential of PDBC and therefore includes all the components of early diagnosis that we examined here.

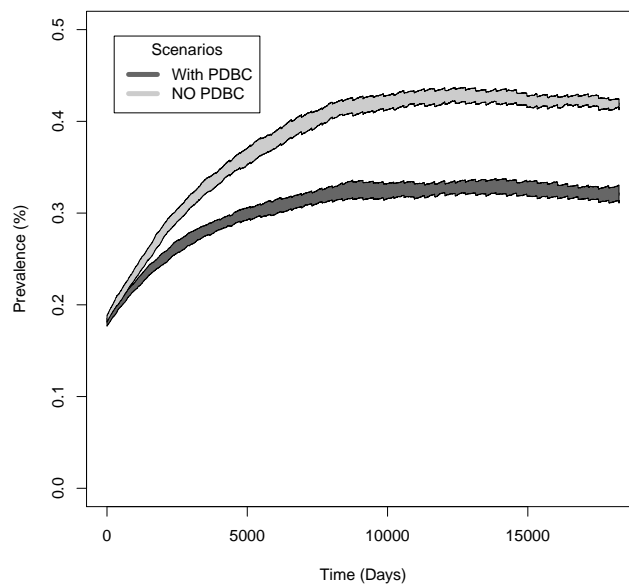


Figure 3.1: HIV prevalence in two different scenarios with and without post-diagnosis behavior change (PDBC) in the baseline “testing-frequency” models. The shaded areas correspond to a 95% confidence interval.

Table 3.1: Proportion of infected individuals who are diagnosed early for the intermediate cases shown in Figure 3.3 with and without PDBC. The column on the left gives the average number of tests every 10 years, and the “With PDBC” and “No PDBC” columns show the outcome of interest in each case.

Average Rate of Testing (10 years)	With PDBC	No PDBC
2	6.9%	7.3%
4	13.7%	13.6%
6	18.3%	19.4%
8	24.1%	24.3%

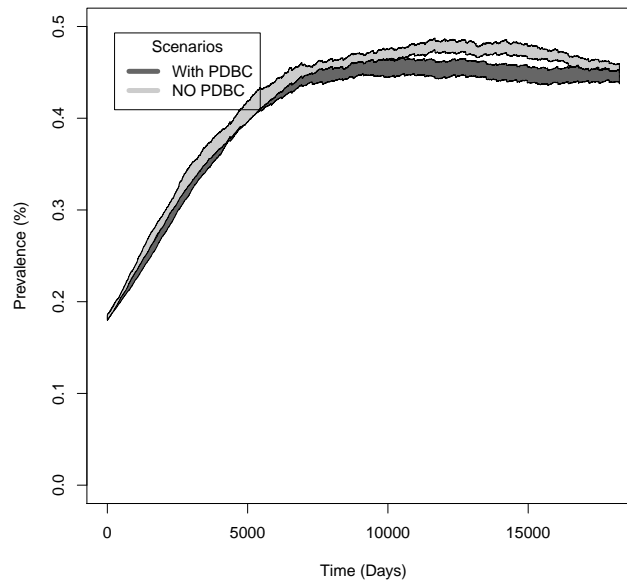


Figure 3.2: HIV prevalence in two different scenarios with and without post-diagnosis behavior change (PDBC) in the “level of awareness” Models. The shaded areas correspond to a 95% confidence interval.

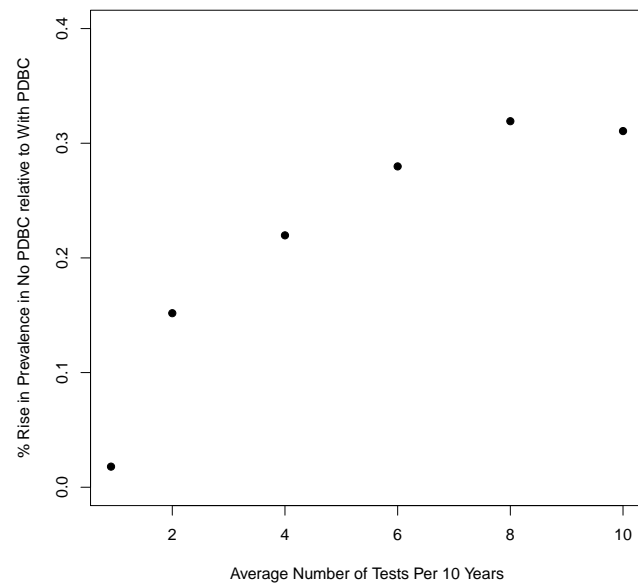


Figure 3.3: Ratio of difference in equilibrium prevalence between the “No-PDBC” and “With PDBC” cases to the equilibrium prevalence in the “With PDBC” case.

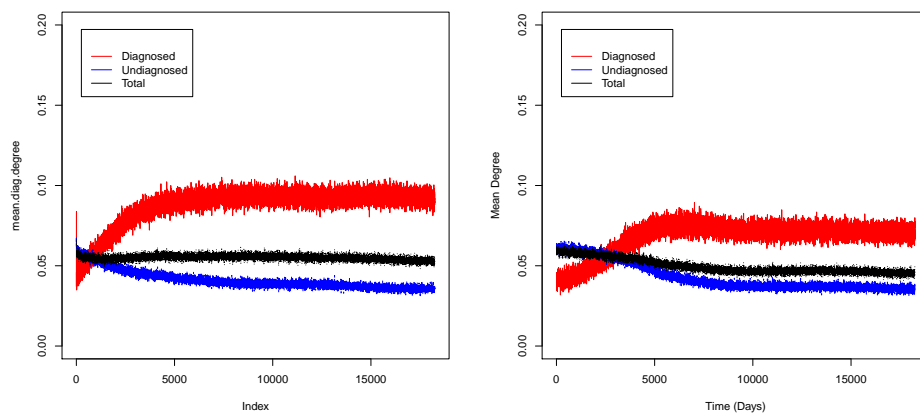


Figure 3.4: Mean degrees of diagnosed and undiagnosed men, and average mean degrees in the “testing frequency” (left figure) and “level of awareness” (right figure) cases when we account for PDBC. Mean degree is the mean number of non-main partners per individual. Undiagnosed includes true negatives, and positive but undiagnosed individuals. Results presented are averaged over 10 experimental repetitions.

Chapter 4

AN EXAMINATION OF DIAGNOSIS STRATEGIES TO MAXIMIZE THE PREVENTIVE POTENTIAL OF POST-DIAGNOSIS BEHAVIOR CHANGE IN SOUTHERN CALIFORNIAN MEN WHO HAVE SEX WITH MEN*Coauthorship*

This is joint work with Dr. Steven M. Goodreau (University of Washington), Dr. Pamina M. Gorbach (University of California Los Angeles), Dr. Eric Daar (University of California Los Angeles), Dr. Susan J. Little (University of California San Diego).

Dr. Goodreau coded the initial versions of network models for HIV transmission in men who have sex with men and much guidance and mentoring along the way; Dr. Gorbach provided the main motivation for modeling the impact of post-diagnosis behavior change and the funding to do this work; Dr. Daar and Dr. Little led the Acute Infection and Early Disease Research Program.

Dr. Martina Morris (University of Washington) gave me the idea of modeling the individualized testing strategies and I communicated with Dan Wohlfeiler (California Department of Public Health) about ideas for implementation of this strategy. However, neither of them are coauthors on this chapter.

Abstract

Objectives: In this chapter we examine various diagnosis strategies to maximize the preventive potential of post-diagnosis behavior change (PDBC) in Southern California men who have sex with men (MSM).

Background: Early diagnosis is a function of the ability of the time it takes tests to detect HIV after infection (“detection window”) and the timing of testing (relative to exposure). Currently used tests – either third or fourth generation enzyme immuno-assays (EIA) are able to detect HIV by the third week. Nucleic Acid Amplification Tests (NAATs) take

about a week less. The focus of this work is to utilize this interaction between the detection window of the test and the frequency testing to prevent new HIV infections.

Methods: Our models follow the same basic structure as Chapter 2. The behavioral data are again primarily from the Southern California Acute Infection and Early Disease Research program and all parameters are explained in Appendix F.

As in Chapter 3, we consider the testing frequency and level of awareness scenarios separately. In particular, we experiment separately with tests that have different detection windows and the frequency of testing. We also consider an individualized testing (IT) strategy, where each undiagnosed man tests after 3 non-main partners or 3 months (the 3/3 strategy), whichever comes first (or a comparable 6/6 strategy that recommends a test every 6 non-main partners or 6 months). Finally we experiment with a hypothetical test that is able to detect HIV one day after infection when every undiagnosed man tests every day (best case).

We repeat each experiment 10 times.

Results: The 6/6 IT strategy appears to be most effective in the testing frequency scenario and the 3/3 IT strategy appears most effective in the level of awareness scenario (ignoring the hypothetical best case) – than the next best strategies in each scenario respectively. This result is consistent across direct comparisons using t-tests, and the proportion of new infections averted (relative to the maximum possible number of infections).

4.1 Introduction

We saw the population-level effects of post-diagnosis behavior change (PDBC) in Chapter 3. In this chapter, we will discuss strategies that maximize the potential of PDBC as an intervention mechanism to reduce the incidence of new HIV infections over a ten year period. We will use the same baseline scenarios that we discussed in Chapter 3 and then introduce the new strategies that reflect current ideas on HIV prevention in MSM. We must clarify that the interventions we discuss here manifest their effects through PDBC; we do explicitly explore recent “treatment-as-prevention” strategies.

4.2 Background

An HIV infected individual is most infectious during the first 6 to 12 weeks of infection [12, 13, 14, 15, 16, 17, 18]. While the precise estimates for number of onward transmission events directly attributable to acute infection are variable [20], it is accepted that early detection of HIV during this phase is critical for infection management [19]. The fourth generation Enzyme Immuno Assay (EIA) has a detection window of 15-20 days [109] and is widely used (especially in industrialized economies) [20, 110, 111]. Other combination methods that combine antigen and antibody testing (for example the Combo RT) are available; however, one study showed that the antigen portion of the test is not successful at detecting acute infection, but the antibody test is [112].

An alternative is the rapid nucleic acid amplification tests (NAATs) and antigen tests that can detect HIV as early as 10 to 15 days after infection [109, 113]. Since the HIV NAATs detect the presence of HIV RNA they do not rely on the presence of antibodies to diagnose HIV infection [114]. However, HIV RNA tests are much more costly than the antibody tests [115]. Current recommendation on tests in high incidence populations include “pooled NAAT or fourth-generation assays” [116]. Pooled RNA tests involve combining sera from antibody negative individuals using certain algorithms into a distinct number of pools. Aliquots from these pools are then combined into master pools, and dilutions from these master pools are then tested for HIV RNA. If a master pool is positive, each pool contributing to the master pool is tested, then individual specimens from the positive pool(s) are tested [117]. Adoption of pooled NAATs by public health programs increased “costs, laboratory complexity and turnaround times for test results” [116], and fourth generation EIAs are “cheaper, have shorter turnaround times, are less labor-intensive” (David Katz, personal communication) and have only a slightly longer detection window.

Early diagnosis depends on the interaction of the test itself and the testing behavior of individuals at risk of infection. Current estimates reveal that MSM tend to test on average about once a year [106], but other data show a high proportion of MSM with undiagnosed infection [92]. Nationally, it is estimated that a fourth of the HIV-infected US population are unaware of their infection [113] and that 32.3% of MSM are diagnosed late [91], i.e.

within 12 months of developing AIDS. Thus the frequency of testing itself is a critical piece in diagnosing HIV early.

It is also hypothesized that the period when HIV is present but not detectable (“eclipse phase”) may be important in new infections – potentially as much as the detectable phase in prevention of new infections [109]. However, currently RNA tests have the shortest detection windows, but theoretical models that assume detection during this eclipse phase can help examine its importance.

Currently, the CDC recommends that all MSM get tested for HIV every year, and sexually active MSM get tested every 3 to 6 months [118]. There is interest in exploring more individually tailored testing strategies. For example, King, Pierce and Snohomish counties in Washington state have recently implemented the “find your frequency” program that encourages MSM to evaluate their risk based on the status of partners, number of partners, type of sex, drug use and history of sexually transmitted diseases. The website has a simple module that anyone can go through and makes a recommendation to test every 3 months or every 12 months [119]. The California Department of Public Health is interested in an individualized testing (IT) strategies (colloquially known as “oil change”) that propose MSM should test after a certain number of partners, or a certain period of time (Dan Wohlfeiler, personal communication). Two particular forms of this strategy are the “3/3” (where men test every 3 months or every 3 partners) and the “6/6” (where men test every 6 months or every 6 partners).

4.3 Methods

4.3.1 Overview

Our models follow the same basic structure as Chapter 3. The behavioral data are again primarily from the Southern California Acute Infection and Early Disease Research program [2, 3] and other published studies. Several other parameters primarily come from [27]; further details are in Appendix F.

Recall the mechanisms of PDBC we model (as explained in Section 3.3):

1. Reduction in total number of non-main partnerships: We model a 25.6% reduction in

mean-degree of non-main partnerships upon diagnosis in accordance with AIEDRP data (as seen in Figure 1 of [3]) where the mean number of partners before diagnosis is 8.37 in the three months prior to diagnosis, and 6.22 post-diagnosis, taking the mean over the four follow-up periods.

2. Selection of non-main partners by diagnosis status: We model an approximate reduction of 40% in the proportion of discordant non-main ties compared to the number expected if there was proportional mixing by diagnosis status.
3. UAI in Main Partnerships: Within main partnerships, a reduction in daily probability of UAI from 0.156 to 0.109. As explained above, the former is obtained from AIEDRP data and the latter is obtained from [27].

In our baseline models we assume a 22-day detection window for tests [27]. Our baseline models are built around a daily testing probability of $1/351$ in accordance with Helms et al. [106].

As we saw in Chapter 3, due to inconsistencies between data on testing frequency and awareness of infection we examined two separate scenarios: the “testing frequency” models and the “level of awareness” models. We examine various strategies for diagnosis under each of these scenarios.

We select one 50-year “burnin” case from Chapter 3 under the testing frequency and one from the level of awareness scenario to serve as a starting point for evaluating various diagnosis strategies. The simulation models follow the same structure as described in Chapter 3. We evolve our models in daily time steps, and each condition we test is over a 10 year period (3650 days). We repeat each experiment 10 times. Starting prevalence in the testing frequency case is 34.3%, and 44.1% in the level of awareness case.

4.3.2 Detection Windows and Testing Frequency

We experiment with detection windows of 1 day and 43 days, when the test has a 22-day detection window. This estimate is consistent with a fourth generation EIA, and in the middle of a third-generation EIA [109]. The Prevention Umbrella for MSM in Americas

team also used 22 days as a reasonable detection window [27]. We then experiment with testing every 3 months and 6 months with tests that have a detection window of 22 days. We assume that testing every undiagnosed man in these cases tests every 3 months, or every 6 months, respectively.

4.3.3 Individualized Testing (IT)

We perform 10 experiments for both the “3/3” and the “6/6” strategies under both testing frequency and level of awareness scenarios. In the 3/3 strategy, a test occurs for an undiagnosed individual at 3 non-main partners, or 3 months (90 days) whichever comes first. In the 6/6 strategy, a test occurs at 6 months (180 days) or 6 non-main partners, whichever comes first.

4.3.4 Best Case (Hypothetical)

As a hypothetical scenario, we experiment with a “best case” scenario where every undiagnosed individual tests every day and the test can detect serostatus one day after transmission. This set of simulations helps us analyze the limits of these strategies (given our models) and helps us experiment with hypothetical detection during the eclipse phase.

4.3.5 Comparing the Strategies

To compare the diagnosis strategies, we hypothesize their effectiveness in the following sequence (in increasing order of effectiveness): baseline models (detection window of 22 days), detection window of 1 day, half-yearly and quarterly testing (coupled with tests of a 22-day detection window), followed by the 6/6 and 3/3 IT strategies, and finally the hypothetical best case. We perform two-sample t-tests to compare consecutive strategies in this sequence: comparison of number of new infections produced with a test that has a one-day detection window with a test that has a 22-day detection window, half-yearly testing with a test that has a one day detection window, quarterly and half-yearly testing, the 6/6 IT strategy and quarterly testing, the 3/3 and 6/6 IT strategies, and the hypothetical best case with the 3/3 IT strategy.

We also quantify and present the number of infections averted relative by each strategy relative to baseline.

For a different measure to compare effectiveness of tests, we define the maximum number of infections averted as the absolute value of the difference in the mean number of infections obtained in the baseline models and the hypothetical best case models. We then evaluate the proportion of cases prevented in each category relative to this maximum number of infections averted.

We perform these comparisons in the testing frequency and the level of awareness scenarios. We assume a daily testing probability of $1/351$ in the testing frequency models and a daily testing probability of $1/4000$ in the level of awareness models (in conjunction with a 22-day detection window) to be consistent with the assumptions in Chapter 3. In both scenarios, a 1-day detection window is used in conjunction with the respective probabilities above in the testing frequency and level of awareness cases.

4.3.6 Testing Every 2 years

A study on MSM in Seattle found that increasing testing frequency from every 2 years to significantly reduced overall HIV prevalence [120]. To verify whether our results here are consistent with this finding, we experiment with a scenario where men test every two years and compare it to the baseline average frequency of one test every 351 days in the testing frequency scenario. We model testing as a memoryless process, where the waiting time until the next test is geometrically distributed, for the sake of comparison with the baseline models in the testing frequency case. We repeat each experiment 10 times, and conduct a t-test to compare if the the number of new infections in two-year testing are significantly different from the number of new infection in our baseline models over a 10 year period.

4.4 Results

4.4.1 Testing Frequency Scenario

Table 4.1 shows the number of new infections under the various strategies we examine over a 10-year period and Table 4.2 presents the number of infections averted relative to the

maximum possible preventions in the testing frequency scenario. We take a daily testing probability of $1/351$ (where the event of testing is a geometrically distributed memoryless process) with a detection window of 22 days as the baseline scenario.

NB It is counter-intuitive that the 6/6 IT strategy produces slightly fewer infections than the 3/3 IT strategy. But it is important to note that this difference is not statistically significant (p-value=0.53) and this slight difference can be attributed to the stochastic nature of these models.

Two sample t-tests at the 0.05 level showed that tests with a detection window of 1 day do not result in significantly fewer infections (p-value=0.94) over a 10 year period than tests with a detection window of 22 days (when the daily probability of testing is $1/351$). We also do not have evidence that half-yearly testing is significantly different than a detection window of 1 day (p-value=0.52), or quarterly testing (p-value=0.54). However, the 6/6 IT strategy is significantly better than quarterly testing (p-value=0.001). The 3/3 IT strategy does not perform significantly better than the 6/6 IT strategy (p-value=0.53), and the hypothetical best case is not significantly better than the 3/3 IT strategy (p-value=0.26).

The 6/6 IT strategy produces the biggest increase in number of infections prevented relative to the next best case. The 3/3 IT strategy does not perform significantly differently than the 6/6 strategy.

We also observe that testing every two years produces a mean of 781.0 (95% CI: 763.4,798.6) new infections over 10 years. This number is significantly different from the number of new infections in our baseline models (p-value = 0.01). Thus 2-year testing produced on average 5.6% more infections than our baseline scenario.

4.4.2 *Level of Awareness Scenario*

Table 4.3 shows the number of new infections under the various strategies we examine over a 10-year period and Table 4.4 presents the number of infections averted relative to the maximum possible preventions in the testing frequency scenario. We take testing once every 4000 days (where the event of testing is a geometrically distributed memoryless process) with a detection window of 22 days as the baseline scenario.

We find no significant difference (with regard to number of new infections in 10 years) between baseline models and tests with a detection window of 1 day (p-value=0.62). Half-yearly testing (using tests with a detection window of 22 days) is significantly better than using a test with detection window of 1 day (with one test every 4000 days on average) (p-value=0.002). Quarterly and half-yearly testing are not significantly different (p-value=0.21).

The 6/6 IT strategy, however, shows significantly fewer new infections over 10 years than quarterly testing (p-value < 0.001). The 3/3 IT strategy is significantly better than the 6/6 ITs strategy (p-value=0.009) and the hypothetical best case is significantly better than the 3/3 IT strategy (p-value=0.002).

The 6/6 and 3/3 IT strategies, and the hypothetical best case all produce significantly lower number of new infections than their respective next best strategies.

Graphical comparisons of various strategies in the two scenarios are presented in Figure 4.1. All three measures of effectiveness – number of new infections over 10 years, number of infections averted relative to baseline, and number of infections averted relative to the maximum number possible – demonstrate that the IT strategies are much more effective than either instituting tests with shorter detection windows or more frequent testing. This result holds in both the testing frequency and level-of-awareness scenarios.

4.5 Discussion

Our comparison of the various strategies according to the sequence in Section 4.3.5 suggests that in terms of preventing more new infections over a 10 year period, both the IT strategies appear to be most effective in the testing frequency scenario (when we ignore the hypothetical best case which is unrealistic to implement). Of the two implementation strategies, the 6/6 IT strategy presumably costs less to implement. Since we do not see a significant difference in the number of new infections produced between the two IT strategies, we recommend the 6/6 IT strategy as the most optimal. The 3/3 IT strategy appears most effective in the level of awareness scenario (ignoring the hypothetical best case). Thus the 6/6 IT strategy in testing frequency and the 3/3 IT strategy in level-of-awareness are most effective (relative to the next best strategies in each scenario respectively).

We find this conclusion holds both when we compare the number of new infections directly with the next best strategy, and in terms of the relative proportion of the number of infections averted relative to the maximum possible preventions in that scenario.

We notice a counter-intuitive result in Table 4.2 – that the 6/6 IT strategy averts more infections (relative to the maximum possible) than the 3/3 IT strategy. However, notice in Table 4.1 that the mean number of infections produced by the two strategies are close to each other, and not (statistically) significantly different. However, this difference appears magnified when we compare it to the maximum possible number of preventions.

We found that an average testing frequency of once every two years produced a significantly greater number of new infections over 10 years than the baseline testing frequency models of one test every 1/351 days. This result is consistent with a conclusion of a study of high risk MSM in Seattle [120]. We assumed a memoryless process of testing where the waiting times between tests are geometrically distributed to be consistent with testing assumptions in the baseline models. We only conducted this comparison in the testing frequency scenario, because in the level of awareness scenario there is no baseline model with annual testing.

One consistent theme through Chapters 3 and 4 of this work is the incompatibility between the testing frequency scenario and the level of awareness scenario. In Chapter 3, we saw that our dynamical models indicated that at close to yearly testing on average greater than 90% of infected should be diagnosed. We also discussed how HIV stigma potentially contributes to under-reporting of HIV status (Patrick Sullivan, personal communication). Moreover, unpublished data from other studies (NHBS – San Francisco, NHBS – New York, a clinical study in Massachusetts [108]) have indicated testing levels close to our assumption of one test every 351 days through [106] (Steven Goodreau, personal communication). Therefore it seems likely that the testing frequency scenario is closer to reality than the level of awareness scenario. However, analyzing both scenarios separately helps us put bounds around the effectiveness of strategies we studied in this chapter.

Given that the testing frequency scenario seems more plausible, our recommendation is that a 6/6 strategy might be most effective in preventing new infections over a 10 year period, given current trends in sexual behavior of Southern California MSM, as documented

in the AIEDRP study.

The hypothetical best case model under the testing frequency scenario is not significantly better than the 3/3 IT strategy (or the 6/6 IT strategy, though that comparison is not shown). While the best case is impossible to implement, the IT strategies are both being considered by the California Department of Health (Dan Wohlfeiler, personal communication). Our work suggests that these IT strategies have much prevention potential, and they should be utilized.

To simplify computation, in our models for IT we only counted the number of non-main partners. The preventive potential of these strategies may exceed our estimates if recommendations include counting main partners as well.

Currently, the “find your frequency” campaign in Washington state allows MSM to sign up for 3-monthly or 12-monthly reminders (via text messages) to test for HIV. High-risk MSM may find it more difficult to adhere to IT strategies than a simple message to test every so many days. However, how frequently one should test is a highly individualized quantity, and clearly depends on the level of risk that one engages in. Therefore, developing strategies to successfully implement the 6/6 IT strategy may be beneficial.

We did not conduct a formal cost-benefit analysis to compare these strategies; our analysis is clearly more focused on the benefits than the costs. As we mentioned above, one potential “cost” of the IT strategies may be lower adherence since it is easier to count the number of days since the last test (especially with a periodic reminder) than the number of partners. Similarly, we do not know (through this work) what the above testing strategies will demand in terms of money, human capital and other public-health resources.

Another limitation is that under the individualized testing strategies, people can potentially test an unlimited number of times (if they have a high enough number of non-main partners in a given period of time). In the real world, it is likely that men may suffer from “testing fatigue” and not test more than a certain number of times in a given time period. We also did not consider dependence between risk behavior and testing behavior when it is quite likely that men are possibly more likely to test after an episode of risky sex.

As we saw there is some speculation that shortening the eclipse phase may result in more effective prevention [109]. Our testing frequency models do not show that diagnosing people

even 1 day after infection makes much more difference than the IT strategies. The level of awareness cases show that under the extreme assumption of daily testing and a detection window of 1 day significantly more infections are averted. However, as we have seen, there is reason to believe that the testing frequency models are closer to reality than the level of awareness models, and therefore we speculate that simply shortening the eclipse phase (without altering testing behavior) may be of limited benefit.

Scientists are also working on determining “a specific intervention to demonstrate whether knowing that one has acute HIV infection specifically translates into more effective prevention than simply knowing that one is infected” [109]. We believe our work demonstrates that given the population-level effects of PDBC (as we saw in Chapter 3), and the effectiveness of IT strategies, knowledge of infection status (especially during acute infection) may result in averting a significant number of new infections via the mechanism of PDBC. Thus besides benefit to the infected individual (in terms of starting treatment), knowledge of acute infection also has potential population-level public health benefits.

We end by stating that this work does not address the preventive benefits or recent developments in “treatment as prevention” strategies. Our focus is on behavioral benefits, though starting early diagnosis has the potential to result in early treatment, and consequent prevention through this other route too.

Table 4.1: Number of new infections produced with different strategies in the testing frequency scenario. The error in the number of new infections is the 95% confidence interval about the mean. The number of new infections are over 10 repetitions of each experiment. “DPT” is daily probability of testing.

Model	Detection Window	Testing Pattern	New Infections
Baseline	22	DPT: 1/351	738.9 ± 20.1
Detection-Window	43	DPT: 1/351	757.3 ± 14.8
	1	DPT: 1/351	738.0 ± 10.2
Testing Behavior	22	Every 180 days	734.4 ± 3.4
	22	Every 90 days	730.1 ± 11.6
IT: 6/6	22	Every 180 days or 6 partners	696.5 ± 10.9
3/3	22	Every 90 days or 3 partners	704.9 ± 21.3
Best Case	1	1	689.9 ± 10.1

Table 4.2: Proportion of cases prevented by each strategy (relative to maximum possible) in testing frequency scenario. Maximum preventions is defined as the difference between the number of new infections in baseline and best-case scenarios.

Detection Window 1 day	1.8%
Half-Yearly Testing	9.2%
Quarterly Testing	17.9%
6/6 IT	86.5
3/3 IT	69.4

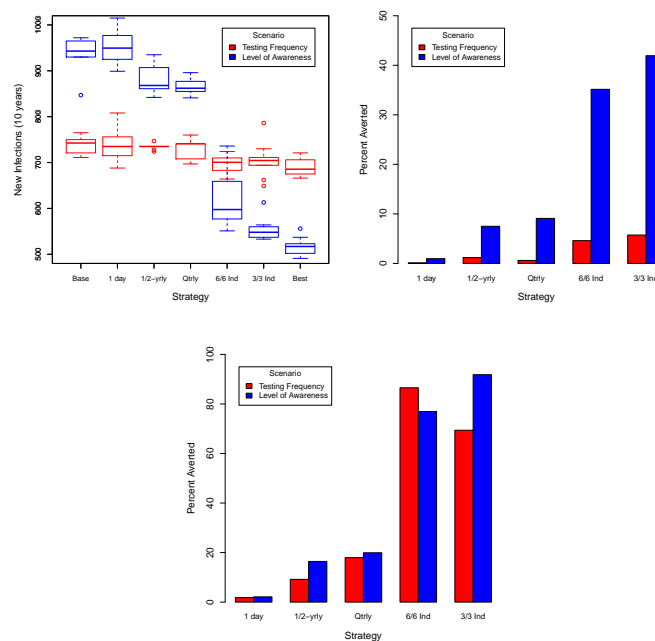
Table 4.3: Number of new infections produced with different strategies in the level-of-awareness scenario. The error in the number of new infections is the 95% confidence interval about the mean. The number of new infections are over 10 repetitions of each experiment. “DPT” is daily probability of testing.

Model	Detection Window	Testing Pattern	New Infections
Level of Awareness (baseline)	43	DPT: 1/4000	939.5 ± 12.6
	22	DPT: 1/4000	955.3 ± 21.1
	1	DPT: 1/4000	944.1 ± 25.7
Testing Behavior	22	Every 180 days	881.8 ± 18.5
	1	Every 180 days	866.6 ± 10.3
IT: 6/6	22	Every 180 days or 6 partners	618.2 ± 35.5
3/3	22	Every 90 days or 3 partners	553.6 ± 15.2
Perfect World	1	1	517.9 ± 10.8

Table 4.4: Proportion of cases prevented by each strategy (relative to maximum possible) in the level of awareness scenario. Maximum preventions is defined as the difference between the number of new infections in baseline and best-case scenarios.

Detection Window 1 day	2.1%
Detection Window 43 days	3.2%
Half-Yearly Testing	16.4%
Quarterly Testing	19.9%
6/6 IT	76.9%
3/3 IT	96.1%

Figure 4.1: Top Left: Number of new infections produced with different strategies in the testing frequency and level-of-awareness scenarios. The solid line shows the median value; the lower and upper ends of the box represent the 25th and 75th percentiles respectively and the whiskers extend to 1.75 times these percentiles in either direction. Top Right: Mean number of new infections averted by each strategy in each scenario relative to the number of new infections in baseline. Bottom: Mean number of new infections averted by each strategy in each scenario relative to the “maximum possible” number of new infections, where “maximum possible” is the difference in the number of new infections in the baseline and the best case. These distributions for number of new infections in the boxplots and the mean values in the bar-plots are over 10 repetitions of each experiment. On the x-axis, “1 day” shows number of new infections (over 10 years) with a test with detection window of 1 day, “1/2-yrly” shows clockwork testing every 6 months, Qtrly shows clockwork testing every 3 months, “6/6 Ind” shows the 6/6 Individualized Testing Strategy, “3/3 Ind” shows the 3/3 Individualized Testing Strategy, and “Best” represents the Best case.



Chapter 5

CONCLUDING THOUGHTS

As HIV progresses through its fourth decade, we continue to look for answers to control its transmission. Many successes have been gained through collaborative effort among many different communities over the past few decades. This dissertation is a humble effort that hopes to add something to our understanding of the dynamics of this infection, and how we can prevent it.

Here I examine the main conclusions that we saw through this work, and identify future directions that can be undertaken.

5.1 Circular Migrations and HIV Transmission

Circular migrations play a complex role in HIV transmission. Prior epidemiological work has identified the populations at risk, presented data on HIV prevalence in these populations, and described possible mechanisms (such as acute infection) that make migrations inextricably linked with HIV transmission. In this work, we examined the problem of understanding *how* migrations impact HIV. We found that to model this critical interaction (between migrations and acute infection) we need to account for actors and their partnerships explicitly, and model the necessary dependence in the structure – for example, our model should account for the fact that multiple partners of migrant men should be in two locations. We also found that “frequent migrations” that took place at intervals shorter than the window of acute infection produced a larger epidemic than infrequent migrations. Compartmental models that aggregate actors and their partnerships, and flows between the various compartments, do not capture the interaction between acute infection and HIV and the consequent impact of migration on transmission.

While this work is largely theoretical, we have built machinery (through the course of this project) that can now be adapted to study more population-specific questions on

migrations and HIV, and possibly inform public health practice. At least one such project is in progress in the Network Modeling Group at the University of Washington (PI: Dr. Susan Cassels) and this work may form a starting point for examining modeling questions that arise out of this project. Other researchers pursuing projects on migrations and HIV in Honduras (Anisha Desai, personal communication) and Kenya (Antoine Dany, personal communication) have also contacted me for possible collaborations that involve network modeling of migrations and HIV.

Given recent developments in methodologies of network modeling, there was also a growing demand for an explicit example of using compartmental and network models to study the same system using both frameworks of modeling. Fulfilling this demand was part of the motivation for this work.

5.2 Individual-Level Post Diagnosis Behavior Change and Population-Level Effects

One effective use mathematical models in epidemiology are useful for is describing connections between individual-level phenomena and population-level effects. Chapter 3 in this thesis is centered around this very theme, and shows how the desire to protect one's partners (even when one is infected) has much prevention potential.

The Acute Infection and Early Disease Research Program (AIEDRP) is a unique dataset in that it provides longitudinal information on the behavior of infected men who have sex with men (MSM), *after diagnosis*. Such data are needed to evaluate how infected MSM in this cohort behaved upon diagnosis. Recent epidemiological work has questioned the population-level protective benefits of such post-diagnosis behavior change (PDBC); however, our work demonstrates that to truly evaluate a temporal phenomenon such as PDBC, we need longitudinal data.

A much bigger study evaluating post-diagnosis behavior change (titled *Recent HIV Infection: New Prevention Challenges and Opportunities* and colloquially known as *Metromates*; PI: Dr. Pamina Gorbach) is in progress in Southern California. In fact, this project was undertaken with the explicit goal of working with *Metromates* data, but due to delays in the collection of those data we went back to AIEDRP. *Metromates* will infact have a richer set

of metrics on PDBC and the methodologies we developed for Chapter 3 will hopefully be adaptable to answer relevant questions in the Metromates project.

5.3 Examination of Diagnosis Strategies

Testing and early diagnosis of HIV are critical components of behavioral interventions. In Chapter 4 we compared various strategies commonly used to diagnose HIV in men, and certain strategies that are still under development. Our focus in Chapter 4 was on the relative benefits of each of these strategies (as measured by number of new infections over a 10 year period), and not on the costs themselves. However, aside from the hypothetical best case, other strategies still being developed that we examined are part of new testing packages that public health programs are considering. In fact, the individualized testing (IT) strategy (3 months/3 partners or 6 months/6 partners) was communicated to us from the California Department of Public Health. Dr. Patrick Sullivan at Emory University and Dr. Joanne Stekler at the University of Washington are both interested in formal analysis of various strategies using network models, and this work will hopefully provide at least a starting point for such projects.

Our work suggests that if MSM adhere to the regimen of an IT strategy then we can get close to preventing the “maximum” possible number of new infections. Of course, adherence is the big unknown here, and a possible motivator for future research work in this area. Explicit accounting of the costs is another area that should be developed as is accounting of possible dependencies between risk and testing behavior.

Current developments in HIV epidemiology are focused on using “treatment as prevention.” The preventive mechanisms we focus on here are behavioral, but strategies that rely on treatment can certainly act in synergy with these behavioral interventions and help curtail the impact of HIV. As a post-doctoral fellow in the International Clinical Center (Department of Global Health, University Washington), I will be adapting some of these programs and developing new ones to study the effectiveness of home-based counseling and testing strategies in prevention of new infections among serodiscordant couples and pregnant women in South Africa and Uganda. These strategies utilize “combination treatment regimens” that have shown much promise.

Appendix A

 **R_0 COMPUTATIONS FOR COMPARTMENTAL MODELS IN
CHAPTER 2**

We follow the algorithm outlined in [89] to compute the reproduction number R_0 . Our system consists of three infectious states for migrant-men (A_{MM}, C_{MM}, L_{MM}), women (A_{NF}, C_{NF}, L_{NF}), and non-migrant men (A_{NM}, C_{NM}, L_{NM}), giving a total of nine infected states in each location. Each of these groups (migrant-men, women and non-migrant men) have a population of susceptibles that form the uninfected state in each group. In the problem we really have two locations, but since the two locations are structurally identical, we base our calculations on one location and nine infected states. We ignore the migration parameter δ because we only consider one location. As our analysis has shown, the migration parameter has no effect on infection transmission in either interpretation of “contact” in the framework of compartmental models.

A.0.1 Notation and Mathematics

We define a column-vector \mathcal{F} as the rate of production of new infections in each of the nine infectious states outlined above. Thus \mathcal{F} consists of nine entries. Since the chronic and late-stages consist only of flows from the acute stage, no new infections are produced in the states. Therefore, vector \mathcal{F} consists of non-zero entries in the first, fourth and seventh positions, and zeros everywhere else.

We then define a new matrix \mathbf{F} as $\frac{\partial \mathcal{F}}{\partial \mathbf{x}}$ where x is each of the nine infectious states. Matrix \mathbf{F} then is a 9×9 matrix.

Our next step is to define a vector \mathbf{V} that consists of everything except the new infections in the nine infectious states. We then define matrix $\mathbf{V} = \frac{\partial \mathcal{V}}{\partial \mathbf{x}}$ where x is as defined above.

We then compute the dominant eigenvalues for \mathbf{FV}^{-1} at the disease-free equilibria to obtain R_0 .

A.0.2 Computation

Our computer-code to perform these calculations in Mathematica 8 is below.

```

Clear["Global`*"] (*Blank Slate*)

(* New Infections F*)

NMMU := SNMU + ANMU + CNMU + LNMU;
NNFU := SNFU + ANFU + CNFU + LNFU;
NNMU := SNMU + ANMU + CNMU + LNMU;

F1 := SMMU*tMMU*(ANFU/NNFU)*betaAM + SMMU*tMMU*(CNFU/NNFU)*betaCM +
      SMMU*tMMU*(LNFU/NNFU)*betaLM
F2 := SNFU*tNFUM*(AMMU/NMMU)*betaAM + SNFU*tNFUM*(CMMU/NMMU)*betaCM +
      SNFU*tNFUM*(LMMU/NMMU)*betaLM +
      SNFU*tNFUNM*(ANMU/NMMU)*betaAM + SNFU*tNFUNM*(CNMU/NMMU)*betaCN +
      SNFU*tNFUNM*(LNMU/NMMU)*betaLN
F3 := SNMU*tNMU*(ANFU/NNFU)*betaAM + SNMU*tNMU*(CNFU/NNFU)*betaCN +
      SNMU*tNMU*(LNFU/NNFU)*betaLN

(* Now create the matrix F*)

F := {{D[F1, AMMU], D[F1, CMMU], D[F1, LMMU], D[F1, ANFU],
      D[F1, CNFU], D[F1, LNFU], D[F1, ANMU], D[F1, CNMU], D[F1, LNMU]},
      {0, 0, 0, 0, 0, 0, 0, 0, 0},
      {0, 0, 0, 0, 0, 0, 0, 0, 0},
      {D[F2, AMMU], D[F2, CMMU], D[F2, LMMU], D[F2, ANFU], D[F2, CNFU],
      D[F2, LNFU], D[F2, ANMU], D[F2, CNMU], D[F2, LNMU]},
      {0, 0, 0, 0, 0, 0, 0, 0, 0},
      {0, 0, 0, 0, 0, 0, 0, 0, 0},
      {D[F3, AMMU], D[F3, CMMU], D[F3, LMMU], D[F3, ANFU], D[F3, CNFU],
      D[F3, LNFU], D[F3, ANMU], D[F3, CNMU], D[F3, LNMU]},
      {0, 0, 0, 0, 0, 0, 0, 0, 0},
      {0, 0, 0, 0, 0, 0, 0, 0, 0}

```

```

}
G = F;

Simplify[F];

(* Jacobian V *)

V1 := - gamma*AMMU - mu*AMMU; (*Took delta terms out here *)

V2 := gamma*AMMU - eta*CMMU - mu*CMMU; (*Took delta terms out here *)

V3 := eta*CMMU - (mu + mud)*LMMU;(*Took delta terms out here *)

V4 := -gamma*ANFU - mu*ANFU;
V5 := gamma*ANFU - eta*CNFU - mu*CNFU;
V6 := eta*CNFU - (mu + mud)*LNFU;
V7 := -gamma*ANMU - mu*ANMU;
V8 := gamma*ANMU - eta*CNMU - mu*CNMU
V9 := eta*CNMU - (mu + mud)*LNMU

V := {{D[V1, AMMU], D[V1, CMMU], D[V1, LMMU], D[V1, ANFU],
      D[V1, CNFU], D[V1, LNFU], D[V1, ANMU], D[V1, CNMU], D[V1, LNMU]},
      {D[V2, AMMU], D[V2, CMMU], D[V2, LMMU], D[V2, ANFU], D[V2, CNFU],
      D[V2, LNFU], D[V2, ANMU], D[V2, CNMU], D[V2, LNMU]},
      {D[V3, AMMU], D[V3, CMMU], D[V3, LMMU], D[V3, ANFU], D[V3, CNFU],
      D[V3, LNFU], D[V3, ANMU], D[V3, CNMU], D[V3, LNMU]},
      {D[V4, AMMU], D[V4, CMMU], D[V4, LMMU], D[V4, ANFU], D[V4, CNFU],
      D[V4, LNFU], D[V4, ANMU], D[V4, CNMU], D[V4, LNMU]},
      {D[V5, AMMU], D[V5, CMMU], D[V5, LMMU], D[V5, ANFU], D[V5, CNFU],
      D[V5, LNFU], D[V5, ANMU], D[V5, CNMU], D[V5, LNMU]},
      {D[V6, AMMU], D[V6, CMMU], D[V6, LMMU], D[V6, ANFU], D[V6, CNFU],
      D[V6, LNFU], D[V6, ANMU], D[V6, CNMU], D[V6, LNMU]},
      {D[V7, AMMU], D[V7, CMMU], D[V7, LMMU], D[V7, ANFU], D[V7, CNFU],
      D[V7, LNFU], D[V7, ANMU], D[V7, CNMU], D[V7, LNMU]},
      {D[V8, AMMU], D[V8, CMMU], D[V8, LMMU], D[V8, ANFU], D[V8, CNFU],
      D[V8, LNFU], D[V8, ANMU], D[V8, CNMU], D[V8, LNMU]}}

```

```
{D[V9, AMMU], D[V9, CMMU], D[V9, LMMU], D[V9, ANFU], D[V9, CNFU],
  D[V9, LNFU], D[V9, ANMU], D[V9, CNMU], D[V9, LNMU]}
}
```

```
Simplify[V];
```

```
VInv := Inverse[V]
```

```
Simplify[VInv];
```

```
FVInv := F.VInv
```

```
Simplify[FVInv];
```

```
finalmatrix = FVInv;
```

```
Eigenvalues[FVInv];
```

```
(*Disease Free Equilibrium*)
```

```
SMMU := 625; AMMU := 0; CMMU := 0; LMMU := 0;
```

```
SNFU := 2*SMMU; ANFU := AMMU; CNFU := CMMU; LNFU := LMMU;
```

```
SNMU := SMMU; ANMU := AMMU; CNMU := CMMU; LNMU := LMMU;
```

```
AMMR := AMMU; CMMR := CMMU; LMMR := LMMU;
```

```
(*****)
```

```
(*Biological Conditions *)
```

```
(*****)
```

```
(*Time Spent in Various Stages of Infection*)
```

```
mu := 1/(45*52);
```

```
mud := 1/40;
```

```
nu := 625/(425*52);
```

```
eta := 1/500;
```

```
gamma := 1/12;
```

```

(*Transmission Probabilities*)
chronicprb := 0.0007;
acuteprb := 0.0007*26;
lateprb := 0.0007*7;
(*****

(*****
(*Contact as Partnership Model*)
(*****

(* Migration parameter delta is not needed *)
c := 3;
tMMU := 0.016;
tMMR := tMMU;
tNMU := 0.0080;
tNMR := tNMU;
tNFUM := tMMU*NMMU/NNFU;
tNFUNM := tMMU*NMMU/NNFU;
dur100 := 100;
betaAM := 1 - (1 - acuteprb)^(dur100*c/2);
betaAN := 1 - (1 - acuteprb)^(dur100*c);
betaCM := 1 - (1 - chronicprb)^(dur100*c/2);
betaCN := 1 - (1 - chronicprb)^(dur100*c);
betaLM := 1 - (1 - lateprb)^(dur100*c/2);
betaLN := 1 - (1 - lateprb)^(dur100*c)

finalmatrix

{{0, 0, 0, -0.550839, -0.46377, -0.164031, 0, 0, 0}, {0, 0, 0, 0, 0,
  0, 0, 0, 0}, {0, 0, 0, 0, 0, 0, 0, 0, 0},
  {0, 0, 0, 0, 0, 0, 0, 0, 0}, {-1.10168, -0.927539, -0.328063, 0, 0,
  0, -1.81908, -1.64862, -0.485086}, {0, 0, 0, 0, 0, 0, 0, 0, 0}, {0,
  0, 0, 0, 0, 0, 0, 0, 0}, {0, 0, 0, -0.45477, -0.412155, -0.121272,
  0, 0, 0}, {0, 0, 0, 0, 0, 0, 0, 0, 0}, {0, 0, 0, 0, 0, 0, 0, 0, 0}}

```

```
Eigenvalues[finalmatrix]
```

```
{-1.19754, 1.19754, -4.68375*10^-17, 0., 0., 0., 0., 0., 0.}
```

```
(*Contact as Act Model*)
```

```
tMMU := 2.4;
```

```
tNMU := tMMU;
```

```
tNFU := tMMU;
```

```
tNFUM := tMMU*NMMU/NNFU;
```

```
tNFUNM := tMMU*NMMU/NNFU;
```

```
betaAM := acuteprb;
```

```
betaAN := acuteprb;
```

```
betaCM := chronicprb;
```

```
betaCN := chronicprb;
```

```
betaLM := lateprb;
```

```
betaLN := lateprb;
```

```
finalmatrix
```

```
{0, 0, 0, -0.794596, -0.536591, -0.231247, 0, 0, 0}, {0, 0, 0, 0, 0, 0, 0, 0, 0},
{0, 0, 0, 0}, {0, 0, 0, 0, 0, 0, 0, 0, 0},
{0}, {-1.58919, -1.07318, -0.462494, 0, 0, 0, 0, 0, 0},
{-1.58919, -1.07318, -0.462494}, {0, 0, 0, 0, 0, 0, 0, 0, 0}, {0, 0, 0, 0, 0, 0, 0, 0, 0},
{0, 0, 0, 0, 0, 0, 0, 0, 0}, {0, 0, 0, -0.794596, -0.536591, -0.231247, 0, 0, 0},
{0, 0, 0, 0, 0, 0, 0, 0, 0}, {0, 0, 0, 0, 0, 0, 0, 0, 0}, {0, 0, 0, 0, 0, 0, 0, 0, 0}}
```

```
Eigenvalues[finalmatrix]
```

```
{-1.58919, 1.58919, 9.04767*10^-17, 0., 0., 0., 0., 0., 0.}
```

Thus we obtain R_0 estimates of 1.19 and 1.58 when contact is defined as partnership and act respectively. In Figure 2.3, the former interpretation of contact yielded an equilibrium prevalence of 1.6% and the latter interpretation yielded an equilibrium prevalence of 36.7%. Given these prevalences, the estimate for R_0 appears high, especially in the contact-as-

partnership case. But these estimates are approximations, and the important result for us is the relative differences in the magnitude of the R_0 estimates in the two interpretations of contact, and the consistency of the higher estimate with the higher prevalence (as seen in the contact-as-act case).

Appendix B

THE ROLE OF ACUTE INFECTION IN CHAPTER 2

B.1 Introduction

The interaction between acute infection and the rate of migrations is a major theme in Chapter 2. Here, we quantify the population-level impact of acute infection as measured by overall HIV prevalence.

B.2 Methods

We assume only one infectious stage, with the probability of transmission per act equal to the weighted average of probabilities over all three stages, as considered in Chapter 2. This probability is therefore

$$\frac{12}{552}\beta_A + \frac{500}{552}\beta_C + \frac{40}{552}\beta_L = 0.0014 \quad (\text{B.1})$$

where β_A , β_C and β_L are transmission probabilities per coital act as defined in Chapter 2. We keep the duration of infection constant at 552 weeks, and assume that the rate of mortality during the infectious stage is a weighted average of non-AIDS and AIDS related rates of mortality. This mortality rate (μ_I) is

$$\frac{512}{552}\mu + \frac{40}{552}\mu_d$$

where μ and μ_d are as defined in Chapter 2.

We simulate the system of Equations in B.2.1 with the “contact-as-partnership” and “contact-as-act” interpretations and present the R_0 estimates for each case.

B.2.1 Compartmental Models

Our system of 12 ordinary differential equations (ODE’s) in any one location is now reduced to 6. We continue to use the same notation for all states and parameters as in Section 2.3.1, except we label the infectives I .

The set of equations for migrant men in the urban area then is

$$\begin{aligned} \frac{dS_{MMU}}{dt} &= \nu/8 - S_{MMU}t_{MMU} \frac{I_{NFU}}{N_{NFU}} \beta_{I,M^-} \\ &\quad \delta S_{MMU} + \delta S_{MMR} - \mu S_{MMU} \end{aligned} \quad (\text{B.2})$$

where $\beta_{I,M}$ is the probability of transmission between an infected migrant man and an uninfected woman, or vice-versa.

The change in population of infected migrating urban males is

$$\begin{aligned} \frac{dI_{MMU}}{dt} &= S_{MMU}t_{MMU} \frac{I_{NFU}}{N_{NFU}} \beta_{I,M^-} \\ &\quad \delta I_{MMU} + \delta I_{MMR} - \mu_I I_{MMU} \end{aligned} \quad (\text{B.3})$$

where μ_I is as defined above.

The set of equations for urban females is

$$\begin{aligned} \frac{dS_{NFU}}{dt} &= \nu/4 - S_{NFU}t_{NFU} \frac{I_{MMU}}{N_{MMU}} \beta_{I,M^-} \\ &\quad S_{NFU}t_{NFU} \frac{I_{NMU}}{N_{MMU}} \beta_{I,N} - \mu S_{NFU}, \end{aligned} \quad (\text{B.4})$$

$$\begin{aligned} \frac{dI_{NFU}}{dt} &= S_{NFU}t_{NFU,M} \frac{I_{MMU}}{N_{MMU}} \beta_{I,M^+} \\ &\quad S_{NFU}t_{NFU,NM} \frac{I_{NMU}}{N_{MMU}} \beta_{I,N} - \mu I_{NFU}. \end{aligned} \quad (\text{B.5})$$

The interactions of non-migrant urban males is give by

$$\frac{dS_{NMU}}{dt} = \nu - S_{NMU}t_{NMU} \frac{I_{NFU}}{N_{NFU}} \beta_{I,N} - \mu_I S_{NMU}, \quad (\text{B.6})$$

$$\frac{dI_{NMU}}{dt} = S_{NMU}t_{NMU} \frac{I_{NFU}}{N_{NFU}} \beta_{I,N} + \mu I_{NMU} - \mu_I I_{NMU}. \quad (\text{B.7})$$

This basic structure of equations (B.2) to (B.7) is the same for the various sub-populations in the rural area with the location sub-script U replaced by R (for rural). Thus, we now have a simplified two-location system of 12 interacting equations.

B.2.2 Network Models

The network models retain all the same features that are described in Chapter 2 except that all the stage-specific probabilities are replaced by a constant transmission probability as given in Equation B.1 above. We restrict our analysis to restricted network models (with deterministic migrations) here and experiment with migration intervals of 3 weeks and 30 weeks. We perform 10 repetitions at each migration frequency.

B.3 Results

B.3.1 Compartmental Models

Figure B.1 shows HIV prevalence in compartmental models with a single infectious stage. We obtain equilibrium prevalences of 30.1% and 0% with contact defined as partnership and as act respectively. The estimates for R_0 in the contact-as-partnership and contact-as-act cases are 1.16 and 1.48 respectively.

B.3.2 Network Models

In Figure B.2 we see that the prevalence trajectories at 3 weeks and 30 weeks are not significantly different. The average number of new infections obtained at the 3-week and 30-week migration frequencies are 3924 (95% CI: 3626.7, 4221.3) and 3788 (95% CI: 3538.0, 4038.0) respectively. The number of new infections over the 10 experimental repetitions are not significantly different (p-value=0.53).

B.4 Discussion

We find that in the contact-as-act case, equilibrium prevalence reduces by 17.8% relative to the model with acute infection in Chapter 2 and the R_0 estimate reduces by 6.3%. Equilibrium prevalence in the contact-as-partnership models – negligible (1.6%) with acute infection – drops to 0 and the R_0 estimate reduces by 2.5% when there is no acute infection.

In the network models, we see that the prevalence trajectories at migration frequencies of 3 weeks and 30 weeks are not statistically significantly different from each other. The number of new infections obtained over the 10 experimental runs are also not significantly different.

The maximum prevalence with acute infection for migrations at 3 weeks in Chapter 2 was seen to be about 15%, with no acute infection we see that this maximum prevalence drops to about 10%.

The interaction between acute infection and migration frequency is a critical focus of this work. We find that without acute infection the magnitude of HIV prevalence in both network and compartmental models reduces. We also observe that the significant difference between frequent (3-week) and infrequent (30-week) migrations in the restricted network models vanishes when there is no acute infection. Thus acute infection, and its interaction with migration frequency, play a critical role in the transmission of HIV in the context of circular migrations.

Figure B.1: Comparison of HIV prevalence as a function of time using compartmental models when there is only one infectious stage. The left-hand figure shows the result for sexual-contact defined as a long term partnership, and the right-hand figure shows the result for sexual-contact defined as a sexual episode. Frequent migrations have a migration-interval shorter than the acute phase of HIV infection (we consider migration-intervals of 1, **3**, 4, 7, and 10 weeks.) Infrequent migrations show migration-intervals longer than the acute-phase (we consider 20, **30**, 40, 60, 80 and 100 weeks).

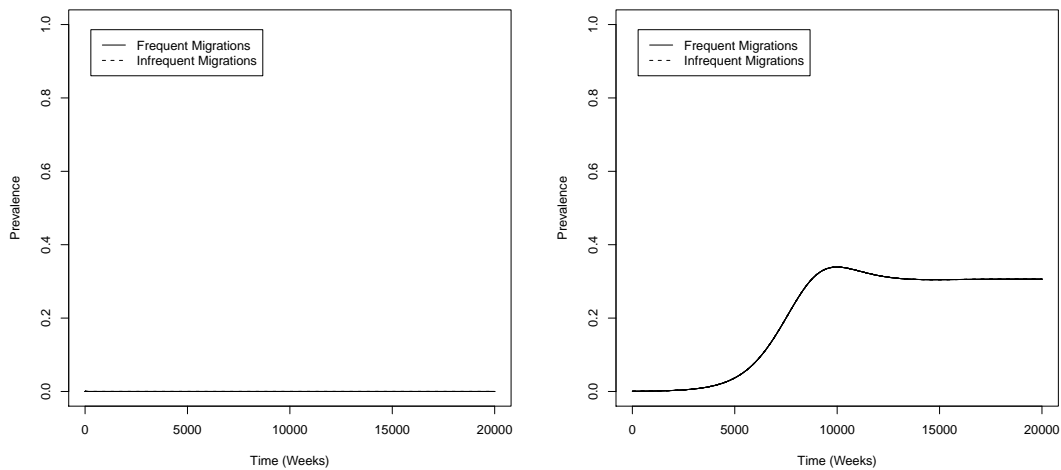
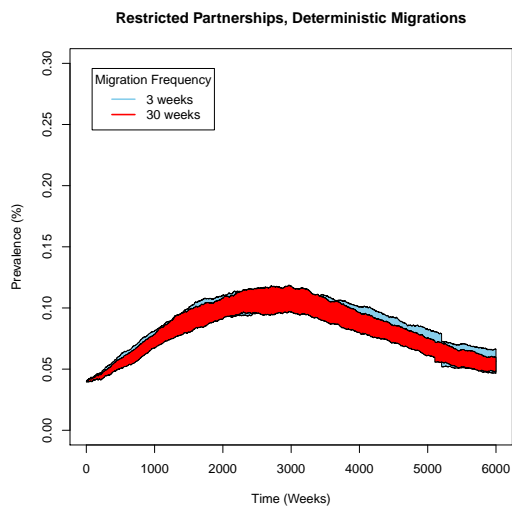


Figure B.2: Comparison of HIV prevalence as a function of time using network models without acute infection. We consider 10 repetitions each of migration-intervals of 3 and 30 weeks. The colored regions correspond to 95% confidence intervals about the mean.



Appendix C

**RESTRICTED NETWORK MODELS WITH STOCHASTIC
MIGRATIONS IN CHAPTER 2*****C.1 Introduction***

Our network-based models in Chapter 2 modeled migrations as a deterministic process in the “restricted” version. Thus our restricted models differed from the basic models in two ways: the restricted structure of the partnership structure of migrant men, and the model of the migration process. To investigate whether the differences we see between the basic and restricted models are due to the different partnership structure or due to the different migration model, we use restricted models but with the process of migration modeled as a stochastic process (identical to the migration model in the basic models in Chapter 2).

C.2 Methods

We create restricted network models as given by Equation 2.18 and model migrations as a stochastic process as described in Chapter 2:

Migrant men spend a given amount of time (on average) in one location. The inverse of this quantity gives the weekly probability of migrating to the other location in any particular week (time steps are in weekly units). Migration to the other location is thus a binary event.

As in Chapter 2, we experiment with migrations at intervals of 3 and 30 weeks with each experiment repeated ten times.

C.3 Results

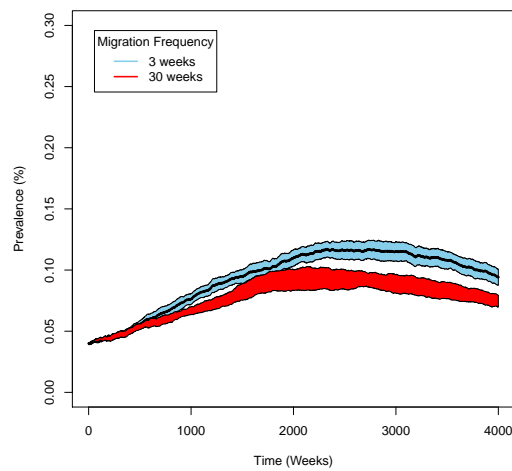
Figure C.1 shows mean prevalence (and 95% confidence intervals) for 10 experimental repetitions at migration frequencies of 3 weeks and 30 weeks over 4000 weeks. We found a

mean of 2779 (95% CI: 2652.5, 2905.5) new infections at 3-week intervals and 2308 (95% CI: 2159.5, 2456.5) new infections at 30-week intervals.

C.4 Discussion

We find that the prevalence trajectories at 3 and 30-week intervals are significantly different, with the more frequent migrations producing a greater prevalence and number of new infections through the course of the simulation. This result is qualitatively similar to the result with restricted models and deterministic (clockwork) migrations in Chapter 2. Thus, the significantly greater prevalence and number of new infections with migrations at 3-week intervals are attributable to the restricted partnership structure of migrant men, rather than the choice of framework (stochastic vs clockwork) for modeling the migration process itself.

Figure C.1: Comparison of HIV prevalence as a function of time using network models with “restricted” network models and migrations as stochastic processes. We consider 10 repetitions each of migration-intervals of 3 and 30 weeks. The colored regions correspond to 95% confidence intervals about the mean.



Appendix D

DEGREE DISTRIBUTIONS IN NETWORK MODELS IN CHAPTER 2

D.1 Introduction

To explain the differences in the overall prevalence between the basic and restricted network models, we investigate the structures of the distribution of partnerships cumulatively and cross-sectionally. These statistics are called “cumulative degree distributions” and “momentary degree distributions” respectively.

D.2 Methods

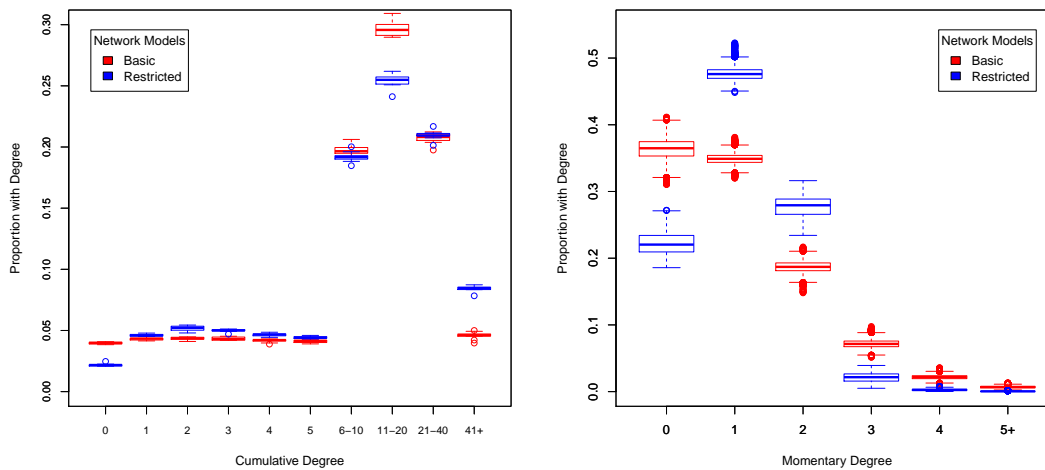
We compare cumulative and momentary degree distributions for migration intervals of 3 weeks in the basic and restricted network models (as shown in Section 2.5.2). The cumulative degree distributions are over 10 repetitions of each experiment. The momentary degree distributions for each time step are averaged over the length of the simulation and then over all 10 repetitions of each experiment.

D.3 Results

The cumulative and momentary degree distributions are in Figure E.2. We find that across the 10 experimental repetitions, the *maximum* cumulative degree (number of lifetime partners over a 45 year sexual lifespan) ranges between 92 and 108 in the basic network models and between 98 and 108 in the restricted models.

We notice that the restricted models have a longer tail in the cumulative degree distributions and the basic models have a longer tail in the momentary degree distributions. We also notice that the restricted models have fewer proportions of isolates both cumulatively and cross-sectionally. We also saw in Figure 2.4 in Section 2.5.2 that prevalence is greater in the basic models than in the restricted models.

Figure D.1: Cumulative (left) and momentary (right) degree distributions in network models with a migration interval of 3 weeks. The solid line in the box shows the median of the distribution and the ends of box are the first and third quartiles. The whiskers extend to 1.5 times the interquartile range in each direction. Note the ranges of the y-axes in the two plots are different because they measure different quantities.



D.4 Discussion

The lifetime distribution of partnerships has a longer tail in the restricted models while the momentary distribution of partnerships has a longer tail in the basic models. The higher prevalence in the basic models (Section 2.5.2) is due to the greater number of multiple partners at any given time (momentary degree) in the basic models. This impact of momentary degree on overall prevalence is an important result. The difference between the cumulative and momentary degree distributions (i.e. the distribution of lifetime and current partnerships) is often misunderstood and this work thus illustrates the epidemiological impact of these two related but distinct phenomena.

In Section 2.5.2 we investigated two migration intervals: 3 weeks and 30 weeks. We selected 3-week migrations in our comparisons for consistency and to isolate the effects of partnership structure by controlling for migration frequency. As we saw above, this choice enables us to account for the impact of the momentary degree distribution on prevalence trajectories. It is important to note the migration frequency and the structure of the network are independent of each other and therefore we could have chosen either migration frequency for our comparison.

One method to check the realism of behavioral assumptions in a model for sexually transmitted diseases is to analyze the number of partners over the sexual lifetime. Both the distribution of lifetime partnerships and the maximum number of lifetime partners in our models seem realistic over a 45-year sexual lifespan. We notice that we obtain a realistic equilibrium prevalence in the basic models, but the prevalence in the restricted models is lower than expected. This may be due to the maximum number of multiple partnerships for migrant men being restricted at 2 (as discussed in Section 2.6).

Appendix E

**ISOLATING THE IMPACT OF DELAYED PATH-ACCELERATION
AND MOMENTARY DEGREE DISTRIBUTION IN THE
RESTRICTED NETWORK MODELS IN CHAPTER 2**

E.1 Introduction

In Chapter 2, we explore two variations of network models: the “basic” models that are as close as possible to compartmental models, and the “restricted” models where migrant men are required to have multiple partners in both locations. The effect of delayed path acceleration is only created in the restricted models. However, an artifact of how we model the restricted networks is that the maximum momentary degree of migrant men is 2. Therefore, as explained in Appendix D, the momentary degree distributions in the basic and restricted models are qualitatively different.

Our objective here is to examine whether the differences we observed in the basic and restricted models (Section 2.3.2) are due to delayed path acceleration only being accounted for in the restricted models, or differences in the momentary degree distributions in the two versions. Therefore, we create “constrained models” where the momentary degree of migrant men is *constrained* to a maximum of 2, but multiple partners of migrant men are not required to be in separate locations. (Hence there is no “restriction” on the partnership structure of migrant men in these constrained models.) These constrained models will therefore help us tease apart the effects of delayed path acceleration and momentary degree distributions on the epidemic trajectories in the basic and restricted models.

E.2 Methods

We create network models using the exponential family random graph models (ERGMs) as explained in Chapter 2. We add a constraint to the basic models as defined by Equation 2.16; this modification limits the maximum momentary degree of migrant men to 2. We then experiment with 3-week and 30-week migrations and repeat each experiment 10 times.

We compare the mean prevalence trajectories (and 95% confidence intervals around them).

We analyze the cumulative and momentary degree distributions for the 10 repetitions at 3-week intervals to compare with the results in Appendix D.

E.3 Results

Figure E.1 shows the mean prevalence trajectories (and 95% confidence intervals) for the constrained models over 10 experimental repetitions at migration intervals of 3 and 30 weeks. We notice the overlap in these trajectories over much of the course of the simulation. With 3-week migrations, we also found a mean of 7646 (95% CI: 7509.5, 7781.5) new infections over 10 repetitions of the experiment. With 30-week migrations, we obtained 7922 (95% CI: 7810.7, 8033.3) new infections over 6000 weeks.

Figure E.2 shows the cumulative and momentary degree distributions over 10 repetitions of migrations at 3-week intervals. We also find that the maximum lifetime number of partners over a 45-year sexual lifespan has mean 118.4 (median: 119.0) over the 10 repetitions.

E.4 Discussion

We find that the constrained models do not show a significant difference either in the prevalence trajectories through the simulation between migrations at 3-week and 30-week intervals. The peak for 30-week migrations appears greater than the peak for 3-week migrations and therefore the number of new infections is greater at the longer migration interval. We investigated this finding further with taking an even slower migration frequency (100 weeks) in Figure E.3 and find that the prevalence trajectory at this frequency of 100 weeks closely aligns with the trajectory we see at 3 weeks. The 100-step migrations produced a mean of 7470.6 (95% CI: 7360.4, 7580.8) new infections over the course of the simulation (6000 weeks).

Thus the significant differences in the prevalence trajectories and number of new infections between frequent and infrequent migrations in the restricted models (of Section 2.3.2) are a consequence of the partnership structure of migrant men, and not an artifact of the maximum number of partnerships being constrained to 2. This finding emphasizes the importance of the partnership structure of migrant men, and the timing with which

they contact partners *in both locations*. Thus the significant differences in the outcomes of epidemic size and number of new infections are a consequence of appropriately accounting for the partnership structure of migrant men in the restricted models of Section 2.3.2, not a consequence of the artifact that the maximum degree of migrant men is constrained to a maximum of 2.

Figure E.2 shows that the cumulative degree distribution in the constrained models is more similar to the cumulative distribution of the restricted models (than the basic models) as seen by the proportion of actors with no partners, and a high number of lifetime partners (11-20 and greater). This structure is likely a consequence of the fact that the constrained and restricted models share the feature that the momentary degree of migrant men is capped at 2.

We also notice that equilibrium prevalence is higher in the constrained models (and basic models) relative to restricted models (as shown in Section 2.5.2). This difference in prevalence at equilibrium, and over the course of the simulation, is attributable to the momentary degree distribution. A greater proportion of individuals have greater than more than 1 partner at any given time in the basic and constrained models, relative to the restricted models, even though the mean number of partnerships is (statistically) equal. We have already seen the importance of the number of concurrent partners at any moment in time in overall HIV prevalence, and this difference in epidemic size between the constrained (and basic) models, and restricted models, is a consequence of the greater proportion of multiple partners at any given time in the constrained models.

Figure E.1: Comparison of HIV prevalence as a function of time using “constrained” network models, where there is no restriction on location of multiple partners of migrant men but the maximum number of partners at any time is constrained to 2. We consider 10 repetitions at migration intervals of 3 and 30 weeks. The colored regions correspond to 95% confidence intervals about the mean.

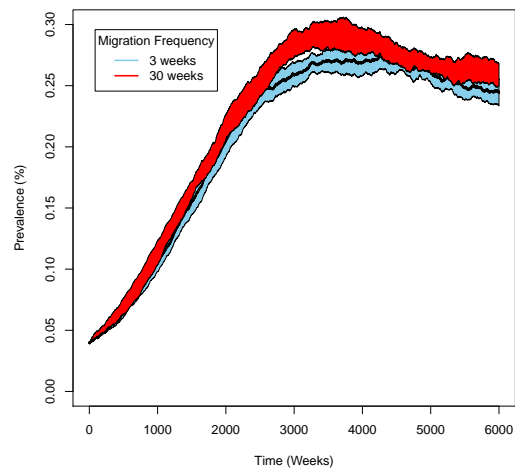


Figure E.2: The figures on the top show cumulative degree distributions (lifetime distribution of number of partnerships) in the entire population (left), and only for migrant men (right) in basic, restricted and constrained network models with a migration interval of 3 weeks. The figures on the bottom show momentary degree distributions (cross-sectional distribution of number of partnerships) in the entire population (left), and only for migrant men (right) in the three versions of network models, with a migration interval of 3 weeks. The solid line in the box shows the median of the distribution and the ends of box are the first and third quartile. The whiskers extend to 1.5 times the interquartile range in each direction. Note the ranges of the y-axes in the two plots are different because they are measuring different quantities.

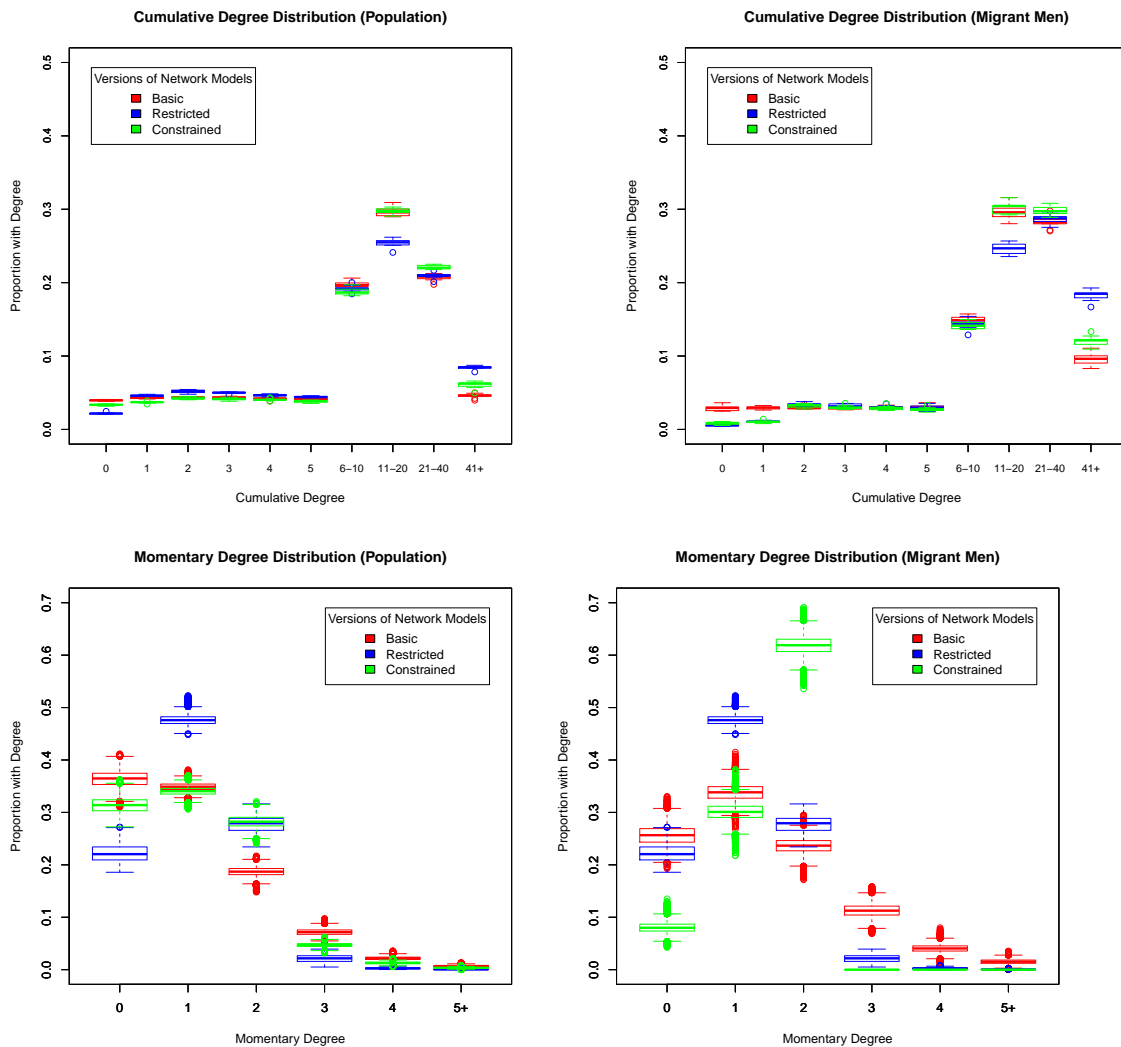
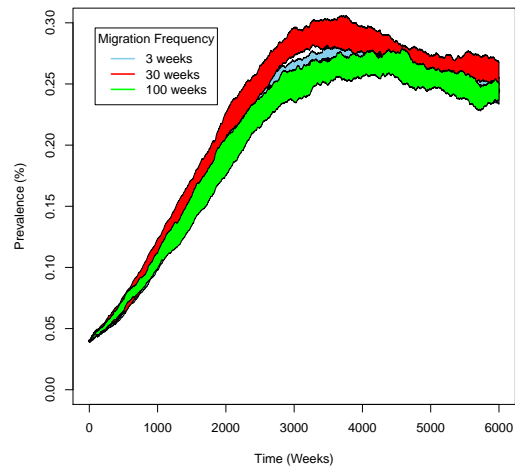


Figure E.3: Comparison of HIV prevalence as a function of time using “constrained” network models, where there is no restriction on location of multiple partners of migrant men but the maximum number of partners at any time is constrained to 2. We add a 100-week migration for the sake of comparison. We consider 10 repetitions at migration intervals of 3 and 30 weeks. The colored regions correspond to 95% confidence intervals about the mean.



Appendix F

**TECHNICAL APPENDIX TO CHAPTER 3: MODEL
FORMULATION, PARAMETER ESTIMATION, SIMULATION, AND
COMPREHENSIVE LISTING OF DATA SOURCES**

In this Appendix, we describe several technical details that are not in the main body of the chapter. We start with an exposition of the non-main network, parameters used to model this partnership network, followed by the estimation procedure for the statistics for these parameters. We repeat the same process for the network model for the main network. We follow with the derivation of statistics for the no post-diagnosis behavior change (no PDBC) model. We describe the various biological and demographic processes that are used in the simulation to model disease transmission and evolution of the network over time (while keeping the statistics of the main and non-main networks within stochastic variation of the mean). Finally, we list our sources for the various behavioral, biological and demographic parameters.

Any item marked “AIEDRP” represents data from the Southern California Acute Infection and Early Disease (AIEDRP) study [2, 3]. A number of biological and demographic parameters are selected from recent modeling work on HIV in MSM called the Prevention Umbrella for MSM in the Americas (henceforth PUMA) [27] and the sources for those parameters are marked “PUMA.”

F.1 Model for Non-Main Partnership Network

The network for non-main partnerships in our network is cross-sectional. We estimate this network once at the start, and then simulate a cross-sectional network from it at each timestep using exponential random graph models in the *statnet* package [24].

F.1.1 Model and Parameters

This non-main network is

$$\begin{aligned}
\text{logit}(p(y_{ij} = 1|Y_{ij}^c)) &= \theta\delta(e) + \theta_a\delta_a e(|\sqrt{a_i} - \sqrt{a_j}|) + \\
&\sum_{c=0, c \neq 3}^5 \theta_c\delta_c(e(c)) + \sum_{r_1, r_2} \theta_{r_1, r_2}\delta_{r_1, r_2}m(r_1, r_2) + \sum_{d_1, d_2} \theta_{m, d_1, d_2}\delta_{d_1, d_2}m(d_1, d_2) + \\
&\theta_{I, R}\delta_{I, R}m(I, R) + \theta_A\delta_A e(A)
\end{aligned} \tag{F.1}$$

where $\delta(e)$ is the total number of non-main ties, $e(|\sqrt{a_i} - \sqrt{a_j}|)$ is the mean difference in the square root of ages of all partners multiplied by the number of edges, $m(r_1, r_2)$ is the mean number of ties between all combinations of races (one is left out to avoid collinearity with edges), $m(d_1, d_2)$ is the mean number of ties between men of the same diagnosis status (discordant is left out to avoid collinearity with edges), $m(I, R)$ is the average number of ties between strictly insertive and receptive (should be 0 but set to 1 for model to converge, though the coefficient corresponding to this statistic is set to $-\infty$), $e(A)$ is the number of ties of men who have progressed to AIDS. The δ terms corresponding to each of these statistics are the changes in their values associated with a toggle of a dyad (i.e. switching the value of any dyad) [51]. This list describes the parameters in our model, the θ terms are the coefficients associated with each of these parameters.

F.1.2 Parameter Estimation

1. From AIEDRP data, we know the number of non-main partners reported over the past three months. We divide this number by 90 days to get the mean number of non-main partners per day as 0.10. However, this number is not limited to unprotected intercourse, therefore, we divide this mean by the proportion of times unprotected sex occurred (from the computation for main network above) to obtain a mean number of non-main partners of 0.075.

However, this number is restricted to negative/undiagnosed men. From AIEDRP, we know that the number of non-main partners reported over the past three months at baseline and follow-up differ by about 25.6%. We also know from NHBS [92] that about half of MSM nationally are unaware of their status, and about 25% of MSM are

positive. Therefore, we roughly estimate about 13% of MSM nationally to be unaware of their status. The weighted mean degree for MSM, accounting for both positively diagnosed and negative/undiagnosed men, therefore, is

$$0.075 - (0.256 \times 0.075) = 0.056.$$

It follows that the expected number of non-main ties for our network of 5000 men is

$$\frac{0.075(0.87)(5000) + 0.056 \times (0.13)(5000)}{2} = 181.3$$

2. Casual Activity Classes: To compute the activity levels of each class, we divide the rates on UAI in non-main partnerships in to five groups by percentile. Then we compute the mean for each of these groups. However, this rate of non-main UAI applies to an entirely negative/undiagnosed population. To make these means compatible with a population of positively diagnosed individuals, we reduce the means for each of the groups by 25.6%. We then obtain the means for the five groups as 0.00017, 0.0119, 0.028, 0.0623, 0.1705.
3. Age: The difference in square roots of ages for the non-main partners is 0.75 years. Therefore the mean statistic is $0.075 \times 181.3 = 135.9$.
4. Mixing Matrix for Race: The AIEDRP data set was too sparse to reliably estimate race mixing in the population. Therefore, we used a study of MSM in San Francisco from 2009 [104] to estimate the “degree of mixing” between any two races (There are slight inconsistencies in the number of partnerships between X and Y as reported by X and Y ; these are resolved by taking the mean of the two reports.). We do not assume the the race distribution in LA County and San Francisco is the same; our data on the the number of individuals in each race for our population of 5000 men come from LA County [121].

We use the number of individuals in each race to compute the total number of possible “dyads” between all combinations of races (i.e. the maximum number of ties between

two race groups). The San Francisco data [104] give us the proportion of partnerships in any combination of races that are actually realized (relative to the maximum possible number of dyads). We do assume that the mean proportion of possible ties between any two races that are realized in the two locations is the same.

Given the number of individuals in each race in our hypothetical population we know the number of possible dyads for each combination of races, and the proportion of possible ties between any two races that are realized. We also have an estimate for the total number of non-main ties in our population (from AIEDRP), and the distribution of these ties across the four races from [104]. Thus we constrain the total number of non-main ties in our mixing matrix to equal our estimate on non-main ties in the model from AIEDRP.

Thus we obtain the mixing matrix for the non-main network in Table F.1.

5. **Diagnosis Status:** From our AIEDRP data set we know that of all the ties that belong to diagnosed individuals, 35.6% are between positively concordant partners who are aware of their statuses (averaging the proportion of partners non-main partners who are positive at each of the four follow-up visits). Since the total degree for positively diagnosed individuals is 36.4, the number of positively concordant ties is

$$\frac{36.4 \times 0.356}{2} = 6.5.$$

6. **Meth Use:** Meth-users in our AIEDRP sample (defined as respondents who have used meth with at least one of the last three partners at baseline) have a mean rate of UAI of 0.101. Of the 194 respondents, $62/194 = 31.9\%$ have used meth. Therefore, the total number of degrees for meth users is

$$5000 \times 0.319 \times 0.101 = 161.3.$$

7. **Sexual Role:** We set the number of ties between strictly insertive and strictly receptive partners as 1 but then set the coefficient for the estimate to $-\infty$ to not have any possible ties between strictly versatile groups. We had to set the number to 1 initially for the models to converge.

8. Reduction in UAI with non-main partners on transition to AIDS: 40%.

F.2 Model for Main Partnership Network

The network of main partnerships is estimated using separable temporal ERGMs [51] also implemented in *statnet* [24]. The variable $y_{ij,t}$ (i.e. the value of the dyad y_{ij} at time t) takes on the value 1 if nodes i and j have a main partnership at time t , and it is 0 otherwise. The variable $Y_{ij,t}^c$ is the rest of the network excluding the tie information of i and j . The variable u is the total number of ties of a particular type, and m represents mixing between two types. The variable v represents nodes with at least two main partnerships.

F.2.1 Model and Parameters

The model is

$$\begin{aligned} \text{logit}(p(y_{ij,t} = 1 | y_{ij,t-1} = 0, Y_{ij,t}^c)) &= \theta \delta(e) + \theta_a \delta_a(e(|\sqrt{a_i} - \sqrt{a_j}|)) + \sum_{d=0}^2 \theta_d \delta_d(\eta(d)) + \\ &\sum_r \theta_r \delta_r(e(r)) + \sum_r \theta_{M,r} \delta_{M,r}(M(r)) \end{aligned} \quad (\text{F.2})$$

where e is the number of edges in the network, $e(|\sqrt{a_i} - \sqrt{a_j}|)$ is the difference in the square roots of the ages of all main partners, times the number of edges, $\eta(d)$ is the total number of nodes with degree 0, 1 and 2 in the network, $e(r)$ is the mean number of ties for each of the four races, $M(r)$ is the mean number of ties between partners belonging to the same race. All the above statistics are measured at time t . The δ functions are “change statistics” defined as the change in the statistics associated with a toggle of a dyad (i.e. switching the value of any dyad) [51]. The θ terms are the coefficients for each of these statistics. The network $Y_{ij,t-1}^c$ is the network at the time step $t - 1$ without the dyad between the particular pair of actors i and j .

The dissolution model is a Bernoulli model of the form

$$\text{logit}(p(y_{ij,t} = 0 | y_{ij,t-1} = 1)) = \theta_{\text{diss}} \quad (\text{F.3})$$

where θ_{diss} is the coefficient associated with the dissolution of one tie (which corresponds to a change statistic of 1, in absolute value).

F.2.2 Statistics for Network of Main Partnerships

1. **Momentary Degree Distributions:** From the AIEDRP questionnaire, we have information on the number of ongoing main partnerships in a sample of 194 men at baseline. This information is scaled to a set of 5000 men in Table F.2.

We define the last category as “2+” and constrain the maximum number of partnerships in our main network (F.2) to 2.

2. **Total Number of Partnerships:** From Table we estimate a mean degree of 0.36. Therefore, the total number of partnerships in our population is $5000 \times 0.36/2 = 900$.
3. **Mean Difference in Ages:** We compute the mean difference in the square roots of the ages of main partners in the AIEDRP data set. This difference is 0.61. Since the modeled population has 900 main edges on average the mean statistic for difference in ages is $0.61 \times 900 = 549$.
4. **Mixing Matrices by Race:** As in non-main partnerships, we again use the data from San Francisco [104] in conjunction with the race distributions from LA County [9] to estimate mixing by race. This estimated mixing matrix for the main network is in Table F.3.
5. **Mean Partnership Duration:** In the AIEDRP data set, the mean partnership durations are 906.7 days at baseline and 921 days at follow-up. Since the follow-up data set consists mostly of last partner information, we assume that each of the main partnerships in the follow-up data set are existent. Therefore the mean partnership duration in the follow-up data set is

$$(906.7 \times 0.87 + 921 \times 0.13) = 908.6$$

days assuming that our population consists of 87% individuals who are aware of their diagnosis status and 13% who are unaware. We estimate durations from the extant ties because our assumption is that the time for which a tie exists is geometrically distributed, and given this assumption, the mean of extant partnerships at any given time gives an unbiased estimate of the duration. Complete mathematical details are in [122].

6. Probability for UAI in Main Partnerships: From the baseline information in AIEDRP we know the number of sex acts in main partnerships, the length of main partnerships, and the proportion of times unprotected intercourse occurred. From this information, we compute the mean probability of UAI in main partnerships for negative/undiagnosed individuals as 0.156 (almost identical to the estimate found in PUMA through other studies).

We also need the mean probability of UAI for positive and diagnosed individuals with negative/undiagnosed individuals. A similar method applied to follow-up information in AIEDRP yields a mean probability of 0.311. However, in the follow-up information in AIEDRP it is impossible to distinguish between unprotected oral and anal acts. It is very likely that upon diagnosis individuals switch from oral to anal intercourse. However, we only model transmission via UAI; therefore, this mean probability cannot be used. We therefore use the mean probability 0.109 of UAI for negative/undiagnosed individuals with positive individuals from [27].

7. Sexual Role: We set the number of ties between strictly insertive and strictly receptive partners as 1. Theoretically, this number should be zero, but we specified it as 1 for our models to converge, but as explained above, the coefficient is set to $-\infty$.
8. Reduction in UAI with Main Partners on transition to AIDS: 40% from the Rakai Study of discordant heterosexuals [15]. ¹

¹While we assume an equivalent reduction in daily probability of UAI after transition to AIDS in non-main and main networks, the modeling method itself is different. In the non-main network, daily UAI defines the network so this 40% reduction is modeled as a reduced propensity to form ties. In the main

F.3 Statistics to Model Counter-Factual with No PDBC

1. **Non-Main Network:** The mean degree is only based on the behavior of negative/undiagnosed men. This mean degree is 0.075, translating to an average of 188 ties in the non-main network. Correspondingly, the casual activity classes change to 0.0023, 0.01595, 0.03798, 0.08415 and 0.2292. The mean statistic for the difference in ages is therefore $0.075 \times 188 = 141$. The cells in the race mixing matrix now sum to 188. This matrix is in Table F.4.
2. **Main Network:** We keep the probabilities for UAI in discordant partnerships the same as that in negatively concordant partnerships.

The statistic corresponding to average number of non-main ties for meth users increases by 4% consistent with the per cent increase in the total number of non-main ties.

We do not consider mixing by diagnosis status since there is no diagnosis induced behavior change in this scenario and therefore mixing by diagnosis status is contingent simply upon the proportions of the two diagnosis states that are present in the population.

F.4 Other Biological and Demographic Processes

1. **Infection Transmission:** In accordance with Model 2 in [27], we model transmission events and their probabilities broken down by roles of the two partners, and their circumcision states. These relative transmission probabilities in UAI are from [123] – a study of heterosexuals. We used data from a cohort of heterosexuals because it presented time varying risk of UAI (as opposed to other widely cited studies on the topic [15, 16, 18, 17]).

network, daily UAI (and consequent possible disease transmission) are modeled as events distinct from the network; therefore, upon transition to AIDS we model a 40% reduction in daily probability of UAI that does not affect the partnership network itself.

2. Viral Load in absence of treatment: We model the viral load trajectory in each man as a six-parameter curve in the following manner [27]:
 - (a) Days 0-21: rises linearly from 0 to 6.886
 - (b) Days 21-40: declines linearly from 6.886 to 4.5 Days 40-3650: assumes a set point of 4.5 [124] that lasts until the onset of AIDS 10 years post-infection [125]
 - (c) Days 3650-4380: linear rise from 4.5 to 7.0
 - (d) Day 4380: death
3. Treatment Trajectories: We follow the basic structure of PUMA. We assume that treatment results in either no suppression, partially effective suppression and full suppression. For each race, we assume that for 15.0 % men there is no suppression, for 29.8% men there is partial suppression, and for 55.3% men there is no suppression [27].
4. As per [27], we assume a 40% reduction in UAI once AIDS stage is reached. This reduction is modeled in the form of a reduction of 40% reduction in daily probability of UAI in main partnerships and a 40% reduction in men number of non-main partners per day.
5. Treatment Initiation: In accordance with PUMA, we model treatment initiation as 4.1 years for Blacks, 5.0 years for Latinos, and 3.6 years for Whites and Others, measured since time of of infection [27].
6. Reduction in UAI upon transition to AIDS stage: We model a 40% reduction in probability of UAI with main partners and a 40% reduction in mean number of daily UAI events with non-main partners [27].
7. Circumcision: Circumcision rates are parametrized in accordance with PUMA [27]. Xu (2007) reports the rates of circumcision in various racial groups in the National Health and Nutrition Examination Surveys (NHANES) as 88% in non-Hispanic whites,

73% in non-Hispanic blacks, 42% in Mexican-Americans, and 50% in others [126]. We use these rates for Whites, Blacks, Latinos and Others in our population respectively. We assume that for a fully recovered circumcised individual infectivity reduces to 40% of the level initially. If full recovery from the surgery has not occurred, infectivity rises to 130%. Both these changes in infectivity are in accordance with PUMA.

8. Role: As per PUMA we assume that 7% of Other, 13% of Black, and 9% of Latino MSM are strictly insertive and strictly receptive. PUMA does not report the proportion of Whites who are strictly insertive and strictly receptive; but the “Other” category in PUMA includes Whites. We assume that the proportion of strictly insertive or strictly receptive MSM for Whites is the same as the proportion of Other in this work.
9. Testing: We based our baseline testing models on PUMA. A clinical study in four major US metropolitan centers [106] reports a median inter-test interval of 243 days. If we assume a simple exponential distribution for time until testing, this median measure corresponds to a mean inter-test interval of 351 days. As we discussed this mean rate of testing corresponds to a mean of 95% of men who have ever tested. Another study [127] that presents NHBS data (from June 2004 to April 2005) reports that 92% of men have ever tested. Thus our baseline testing models are consistent with NHBS data from 2004 to 2005, though not so with NHBS data from 2008 (as we discussed in the main body of the paper).

F.5 Itemized List for Data Sources

1. Initial prevalence: NHBS (2008) [92]
2. Age-specific non-AIDS mortality rates: National Vital Statistics Reports [128]
3. Race distributions: LA County Department of Health Services [121]
4. Circumcision rates by race: National Health and Nutrition Examination Surveys (1999–2004) [126]

5. Prevalence of role exclusivity: PUMA
6. Proportion of positive men who never receive treatment: PUMA
7. Proportion of Treated Men who Achieve Full Suppression: PUMA
8. Role versatility in non-main partnerships: PUMA
9. Daily probability of HIV testing: Clinical study from 4 major US urban centers (Seattle-King County, San Francisco, Denver and District of Columbia) [106]
10. HIV Test Window Period: 22 days, and treated as a sensitivity parameter [27]. A detection window of 22 days is reasonable for fourth generation and RNA tests [109].
11. Initial age distribution: Uniformly distributed between 18 and 65 years for every man in the population.
12. Non-Main Partnerships
 - (a) Race Mixing matrix: Study of MSM in San Francisco [104] in combination with race distribution from LA County Department of Health Services [121]
 - (b) Mean age difference of main partners: AIEDRP
 - (c) Quintiles for activity classes in Non-Main Partnerships: AIEDRP
 - (d) Number of Partnerships for meth users and non-users: AIEDRP
 - (e) Mixing by diagnosis status: AIEDRP
13. Main Partnerships:
 - (a) Baseline daily probability of UAI: AIEDRP for negatively concordant partnerships, PUMA for partnerships discordant by diagnosis status
 - (b) Momentary (cross-sectional) degree distributions for main partnerships: AIEDRP

- (c) Race mixing matrix: Study of MSM in San Francisco [104] in combination with race distribution from LA County Department of Health Services [121]
- (d) Age mixing in the main partnerships: Mean difference in ages of main partners in AIEDRP
- (e) Distribution for number of partnerships at cross-section: AIEDRP
- (f) Average Partnership Duration: AIEDRP
- (g) Reduction in UAI with Main Partners in AIDS stage: Study of HIV discordant couples in Uganda [15]

Biological Data: Estimates for all of the following parameters are from PUMA.

1. Time until peak of acute viremia
2. Time from peak viremia until set point
3. Set point viral load
4. Time from onset of set point until AIDS-related viral increase
5. Time from onset of AIDS-related viral increase until death
6. Slope of viral load during AIDS
7. Viral load at full suppression
8. Mean viral load for those partially suppressed
9. Time from initiation of treatment until partial suppression
10. Time until partial suppression escape
11. Partial escape slope
12. Probability of transmission by act

Table F.1: Race Mixing Matrix for the Non-Main Network. The rows represent respondents, and the columns represent the partners.

	Race Groups				Total Degree	Mean Degree
	W	B	L	O		
W	24.4	8.2	51.7	20.0	128.6	0.0715
B		4.0	10.8	2.0	29.0	0.0724
L			38.2	16.6	155.5	0.0723
O				4.2	47.1	0.0725

Table F.2: Degree Distributions for Main Partnerships for network of size 5000.

Number of Partners	0	1	2	3	Not Reported
Number of Actors	3351	1546	52	51	0

Table F.3: Race Mixing Matrix for the Main Network. The rows represent respondents, and the columns represent the partners.

	Race Groups				Total Degree	Mean Degree
	W	B	L	O		
W	121.0	40.5	259.5	99.1	641.1	0.36
B		19.8	54.2	10.0	144.3	0.36
L			194.0	82.6	784.9	0.36
O				20.9	233.6	0.36

Table F.4: Race Mixing Matrix for the Non-Main Network in the counter-factual that does not include PDBC. The rows represent respondents, and the columns represent the partners.

	Race Groups				Total Degree	Mean Degree
	W	B	L	O		
W	25.2	8.4	54.6	20.3	133.7	0.0743
B		4.1	11.3	2.0	29.9	0.0748
L			41.3	17.0	165.5	0.077
O				4.2	47.7	0.0734

BIBLIOGRAPHY

- [1] Susan Cassels, Samuel J. Clark, and Martina Morris. Mathematical models for HIV transmission dynamics. *JAIDS- Journal of Acquired Immune Deficiency Syndrome*, 47(Supplement 1), March 2008.
- [2] P. M. Gorbach, L. N. Drumright, E. S. Daar, and S. J. Little. Transmission behaviors of recently HIV-infected men who have sex with men. *Journal of Acquired Immune Deficiency Syndromes*, 42(1):80–85, May 2006.
- [3] Pamina M. Gorbach, Robert E. Weiss, Robin Jeffries, Marjan Javanbakht, Lydia N. Drumright, Eric S. Daar, and Susan J. Little. Behaviors of recently HIV-infected men who have sex with men in the year postdiagnosis: Effects of drug use and partner types. *Journal of Acquired Immune Deficiency Syndromes*, 56:176 – 182, February 2011.
- [4] UNAIDS. UNAIDS World AIDS Day Report. Retrieved August 16, 2012 from the World Wide Web: http://www.unaids.org/en/media/unaids/contentassets/documents/unaidspublication/2011/JC2216_WorldAIDSday_report_2011_en.pdf, 2011.
- [5] Department of Health and Human Services, Centers for Disease Control and Prevention. Basic Statistics. Retrieved August 16, 2012 from the World Wide Web: <http://www.cdc.gov/hiv/topics/surveillance/print/basic.htm>, 2011.
- [6] Department of Health and Human Services, Centers for Disease Control and Prevention. HIV and AIDS among gay and bisexual men. Retrieved December 6, 2010 from the World Wide Web: <http://www.cdc.gov/nchhstp/newsroom/docs/FastFacts-MSM-FINAL508COMP.pdf>, September 2010.
- [7] County of San Diego Health and Human Services Agency HIV/AIDS Epidemiology Unit. AIDS in men who have sex with men (MSM). County of San Diego, 2004. Retrieved June 10, 2012 from the World Wide Web: http://www.sdcounty.ca.gov/hhsa/programs/phs/documents/AIDS_in_MSM.pdf, 2004.
- [8] J. L. Chen, D. Kodagoda, A. M. Lawrence, and P. R. Kerndt. Rapid public health interventions in response to an outbreak of syphilis in Los Angeles. *Sexually Transmitted Diseases*, 29(5):277–284, May 2002.

- [9] HIV Epidemiology Program, Los Angeles County Department of Health Services. HIV/AIDS Semi-Annual Surveillance Summary – Cases Reported as of June 2005. Los Angeles: Los Angeles County Department of Health Services, 2005.
- [10] Office of AIDS, Center for Infectious Diseases, California Department of Public Health. Integrated Epidemiologic Profile of HIV/AIDS in California, 2001-2005 with 2007 Update. Retrieved April 27, 2010 from the World Wide Web: www.cdph.ca.gov/programs/aids/Documents/EPIProfile.pdf, February 2010.
- [11] County of Los Angeles. Public Health: HIV Epi. An Epidemiologic Profile of HIV and AIDS in Los Angeles County 2009, 2009.
- [12] E. S. Daar, T. Moudgil, R. D. Meyer, and D. D. Ho. Transient high-levels of viremia in patients with primary human-immunodeficiency-virus type-1 infection. *New England Journal of Medicine*, 324(14):961–964, April 1991.
- [13] Sarah L. Rowland-Jones. AIDS pathogenesis: What have two decades of HIV research taught us? *Nature Reviews*, 3:343–348, April 2003.
- [14] V. Simon, D. D. Ho, and Q. A. Karim. HIV/AIDS epidemiology, pathogenesis, prevention, and treatment. *Lancet*, 368(9534):489–504, August 2006.
- [15] M. J. Wawer, R. H. Gray, N. K. Sewankambo, D. Serwadda, X. B. Li, O. Laeyendecker, N. Kiwanuka, G. Kigozi, M. Kiddugavu, T. Lutalo, F. Nalugoda, F. Wabwire-Mangen, M. P. Meehan, and T. C. Quinn. Rates of HIV-1 transmission per coital act, by stage of HIV-1 infection, in Rakai, Uganda. *Journal Of Infectious Diseases*, 191(9):1403–1409, May 2005.
- [16] T. D. Hollingsworth, R. M. Anderson, and C. Fraser. HIV-1 transmission, by stage of infection. *Journal Of Infectious Diseases*, 198(5):687–693, September 2008.
- [17] S. D. Pinkerton. Probability of HIV transmission during acute infection in Rakai, Uganda. *AIDS And Behavior*, 12(5):677–684, September 2008.
- [18] L. J. Abu-Raddad and I. M. Longini. No HIV stage is dominant in driving the HIV epidemic in sub-Saharan Africa. *AIDS*, 22(9):1055–1061, May 2008.
- [19] T. Cohen and E. L. Corbett. Test and treat in HIV: success could depend on rapid detection. *Lancet*, 378(9787):204–206, Jul 2011.
- [20] M. S. Cohen, G. M. Shaw, A. J. McMichael, and B. F. Haynes. Acute HIV-1 Infection. *New England Journal of Medicine*, 364(20):1943–1954, May 2011.

- [21] W. T. Steward, R. H. Remien, J. A. Higgins, R. Dubrow, S. D. Pinkerton, K. J. Sikkema, H. H. M. Truong, M. O. Johnson, J. Hirsch, R. A. Brooks, and S. F. Morin. Behavior change following diagnosis with acute/early HIV infection – a move to serosorting with other HIV-infected individuals. The NIMH multisite acute HIV infection study: III. *AIDS and Behavior*, 13(6):1054–1060, December 2009.
- [22] Martina Morris, Steven Goodreau, and James Moody. *Sexual Networks, Concurrency, and STD/HIV*. McGraw-Hill Medical, New York [u.a.], 4 edition, 2008. In *Sexually Transmitted Diseases*, Editor: King Holmes.
- [23] Douglas D Heckathorn. Respondent-driven sampling: A new approach to the study of hidden populations. *Social Problems*, 44(2):174–99, 1997.
- [24] Mark S. Handcock, David R. Hunter, Carter T. Butts, Steven M. Goodreau, and Martina Morris. *statnet: Software tools for the Statistical Modeling of Network Data*. Seattle, WA, 2003. Version 2.0.
- [25] Martina Morris, Ann E. Kurth, Deven T. Hamilton, James Moody, and Steve Wakefield. Concurrent partnerships and HIV prevalence disparities by race: Linking science and public health practice. *American Journal of Public Health*, 99(6):1023–1031, Jun 2009.
- [26] SM Goodreau, S Cassels, D Kasprzyk, DE Montao, A Greek, and M Morris. Concurrent partnerships, acute infection and HIV epidemic dynamics among young adults in Zimbabwe. *AIDS and Behavior*, 16(2):312–322, 2012.
- [27] Steven M. Goodreau, Nicole B Carnegie, Eric Vittinghoff, Javier R Lama, Jorge Sanchez, Beatriz Grinsztejn, Beryl A Koblin, Kenneth H Mayer, and Susan P Buchbinder. What drives the US and Peruvian HIV epidemics in men who have sex with men (MSM)? Under Review.
- [28] Thomas C. Quinn. Population migration and the spread of types 1 and 2 human immunodeficiency viruses. *Proceedings of the National Academy of Sciences*, 91:2407 – 2414, March 1994. Colloquium Paper.
- [29] K Zuma, E Gouws, B Williams, and M Lurie. Risk factors for HIV infection among women in carletonville, South Africa: migration, demography and sexually transmitted diseases. *International Journal of STD and AIDS*, 14:814–817, 2003.
- [30] Martina Morris and Mirjam Kretzschmar. Concurrent partnerships and transmission dynamics in networks. *Social Networks*, 17:299–318, 1995.
- [31] Mirjam Kretzschmar and Martina Morris. Measures of concurrency in networks and the spread of infectious disease. *Mathematical Biosciences*, 133:165 – 195, 1996.

- [32] Martina Morris and Mirjam Kretzschmar. Concurrent partnerships in the spread of HIV. *AIDS*, 133:165 – 195, 1997.
- [33] Steven M Goodreau. Accounting for epidemics: Mathematical modeling and anthropology. In Ann Herring and Alan C Swedlund, editors, *Plagues and Epidemics*, chapter 12, pages 213 – 230. Berg, 2010.
- [34] Larry Sawers and Eileen Stillwaggon. Concurrent sexual partnerships do not explain the HIV epidemics in Africa: a systematic review of the evidence. *Journal of the International AIDS Society*, 13(34), 2010.
- [35] W. O. Kermack and A. G. McKendrick. Contributions to the mathematical theory of epidemics: IV. Analysis of experimental epidemics of the virus disease mouse ectromelia. *J Hyg (Lond)*, 37(2):172–187, Apr 1937.
- [36] N. Becker. An epidemic chain model. *Biometrics*, 36(2):249–254, Jun 1980.
- [37] R. M. May and R. M. Anderson. Transmission dynamics of HIV-infection. *Nature*, 326(6109):137–142, March 1987.
- [38] R. M. May and R. M. Anderson. The transmission dynamics of human immunodeficiency virus (HIV). *Philosophical Transactions Of The Royal Society Of London Series B-Biological Sciences*, 321(1207):565–607, October 1988.
- [39] R. M. May, R. M. Anderson, and A. R. Mclean. Possible demographic consequences of HIV/AIDS epidemics. I. Assuming HIV infection always leads to AIDS. *Mathematical Biosciences*, 90(1-2):475–505, July 1988.
- [40] J. W. Eaton, L. F. Johnson, J. A. Salomon, T. Barnighausen, E. Bendavid, A. Bershteyn, D. E. Bloom, V. Cambiano, C. Fraser, J. A. Hontelez, S. Humair, D. J. Klein, E. F. Long, A. N. Phillips, C. Pretorius, J. Stover, E. A. Wenger, B. G. Williams, and T. B. Hallett. HIV Treatment as Prevention: Systematic Comparison of Mathematical Models of the Potential Impact of Antiretroviral Therapy on HIV Incidence in South Africa. *PLoS Medicine*, 9(7):e1001245, Jul 2012.
- [41] A. L. Bárabasi, R. Albert, and H. Jeong. Mean-field theory for scale-free random networks. *Physica A*, 272(1-2):173–187, October 1999.
- [42] F. Liljeros, C. R. Edling, L. A. N. Amaral, H. E. Stanley, and Y. Aberg. The web of human sexual contacts. *Nature*, 411(6840):907–908, June 2001.
- [43] ME. Newman. Spread of epidemic disease on networks. *Physical Review E*, 66:016128–1 – 016128–11, 2002.

- [44] James Holland Jones and Mark S. Handcock. An assessment of preferential attachment as a mechanism for human sexual network formation. *Proceedings of the Royal Society B: Biological Sciences*, 270(1520):1123–1128, June 2003.
- [45] James Holland Jones and Mark S. Handcock. Social networks (communication arising): Sexual contacts and epidemic thresholds. *Nature*, 423:605–606, 2003.
- [46] H.A. Simon. On a class of skew distribution functions. *Biometrika*, 42(3-4):425–440, 1955.
- [47] D Price. A general theory of bibliometric and other cumulative advantage processes. *Journal of the American Society for Information Science*, 27(5):292306, 1976.
- [48] Steven M. Goodreau. A decade of modelling research yields considerable evidence for the importance of concurrency: A response to Sawers and Stillwaggon. *Journal of the International Aids Society*, 14:12, March 2011.
- [49] S. Wasserman and P. Pattison. Logit models and logistic regressions for social networks. 1. an introduction to Markov graphs and p^* . *Psychometrika*, 61(3):401–425, September 1996.
- [50] David R. Hunter, Steven M. Goodreau, and Mark S. Handcock. Goodness of fit of social network models. *Journal of the American Statistical Association*, 103(481):248–258, March 2008.
- [51] Pavel N. Krivitsky, Mark S. Handcock, and Martina Morris. Adjusting for network size and composition effects in exponential-family random graph models. *Statistical Methodology*, 8(4):319339, 2011.
- [52] Mark N Lurie. The epidemiology of migration and HIV/AIDS in South Africa. *Journal of Ethnic and Migration Studies*, 32(4):649 – 666, May 2006.
- [53] K. D. Deane, J. O. Parkhurst, and D. Johnston. Linking migration, mobility and HIV. *Tropical Medicine & International Health*, 15(12):1458–1463, December 2010.
- [54] M. Lurie. Migration and AIDS in Southern Africa: a review. *South African Journal of Science*, 96:343–347, June 2000.
- [55] Mark Lurie, Abigail Harrison, David Wilkinson, and Salim Abdool Karim. Circular migration and sexual networking in KwaZulu/Natal: implications for the spread of HIV and other sexually transmitted diseases. *Health Transition Review*, Supplement 3 to Volume 7:17–27, 1997.

- [56] Mark Lurie, Brian G. Williams, Khangleani Zuma, David Mkaya-Mwamburi, Geoff P. Garnett, Michael D. Swat, Joel Gittelsohn, and Salim Abdool Karim. Who infects whom? HIV-1 concordance and discordance among migrant and non-migrant couples in South Africa. *AIDS*, 17:2245–2252, 2003.
- [57] Mark Lurie, Brian G. Williams, Khangleani Zuma, David Mkaya-Mwamburi, Geoff P. Garnett, Michael D. Swat, Joel Gittelsohn, and Salim Abdool Karim. The impact of migration on HIV-1 transmission in South Africa. *Sexually Transmitted Diseases*, 30(2):149–156, February 2003.
- [58] Martin Brouckhoff and Ann E. Biddlecom. Migration, behavior and the risk of HIV in Kenya. *International Migration Review*, 33(4):833–856, 1999.
- [59] M. P. Coffee, G. P. Garnett, M. Mlilo, H. A. C. M. Voeten, S. Chandiwana, and S. Gregson. Patterns of movement and risk of HIV infection in rural Zimbabwe. *Journal of Infectious Diseases*, 191:S159–S167, February 2005.
- [60] C. Mundandi, D. Vissers, H. Voeten, D. Habbema, and S. Gregson. No difference in HIV incidence and sexual behaviour between out-migrants and residents in rural Manicaland, Zimbabwe. *Tropical Medicine & International Health*, 11(5):705–711, May 2006.
- [61] E. Coast. Local understandings of, and responses to, HIV: Rural-urban migrants in Tanzania. *Social Science & Medicine*, 63(4):1000–1010, August 2006.
- [62] Jeffrey W. Eaton, Timothy B. Hallett, and Geoffrey P. Garnett. Concurrent Sexual Partnerships and Primary HIV Infection: A Critical Interaction. *AIDS and Behavior*, 15(4):687–692, MAY 2011.
- [63] S. W. Sorensen, S. L. Sansom, J. T. Brooks, G. Marks, E. M. Begier, K. Buchacz, E. A. Dinunno, J. H. Mermin, and P. H. Kilmarx. A mathematical model of comprehensive test-and-treat services and HIV incidence among men who have sex with men in the United States. *PLoS ONE*, 7(2):e29098, 2012.
- [64] G. B. Gomez, A. Borquez, C. F. Caceres, E. R. Segura, R. M. Grant, G. P. Garnett, and T. B. Hallett. The Potential Impact of Pre-Exposure Prophylaxis for HIV Prevention among Men Who Have Sex with Men and Transwomen in Lima, Peru: A Mathematical Modelling Study. *PLoS Med.*, 9(10):e1001323, Oct 2012.
- [65] B. G. Wagner and S. Blower. Universal Access to HIV Treatment versus Universal 'Test and Treat': Transmission, Drug Resistance & Treatment Costs. *PLoS ONE*, 7(9):e41212, 2012.
- [66] J. H. Kim and J. S. Koopman. HIV transmissions by stage in dynamic sexual partnerships. *J. Theor. Biol.*, 298:147–153, Apr 2012.

- [67] N. J. Nagelkerke, J. A. Hontelez, and S. J. de Vlas. The potential impact of an HIV vaccine with limited protection on HIV incidence in Thailand: a modeling study. *Vaccine*, 29(36):6079–6085, Aug 2011.
- [68] V. Cambiano and A. N. Phillips. Modelling the impact of treatment with individual antiretrovirals. *Curr Opin HIV AIDS*, 6(2):124–130, Mar 2011.
- [69] S. S. Alistar, D. K. Owens, and M. L. Brandeau. Effectiveness and cost effectiveness of expanding harm reduction and antiretroviral therapy in a mixed HIV epidemic: a modeling analysis for Ukraine. *PLoS Med.*, 8(3):e1000423, Mar 2011.
- [70] M. C. Boily, D. Dimitrov, S. S. Abdool Karim, and B. Masse. The future role of rectal and vaginal microbicides to prevent HIV infection in heterosexual populations: implications for product development and prevention. *Sex Transm Infect*, 87(7):646–653, Dec 2011.
- [71] W. J. Reidy and S. M. Goodreau. The role of commercial sex venues in the HIV epidemic among men who have sex with men. *Epidemiology*, 21(3):349–359, May 2010.
- [72] G. P. Garnett, P. J. White, and H. Ward. Fewer partners or more condoms? Modelling the effectiveness of STI prevention interventions. *Sex Transm Infect*, 84 Suppl 2:4–11, Oct 2008.
- [73] S. M. Goodreau and M. R. Golden. Biological and demographic causes of high HIV and sexually transmitted disease prevalence in men who have sex with men. *Sex Transm Infect*, 83(6):458–462, Oct 2007.
- [74] Herbert W Hethcote and James W van Ark. *Modeling HIV Transmission and AIDS in the United States*. Springer-Verlag Telos, 1992.
- [75] Joint United Nations Programme on HIV/AIDS (UNAIDS). Global report: UNAIDS report on the global AIDS epidemic. Retrieved January 29, 2011 from the World Wide Web: http://www.unaids.org/globalreport/Global_report.htm, 2010.
- [76] Joint United Nations Programme on HIV/AIDS (UNAIDS). South Africa. Retrieved March 28, 2012 from the World Wide Web: <http://www.unaids.org/en/regionscountries/countries/southafrica/>, 2009.
- [77] J. Hargrove. Migration, mines and mores: The HIV epidemic in Southern Africa. *South African Journal of Science*, 104(1-2):53–61, January 2008.
- [78] Charles W. Hunt. Migrant labor and sexually transmitted diseases. *Journal of Health and Social Behavior*, 30(4):353 – 373, December 1989. Theme: Sociological Studies of Third World Health and Health Care.

- [79] S. J. Clark, M. A. Collinson, K. Kahn, K. Drullinger, and S. M. Tollman. Returning home to die: Circular labour migration and mortality in South Africa. *Scandinavian Journal of Public Health*, 35:35–44, August 2007.
- [80] Richard G White. Commentary: What can we make of an association between human immunodeficiency virus prevalence and population mobility? *International Journal of Epidemiology*, 32:753 – 754, 2003.
- [81] M. A. Collinson, B. Wolff, S. M. Tollman, and K. Kahn. Trends in internal labour migration from rural Limpopo Province, male risk behaviour, and implications for the spread of HIV/AIDS in rural South Africa. *Journal of Ethnic and Migration Studies*, 32(4):633–648, May 2006.
- [82] AVERT. The South African Department of Health Study, 2009. Retrieved January 29, 2011 from the World Wide Web: <http://www.avert.org/safricastats.htm>, 2009.
- [83] Megan Coffee, Mark N Lurie, and Geoff P Garnett. Modelling the impact of migration on the HIV epidemic in South Africa. *AIDS*, 21:343 – 350, 2007.
- [84] K Zuma, M.N. Lurie, B.G. Williams, D. Mkaya-Mwamburi, G.P. Garnett, and A.W. Sturm. Risk factors of sexually transmitted infections among migrant and non-migrant sexual partnerships among rural South Africa. *Epidemiology and Infection*, 133:421–428, 2005.
- [85] L. Sawers, A. G. Isaac, and E. Stillwaggon. HIV and concurrent sexual partnerships: modelling the role of coital dilution. *J Int AIDS Soc*, 14:44, 2011.
- [86] R. Woodrow Setzer. *odesolve: Solvers for Ordinary Differential Equations*, 2012. Version 0.9-9.
- [87] Thomas Petzoldt. R as a simulation platform in ecological modeling. *R News*, 3(3):8 – 16, 2003.
- [88] Matthew James Keeling and Pejman Rohani. *Modeling infectious diseases in humans and animals*. Princeton University Press, Princeton, 2008.
- [89] P. van den Driessche and J. Watmough. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Mathematical Biosciences*, 180:29–48, November 2002.
- [90] Martina Morris, Mark S. Handcock, and David R. Hunter. Specification of exponential-family random graph models: Terms and computational aspects. *Journal of Statistical Software*, 24(4):1–24, 12 2007.

- [91] Department of Health and Human Services, Centers for Disease Control and Prevention. Vital signs: HIV testing and diagnosis among adults - United States, 2001-2009. *Morbidity and Mortality Weekly Report*, 59(47):1550–1555, December 3 2010.
- [92] Department of Health and Human Services, Centers for Disease Control and Prevention. Prevalence and awareness of HIV infection among men who have sex with men - 21 cities, United States, 2008. *Morbidity and Mortality Weekly Report*, 59(37):1201–1207, September 24 2010.
- [93] P. S. Sullivan, A. Carballo-Dieguez, T. Coates, S. M. Goodreau, I. McGowan, E. J. Sanders, A. Smith, P. Goswami, and J. Sanchez. Successes and challenges of HIV prevention in men who have sex with men. *Lancet*, Jul 2012.
- [94] N. M. van Kesteren, H. J. Hospers, and G. Kok. Sexual risk behavior among HIV-positive men who have sex with men: A literature review. *Patient Education and Counseling*, 65(1):5–20, Jan 2007.
- [95] N. M. van Kesteren, H. J. Hospers, P. van Empelen, G. van Breukelen, and G. Kok. Sexual decision-making in HIV-positive men who have sex with men: How moral concerns and sexual motives guide intended condom use with steady and casual sex partners. *Archives of Sexual Behavior*, 36(3):437–449, Jun 2007.
- [96] J. Fox, P. J. White, N. Macdonald, J. Weber, M. McClure, S. Fidler, and H. Ward. Reductions in HIV transmission risk behaviour following diagnosis of primary HIV infection: a cohort of high-risk men who have sex with men. *HIV Med.*, 10(7):432–438, Aug 2009.
- [97] S. P. Kurtz, M. E. Buttram, H. L. Surratt, and R. D. Stall. Resilience, syndemic factors, and serosorting behaviors among HIV-positive and HIV-negative substance-using MSM. *AIDS Educ Prev*, 24(3):193–205, Jun 2012.
- [98] E. Vittinghoff, J. Douglas, F. Judson, D. McKirnan, K. MacQueen, and S. P. Buchbinder. Per-contact risk of human immunodeficiency virus transmission between male sexual partners. *American Journal Of Epidemiology*, 150(3):306–311, August 1999.
- [99] F. Molitor, S. R. Truax, J. D. Ruiz, and R. K. Sun. Association of methamphetamine use during sex with risky sexual behaviors and HIV infection among non-injection drug users. *West. J. Med.*, 168(2):93–97, Feb 1998.
- [100] R. Rajasingham, M. J. Mimiaga, J. M. White, M. M. Pinkston, R. P. Baden, and J. A. Mitty. A systematic review of behavioral and treatment outcome studies among HIV-infected men who have sex with men who abuse crystal methamphetamine. *AIDS Patient Care STDS*, 26(1):36–52, Jan 2012.

- [101] S. J. Semple, T. L. Patterson, and I. Grant. Motivations associated with methamphetamine use among HIV plus men who have sex with men. *Journal of Substance Abuse Treatment*, 22(3):149–156, April 2002.
- [102] S. Shoptaw and C. J. Reback. Associations between methamphetamine use and HIV among men who have sex with men: A model for guiding public policy. *Journal Of Urban Health-Bulletin Of The New York Academy Of Medicine*, 83(6):1151–1157, November 2006.
- [103] P. M. Gorbach and K. K. Holmes. Transmission of STIs/HIV at the partnership level: beyond individual-level analyses. *J Urban Health*, 80(4 Suppl 3):15–25, Dec 2003.
- [104] H. F. Raymond and W. McFarland. Racial mixing and HIV risk among men who have sex with men. *AIDS and Behavior*, 13(4):630–637, August 2009.
- [105] A. R. Wohl, D. M. Frye, and D. F. Johnson. Demographic characteristics and sexual behaviors associated with methamphetamine use among msm and non-msm diagnosed with AIDS in Los Angeles county. *AIDS and Behavior*, 12(5):705–712, September 2008.
- [106] D. J. Helms, H. S. Weinstock, K. C. Mahle, K. T. Bernstein, B. W. Furness, C. K. Kent, C. A. Rietmeijer, A. M. Shahkolahi, J. P. Hughes, and M. R. Golden. HIV testing frequency among men who have sex with men attending sexually transmitted disease clinics: Implications for HIV prevention and surveillance. *Journal Of Acquired Immune Deficiency Syndromes*, 50(3):320–326, 2009.
- [107] Centers for Disease Control and Prevention. HIV among gay, bisexual and other men who have sex with men (msm). Retrieved December 6, 2010 from the World Wide Web: <http://www.cdc.gov/hiv/topics/msm/index.htm>, September 2010.
- [108] M J Mimiaga, S L Reisner, S Bland, M Skeer, K Cranston, D Isenberg, B A Vega, and K H Mayer. Health system and personal barriers resulting in decreased utilization of HIV and STD testing services among at-risk black men who have sex with men in Massachusetts. *AIDS Patient Care STDS*, 23(10):825–835, Oct 2009.
- [109] B. M. Branson and J. D. Stekler. Detection of acute HIV infection: We can’t close the window. *Journal of Infectious Diseases*, 205(4):521–524, Feb 2012.
- [110] B. M. Branson. State of the art for diagnosis of HIV infection. *Clinical Infectious Diseases*, 45 Suppl 4:S221–225, Dec 2007.
- [111] S. H. Eshleman, L. Khaki, O. Laeyendecker, E. Piwowar-Manning, L. Johnson-Lewis, M. Husnik, B. Koblin, T. Coates, M. Chesney, A. Vallari, S. G. Devare, and J. Hackett. Detection of individuals with acute HIV-1 infection using the ARCHITECT HIV

- Ag/Ab Combo assay. *Journal of Acquired Immune Deficiency Syndromes*, 52(1):121–124, Sep 2009.
- [112] N. E. Rosenberg, G. Kamanga, S. Phiri, D. Nsona, A. Pettifor, S. E. Rutstein, D. Kamwendo, I. F. Hoffman, M. Keating, L. B. Brown, B. Ndalama, S. A. Fiscus, S. Congdon, M. S. Cohen, and W. C. Miller. Detection of acute HIV infection: A field evaluation of the Determine HIV-1/2 Ag/Ab combo test. *Journal of Infectious Diseases*, 205(4):528–534, Feb 2012.
- [113] C. D. Pilcher, K. A. Christopoulos, and M. Golden. Public health rationale for rapid nucleic acid or p24 antigen tests for HIV. *Journal of Infectious Diseases*, 201:S7–S15, April 2010.
- [114] P. R. Kerndt, R. Dubrow, G. Aynalem, K. H. Mayer, C. Beckwith, R. H. Remien, H. H. M. Truong, A. Uniyal, M. Chien, R. A. Brooks, O. R. Vigil, W. T. Steward, M. Merson, M. J. Rotheram-Borus, and S. F. Morin. Strategies used in the detection of acute/early HIV infections. The NIMH multisite acute HIV infection study: I. *AIDS and Behavior*, 13(6):1037–1045, December 2009.
- [115] J. A. Kelly, S. F. Morin, R. H. Remien, W. T. Steward, J. A. Higgins, D. W. Seal, R. Dubrow, J. H. Atkinson, P. R. Kerndt, S. D. Pinkerton, K. Mayer, and K. J. Sikkema. Lessons learned about behavioral science and acute/early HIV infection. The NIMH multisite acute HIV infection study: V. *AIDS and Behavior*, 13(6):1068–1074, December 2009.
- [116] J. D. O’neal, M. R. Golden, B. M. Branson, and J. D. Stekler. HIV Nucleic Acid Amplification Testing Versus Rapid Testing: It Is Worth the Wait. Testing Preferences of Men Who Have Sex With Men. *Journal of Acquired Immune Deficiency Syndromes*, 60(4):e119–122, Aug 2012.
- [117] J. Stekler, P. D. Swenson, R. W. Wood, H. H. Handsfield, and M. R. Golden. Targeted screening for primary HIV infection through pooled HIV-RNA testing in men who have sex with men. *AIDS*, 19(12):1323–1325, Aug 2005.
- [118] Department of Health and Human Services, Centers for Disease Control and Prevention. HIV among gay and bisexual men. Retrieved August 20, 2012 from the World Wide Web: <http://www.cdc.gov/hiv/topics/msm/>, 2012.
- [119] It feels better when you know. Test often for HIV. Retrieved August 20, 2012 from the World Wide Web: <http://www.findyourfrequency.com/>, 2012.
- [120] S. Cassels, T. W. Menza, S. A. Goodreau, and M. R. Golden. HIV serosorting as a harm reduction strategy: evidence from seattle, washington. *AIDS*, 23(18):2497–2506, November 2009.

- [121] County of Los Angeles Department of Health Services Office of AIDS Programs and Policy. County of los angeles HIV prevention plan 2004 - 2008. Retrieved From the World Wide Web.
- [122] Pavel Krivitsky. *Statistical Models for Social Network Data and Processes*. PhD thesis, University of Washington, 2009.
- [123] A M Leynaert B FAU Downs, I Downs AM FAU de Vincenzi, and de Vincenzi I. Heterosexual transmission of human immunodeficiency virus: variability of infectivity throughout the course of infection. european study group on heterosexual transmission of HIV. *American Journal of Epidemiology*, 148(1):88–96, 1998.
- [124] S. J. Little, A. R. McLean, C. A. Spina, D. D. Richman, and D. V. Havlir. Viral dynamics of acute HIV-1 infection. *Journal Of Experimental Medicine*, 190(6):841–850, September 1999.
- [125] S. P. Buchbinder, M. H. Katz, N. A. Hessol, P. M. O’Malley, and S. D. Holmberg. Long-term HIV-1 infection without immunologic progression. *AIDS*, 8(8):1123–1128, Aug 1994.
- [126] Fujie Xu, Lauri E. Markowitz, Maya R. Sternberg, and Sevgi O. Aral. Prevalence of circumcision and herpes simplex virus type 2 infection in men in the United States: The national health and nutrition examination survey (NHANES), 1999-2004. *Sexually Transmitted Diseases*, 34(7):479–484, Jul 2007.
- [127] Department of Health and Human Services, Centers for Disease Control and Prevention. HIV prevalence, unrecognized infection, and hiv testing among men who have sex with men - five U.S. cities, June 24-April 2005. *Morbidity and Mortality Weekly Report*, 54(24):597–600, June 24 2005.
- [128] E Arias. United states life tables, 2004. Atlanta: US Centers for Disease Control and Prevention, Division of Vital Statistics, 2007.