AN ABSTRACT OF THE DISSERTATION OF

Gordon Amed Akudibillah for the degree of Doctor of Philosophy in Environmental Science presented on March 17, 2016.

Title: Optimizing HIV Treatment in Resource-Limited Settings

Abstract approved: _____

Jan Medlock

Apart from the traditional role of preventing progression from HIV to AIDS, antiretroviral drug therapy (ART) has an additional benefit of substantially reducing infectiousness, making them potentially an important strategy in the fight against HIV. Recent advances in drug therapy have also seen the use of antiretroviral drugs as a prophylaxis, administered either as post-exposure prophylaxis (PEP) after high-risk exposure or as pre-exposure prophylaxis (PrEP) in those with ongoing HIV exposure. In this dissertation I developed two models for HIV transmission and parameterized them with data from South Africa to study governmental-level intervention programs in which antiretroviral drugs are given as treatment and prophylaxis.

The first model is based on the dynamics of HIV in heterosexual population in Sub-Saharan Africa. The model classifies the male and female adult populations by HIV risk into three categories (low, medium and high) according to their sexual preferences. I used a non-linear optimization method to determine the optimal population-level allocation of ART and PrEP allocations required to minimize four objectives: new infections, infection-years, deaths and cost. I considered several strategies for allocating ART and PrEP. I found that

generally for low treatment availability, prevention through PrEP to the general population or PrEP and ART to high-risk females is key to optimize all objectives, while for higher drug availability, an all-ART treatment is optimal. At South Africa's current level of treatment availability, using prevention is most effective at reducing new infections, infection-years, and cost, while using the treatment as ART to the general population best reduces deaths. At treatment levels that meet the UNAIDS's ambitious new 90–90–90 target in South Africa, using all or almost all treatment as ART to the general population best reduces all four objectives considered.

The second model is based on the WHO's five-stage classification of HIV/AIDS disease progression. This models stratified the population by disease status, whether diagnosed and whether on treatment. I used optimal control methods to determine the best time-dependent treatment allocation required to minimize new infections, infection-years, deaths and cost. My results indicated that the treatment strategy to minimize infection-years and new infections is to place emphasis on early treatment (i.e. treatment in Stage II & III) while to minimize cost and death, the emphasis should be on late treatment (i.e. Stage III & IV). Applying the optimal treatment strategy also leads to a substantial reduction in disease incidence and prevalence.

The results of this study will hopefully provides some guidance for policymakers in determining how to allocate antiretroviral drugs in order to maximize the benefit of treatment.

Optimizing HIV Treatment in Resource-Limited Settings

by

Gordon Amed Akudibillah

A Dissertation

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

Presented March 17, 2016
Commencement June 2016

Doctor of Philosophy dissertation of Gordon Amed Akudibillah presented on March 17, 2016

APPROVED:

Major Professor, representing Environmental Science

Director of the Environmental Sciences Program

Dean of the Graduate School

I understand that my dissertation will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my dissertation to any reader upon request.

Gordon Amed Akudibillah, Author

ACKNOWLEDGEMENTS

Academic

I cannot express enough thanks to my committee for their continued support and encouragement: Special thanks should be given to Dr. Jan Medlock, my research project supervisor for his professional guidance and valuable support. To my committee members, Dr. John Bolte, Dr. Charlotte Wickham, Dr. Vrushali Bokil and Dr. Susan Carozza, I offer my sincere appreciation for the learning opportunities each of you provided. I also want to extend my sincere gratitude to Dr. Suzanne Lenhart of National Institute for Mathematical and Biological Synthesis (NIMBioS) for her guidance on my optimal control method.

Personal

My completion of this project could not have been accomplished without the support of Dr. Abishek Pandey: your help was invaluable. To my three sisters, Charity Adamah, Linda Marfoh, and Barbara Akudibillah, thanks for your support and encouragement. To my mother in-law, Ms. Margrett Taylor, thanks you for your laughter and comforting words of wisdom. Special thanks to my parents, Hon. Joseph Akudibillah and Mrs. Margaret Minyila. The two of you instilled the importance of both education and perserverance.

Finally, to my beloved, Kenyetta, I owe you my deepest gratitude. Your encouragement when the times got rough are much appreciated and duly noted. Thank you for allowing me time away from you to research and write. Your support has been selfless and enduring. I am honored to have you as my wife!

Contents

TABLE OF CONTENTS (Continued)

List of Figures

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

List of Tables

Optimizing HIV Treatment in Resource-Limited Settings

1. GENERAL INTRODUCTION

The Human Immunodeficiency Virus (HIV), the virus that causes Acquired Immunodeficiency Syndrome (AIDS) was first identified in 1981 [77]. Since its discovery, the reach of HIV/AIDS has grown steadily, engulfing every continent. Although resource-constrained countries in Sub-Saharan African are the hardest hit, HIV/AIDS continues to be one of the most serious public health problems worldwide.

Upon introduction into the body, the HIV virus attacks the T helper cells (CD4+ T cells), which are crucial to the immune response for fighting off invading pathogens. The ultimate destruction of T helper cells leads to the body becoming immunodeficient i.e. open to opportunistic infections, a condition that characterizes Acquired Immunodeficiency Syndrome (AIDS).

Once a person becomes infected, the World Health Organization (WHO) has defined five clinical stages of infection progression (Figure 1.1) [78]. The first stage, the Acute Stage, lasts a few week following the initial introduction of the virus into the body. A spike in virus load occurs in the Acute Stage, but this stage is largely asymptomatic with no significant immunosuppression (CD4 count $> 500 \frac{\text{cells}}{\text{mm}^3}$) [73, 78]. Seroconversion occurs with the first three months, and after seroconversion the infected person the goes through the four other stages. Stage I lasts for an average of ten years and is also asymptomatic with no significant immunosuppression (CD4 count $> 500 \frac{\text{cells}}{\text{mm}^3}$) and a low viral load. Stage II is

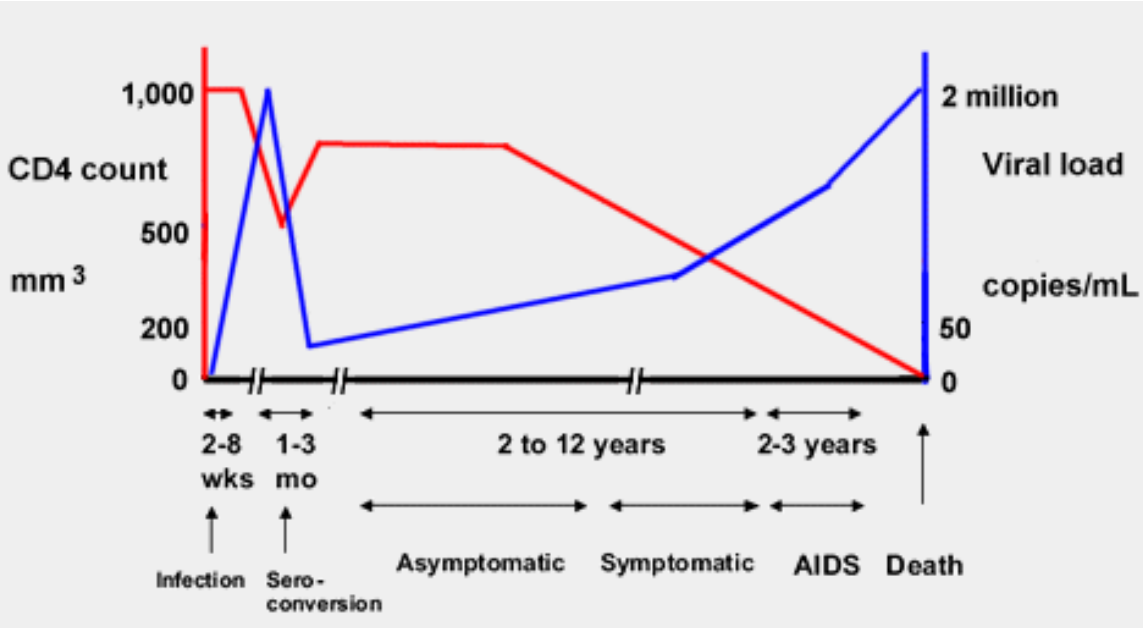

Figure 1.1: Stages of HIV/AIDS.

symptomatic: the infected person exhibits moderate weight loss, amongst other symptoms, and mild immunosuppression (CD4 count 350–499$\frac{\text{cells}}{\text{mm}^3}$). Stage III is also symptomatic and is characterized by advanced immunosuppression (CD4 count 200–349$\frac{\text{cells}}{\text{mm}^3}$). Stage IV is the AIDS phase, where the infected person exhibits HIV wasting syndrome, amongst other symptoms, and severe immunosuppression (CD4 count $< 200\frac{\text{cells}}{\text{mm}^3}$). This classification based on clinical symptoms is advantageous because it allows for areas with limited laboratory facilities to estimate HIV progression and aid in the management of HIV/AIDS patients, e.g. when to start treatment.

The Acute Stage is characterized by high rates of transmission because of the high viral loads observed, albeit for a short period. The infected person then goes through a extended period (Stages I, II & III) where the viral load is lower and thus transmission is lower. Without treatment, this person moves to Stage IV, which is characterized by high

transmission rates due to high viral loads [5, 44, 46]

HIV is transmitted through the exchange of infected bodily fluids such as semen, vaginal secretions, blood and breast milk [25, 56]. An infected person can infect a susceptible person via unprotected vaginal or anal intercourse. Sharing or accidentally being stuck with contaminated needles, syringes or other injection equipment can also lead to an infection [56]. Vertical transmission for mother to child is also possible during birth or through breastfeeding [25]. Transmission has also been documented through transfusion of infected blood [56]

Although HIV has no cure, several interventions can be used to help control its spread and mitigate the disease it causes. UNAIDS defines the core responses to control HIV as prioritizing interventions for key populations, communication for behavior-change programs, condom promotion, prevention of mother-to-child transmission, male circumcision, and antiretroviral drugs for both infected and uninfected people [62].

Globally, an estimated 38 million people are living with HIV/AIDS in 2013, with 2.1 million (1.9–2.4 million) new HIV infections and 1.5 million (1.4–1.7 million) AIDS-related deaths recorded in 2013 [70]. Sub-Saharan Africa is the worst-affected region with 71% of the world's people living with HIV, 70% of new HIV infections and 73% of the global AIDS-related deaths in 2013.

Every geographical area, class, and cultural group has been affected by the HIV/AIDS pandemic (Figure 1.2). Whilst HIV/AIDS in developed countries remains most prevalent amongst homosexuals, hemophiliacs, and injection-drug users, it is most prevalent amongst heterosexuals in under-developed countries [16, 41].

Asymmetries in the prevalence between developed countries and under-developed countries, especially those in Sub-Saharan Africa, are fundamental to making HIV a huge public-health burden for the under-developed countries. Additionally, the economic

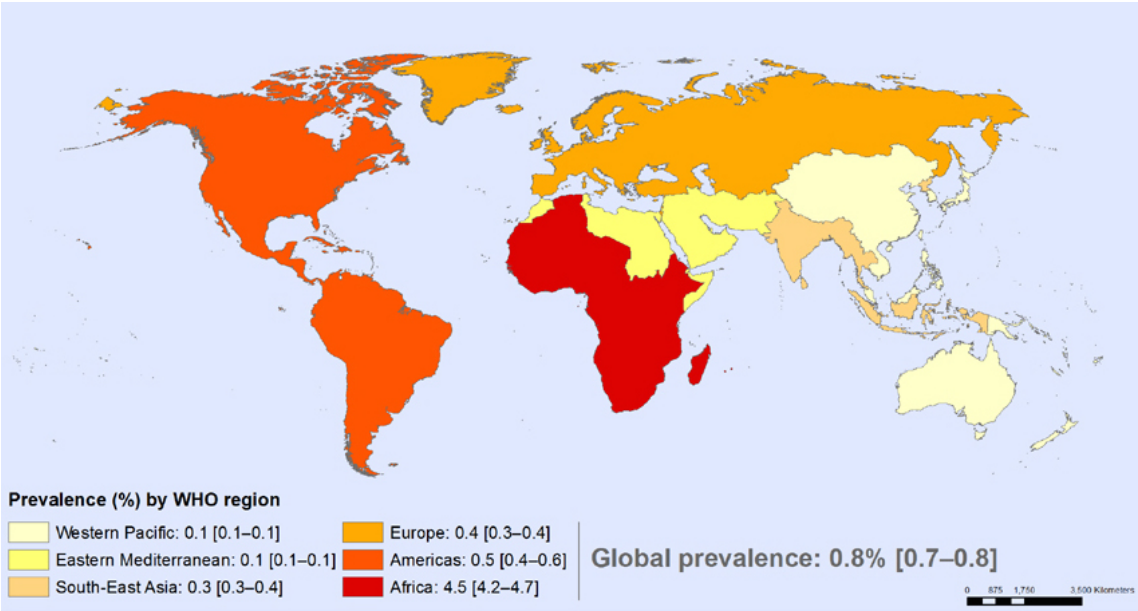

Figure 1.2: Global Distribution of HIV Adult Prevalence [81]

.

consequences of the HIV epidemic in under-developed countries are exacerbated by the disproportionate impact on young adults. The HIV epidemic is resulting in the death of the most productive members of society and the orphaning of millions of children [84].

In the 1990s, highly active antiretroviral therapy (HAART) was used mainly to treat opportunistic infections. Due partly to toxicity HAART, the WHO recommendation then was give treat only infected people with CD4 count $< 200 \frac{\text{cells}}{\text{mm}^3}$ [78]. Over the years, the benefits of antiretroviral therapy became apparent and the toxicity has decreased in new generations of drugs, so treatment recommendations have changed. The first randomized controlled trial ever to test the efficacy of early use of ART, HIV Prevention Trials Network 052 (HPTN 052), conclusively revealed that when administered early, ART leads to a substantial decrease in sexual transmission of HIV among HIV serodiscordant couples [18]. The WHO issued "Consolidated Guidelines on the use of Antiretroviral Drugs for Treating and Preventing HIV Infection" recommending treatment to infected people with CD4 count of $> 350 \frac{\text{cells}}{\text{mm}^3}$ in 2010 [79] and $> 500 \frac{\text{cells}}{\text{mm}^3}$ in 2013 [53]. In 2015, the WHO finally recommended that all infected people infected people should be given treatment regardless of CD4 count [55].

Early treatment and strict adherence to treatment by infected people is the best option for sustained viral suppression leading to a reduction of morbidity and mortality [47, 68]. Apart from improving the health and prolonging the life span of infected people, viral suppression also substantially reduces their infectiousness [18], thus reducing chances of onward transmission [60]. In addition, uninfected people taking antiretrovirals as prophylaxis also have a substantially reduced chances of acquiring HIV [1, 11]. Currently, there are three types of treatment: antiretroviral therapy (ART) given to infected people to prevent disease progression, pre-exposure prophylaxis (PrEP) administered to uninfected people with ongoing HIV risk to reduce their chances of acquiring the virus and post-exposure

prophylaxis (PEP) given to uninfected people immediately after exposure to HIV, usually perinatal or occupational exposures [19, 45, 63, 66, 74, 85].

Due to the benefits of antiretroviral therapy both as treatment and as prophylaxis at blocking transmission, treatment as prevention is increasing becoming an HIV/AIDS prevention strategy [17, 48, 55, 69]. The WHO introduced an ambitious treatment target called 90–90–90 aimed at ending the AIDS epidemic as a public health threat by 2030 [72, 84]. The 90–90–90 strategy is for 90% of all people living with HIV know their HIV status, 90% of all diagnosed people receive sustained antiretroviral therapy and 90% of all people receiving antiretroviral therapy have viral suppression, all by 2020. Countries worldwide are scaling up ART coverage to achieve the 90–90–90 target in order to halt the HIV/AIDS pandemic.

Due to the devastation caused by HIV/AIDS on countries, there is not only a moral but also an economic impetus to halt the epidemic. However, implementing control measures, including the scaling-up ART coverage requires a lot of resources. Decreases in international funding to fight HIV [23] coupled with severely limited health budgets, resource-constrained countries must find smarter ways to use their limited resources to get more "bang for the buck."

For this dissertation, I used techniques from optimization to model and analyze the allocation of resources to combat HIV. Optimization is a useful tool that can be used to decide how to optimally appropriate limited resources. The general optimization problem is: given an objective function $f(x)$ that maps controls x to a quantification of the impact (a real number), find the control x^* that maximizes (or minimizes) the objective f. If there are constraints on the controls x then we have a "constrained optimization problem," otherwise we have an "unconstrained optimization problem." For the problems in this dissertation, optimizing the objective $f(x)$ will involve solving a system of ordinary differential equations

(ODE) that models HIV transmission in the presence of control levels x. Such ODE models have long been used to model the transmission of infectious diseases [33, 39], including HIV [5, 9, 27]. The objective function, through solving the ODEs, calculates the simulated number of new infections, deaths, or similar at control level x, and thus the best control values x^* are sought. Many methods have been developed to solve optimization problems: in this dissertation we will explore two search methods, one for standard optimization problems, where x is a vector of real numbers, and another for optimal control problems, where x is a vector of time-dependent control functions.

1.1. Outline of Dissertation

In Chapter 2, a model of heterosexual transmission of HIV with particular characteristics of Sub-Saharan Africa is presented. We used a modeling approach to determine the optimal population-level combination of ART and PrEP allocations required in South Africa to maximize program effectiveness for four objectives: new infections, infection-years, deaths and cost. We considered two different strategies for allocating treatment, one that selectively allocates drugs to sex workers and the other that does not. We also considered the treatment levels that meet the UNAIDS's ambitious new 90–90–90 target.

In Chapter 3, we used the same model and expanded on the methods to examine three further treatment strategies. These are PrEP and ART allocation based on gender, PrEP and ART allocation based on risk and PrEP and ART allocation to people based on both gender and risk group. For each strategy, we determined the optimal allocation of treatment needed to minimize the same four objectives in Chapter 2.

In Chapter 4, we introduce a model based on the WHO five-stage classification of HIV progression and the WHO recommendations on when to commence treatment. We

calibrated our model with data from South Africa and used optimal control methods to determine the time-dependent stage-based allocation need to minimize new infections, infection-years, deaths and cost.

Finally, in Chapter 5, general conclusions are presented.

Detailed methods are provided in the Appendices.

1.2. Resulting Publications

This dissertation resulted in the following manuscripts that have been submitted to peer-reviewed journals:

G. Akudibillah, A. Pandey, M. Medlock Maximizing the benefits of ART and PrEP in resource-limited settings. Submitted Oct 2015 to *The Journal of Epidemiology and Infection*.

G. Akudibillah, A. Pandey, M. Medlock Optimal HIV treatment based disease stage. To be submitted in March 2016 to *Mathematical Biosciences*.

2. MAXIMIZING THE BENEFITS OF ART AND PREP IN RESOURCE-LIMITED SETTINGS

2.1. Introduction

The Joint United Nations Programme on HIV and AIDS (UNAIDS) estimates that 31.6–35.2 million people are living with HIV worldwide, with 2.4–2.9 million new HIV infections and 1.6–1.9 million AIDS-related deaths in 2010 [71]. HIV/AIDS is a major problem in developing countries especially in Sub-Saharan Africa. This region constitutes 12% of the world's population, yet accounts for 72% of AIDS-related deaths [71]. In Sub-Saharan Africa, most HIV transmission in adults occurs from heterosexual sex, with commercial sex [3] and multiple partners being key drivers of HIV transmission [31, 64].

Antiretroviral therapy (ART) has been traditionally used to prevent progression from HIV to AIDS [10, 73]. Recently, antiretroviral drugs have been shown to have the additional benefit of substantially reducing the infectiousness of infected people, leading to reduced transmission [30, 58, 60]. The HPTN052 trials showed that early treatment is able to reduce hetrosexual transmission of HIV in serodiscordant couples by as much as 96% [17].

Antiretroviral therapy is also administered to uninfected people either as post-exposure prophylaxis (PeP) after high-risk exposure or as pre-exposure prophylaxis (PrEP) for people with ongoing HIV exposure [1, 6, 29]. Giving ART as PrEP to uninfected heterosexual people reduced chance of acquiring HIV by 63–73% [6]. The US Centers for Disease Control and Prevention (CDC) recommends the use of ART as pre-exposure prophylaxis to, among others, serodiscordant couples and homosexual men or women who do not always use condoms when having sex with partners known to be at risk for HIV [12].

Whether as treatment to reduce the infectivity of infected people or as prophylaxis

to reduce the susceptibility of uninfected people, antiretroviral therapy has thus become an important strategy in the fight against HIV transmission [17, 20, 21, 50]. To take advantage of the benefits of antiretroviral theraphy, the World Health Organization (WHO) revised its treatment guidelines in 2013, recommending that treatment be initiated in infected adults and adolescents with a CD4 count of 500 cells/mm^3 or less [55]. UNAIDS also recently introduced an ambitious treatment target called 90–90–90, which by the year 2020 aims to have 90% of all people living with HIV know their HIV status, 90% of all people with diagnosed HIV infection receive sustained antiretroviral therapy and 90% of all people receiving antiretroviral therapy have viral suppression [72].

In this paper, we aim to inform drug-allocation policy in resource-limited settings by using a compartmental mathematical model for heterosexual transmission of HIV with treatment targeted by infection status, sexual-activity level and gender. Due to the availability of high-quality data, the model was parameterized for South Africa. We used optimization methods to determine the allocations of ART among the target groups that minimize new infections, HIV-related deaths and cost. We examined the effect increasing the amount of drugs available on new infections and prevalence. Also, we compared our optimization results to the 90–90–90 policy.

2.2. Methods

We constructed a mathematical model that incorporates risk, gender and treatment (Figure 2.1). We briefly outline the model here and the Supplementary Material contains a detailed description of the model.

We parameterized the model from from a variety of data sources from South Africa and, more broadly, Sub-Saharan Africa. We used demographic data from 2012, the date

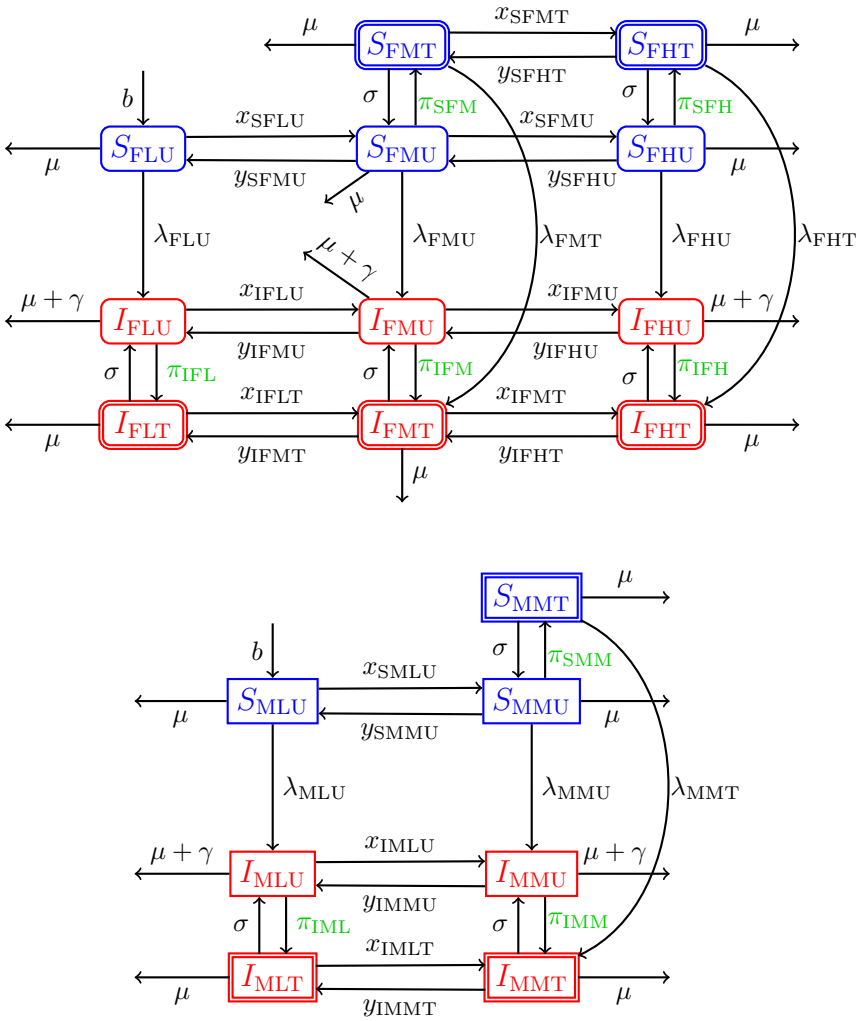

Figure 2.1: Diagram of HIV model. HIV-infected people are in red and susceptible (i.e. uninfected) people are in blue. The first subscript on the state variables S and I denotes gender (female or male), the second denotes risk level (low, medium or high), and the last denotes treatment status (untreated or treated). The subscripts on the other variables and parameters have similar meanings. The variables effected by treatment interventions are in green.

of the last South African National HIV Prevalence, Incidence and Behavior Survey [38]. Our model is initialized for the beginning of the year 2014 with an HIV prevalence of 17% in South African adult population, 42% of infected people on ART and 0% of susceptible people on PrEP [71]. The model included a 96% efficacy of ART [17] and a 71% efficacy of PrEP [6] at preventing HIV transmission.

Over a 10-year time horizon we calculated the effectiveness of different intervention strategies according to four different objectives: new infections, total infection-years, total deaths due to AIDS and total cost. New infections is the total incidence over the 10-year span. Total infection-years is the number of infected people, both treated and untreated, at each time, summed over the time period. Deaths due to AIDS is simply the total number of deaths over the time period. The total cost sums costs of infections, deaths and treatment.

We considered two different ways to model governmental-level intervention programs in which antiretroviral drugs are given as treatment of infected people as ART or as prophylaxis for uninfected people as PrEP. In the first intervention model, Global ART & PrEP, the two control variables are the amount of treatment allocated as ART and the amount allocated as PrEP, independent of the risk level of the people treated. The second intervention model was constructed to evaluate the current CDC recommendations for the use of PrEP in high-risk people [12, 13]. In this second intervention model, high-risk ART & PrEP, the three control variables are amounts of treatment allocated for ART in high-risk people, the amount for PrEP in high-risk people and the amount for ART for low- and medium-risk people, with no PrEP allocated for low-and medium-risk people. For both control models and all four objectives, the levels of the control variables that minimize the objective were found using numerical optimization methods.

2.3. Results

We simulated our model to determine the levels of treatment that minimize each of the objectives and to determine the effect of applying the optimal treatment strategy over a 10-year period. We examined the optimal allocation of ART and PrEP to people independent of their risk status (global ART & PrEP). In order to evaluate the current CDC treatment recommendations, we also evaluated the optimal allocation of ART and PrEP to high-risk people in addition to ART for the remainder of the population (ART & high-risk PrEP). We also compared the effectiveness of global ART & PrEP and ART & high-risk PrEP with the current treatment strategy in South Africa of only giving ART and the UNAIDS's 90–90–90 goals, in terms of reducing new infections, infection-years, deaths due to HIV, and cost.

With the global ART & PrEP control strategy, the optimal treatment allocations that minimized each of the four objectives showed similar patterns below about 7.5M available treatment spots (Figures 2.2A–2.5A), i.e. only PrEP was used at low treatment availability and only ART was used at higher levels of treatment availability, with the threshold level of treatment for switching ranging from about 2.0M to 5.3M, depending on the objective. Above 7.3–7.8M available treatment spots, two patterns emerged, adding PrEP in the cases of minimizing new infections and infection-years, and continuing to add ART, not PrEP, when minimizing deaths and cost.

For the ART & high-risk PrEP control strategy, the minimizing the four objectives again showed similar patterns (Figures 2.2B–2.5B). At low treatment availability, treatment was prioritized to high-risk people, first starting with PrEP and then adding on, or switching to in the case of minimizing deaths, ART. As available treatment increased, the treatment strategy switched to predominantly or entirely ART for non-high-risk people, with the

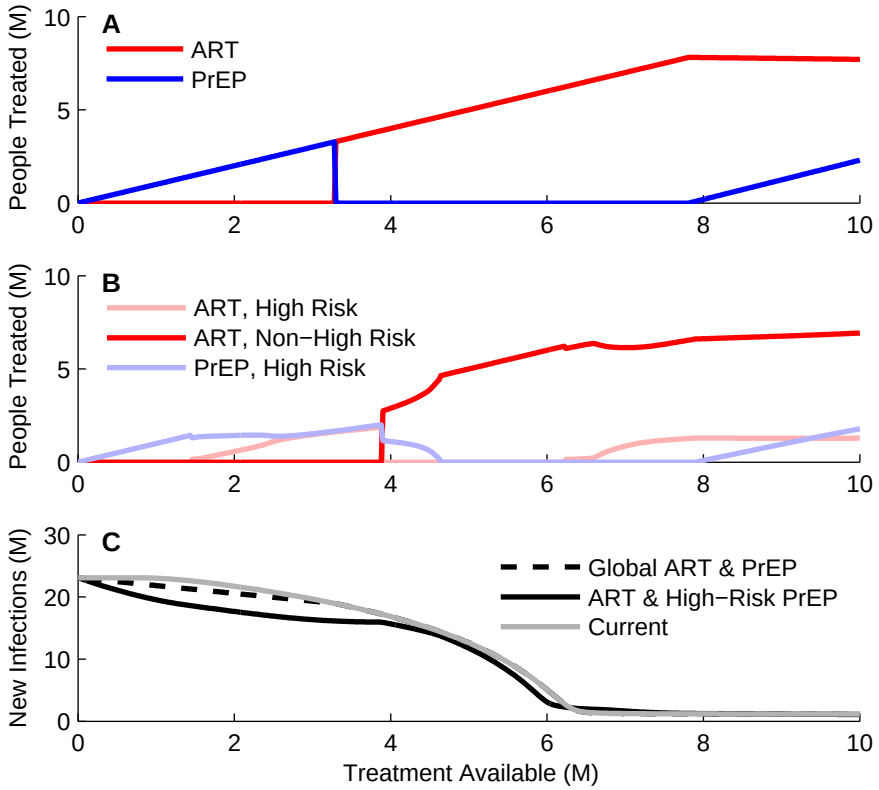

Figure 2.2: Treatment allocations that minimize new infections over 10 years vs. treatment availability.

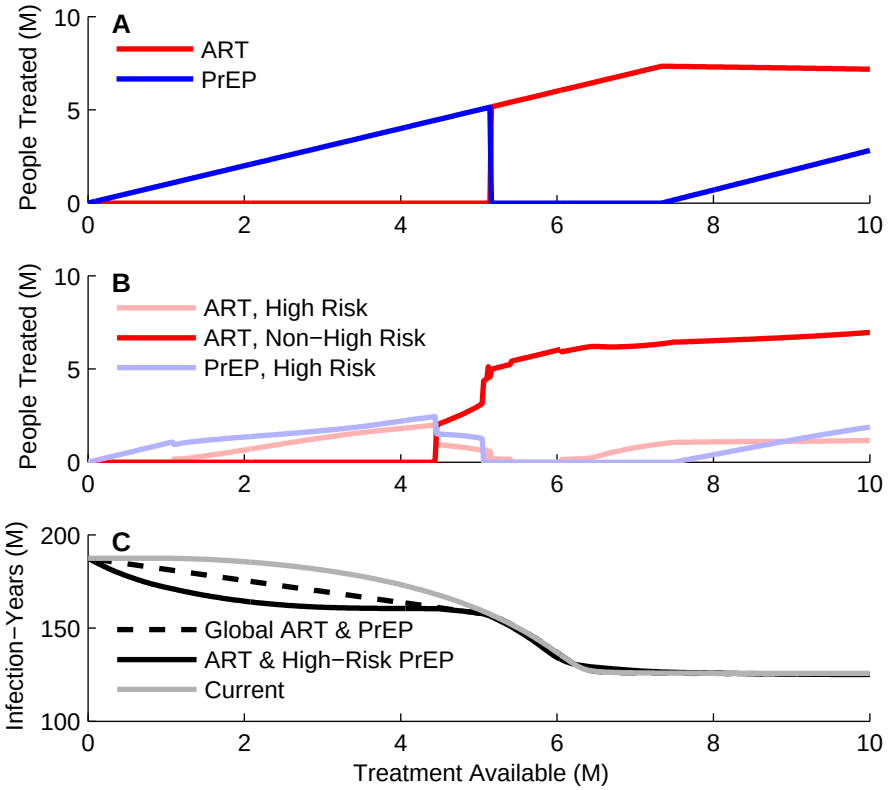

Figure 2.3: Treatment allocations that minimize infections-years over 10 years vs. treatment availability.

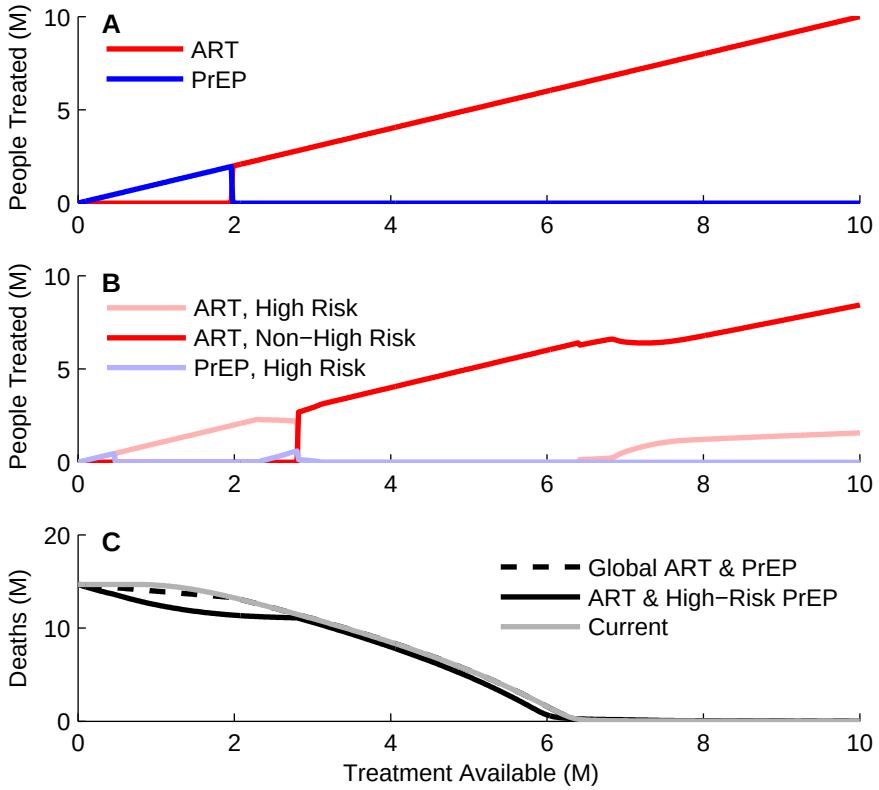

Figure 2.4: Treatment allocations that minimize deaths due to AIDS over 10 years vs. treatment availability.

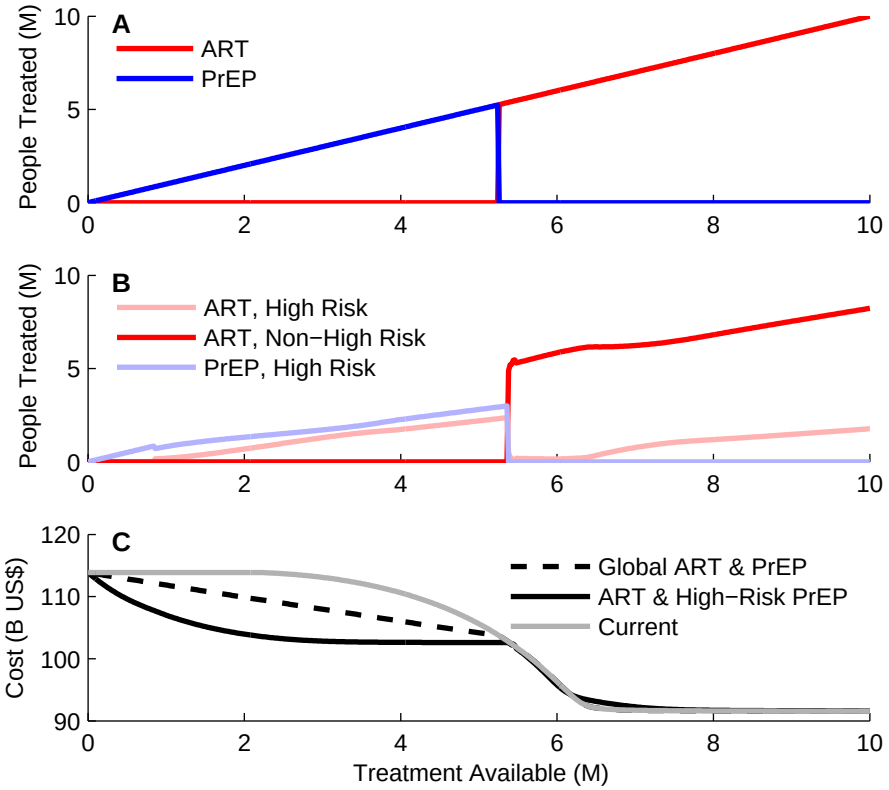

Figure 2.5: Treatment allocations that minimize cost over 10 years vs. treatment availability.

thresholds occurring between 2.8M and 5.8M available treatment spots, depending on the objective. At even higher levels of treatment, above 6.0–6.4M, ART was added for high-risk people for all objectives. Finally, above 7.5–7.9M available treatment spots, high-risk PrEP was added when minimizing new infections and infection-years.

Over the 10-year model period, increasing the amount of available treatment led to reductions in all four objectives (Figures 2.2C–2.5C), as expected. Substantial decreases in all objectives were seen for increasing treatment until around 6M available treatment spots. Increased treatment beyond this level resulted in very small gains to all objectives. For example, the number of deaths in the 10 years fell from 14.7M with no treatment to about 3,000 with 6.5M available treatment spots and increased treatment above 6.5M did not change the number of deaths by much. For low drug availability (i.e. 2.5–5.3M), the ART & high-risk PrEP control strategy decreased all objectives faster than either global ART & PrEP or the current strategy of global ART and, for minimizing new infections and minimizing deaths, the global ART & PrEP strategy preformed only slightly better than the current strategy. For higher levels of treatment, all three strategies performed similarly for all objectives.

In South Africa currently, 42% or 2.7M of the 6.5M HIV-infected people are on ART. To implement the UNAIDS's 90–90–90 goal, 5.8M people would need to be on ART currently. With the global ART & PrEP control strategy, allocating all of the available treatment as ART and none to PrEP minimized all four objectives (Figure 2.1). These optimal allocations, all ART and no PrEP, are exactly the current strategy in South Africa. With the ART & high-risk PrEP strategy, allocating all of the available treatment to non-high-risk ART minimized new infections, infection-years, and deaths, while cost was minimized by allocating 150,000 treatment spots to high-risk ART and the remainder to non-high-risk ART. The ART & high-risk PrEP strategy performed better than global ART

& PrEP for all four objectives, but only slightly so for infection-years and cost.

2.4. Discussion

Using a mathematical model of heterosexual HIV transmission, we determined two treatment strategies that minimize new infections, infection-years, AIDS-related deaths and cost. We found that when PrEP and ART are given to the general population at different levels (our global PrEP & ART strategy), only PrEP should be used at low treatment availability, but only ART should be used at higher levels, with the thresholds for switching from PrEP to ART differing for the four objectives. At very high levels of drug availability, adding PrEP minimizes new infections and infection-years, while continuing to use only ART minimizes deaths and cost. When PrEP and ART are given to high-risk people and ART to non-high-risk people (our ART & high-risk PrEP strategy), we found that at low treatment availability only high-risk people should be treated; non-high-risk people should only be treated above threshold levels of treatment that vary by objective. At high treatment availability, treating both high-risk and non-high risk becomes the optimal allocation for all objectives.

We also evaluated the UNAIDS recent 90–90–90 target and found that the current ART-only treatment strategy of South Africa and indeed most resource-limited counties is optimal or near optimal at 90% treatment coverage for minimizing all four objectives. The one exception was that allocating 150,000 for ART to high-risk people and the remainder to ART for non-high-risk people minimized cost.

Resource-limited countries are likely to have low treatment availability: for example with South Africa's current level of treatment availability (2.7M slots), PrEP to the general population (for global ART & PrEP) or treatment of only sex workers with ART & PrEP

	Global ART & PrEP			ART & high-risk PrEP			
				High-risk	High-risk	Non-high-risk	
Objective minimized	**ART (M)**	**PrEP (M)**	**Objective value**	**ART (M)**	**PrEP (M)**	**ART (M)**	**Objective value**
New infections	5.8	0	6.9M	0	0	5.8	5.0M
Infection-years	5.8	0	142M	0	0	5.8	139M
Deaths	5.8	0	2.4M	0	0	5.8	1.4M
Cost	5.8	0	$98.4B	0.1	0	5.7	$98.0B

Table 2.1: Optimal treatment allocations at 5.8M available treatment spots (i.e. enough treatment for 90% coverage of infected people) and their effectiveness for the four objectives under both the global ART & PrEP and ART & high-risk PrEP control strategies.

(for ART & high-risk PrEP) is most effective at reducing new infections, infection-years, and cost, while using the treatment as ART to either the general population (global ART & PrEP) or ART to non-high-risk people (ART & high-risk PrEP) best reduces deaths.

Allocations that prioritize treatment of commercial sex workers (CSW) may be difficult to implement. CSW are often stigmatized and their actions are frequently illegal, which may present challenges to getting them out of the shadows for treatment. Even when CSW do come forward for treatment, prioritizing treatment to CSW will mean withholding treatment from non-high-risk people, which may be a difficult decision for policymakers to justify. We believe, however, that the savings in HIV mortality, morbidity, and cost clearly justify such policies.

We hope that this work provides some guidance for policymakers in determining how to allocate HIV treatment to maximize their benefit.

3. FURTHER TREATMENT STRATEGIES FOR ART AND PREP

3.1. Introduction

More than three decades on, HIV/AIDS remains a massive public health problem especially in Sub-Saharan Africa. Globally, 33.2–37.2million people are living with HIV, with 1.9–2.4 million new HIV infections and 1.4–1.7 million AIDS-related deaths in 2013 [71]. Antiretroviral therapy (ART) has been traditionally used to prevent progression from HIV to AIDS in infected people [10, 73] while pre-exposure prophylaxis (PrEP) is antiretroviral therapy administered to uninfected people for people with ongoing HIV exposure to reduce the chances of getting infected [1, 6, 29]. The use of PrEP has proven effective in reducing HIV transmission and is increasingly becoming an important biomedical strategy to fight the HIV/AIDS pandemic [1, 6, 29].

While governments worldwide seek judicious use of their health resources, the need for optimal allocation of resources is especially needed in limited-resource settings. To inform governmental-level drug-allocation policy, we used a compartmental mathematical model for heterosexual transmission of HIV stratified by infection status, sexual-activity level, gender and treatment. This model was initially developed in Chapter 2. to study two optimal HIV drug-allocation strategies. We used the same model and expanded on the methods to examine three further treatment strategies. These are PrEP and ART treatment based on gender, PrEP and ART treatment based on risk and PrEP and ART treatments based on both gender and risk. For these three new treatments strategies, we generally followed the same methods as in Chapter 2 to determine how to allocate treatment to minimize new infections, infection-years, deaths and cost.

3.2. Methods

Our model (Figure 3.1) is a heterosexual model for HIV transmission that divides the sexually active population by gender, risk group and treatment status. The risk groups are: low risk are people are in monogamous marriages, single or otherwise have zero or one sexual partner; medium risk are people who have more than one sexual partner and high risk are commercial sex workers, sometimes called transactional sex workers. Movement between risk groups is due to sexual behavioral change.

Treatment in the model is available as PrEP and ART. In our model, PrEP is limited to some medium and high-risk uninfected people while ART can be administered to any infected person in any risk group. The complete description of the model and equations and calibration are provided in the appendix (A1, A2)

We expanded the methods in Chapter 2 to consider three further ways to model government-level intervention programs that use antiretroviral drugs to control HIV. PrEP. In the first intervention model, the gender treatment strategy, there are coverage levels of ART and PrEP for the male and female population. Thus all infected males and females are each assigned a level of ART regardless of risk group and all uninfected males and females are are assigned a level of PrEP regardless of risk group. The second intervention model, the risk treatment strategy, there are unique coverage levels of ART and PrEP for each risk group. Meaning all infected people in each risk group are are each assigned a level of ART regardless of gender and all uninfected people in each risk group are assigned a level of PrEP regardless of gender. In the third intervention model, the gender & risk treatment strategy, infected people of each gender and risk group are allocated a level of ART and uninfected people of each gender and risk group are allocated a level of

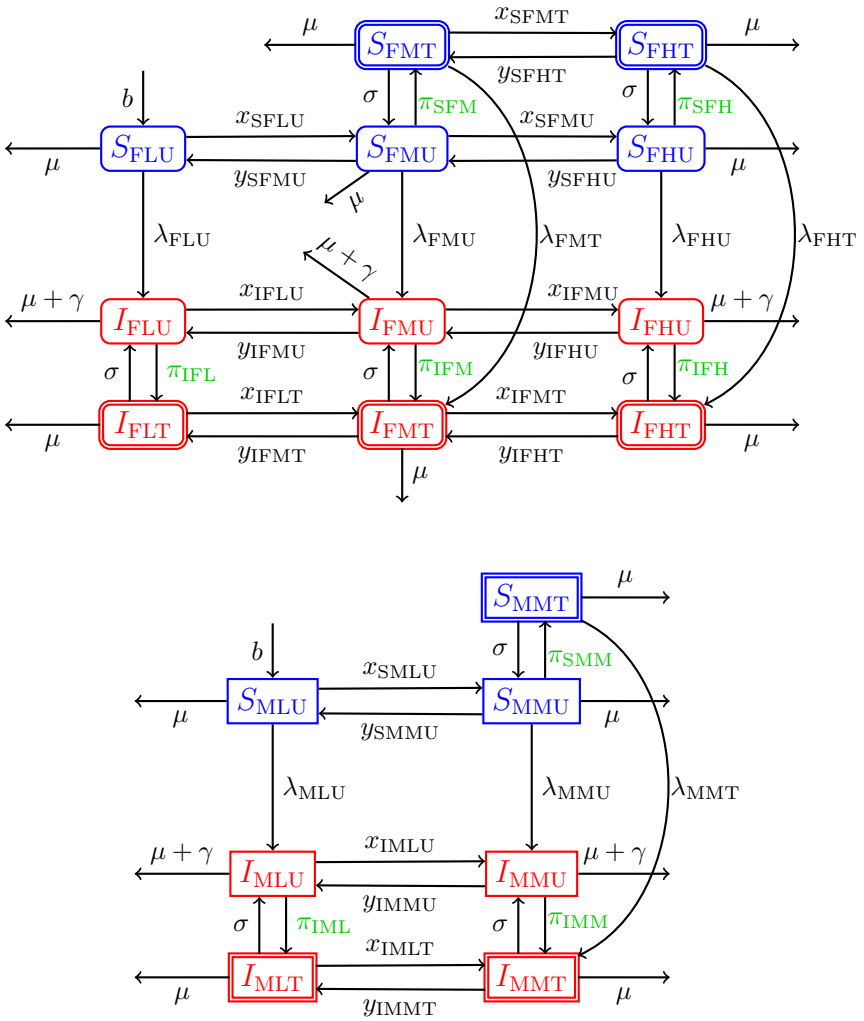

Figure 3.1: Diagram of HIV model. HIV-infected people are in red and susceptible (i.e. uninfected) people are in blue. The first subscript on the state variables S and I denotes gender (female or male), the second denotes risk level (low, medium or high), and the last denotes treatment status (untreated or treated). The subscripts on the other variables and parameters have similar meanings. The variables effected by treatment interventions are in green.

3.2.1 Objectives

We determined the proportion of each group to be treated by minimizing one of four different outcome measures: the number of new infections, the number of total infection-years, the number of deaths and total cost. These were all evaluated over a time horizon of $t_{\text{end}} = 10$ y.

New infections The total number of new infections is the sum of the rates of new infections arising in each of the 8 uninfected classes integrated over the time period,

$$\int_0^{t_{\text{end}}} \sum_{\substack{i\in\{\text{F,M}\}\\ j\in\{\text{L,M,H}\}\\ k\in\{\text{U,T}\}}} \lambda_{ijk}(t) S_{ijk}(t)\,\mathrm{d}t, \tag{3.1}$$

with λ_{ijk} defined to be 0 for the groups not in the model (FLT, MLT, MHU, and MHT).

Total infection-years The total infection-years is the sum of the number of all the infected people (both treated and untreated) in the 10 infected classes at each time, integrated over the time period,

$$\int_0^{t_{\text{end}}} I(t)\,\mathrm{d}t, \tag{3.2}$$

where the number of infected people is

$$I(t) = \sum_{\substack{i\in\{\text{F,M}\}\\ j\in\{\text{L,M,H}\}\\ k\in\{\text{U,T}\}}} I_{ijk}(t). \tag{3.3}$$

Deaths due to AIDS Treatment prevents progression from HIV to AIDS, thus reducing deaths. We assumed that AIDS-related deaths only occur in infected, untreated people. Thus, the total number of deaths due to AIDS is the number of untreated

infected people at each time,

$$I_{\mathrm{U}}(t) = \sum_{\substack{i\in\{\mathrm{F,M}\}\\ j\in\{\mathrm{L,M,H}\}}} I_{ij\mathrm{U}}(t), \tag{3.4}$$

multiplied by the rate of death due to AIDS:

$$\int_0^{t_{\mathrm{end}}} \gamma I_{\mathrm{U}}(t)\mathrm{d}t. \tag{3.5}$$

Total cost The total cost consists of cost of infections, cost of deaths, and cost of treatment. The disease cost per year, which includes monetary equivalent loss of the infected people, like lost productivity etc., is average cost of disease per person per year times the total number of people infected,

$$C_{\mathrm{I}}(t) = c_{\mathrm{I}} I(t). \tag{3.6}$$

The cost per year of deaths is the the cost per death times the number of deaths per year,

$$C_{\mathrm{D}}(t) = c_{\mathrm{D}} \gamma I_{\mathrm{U}}(t). \tag{3.7}$$

The treatment cost per year is the cost per person per year times the number of people treated,

$$C_{\mathrm{T}}(t) = c_{\mathrm{T}} T(t), \tag{3.8}$$

where the number of people treated is

$$T(t) = \sum_{\substack{i\in\{\mathrm{F,M}\}\\ j\in\{\mathrm{L,M,H}\}}} \left[S_{ij\mathrm{T}}(t) + I_{ij\mathrm{T}}(t)\right]. \tag{3.9}$$

The cost objective is discounted sum of these costs, integrated over the time period,

$$\int_0^{t_{\mathrm{end}}} \left[C_{\mathrm{I}}(t) + C_{\mathrm{D}}(t) + C_{\mathrm{T}}(t)\right]\mathrm{e}^{-rt}\mathrm{d}t. \tag{3.10}$$

The total cost is discounted at rate $r = 0.03$, representing the rate a policymaker is willing to pay as trade-off for the cost today versus future cost [24].

3.2.2 Controls and Constraints

Like in in Chapter 2, the control variables are the total number of people targeted to be on treatment in each of the designated groups (V_G) at any one time. From these treatment targets, the flows of people into treatment per unit time were taken to be

$$\pi_G(t) = r_{\max} \max\left(V_G - T_G(t),\ 0\right), \tag{3.11}$$

where $T_G(t)$ is the total number of people in group G currently on treatment. The quantity $\max(V_G - T_G(t), 0)$ is the number of treatment slots available at time t and $r_{\max}$ is the rate of enrolling people on treatment per drug available per untreated person.

Gender Treatment Strategy We modeled this treatment strategy with four control groups, one for each gender and infection group:

infected males on ART,

$$G_{\text{ART,M}} = \{I_{\text{ML}}, I_{\text{MM}}\}, \tag{3.12}$$

uninfected males on PrEP,

$$G_{\text{PrEP,M}} = \{S_{\text{MM}}\}, \tag{3.13}$$

infected females on ART,

$$G_{\text{ART,F}} = \{I_{\text{FL}}, I_{\text{FM}}, I_{\text{FH}}\}, \tag{3.14}$$

and uninfected females on PrEP,

$$G_{\text{PrEP,F}} = \{S_{\text{FM}}, S_{\text{FH}}\}. \tag{3.15}$$

The control variables ($V_{\text{ART,M}}$, $V_{\text{ART,F}}$, $V_{\text{PrEP,M}}$, $V_{\text{PrEP,F}}$) are the total number of people targeted to be on treatment in each of the control groups at any one time.

The total numbers of people in the treatment groups at time t are

$$\begin{aligned} T_{\mathrm{ART,M}}(t) &= I_{\mathrm{MLT}} + I_{\mathrm{MMT}}, \\ T_{\mathrm{ART,LM}}(t) &= I_{\mathrm{FLT}} + I_{\mathrm{FMT}} + I_{\mathrm{FHT}}, \\ T_{\mathrm{PrEP,M}}(t) &= S_{\mathrm{MMT}}, \\ T_{\mathrm{PrEP,F}}(t) &= S_{\mathrm{FMT}} + S_{\mathrm{FHT}}, \end{aligned} \tag{3.16}$$

and the flows into treatment for the control groups are

$$\begin{aligned} \pi_{\mathrm{IML}} = \pi_{\mathrm{IMM}} &= \pi_{\mathrm{ART,M}}, \\ \pi_{\mathrm{IFL}} = \pi_{\mathrm{IFM}} = \pi_{\mathrm{IFH}} &= \pi_{\mathrm{ART,F}}, \\ \pi_{\mathrm{SMM}} &= \pi_{\mathrm{PrEP,M}}, \\ \pi_{\mathrm{SFM}} = \pi_{\mathrm{SFH}} = \pi_{\mathrm{SMM}} &= \pi_{\mathrm{PrEP,F}}. \end{aligned} \tag{3.17}$$

The constraints on the controls are that each of them are positive and that their sum is less than the total amount of treatment available, $T_{\max}$:

$$\begin{aligned} V_{\mathrm{ART,M}} \geq 0, \quad V_{\mathrm{ART,F}} \geq 0, \quad V_{\mathrm{PrEP,M}} \geq 0, \quad V_{\mathrm{PrEP,F}} \geq 0, \\ V_{\mathrm{ART,M}} + V_{\mathrm{ART,F}} + V_{\mathrm{PrEP,M}} + V_{\mathrm{PrEP,F}} \leq T_{\max}. \end{aligned} \tag{3.18}$$

Risk Treatment Strategy This strategy involves unique treatment to each risk group, giving five control groups, one for each of the three infected-treated risk group and two uninfected-treated risk groups:

low-risk people on ART,

$$G_{\mathrm{ART,L}} = \{I_{\mathrm{MLT}}, I_{\mathrm{FLT}}\}, \tag{3.19}$$

medium-risk people on PrEP,

$$G_{\mathrm{PrEP,M}} = \{S_{\mathrm{MMT}}, S_{\mathrm{FMT}}\}, \tag{3.20}$$

medium-risk people on ART,

$$G_{\mathrm{ART,M}} = \{I_{\mathrm{MMT}}, I_{\mathrm{FMT}}\}, \tag{3.21}$$

high-risk people on PrEP,

$$G_{\mathrm{PrEP,H}} = \{S_{\mathrm{FHT}}\}, \tag{3.22}$$

high-risk people on ART,

$$G_{\mathrm{ART,H}} = \{I_{\mathrm{FHT}}\}. \tag{3.23}$$

The control variables ($V_{\mathrm{ART,L}}$, $V_{\mathrm{ART,M}}$, $V_{\mathrm{ART,H}}$ $V_{\mathrm{PrEP,M}}$, $V_{\mathrm{PrEP,H}}$) are the total number of people targeted to be on treatment in each of the control groups at any one time. The total numbers of people in the treatment groups at time t are

$$\begin{aligned} T_{\mathrm{ART,L}}(t) &= I_{\mathrm{FLT}} + I_{\mathrm{MLT}}, \\ T_{\mathrm{ART,M}}(t) &= I_{\mathrm{FMT}} + I_{\mathrm{MMT}}, \\ T_{\mathrm{ART,H}}(t) &= I_{\mathrm{FHT}}, \\ T_{\mathrm{PrEP,M}}(t) &= S_{\mathrm{FMT}} + S_{\mathrm{MMT}}, \\ T_{\mathrm{PrEP,H}}(t) &= S_{\mathrm{FHT}}. \end{aligned} \tag{3.24}$$

and the flows into treatment for the control groups are

$$\begin{aligned} \pi_{\mathrm{IFL}} = \pi_{\mathrm{IML}} &= \pi_{\mathrm{ART,L}}, \\ \pi_{\mathrm{IFM}} = \pi_{\mathrm{IMM}} &= \pi_{\mathrm{ART,M}}, \\ \pi_{\mathrm{IFH}} &= \pi_{\mathrm{ART,H}}, \\ \pi_{\mathrm{SFM}} = \pi_{\mathrm{SMM}} &= \pi_{\mathrm{PrEP,M}}, \\ \pi_{\mathrm{SFH}} &= \pi_{\mathrm{PrEP,H}}. \end{aligned} \tag{3.25}$$

The constraints on the controls are that each of them are positive and that their sum

is less than the total amount of treatment available, $T_{\max}$:

$$\begin{gathered} V_{\mathrm{ART,L}} \geq 0, \quad V_{\mathrm{ART,M}} \geq 0, \quad V_{\mathrm{ART,H}} \geq 0, \quad V_{\mathrm{PrEP,M}} \geq 0, \quad V_{\mathrm{PrEP,H}} \geq 0, \\ V_{\mathrm{ART,L}} + V_{\mathrm{ART,M}} + V_{\mathrm{ART,H}} + V_{\mathrm{PrEP,M}} + V_{\mathrm{PrEP,H}} \leq T_{\max}. \end{gathered} \tag{3.26}$$

Gender & Risk Treatment Strategy This is a strategy to have a unique treatment level for each risk and gender group, giving eight control groups:

low-risk infected males on ART,

$$G_{\mathrm{ART,LM}} = \{I_{\mathrm{ML}}\}, \tag{3.27}$$

low-risk infected females on ART,

$$G_{\mathrm{ART,FL}} = \{I_{\mathrm{FL}}\}, \tag{3.28}$$

medium-risk infected males on ART,

$$G_{\mathrm{ART,MM}} = \{I_{\mathrm{MM}}\}, \tag{3.29}$$

medium-risk infected females on ART,

$$G_{\mathrm{ART,FM}} = \{I_{\mathrm{FM}}\}, \tag{3.30}$$

high-risk infected females on ART,

$$G_{\mathrm{ART,FH}} = \{I_{\mathrm{FH}}\}, \tag{3.31}$$

medium-risk uninfected males on PrEP,

$$G_{\mathrm{PrEP,MM}} = \{S_{\mathrm{MM}}\}, \tag{3.32}$$

medium-risk uninfected females on PrEP,

$$G_{\mathrm{PrEP,FM}} = \{S_{\mathrm{FM}}\}, \tag{3.33}$$

and high-risk uninfected females on PrEP,

$$G_{\mathrm{PrEP,FH}} = \{S_{\mathrm{FH}}\}\,. \tag{3.34}$$

The control variables ($V_{\mathrm{ART,ML}}$, $V_{\mathrm{ART,FL}}$, $V_{\mathrm{ART,MM}}$, $V_{\mathrm{ART,FM}}$, $V_{\mathrm{ART,FH}}$, $V_{\mathrm{PrEP,MM}}$, $V_{\mathrm{PrEP,FM}}$, $V_{\mathrm{PrEP,FH}}$) are the total number of people targeted to be on treatment in each of the control groups at any one time. The total numbers of people in the treatment groups at time t are

$$\begin{aligned} T_{\mathrm{ART,Mj}}(t) &= I_{\mathrm{MjT}} \quad j \in \{\mathrm{L,M}\}, \\ T_{\mathrm{ART,Fj}}(t) &= I_{\mathrm{FjT}} \quad j \in \{\mathrm{L,M,H}\}, \\ T_{\mathrm{PrEP,MM}}(t) &= S_{\mathrm{MMT}}, \\ T_{\mathrm{PrEP,Fj}}(t) &= S_{\mathrm{FjT}} \quad j \in \{\mathrm{M,H}\}, \end{aligned} \tag{3.35}$$

and the flows into treatment for the control groups are

$$\begin{aligned} \pi_{\mathrm{IMj}} &= \pi_{\mathrm{ART,Mj}} \quad j \in \{\mathrm{L,M}\}, \\ \pi_{\mathrm{IFj}} &= \pi_{\mathrm{ART,Fj}} \quad j \in \{\mathrm{L,M,H}\}, \\ \pi_{\mathrm{SMM}} &= \pi_{\mathrm{PrEP,MM}}, \\ \pi_{\mathrm{SFj}} &= \pi_{\mathrm{PrEP,Fj}} \quad j \in \{\mathrm{M,H}\}, \end{aligned} \tag{3.36}$$

The constraints on the controls are that each of them are positive and that their sum is less than the total amount of treatment available, $T_{\max}$:

$$\begin{aligned} V_{\mathrm{ART,Mj}} &\geq 0 \quad j \in \{\mathrm{L,M}\}, \\ V_{\mathrm{ART,Fj}} &\geq 0 \quad j \in \{\mathrm{L,M,H}\}, \\ V_{\mathrm{PrEP,MM}} &\geq 0, \\ V_{\mathrm{PrEP,Fj}} &\geq 0 \quad j \in \{\mathrm{M,H}\}, \end{aligned} \tag{3.37}$$

$$\sum_{i\in\{L,M\}} V_{\mathrm{ART,Mi}} + \sum_{j\in\{L,M,H\}} V_{\mathrm{ART,Fj}} + \sum_{k\in\{M,H\}} V_{\mathrm{PrEP,Fk}} + V_{\mathrm{PrEP,MM}} \leq T_{\max}$$

The COBYLA algorithm (Constrained Optimization BY Linear Approximations) by Powell [59] is our method of choice for the optimization. COBYLA approximates our nonlinear constrained optimization problem to a linear programming problem and solves the linear programming problem iteratively to obtain an optimal solution. We adopted a random restart strategy to increase our chances of obtaining a global solution in each case.

3.3. Results

Like in Chapter 2, we simulated our model to determine the levels of treatment that minimize each of the objectives and to determine the effect of applying the optimal treatment strategy over a 10-year period. We examined the optimal allocation of ART and PrEP to people based on the gender and risk (gender & risk treatment strategy), gender independent of risk group (gender treatment strategy) and risk independent of gender group (risk treatment strategy)

3.3.1 Gender Treatment Strategy

The optimal treatment strategy for minimizing all four objectives showed similar patterns. At low treatment availability, PrEP for uninfected females is prioritized (Figures 3.2A–3.5A). Above a threshold amount of available drugs, priority switches to infected males with the threshold varying by objective from 1M to 3M treatment slots. As available treatment continues to increase, infected females are added, except for under the cost objective. The available treatment spots at which ART to infected females should be considered are 5.8M for new infections, 6.1M for infection-years and cost and 5.6M for deaths. At even higher levels of treatment levels adding PrEP treatment spots to females is required to minimize new infection, female ART treatment spots are turned off in favor

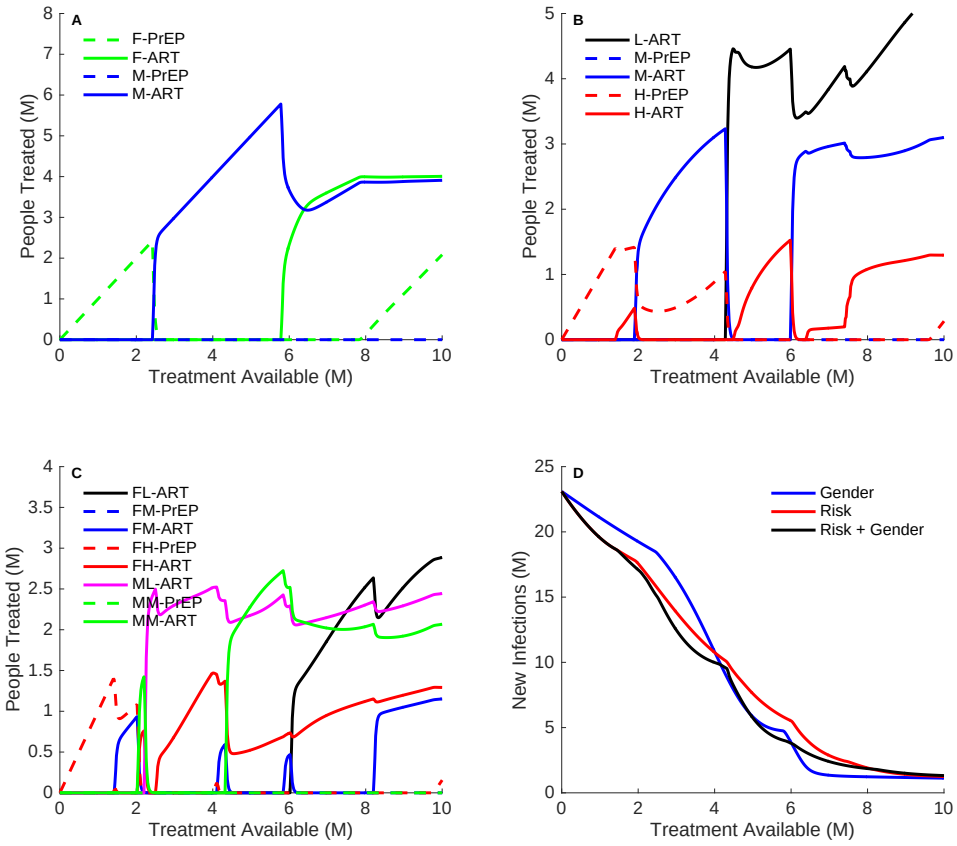

Figure 3.2: Treatment allocations that minimize new infections over 10 years vs. treatment availability: (A) Gender treatment strategy, (B) Risk treatment strategy, (C) Gender & risk treatment Strategy and (D) Effect of drug availability on objective value

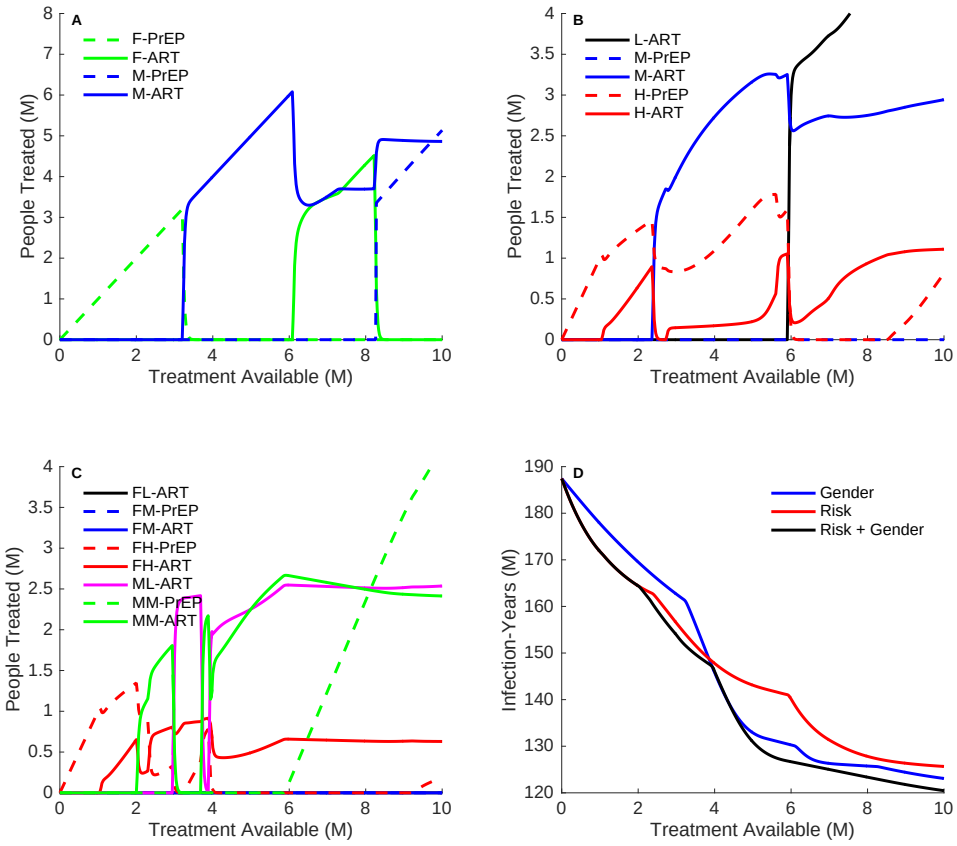

Figure 3.3: Treatment allocations that minimize infections-years over 10 years vs. treatment availability: (A) Gender treatment strategy, (B) Risk treatment strategy, (C) Gender & risk treatment Strategy and (D) Effect of drug availability on objective value

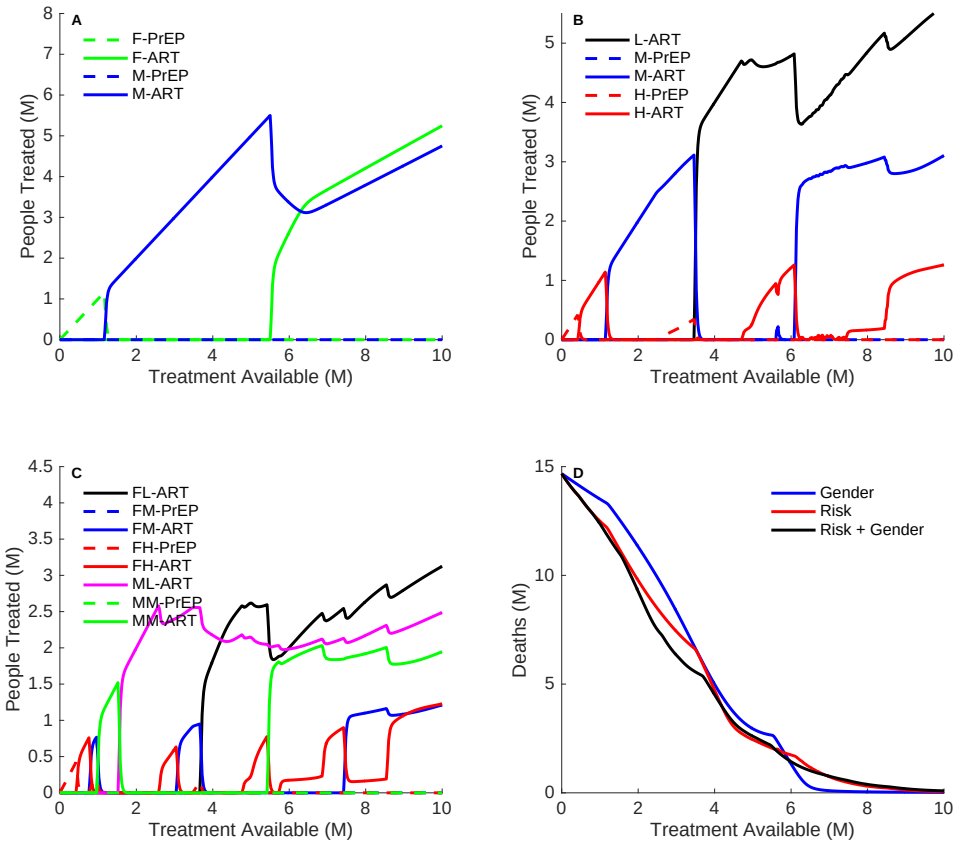

Figure 3.4: Treatment allocations that minimize deaths due to AIDS over 10 years vs. treatment availability: (A) Gender treatment strategy, (B) Risk treatment strategy, (C) Gender & risk treatment Strategy and (D) Effect of drug availability on objective value

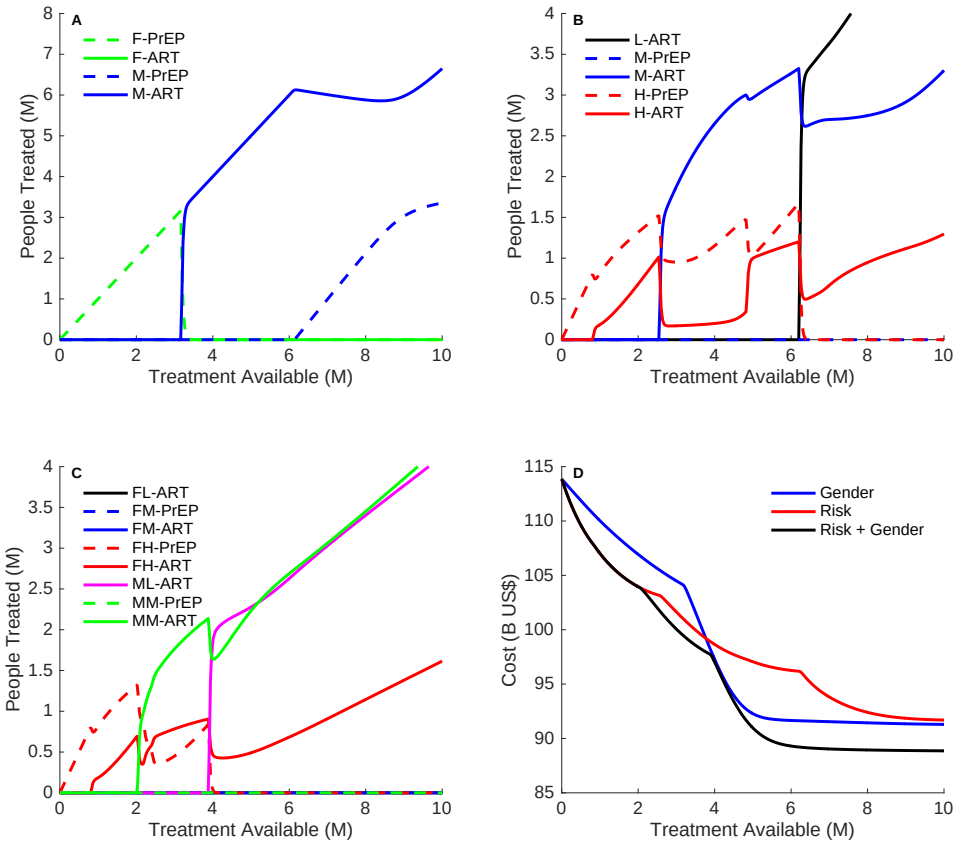

Figure 3.5: Treatment allocations that minimize cost over 10 years vs. treatment availability: (A) Gender treatment strategy, (B) Risk treatment strategy, (C) Gender & risk treatment Strategy and (D) Effect of drug availability on objective value

of male PrEP treatment spots to minimize infection-years, and addition of male PrEp is required to minimize cost and no additional PrEP treatment is need to minimize deaths

3.3.2 Risk Treatment Strategy

To minimize new infections at low available treatment spots (i.e. between 0–2.0M) treatment should be prioritized to high-risk people, first PrEP and then ART (Figure 3.2B). After 2.0M, the optimal strategy prescribes adding on ART for medium-risk people. PrEP for high-risk people should ends at 4M. After 6M treatment spots an all-ART strategy is required to minimize new infections, deaths and cost with most treatment going to low-risk people and the least treatment to high-risk people.

The optimal strategy to minimizing both infection-years and cost are similar to each other (Figures 3.5B, and 3.3B). The start with PrEP and ART treatment for high-risk females at low treatment levels then add on medium risk ART at 2.5M available treatment spots. PrEP for high-risk people is stopped at 6.0M switching to an ART to all risk groups strategy, with high levels allocated to low-risk people.

To minimize deaths the optimal strategy is to prioritize PrEP to high-risk people for the first 0.5M available treatment spots and switching to a mostly all-ART strategy (Figure 3.4B). Starting first with ART to only high-risk people between 0.5–1.0M switching entirely to ART for medium-risk people between 1–3.3M. After 3.3M ART treatment to all three risk groups become the optimal strategy. Generally, the treatment allocations are similar to those for new infections, but with much less high-risk PrEP.

3.3.3 Gender & Risk Treatment Strategy

To minimize new infections at low availability treatment spots (between 0–2.0M), initially PrEP is allocated for high-risk females (Figure 3.2C), which is consistent with the the results from both the gender (Figure 3.2A) and risk treatment strategies (Figure 3.2B).

Above 4.0M available treatment spots, an all-ART treatment becomes the optimal strategy (again consistent both the gender and risk models), starting first with ART to low-risk males and high-risk females and adding on ART to medium-risk males, low-risk females and medium-risk females at 4.2M, 6.0M and 8.1M respectively. There are brief periods near 2.0M, 4M and 6M available treatment spots when ART to medium-risk females should be be considered (Figure 3.2A).

As in the case of minimizing new infection above, to minimize infection-years at low available treatment spots, treatment of high-risk females is prioritized (Figure 3.2C). However, this time a combination of PrEP and ART to high-risk females is optimal at low treatment availability. ART to medium-risk males and low-risk males should begin at 3M and 3.9M available treatment spots respectively, adding PrEP for medium-risk male at 6.0M. No treatment to to medium risk females is required to minimize minimize infection-years.

To minimize deaths between 0–2.0M available treatment spots, four different treatment pattens were observed (Figure 3.4C). Initially, only PrEP for high-risk females is allocated, replaced by only ART for high-risk females at 0.5M available treatment spots, then replaced by only ART for medium-risk females and finally only ART for medium-risk males. At 2.0M there is a switch to entirely ART for low-risk males, adding on ART to high-risk and medium-risk females as more slots become available between 2.0–4.0M. After 4.0M, ART for low-risk females and low-risk males is the optimal strategy, adding on ART to high-risk females as more slot become available before 5.4M. Between 5.4–7.0M the optimal strategy is ART for low-risk females and males, ART for medium-risk males and ART for high-risk females. After 7.0M the optimal strategy is ART for low-risk females and males, ART for medium-risk males and females and ART for high-risk females. In general very little PrEP treatment spots are required to minimize death.

To minimize cost between 0–2.0M available treatment spots treatment of high-

risk females is prioritized (Figure 3.5A). Between 2.0–4.0M available treatment spots, a combination of treatment to high-risk females and ART for medium-risk male is optimal. After 4.0M, an all-ART strategy to infected low and medium-risk males and high-risk females is optimal. The results for both the gender (Figure 3.5A) and risk treatment strategies (Figure 3.5B) mirror that of the gender & risk treatment strategy .

Over the 10-year model period, increasing the amount of available treatment led to reductions in all four objectives (Figures 3.2D–3.5D), as expected. For less than 6M available treatment slots, as expected the gender &risk treatment strategy performed at lease as well as the other two strategies. After 6M, the Gender treatment strategy, performed better than the Gender & Risk Treatment Strategy in minimizing new infections and deaths, this might be due to a local minima.

3.4. Discussion

We used a mathematical model of heterosexual HIV transmission and optimization techniques to determine three HIV treatment strategies that minimize new infections, infection-years, AIDS-related deaths and cost. We found that at low treatment availability the optimal treatment strategy is to prioritize ART and PrEP to high-risk people while at very high levels of drug availability mostly ART treatment is generally optimal.

The optimal gender-based treatment strategy to minimize all four objectives showed very similar patterns: PrEP for females at low availability replaced by ART for males as availability increases and depending the objective, adding ART for females and PrEP for males and females as more treatment spots become available. It interesting to note that when risk is not a factor in determining treatment allocation, all four objectives result in prioritizing PrEP for females at low drug availability and switching to ART to males as

more treatment spots become available. This trend as is confirmed by the risk-based and the gender & risk-based treatment strategies might be from the need to high-risk people (sex workers) in order to prevent them from spreading the disease. Another plausible reason is from the fact that as compared to males, females are twice as likely to contract HIV per coital act so at low drug level it might be beneficial to protect them first.

Expect for a initial brief period where PrEP is preferred, all three treatment strategies showed that very little PrEP is needed to minimize deaths. The predominantly all-ART treatment strategies to minimize deaths is based on the need to provide ART to infected people in order to prevent them from dying from AIDS.

Over the 10-year model period, increasing the amount of available treatment led to substantial reductions in all four objectives. The decrease is sleep initially and however, at high drug availability decreasing returns are observed probably because all anybody requiring treatment is on treatment. After 10 years of applying the respective optimal treatment strategies, the number of people dying from AIDS can be reduced by 14M with 21.8M new infection and 67M infection-years averted and a cost savings of 24B US$.

The results of all three treatment strategies generally agree with of the results of the two previous treatment strategies in Chapter 2. All 5 treatment strategies suggest an only PrEP was used at low treatment availability and only ART was used at higher levels of treatment availability, with the threshold level of treatment for switching depending on the objective.

Despite the new WHO "test and Treat" recommendation i.e to start treatment immediately anybody is diagnosed with HIV regardless of CD4 count, many limited resources countries still prioritize ART treatment to infected people in the advance stages of HIV. However, our results suggest that at low drug availability policy makers should be giving those drugs as PrEP and ART to high-risk people. For example at South Africa's

current level of treatment availability (2.7M slots), allocating these treatment slots as PrEP and ART to low-risk people and ART for medium-risk people minimizes cost and infection years. If the objective is to minimize deaths, then the best strategy is ART to medium-risk people, while PrEP to low-risk people and ART for medium-risk people minimizes new infection.

Ethically speaking, the policy of selectively allocating treatment can be a very difficult to justify because it is is akin to playing "God". One way a health policymaker can justify selectively allocating treatment is when this leads to large savings in HIV mortality, morbidity, and cost. By proving the actual number of savings in HIV mortality, morbidity, and cost we hope this will help policymakers make a their case.

4. OPTIMAL CONTROL FOR HIV TREATMENT

4.1. Introduction

After three decades of HIV/AIDS, it still remains a public health threat, especially in developing countries. The World Health Organization (WHO) estimates that globally over 22 million people have already lost their lives dues to AIDS, and in 2012 there were 35.3 million people living with HIV, 2.3 million new infections and 1.6 million AIDS-related deaths [80].

Once a person becomes infected, the WHO defines five clinical stages as the infection progresses [78]. The acute stage is in the first few months following the initial introduction of the virus into the body. This stage is asymptomatic with no significant immunosuppression (CD4 count $>500\frac{\text{cells}}{\text{mm}^3}$) and a spike in virus titre. Stage I is also asymptomatic with no significant immunosuppression and low viral load. Stage II is symptomatic: the infected person exhibits moderate weight loss amongst other symptoms and mild immunosuppression (CD4 count is 350–$499\frac{\text{cells}}{\text{mm}^3}$). Stage III is also symptomatic and is characterized by advanced immunosuppression (CD4 count is 200–$349\frac{\text{cells}}{\text{mm}^3}$). Stage IV is the AIDS phase, where the infected person exhibits HIV wasting syndrome, amongst other symptoms, and severe immunosuppression (CD4 $<200\frac{\text{cells}}{\text{mm}^3}$). This classification based on clinical symptoms is advantageous because it allows for areas with limited laboratory facilities to estimate HIV progression and aid in the management of HIV/AIDS patients, e.g. when to start treatment.

In terms of transmissibility, the acute stage is characterized by high rates of transmission because of high viral loads. The infected person then goes through a phase (Stages I & II) where the viral load is lower and thus transmission is lower. Without treatment this person moves to Stages III & IV, which is characterized by high transmission rates due to

high viral loads [5, 44, 46]. A study of transmission involving a cohort of stable partnerships between heterosexuals in Rakai, Uganda quantified the relative transmissibility of HIV by stage of infection [76] . The probability of transmission per coital act in the acute stage was estimated to be 8–10 times higher than during asymptomatic (Stages I & II). In the last the 2 years before death (Stages III & IV) the probability of transmission per coital act was estimated to be 4–8 times higher than during asymptomatic infection.

Antiretroviral therapy (ART) are drugs that are used to target the HIV life cycle with the aim of halting HIV replication and restoring immune function, thus slowing the progression from HIV to AIDS [10, 73]. Apart from traditional role of preventing progression from HIV to AIDS, ART has the additional benefit of substantially reducing the infectiousness of infected people leading to reduced transmission [4, 7]. In 2011, the HIV Prevention Trials Network (HPTN) reported in their HPTN 052 trial that early ART reduces HIV transmission amongst serodiscordant couples by 96% [18]. Thanks to the HPTN 052 trial and other studies on the benefits of treatment, the WHO in June 2013 released new guidelines on the use of ART for treating and preventing HIV infection, recommending treatment to infected people with CD4 count of $<500\frac{\text{cells}}{\text{mm}^3}$, i.e. from Stage II, thus broadening the spectrum of people eligible for initiation of ART [54].

Health authorities worldwide are faced with limited resources and must find economical ways to administer ART. In this study, we used optimal control theory to determine time-dependent treatment strategies that maximize the effectiveness of population-scale interventions. We measured the effectiveness by total infection-years, new infections, AIDS-related deaths and cost, and separately found the strategies that optimize each of them.

Optimal control theory, which was developed by Pontryagin and his co-workers in the late 1950s [57], has been applied to many areas including economics, management,

engineering, biology, physiology and medicine [2, 37, 40, 42, 43, 86]. Indeed, optimal control theory has been used to study the treatment of individual HIV patients [32, 52].

In this paper, we first introduce our model that captures the 5 stages of HIV/AIDS infection and incorporates three controls for treatment in Stages II, III & IV. We then define the objective functions and the optimal control problem and follow that by an analysis of the optimal controls. Finally, we present some numerical solutions for the South African HIV epidemic and discuss the results.

4.2. Model and Approach

We developed an HIV model by adopting the model originally developed to study the importance of promoting HIV testing for preventing secondary transmission (Figure 4.1) [83]. The model is for an adult heterosexual population and stratifies the population by HIV status, diagnosis and treatment. People are either susceptible or infected, and then the infected population is divided into five classes based on the WHO HIV/AIDS Staging System. Each of the five stages is further divided in three levels: those who are infected but unaware of their HIV status (Undiagnosed), those who have been diagnosed but are not yet on treatment (Diagnosed) and the those on treatment (Treated). To reflect current WHO guidelines on treatment, a fraction of diagnosed people are on treatment in Stages II, III & IV.

People enter the model as susceptible (S) with a recruitment rate b and leave either through natural death (μ) or death due to AIDS or HIV related symptoms (γ_4). Once infected, people move from the susceptible class to the undiagnosed acute infection class (I_{Ua}). From here, people get tested at rate d and move into the diagnosed acute infection class (I_{Da}). Likewise in the other stages, undiagnosed people are tested at rates d. People

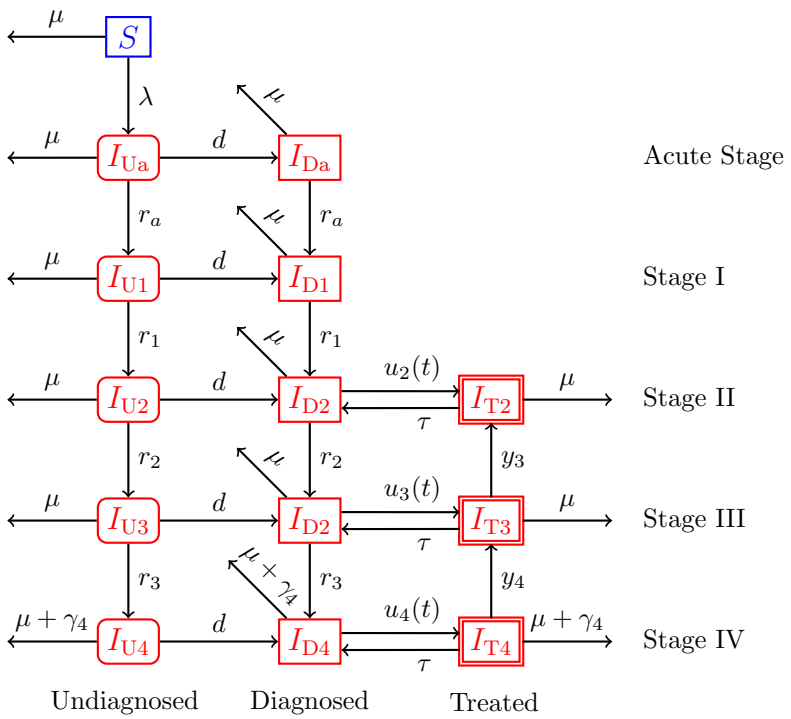

Figure 4.1: Diagram of 5-stage HIV/AIDS infection model. HIV-infected people are in red and susceptible people (S) people in blue. The first subscript on the state variables I denotes diagnosis or treatment status (undiagnosed, diagnosed or treated) and the second denotes stage of infection (acute Stage, Stage I, Stage II, Stage III or Stage IV).

Parameter	Description	Value	Source
b	Birth rate	0.0309 y^{-1}	[67]
μ	Natural death rate	0.0244 y^{-1}	[67]
γ_4	Disease induced death rate	0.9091 y^{-1}	[71]
α	Efficacy of treatment at reducing transmission	0.960	[18]
β_a	HIV Acquisition Risk in Acute Stage	0.6560 y^{-1}	[76]
β_1	HIV Acquisition Risk in Stage I	0.0960 y^{-1}	[76]
β_2	HIV Acquisition Risk in Stage II	0.6540 y^{-1}	[76]
β_3	HIV Acquisition Risk in Stage III	0.2480 y^{-1}	[76]
C_{I}	Cost of an infection	\$1,000 y^{-1}	[15]
C_{D}	Cost of a death	\$100,000	—
C_{T}	Cost of treatment	\$120 y^{-1}	[49]
r	Discount rate for costs	3% y^{-1}	[28]
r_a	Rate of Progression from Acute Stage to Stage I	4.8 y^{-1}	[76]
r_1	Rate of Progression from Stage I to Stage II	0.3235 y^{-1}	[76]
r_2	Rate of Progression from Stage II to Stage III	0.6667 y^{-1}	[76]
r_3	Rate of Progression from Stage III to Stage IV	0.1538 y^{-1}	[76]
y_3	Recovery Rate from Stage III Stage II	1 y^{-1}	—
y_4	Recovery Rate from Stage IV Stage III	1 y^{-1}	—
d	Testing Rates in Acute Stage and Stages I–IV	0.3333 y^{-1}	[35]
τ	Treatment Failure Rates in Stages II–IV	0.2 y^{-1}	[65]

Table 4.1: Model parameters.

in the untreated ($I_{\mathrm{U}i}$) and diagnosed ($I_{\mathrm{D}i}$) classes progress at rate r_i into the next infection stage, for $i = \mathrm{a}, 1, 2, 3$.

A proportion of the diagnosed people in Stages II, III & IV begin treatment at rates $u_2(t)$, $u_3(t)$ and $u_4(t)$, moving into I_{T2}, I_{T3} and I_{T4}, respectively. The treatment rates $u_i(t)$ are functions of the controls $U_i(t)$, which will be described below. The controls ($U_i(t)$) are assumed to be bounded and Lebesgue integrable. Treated people in stage 2 (I_{T2}) stay in that class because their immune system does not deteriorate. Treated people in stage 3 (I_{T3}) transition back to stage 2 (I_{T2}) at rate y_3 due to improvement of their immune system. Likewise, treated people in stage 4 (I_{T4}) transition back to stage 3 (I_{T3}) at rate y_4.

New infections occur from effective contact between susceptible people and infected people. People in stage 4 (I_{U4}, I_{D4} and I_{T4}) have full-blown AIDS and so we assumed that they are too ill to engage in sexual activity. Treatment reduces the probability of transmission by $\alpha = 96\%$ [18]. The force of infection is

$$\lambda = \lambda_a + \lambda_1 + \lambda_2 + \lambda_3, \tag{4.1}$$

with

$$\begin{aligned}
\lambda_{\mathrm{a}} &= \frac{\beta_a}{N}\left[I_{\mathrm{Ua}} + I_{\mathrm{Da}}\right], \\
\lambda_1 &= \frac{\beta_1}{N}\left[I_{\mathrm{U1}} + I_{\mathrm{D1}}\right], \\
\lambda_2 &= \frac{\beta_2}{N}\left[I_{\mathrm{U2}} + I_{\mathrm{D2}} + (1-\alpha)I_{\mathrm{T2}}\right], \\
\lambda_3 &= \frac{\beta_3}{N}\left[I_{\mathrm{U3}} + I_{\mathrm{D3}} + (1-\alpha)I_{\mathrm{T3}}\right],
\end{aligned} \tag{4.2}$$

and

$$N = S + \sum_{i\in\{a,1,2,3,4\}} I_{\mathrm{Ui}} + \sum_{i\in\{a,1,2,3,4\}} I_{\mathrm{Di}} + \sum_{i\in\{2,3,4\}} I_{\mathrm{Ti}}. \tag{4.3}$$

The HIV model is given by the system of differential equations

$$
\begin{aligned}
\frac{\mathrm{d}S}{dt} &= bN - \mu S - \lambda S, \\
\frac{\mathrm{d}I_{\mathrm{Ua}}}{dt} &= \lambda S - (d + r_{\mathrm{a}} + \mu) I_{\mathrm{Ua}}, \\
\frac{\mathrm{d}I_{\mathrm{U1}}}{dt} &= r_{\mathrm{a}} I_{\mathrm{Ua}} - (d + r_1 + \mu) I_{\mathrm{U1}}, \\
\frac{\mathrm{d}I_{\mathrm{U2}}}{dt} &= r_1 I_{\mathrm{U1}} - (d + r_2 + \mu) I_{\mathrm{U2}}, \\
\frac{\mathrm{d}I_{\mathrm{U3}}}{dt} &= r_2 I_{\mathrm{U2}} - (d + r_3 + \mu + \gamma_3) I_{\mathrm{U3}}, \\
\frac{\mathrm{d}I_{\mathrm{U4}}}{dt} &= r_3 I_{\mathrm{U3}} - (d + \mu + \gamma_4) I_{\mathrm{U4}}, \\
\frac{\mathrm{d}I_{\mathrm{Da}}}{dt} &= d I_{\mathrm{Ua}} - (r_{\mathrm{a}} + \mu) I_{\mathrm{Da}}, \\
\frac{\mathrm{d}I_{\mathrm{D1}}}{dt} &= r_{\mathrm{a}} I_{\mathrm{Da}} + d I_{\mathrm{U1}} - (r_1 + \mu) I_{\mathrm{D1}}, \\
\frac{\mathrm{d}I_{\mathrm{D2}}}{dt} &= r_1 I_{\mathrm{D1}} + d I_{\mathrm{U2}} + \tau I_{\mathrm{T2}} - (u_2(t) + r_2 + \mu)\, I_{\mathrm{D2}}, \\
\frac{\mathrm{d}I_{\mathrm{D3}}}{dt} &= r_2 I_{\mathrm{D2}} + d I_{\mathrm{U3}} + \tau I_{\mathrm{T3}} - (u_3(t) + r_3 + \mu + \gamma_3)\, I_{\mathrm{D3}}, \\
\frac{\mathrm{d}I_{\mathrm{D4}}}{dt} &= r_3 I_{\mathrm{D3}} + d I_{\mathrm{U4}} + \tau I_{\mathrm{T4}} - (u_4(t) + \mu + \gamma_4) I_{\mathrm{D4}}, \\
\frac{\mathrm{d}I_{\mathrm{T2}}}{dt} &= u_2(t) I_{\mathrm{D2}} + y_3 I_{\mathrm{T3}} - (\tau + \mu) I_{\mathrm{T2}}, \\
\frac{\mathrm{d}I_{\mathrm{T3}}}{dt} &= u_3(t) I_{\mathrm{D3}} + y_4 I_{\mathrm{T4}} - (\tau + y_3 + \mu) I_{\mathrm{T3}}, \\
\frac{\mathrm{d}I_{\mathrm{T4}}}{dt} &= u_4(t) I_{\mathrm{D4}} - (\tau + y_4 + \mu + \gamma_4) I_{\mathrm{T4}}.
\end{aligned}
\tag{4.4}
$$

4.2.1 Optimal Control Problem Formulation

The control variables U_2, U_3, U_4 are the total number of people targeted to be on treatment in each of the designated Stages II, III & IV at any one time. From these treatment targets, the flows of people into treatment per unit time were taken to be

$$
u_i(t) = r_{\max} \max\left(U_i(t) - I_{Ti},\ 0\right) \quad \text{for } i = 2, 3, 4, \tag{4.5}
$$

where I_{Ti} is the total number of people currently on treatment in Stage i, $i \in \{\mathrm{II}, \mathrm{III}, \mathrm{IV}\}$. $r_{\max}$ is the rate of enrolling people on treatment per drug available per untreated person with units $[\text{time}]^{-1}[\text{drugs}]^{-1}$. If ν is the total number of drugs available, then

$$U_2(t) + U_3(t) + U_4(t) \leq \nu. \tag{4.6}$$

We seek to minimize four objectives:

Total infection-years The total infection-years is the sum of the number of all the infected people (undiagnosed, diagnosed, or treated) in all stages at each time, integrated over the time period,

$$J_{\mathrm{I}}(X, u) = \int_0^T N_I \, \mathrm{d}t, \tag{4.7}$$

where (N_{I}) is the sum of all the infected people (both treated and untreated) in the five stages of infection:

$$N_{\mathrm{I}} = I_{\mathrm{Ua}} + I_{\mathrm{U1}} + I_{\mathrm{U2}} + I_{\mathrm{U3}} + I_{\mathrm{U4}} + I_{\mathrm{Da}} + I_{\mathrm{D1}} + I_{\mathrm{D2}} + I_{\mathrm{D3}} + I_{\mathrm{D4}} + I_{\mathrm{T2}} + I_{\mathrm{T3}}. \tag{4.8}$$

New infections The total number of new infections is the sum of the rates of new infections arising contact of the susceptible people with infected people integrated over the time period,

$$J_{\mathrm{NI}}(X, u) = \int_0^T N_{\mathrm{NI}} \, \mathrm{d}t, \tag{4.9}$$

where (N_{NI}) is the rate of the new infection arising from contacts of uninfected people with infected people,

$$N_{NI} = \lambda S. \tag{4.10}$$

Deaths due to AIDS The total AIDS-related deaths (N_{D}) is the sum of infected people dying from AIDS-related diseases. Our assumption that AIDS-related deaths only occur in stage IV means that this is the sum of infected people in Stage IV dying from AIDS,

$$J_{\mathrm{D}}(X, u) = \int_0^T N_{\mathrm{D}} \, \mathrm{d}t. \tag{4.11}$$

N_{D} is simply the number of people infected people in stage IV at each time multiplied by the disease induced deaths rate γ_4,

$$N_{\mathrm{D}} = \gamma_4 (I_{\mathrm{U}4} + I_{\mathrm{D}4} + I_{\mathrm{T}4}). \tag{4.12}$$

Total cost The total cost (N_{C}) consists of cost of infections, cost of deaths, and cost of treatment. The disease cost per year, which includes monetary equivalent loss of the infected people, like lost productivity etc., is average cost of disease per person per year times the total number of people infected,

$$C_{\mathrm{I}}(t) = c_{\mathrm{I}} N_I(t). \tag{4.13}$$

The cost per year of deaths is the the cost per death times the number of deaths per year,

$$C_{\mathrm{D}}(t) = c_{\mathrm{D}} \gamma N_D(t). \tag{4.14}$$

The treatment cost per year is the cost per person per year times the number of people treated,

$$C_{\mathrm{T}}(t) = c_{\mathrm{T}} N_T(t), \tag{4.15}$$

where the number of people treated (N_T) is the sum of all people in the treatment class,

$$N_{\mathrm{T}} = I_{\mathrm{T}2} + I_{\mathrm{T}3} + I_{\mathrm{T}4}. \tag{4.16}$$

The cost objective is discounted sum of these costs, integrated over the time period:

$$J_{\mathrm{C}}(X,u) = \int_0^T \left[C_{\mathrm{I}}(t) + C_{\mathrm{D}}(t) + C_{\mathrm{T}}(t)\right]\mathrm{e}^{-rt}\mathrm{d}t. \tag{4.17}$$

The total cost is discounted at rate r, representing the rate a policymaker is willing to pay as trade-off for the cost today versus future cost [24].

For ease of analysis, we can define system 4.4 compactly as

$$\dot{X} = g\left(t, X, U\right), \tag{4.18}$$

with $X = (S, I_{\mathrm{Ua}}, I_{\mathrm{U1}}, I_{\mathrm{U2}}, I_{\mathrm{U3}}, I_{\mathrm{U4}}, I_{\mathrm{Da}}, I_{\mathrm{D1}}, I_{\mathrm{D2}}, I_{\mathrm{D3}}, I_{\mathrm{D4}}, I_{\mathrm{T2}}, I_{\mathrm{T3}}, I_{\mathrm{T4}})$.

The optimal control problem is

$$\begin{cases} \text{Minimize} & J_k(X,U) \quad \text{for one of } k \in \{\mathrm{I}, \mathrm{NI}, \mathrm{D}, \mathrm{C}\}, \\ \text{subject to} & \dot{X} = g(t,X,U), \\ & X(0) = X_0, \\ & U_j \geq 0 \qquad \text{for every } j \in \{2,3,4\}, \\ & U_2(t) + U_3(t) + U_4(t) \leq \nu. \end{cases} \tag{4.19}$$

4.2.2 Analysis of Optimal Controls

If we define the integrand of our objective function by $f_k(t,X)$ and $g(t,X,U)$ is the right-hand side of the system of differential equations, then the Hamiltonian is

$$\mathcal{H}(t,X,U,\theta) = f_k(t,X) + \theta^T g(t,X,U). \tag{4.20}$$

Pontryagin's Maximum Principle [57] converts the optimal control problem into a problem of minimizing the Hamiltonian point-wise with respect to U_2, U_3 and U_4. See Appendix B for a detailed characterization of the optimal control problem.

We can characterize the optimal controls as

$$\begin{aligned}\frac{\partial \mathcal{H}}{\partial U_2} &= r_{\max} I_{\mathrm{D2}}(\theta_{12} - \theta_9) H\left(U_2 - I_{\mathrm{T2}}\right) = 0,\\ \frac{\partial \mathcal{H}}{\partial U_3} &= r_{\max} I_{\mathrm{D3}}(\theta_{13} - \theta_{10}) H\left(U_3 - I_{\mathrm{T3}}\right) = 0,\\ \frac{\partial \mathcal{H}}{\partial U_4} &= r_{\max} I_{\mathrm{D4}}(\theta_{14} - \theta_{11}) H\left(U_4 - I_{\mathrm{T4}}\right) = 0,\end{aligned} \tag{4.21}$$

where $H(x)$ is the Heaviside function,

$$H(x) = \begin{cases} 0 & \text{if } x < 0,\\ 1 & \text{if } x > 0.\end{cases} \tag{4.22}$$

The above optimal controls are bounded by the number of total drugs available (i.e. $0 \leq U_i \leq \nu_i$) where ν_2, ν_3 and ν_4 are the total number of drugs allocated to U_2, U_3 and U_4 respectively at each time point. From (4.6)

$$\nu_2 + \nu_3 + \nu_4 \leq \nu \tag{4.23}$$

Applying these bounds to the controls we obtain

$$U_2^*(t) = \begin{cases} 0 & \text{if} \quad r_{\max} I_{\mathrm{D2}}(\theta_{12} - \theta_9) H\left(U_2 - I_{\mathrm{T2}}\right) < 0,\\ \nu_2 & \text{if} \quad r_{\max} I_{\mathrm{D2}}(\theta_{12} - \theta_9) H\left(U_2 - I_{\mathrm{T2}}\right) > 0,\end{cases} \tag{4.24}$$

$$U_3^*(t) = \begin{cases} 0 & \text{if} \quad r_{\max} I_{\mathrm{D3}}(\theta_{13} - \theta_{10}) H\left(U_3 - I_{\mathrm{T3}}\right) < 0,\\ \nu_3 & \text{if} \quad r_{\max} I_{\mathrm{D3}}(\theta_{13} - \theta_{10}) H\left(U_3 - I_{\mathrm{T3}}\right) > 0,\end{cases} \tag{4.25}$$

and

$$U_4^*(t) = \begin{cases} 0 & \text{if} \quad r_{\max} I_{\mathrm{D4}}(\theta_{14} - \theta_{11}) H\left(U_4 - I_{\mathrm{T4}}\right) < 0,\\ \nu_4 & \text{if} \quad r_{\max} I_{\mathrm{D4}}(\theta_{14} - \theta_{11}) H\left(U_4 - I_{\mathrm{T4}}\right) > 0.\end{cases} \tag{4.26}$$

Note that in the numerical results, the singular case (when $\frac{\partial \mathcal{H}}{\partial U_i} = 0$ on a set of positive measure) does not occur and we can restrict our attention to the bang-bang controls

above. We also expect that at all time points $[0, T]$ either $U_2^*(t) + U_3^*(t) + U_4^*(t) = 0$ or $U_2^*(t) + U_3^*(t) + U_4^*(t) = \nu$.

As result of the convexity of the integrand of J with respect to U_2, U_3 and U_4; the priori boundedness and the Lipschitz property with respect to the state variables, the existence of optimal controls with the constraints $U_2 + U_3 + U_4 \leq \nu$ follows by results using minimizing sequences which converge weakly in $L^2(0, T)$ to an optimal triple (which is an extension of the result in [14, 26].

4.2.3 Numerical Simulations

With an initial guess for the control variables U_2, U_3 and U_4, we solve the state equations (4.4) forward in time. Using the solutions of the state equations together with the transversality conditions (A2) we solve the adjoint equations (A14) backward in time. The control is updated after each iteration using the new values of state and adjoint variables into the optimality conditions (4.21), and the process is repeated. Iteration is stopped when the difference between successive iteration meet a predetermined tolerance.

We implemented a numerical algorithm originally developed in [75]. The steps of algorithm are

1. Divide the time $[0, T]$ into W subintervals.

2. We start with an initial guess of the controls U_2^0, U_3^0 and U_4^0

3. Obtain the state variables X_i by integrating the state equation (4.4) from 0 to T using the controls U_2^i, U_3^i and U_4^i and initial condition $X^i(t_0) = X_0^i$.

4. Integrate the adjoint equations (A14) backward in time (from T to 0) to obtain the adjoint variables θ^i, using X^i, U_2^i, U_3^i and U_4^i

5. Stop the algorithm if a $\|P_{i+1} - P_i\| \leq \epsilon_{\text{abs}} + \epsilon_{\text{rel}}\|P_i\|$ for $P \in \{X, U, \theta\}$, where ϵ_{abs} and ϵ_{rel} are predetermined absolute and relative errors respectively.

6. If step 5 is not satisfied, adjust the control functions, by replacing U^i with U^{i+1} and return to step 3. U^{i+1} is calculated by the Newton–Raphson method: $U^{i+1}(t_k) = U^i(t_k) - \Delta \frac{\partial \mathcal{H}^i}{\partial U}(t_k)$, $k = 0, 1, 2, \ldots, W-1$ and Δ being the step size ($\frac{\partial \mathcal{H}}{\partial U}$ is from 4.21). We developed a line search method to find the step size Δ that minimizes J for each iteration.

7. Return to step 3

4.2.4 Parameterization

Our model is initialized for the beginning of the year 2014 ($t = 0$), although we used demographic data from 2012, the date of the last South African National HIV Prevalence, Incidence and Behavior Survey [65]. We parameterized our model with demographic data from South Africa because of the availability of excellent demographic statistics. If N_0 is initial total adult (ages 15+) population size, then using prevalence (ϕ) we determined the infected and the susceptible populations, i.e.

$$N_{\text{I}}(0) = \phi N_0, \qquad S(0) = N_0 - N_{\text{I}}(0). \tag{4.27}$$

The initial total infected population in each stage was determined by proportion of time a person spends in that Stage. If N_{Ia}, N_{I1}, N_{I2}, N_{I3} and N_{I4} represent the total population in the acute, I, II, II and IV stages respectively, then

$$N_{Ij}(0) = \frac{\frac{1}{r_j}}{\frac{1}{r_{\text{a}}} + \frac{1}{r_1} + \frac{1}{r_2} + \frac{1}{r_3}} N_I(0) \qquad \text{for} \quad j \in \{\text{a}, 1, ..., 3\}. \tag{4.28}$$

The acute stage is very short which means generally it is not enough time for an infected person to be diagnosed, so we assumed that the initial diagnosed population in

Variable	**Value**
Variable	Initial Value
S	30700000
I_{Ua}	27283
I_{U1}	94590
I_{U2}	704976
I_{U3}	588584
I_{U4}	53968
I_{Da}	0
I_{D1}	310284
I_{D2}	950229
I_{D3}	793345
I_{D4}	72743
I_{T2}	1362315
I_{T3}	1137396
I_{T4}	104289

Table 4.2: Initial conditions.

the acute stage is zero. For the remaining stages, 76.3% of the infected population were considered diagnosed [35] and of the diagnosed, 42% are on treatment [71]. The initial conditions are summarized in Table 4.2.

4.3. Results

With an initial drug availability of 6.4 million drugs per year, which is enough drugs to treat all of the initial infected population, we simulated our model to determine the levels of treatment that minimize each of the four objectives and to determine the effect of applying the optimal treatment strategy over a 10-year period. The optimal strategies that minimize infection-years and new infections are similar to each other, and the optimal strategies that minimize death and cost are similar to each other.

To minimize infection-years, the optimal strategy is to start off by limiting treatment to only infected people in Stages II and III, with a preference for Stage II (Figure 4.2A). After 4.4 years, some infected people in Stages IV should also be be treated. The optimal strategy prescribes an initial scale-up of treatment in Stage II and III from 42% to about 70% and 90% respectively and a corresponding decrease in treatment in Stage IV from 42% to 0%. After the initial scale-up, a steady decrease in the proportion of people treated in Stage II corresponding with a steady increase in treatment of Stage III is observed (Figure 4.2B). Treatment of people in Stage IV begins after 5.5 years, increasing rapidly to 45% by year 7 and staying at 45% after that. Total treatment coverage also increases rapidly from the current coverage of 42% to about 70% and is maintained at about 70% throughout the period.

To optimal strategy to minimize new infections is similar to that which minimizes infection-years above. Initially treatment is allocated to only infected people in Stages II and III, however more or the treatment is allocated to Stage III instead of Stage II (Figure 4.3B). Treatment for Stage IV starts after 9 years as opposed to after 5.5 years when minimizing infection-years.

The optimal treatment strategies to minimize deaths and cost is to administer late

treatment (i.e treatment to Stages III & IV) with treatment in Stage III being the most favored(Figures 4.4A & 4.5A). An initial scale-up in proportions of people on treatment in all three stages is observed, followed by a decrease in Stage IV (Figures 4.4B & 4.5B). The initial decrease in the proportion of treatment in Stage IV when minimizing both infection-years and new infections is not observed here.

Over the 10-year period, all four optimal strategies resulted in lower prevalence and incidence than the current treatment strategy which is a fixed 42% treatment across all three stages at all times. Under the current treatment strategy, prevalence increases from from 16.8% to 21% and annual incidence from 400,000 to 600,000 in 10 years. The optimal treatment strategies that minimize deaths and cost reduced prevalence from 16.8% to 16.5% (Figure 4.6A) and annual incidence from 400,000 to about 200,000 (Figure 4.6B). The optimal treatment to minimize new infections does the most to reduce both prevalence and incidence, reducing prevalence from 16.8% to 15.3% and annual incidence from 400,000 to 100,000 in 10 years.

As expected, each of the optimal treatment strategies minimizes its own objective (Figure 4.7). Deaths are most impacted by the interventions: 80% of the deaths that would occur using the current treatment strategy can be averted by the optimal strategy. Infection-years are the least impacted: only 15% of the infection-years can be averted.

4.4. Discussion

In this paper, we considered the problem of optimal use of drugs by disease stage to maximize the impact of HIV on the population. Time-dependent optimal control strategies that minimize four objectives, new infection, infection-years, deaths and cost, were presented. Treatment in Stages II, III and IV, consistent with the WHO recommendations, were

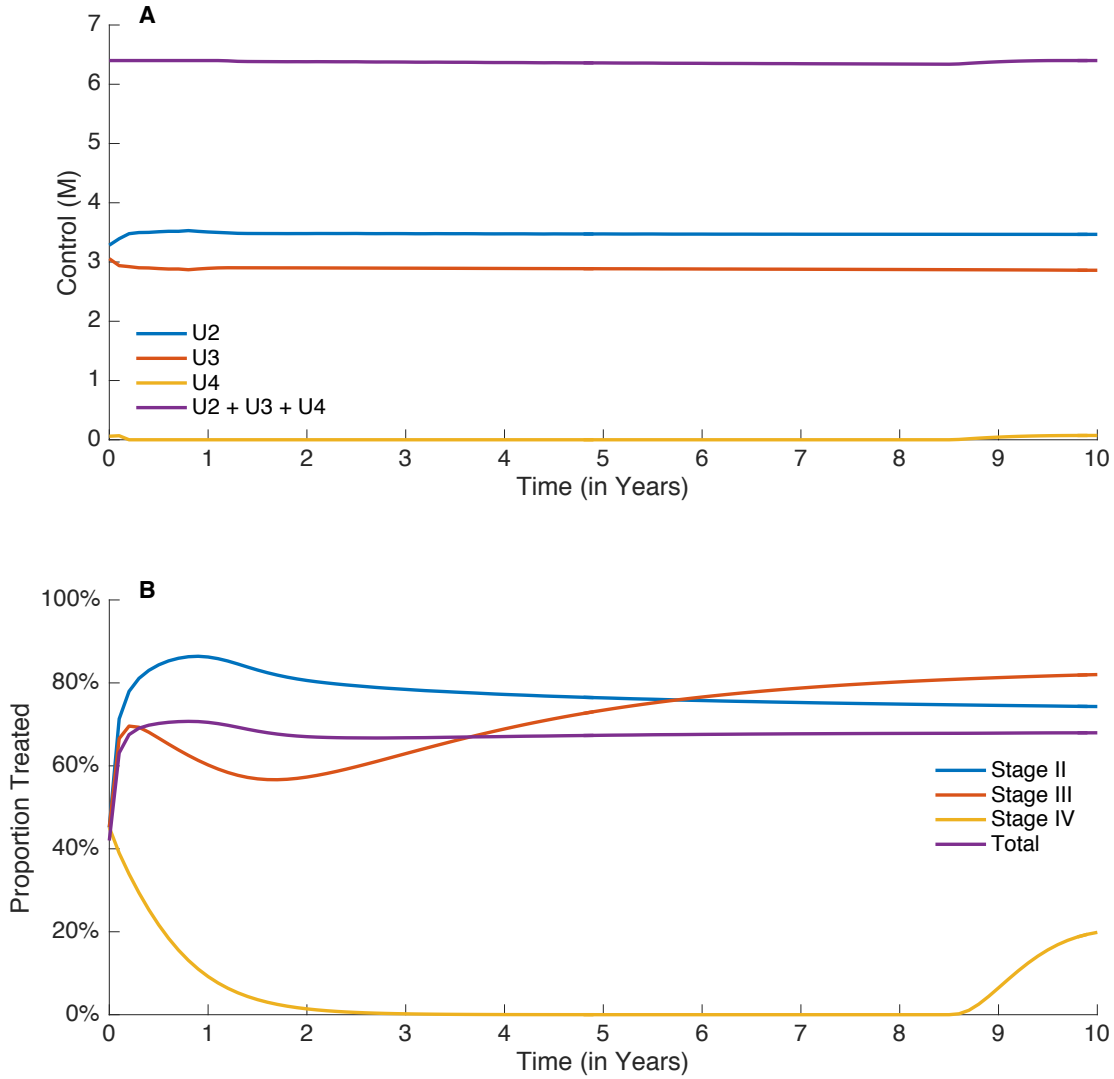

Figure 4.2: Optimal strategy to minimize infection-years: (A) number of drugs allocated to each stage vs time and (B) proportion of people treated in each stage vs time.

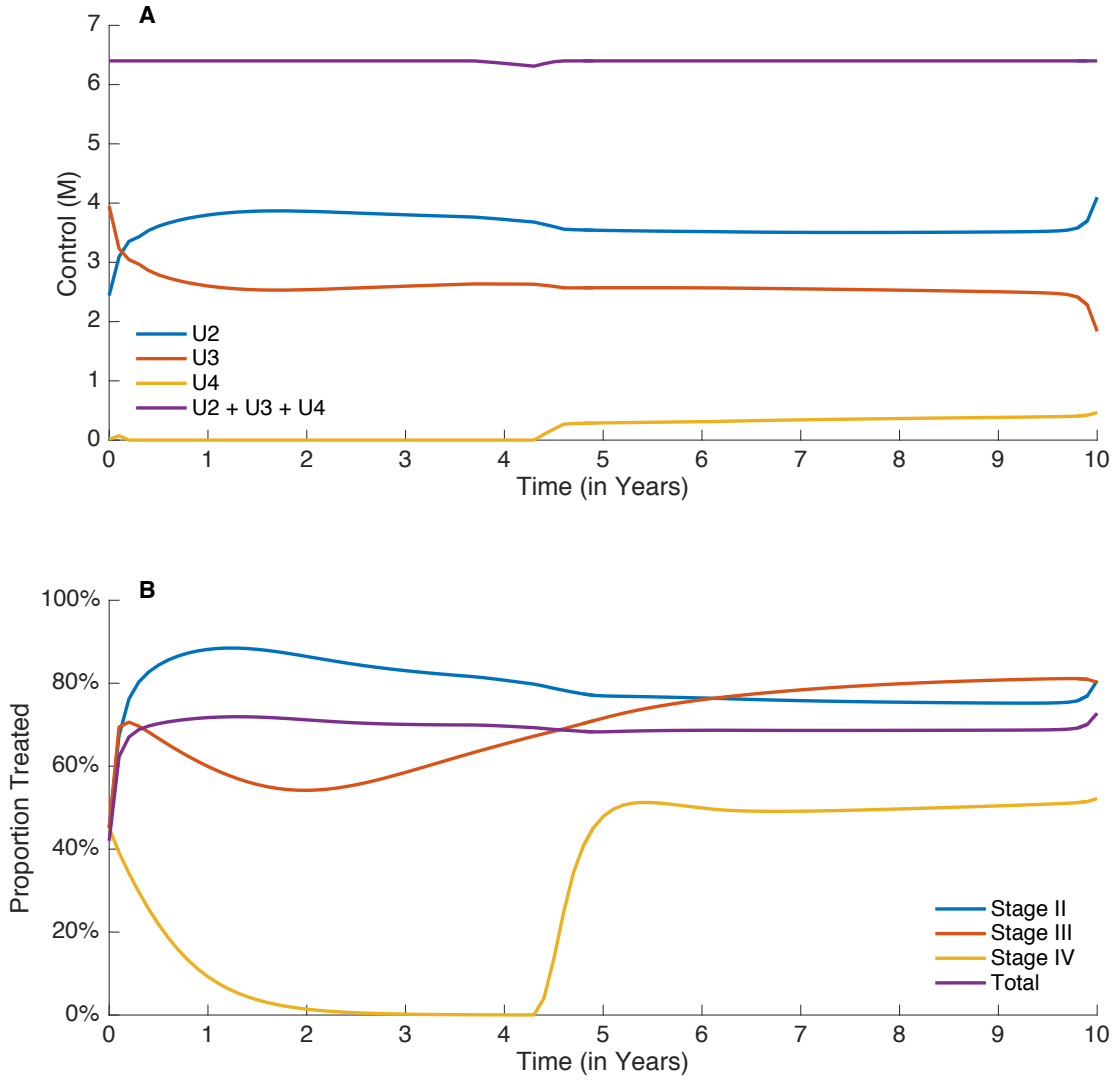

Figure 4.3: Optimal strategy to minimize new infections: (A) number of drugs allocated to each stage vs time and (B) proportion of people treated in each stage vs time.

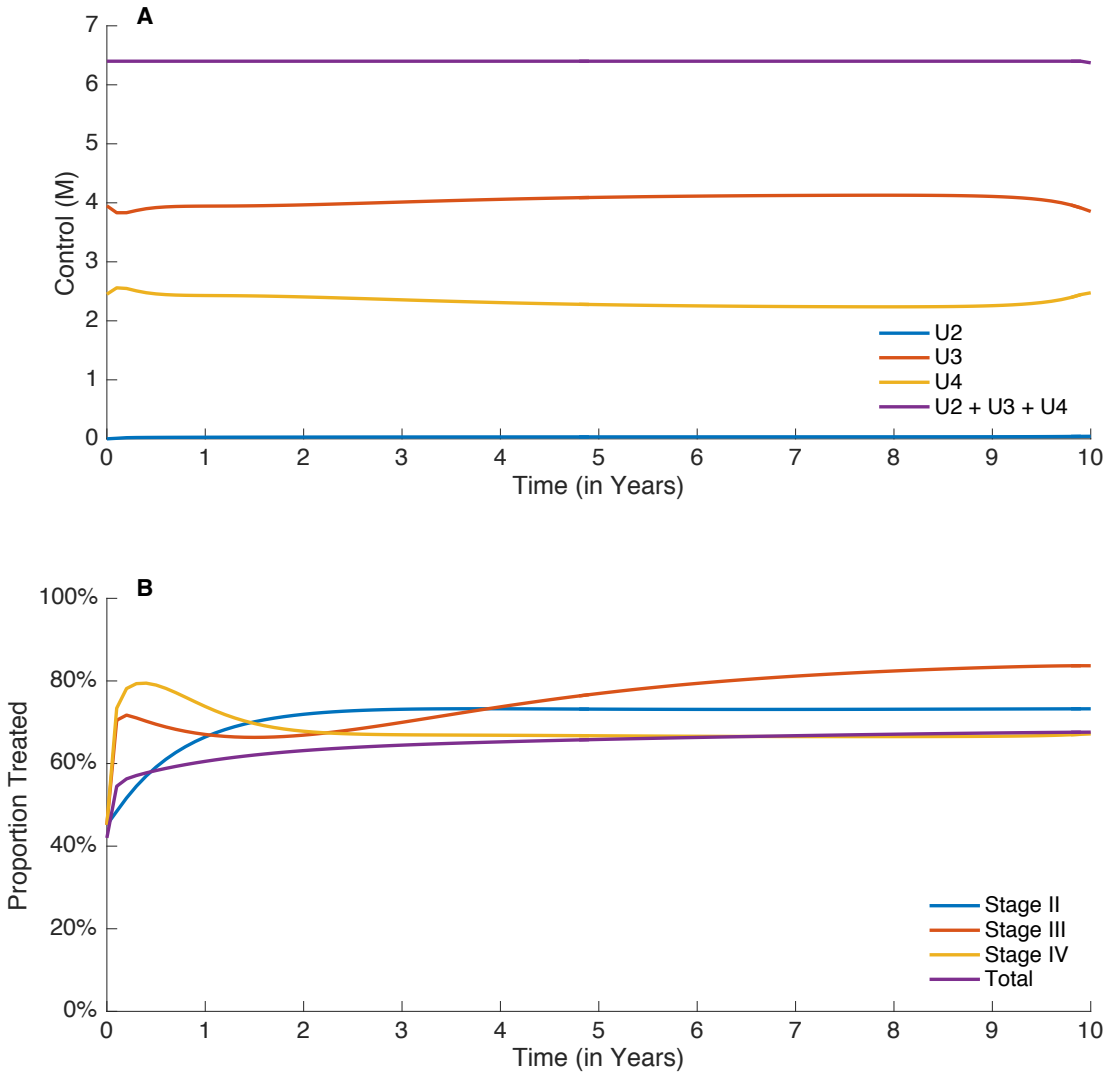

Figure 4.4: Optimal strategy to minimize AIDS-related deaths: (A) number of drugs allocated to each stage vs time and (B) proportion of people treated in each stage vs time.

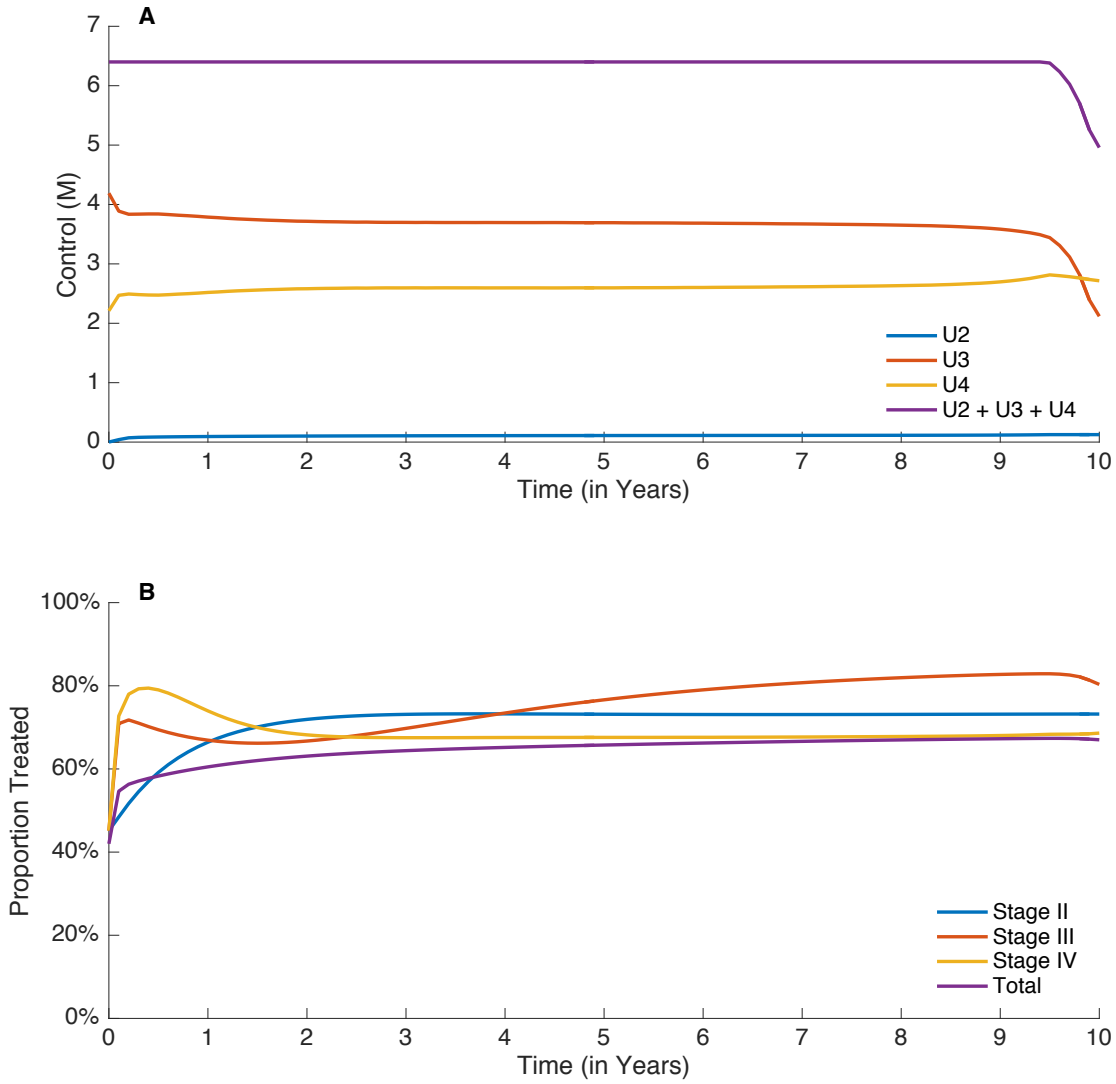

Figure 4.5: Optimal strategy to minimize cost: (A) number of drugs allocated to each stage vs time and (B) proportion of people treated in each stage vs time.

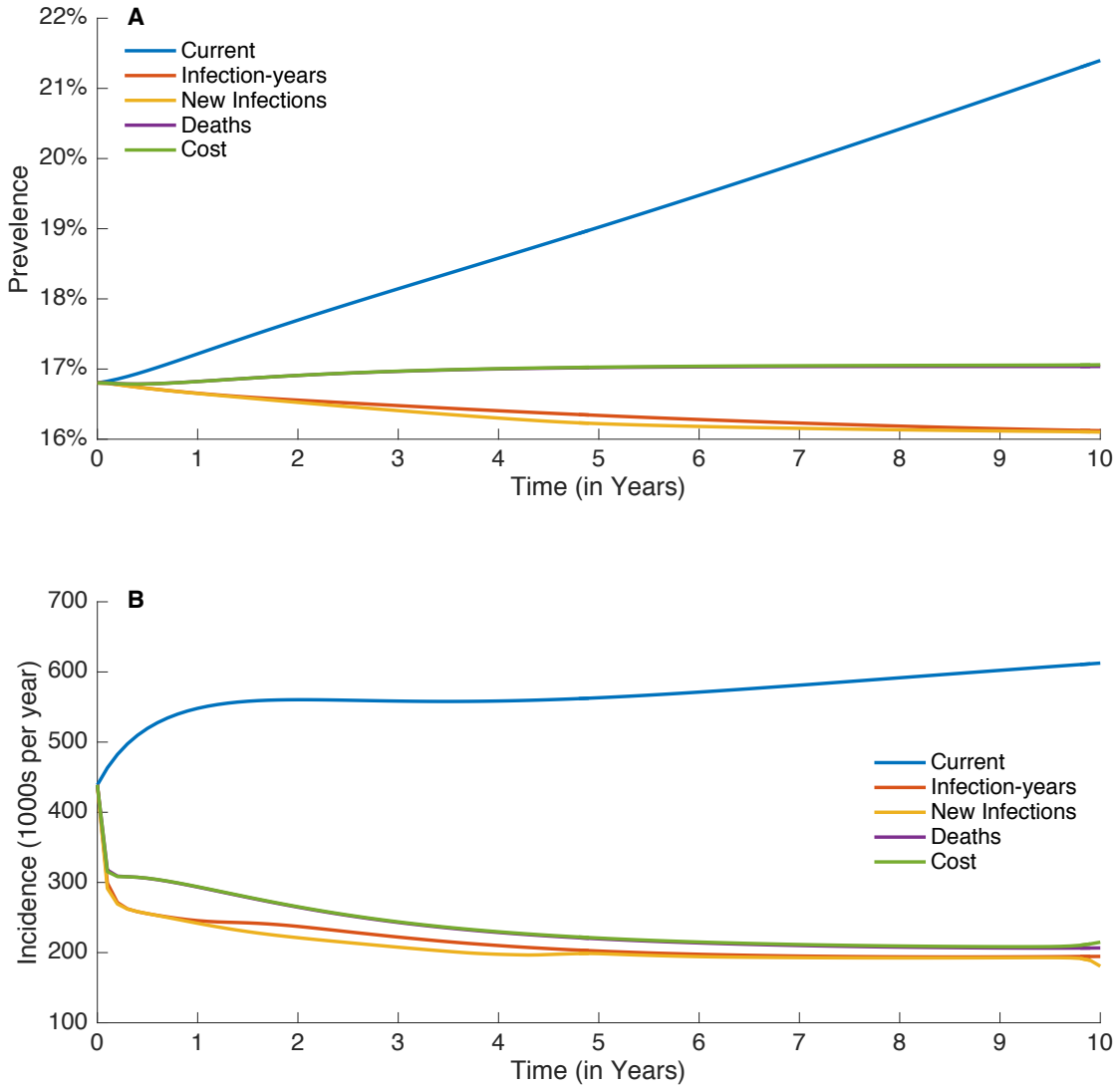

Figure 4.6: Impact on (A) prevalence and (B) incidence of the optimal treatment strategies.

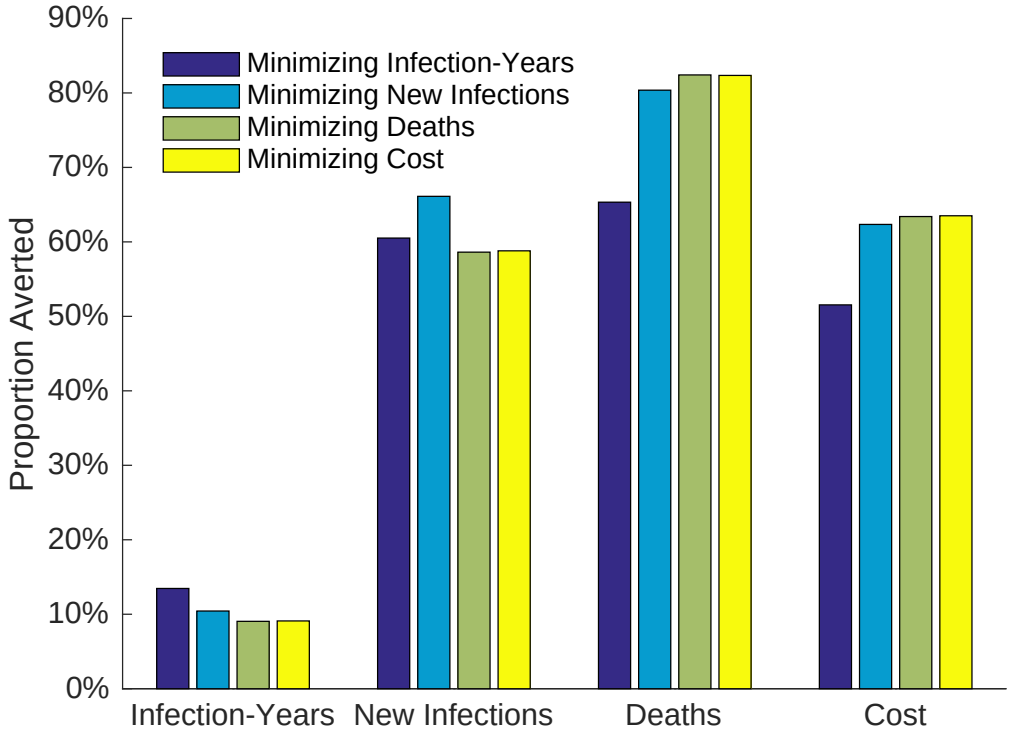

Figure 4.7: Relative benefits from applying the optimal strategies of each outcome for 6.4 million annual drug doses.

considered.

Our simulations indicated that to minimize infection-years and new infections, emphasis should be placed on treatment in the earlier Stages while to minimize cost and death, the emphasis should be on treating people in Stage III & IV. Our numerical simulations illustrate the effectiveness of adopting each treatment strategy to allow policy makers to learn how much saving they will gain.

The optimal treatment strategies to minimize minimize new infections does not provide treatment in Stage IV consistent with the assumption we made that people in Stage IV are too weak to engage in sexual activity. Infection-years is related to the number of new infection and HIV related deaths, which is why the optimal strategy for infection-years has a similar to that of new infection in the first 5.5 years and for treatment in Stage III after 5.5 years to keep them out of Stage III ad preventing an increase in stransmission. The optimal treatment strategies to minimize minimize deaths and cost also place emphasis on treatment in Stage III to prevent infected people from progressing to Stage IV where they die from AIDS. The similarity of treatment for death and cost is because the cost associated with deaths is very high relative to the other costs thus minimizing deaths is also minimizing cost.

Most optimal-control models often include an assumption for a quadratic cost of the controls in the objective function to simplify the solution process. Our model does not make such an assumption and we also impose a cap on the total drugs available each year akin to [8, 22, 51].

Consideration of the long time horizon for control is also important to give more weight in the objective function to earlier rather than later control. In order to account for this we discounted cost by a rate of 3%.

In addition to improving their health directly, treatment of infected people is know to

reduce their ability to transmit the virus to uninfected people. Policy makers, especially in limited-resource settings continually seek better ways of harnessing the benefits of treatment. We hope that the results of this modeling study, despite its necessary simplification, can help to guide policy decisions.

5. GENERAL CONCLUSIONS

HIV treatment is increasingly becoming an important strategy in the fight against HIV/AIDS pandemic because of its well-document effect of reducing disnease morbidity and mortality. Health administrators and policy makers worldwide are therefore seeking ways of harnessing the benefits of HIV treatment. Given that health resources are limited, choices have to be made on the best way to allocate these resources to HIV treatment. Optimal allocation of HIV treatment resources is particularly significant for limited-resource settings that also have high HIV burden, as in much of Sub-Saharan Africa.

In this dissertation, I used different mathematical modeling techniques to study HIV treatment strategies. I developed two distinct models: the first is a heterosexual transmission model that classifies people by gender, risk and treatment status and the second model based on the WHO's five-stage classification of HIV/AIDS disease progression. In the first model, two types of treatment were considered, ART for infected people and PrEP as prophylaxis to uninfected people, while in the second model only ART was considered. Data from South Africa were used to calibrate both models.

Two distinct optimization in methods were deployed to find treatment strategies that minimizes four objective functions: infection-years, new infections, deaths and cost. The first is a non-linear optimization method, which was used to find optimal treatment strategies for people based on gender or risk-group that minimized each of the four objectives. The second method, optimal control, was used to find the time-dependent optimal treatment strategies based on HIV disease stage that minimizes each of the four objectives. One main difference in the two methods is that the optimal control method provides time-dependent optimal treatment allocations, while using optimization provides only optimal allocations

that are constant in time.

Based on the heterosexual transmission model, the results show that at very low drug availability, treating using only PrEP is optimal, while at high drug availably ART is preferred. Resource-limited countries are likely to have low treatment availability and so our results suggest that all treatment should be allocated to prevention, i.e. PrEP to the general population before adding on ART when more drugs become available. In terms of risk groups, PrEP and ART to high-risk people at low drug availability and a mostly all-ART strategy is preferred. The optimal treatment strategy when treatment is based on gender is to prioritize PrEP for uninfected females first, switching to ART for infected males as more drugs become available. Allocating ART to infected females and PrEP to uninfected males should only be an option at high drug availability.

As countries strive to achieve UNAIDS's ambitious 90–90–90 goal, this work can also be used to determine the best treatment strategies. For example, in South Africa, to implement the UNAIDS's 90–90–90 goal, with the global ART & PrEP control strategy, allocating all of the available treatment as ART and none to PrEP minimized all four objectives. These optimal allocations, all ART and no PrEP, are exactly the current strategy in South Africa. With the ART & high-risk PrEP strategy, allocating all of the available treatment to non-high-risk ART minimized new infections, infection-years, and deaths, while cost was minimized by allocating a number of treatment spots to high-risk ART and the remainder to non-high-risk ART.

In the second model, a combination of ART allocations to Stages II & III is the time-dependent optimal strategy to minimize infection-years and new infections. Very little treatment goes to Stage IV because of our assumption that people in Stage IV are too weak to contribute to transmission. To minimize death and cost however, allocation to a combination of all three stages is required. Treatment in Stage IV features predominantly

because only people in Stage IV can die from AIDS: since the cost of death is high, preventing people from dying will greatly reduce cost. Increasing the number of people treated in both models leads to reduction in all four objectives and applying the optimal strategy did result in a substantially decrease in disease prevalence and incidence.

These three Chapters have been put together to help answer the question “when”, “how” and “to whom” to optimally allocate treatment in order to minimize the burden of HIV. I hope results of this study will provides some guidance for policymakers in determining how to allocate treatment in order to maximize the benefit.

References

1. Q. Abdool Karim, S. S. Abdool Karim, J. A. Frohlich, A. C. Grobler, C. Baxter, L. E. Mansoor, A. B. M. Kharsany, S. Sibeko, K. P. Mlisana, and Z. et al. Omar. Effectiveness and safety of tenofovir gel, an antiretroviral microbicide, for the prevention of HIV infection in women. *Science*, 329(5996):1168–1174, 2010. doi: 10.1126/science.1193748.

2. Folashade B Agusto, Nizar Marcus, and Kazeem O Okosun. Application of optimal control to the epidemiology of malaria. *Electronic Journal of Differential Equations*, 2012(81):1–22, 2012.

3. Michel Alary and Catherine M Lowndes. The central role of clients of female sex workers in the dynamics of heterosexual HIV transmission in sub-Saharan Africa. *AIDS*, 18(6):945–947, 2004. doi: 10.1097/00002030-200404090-00013.

4. Michel Alary, Leonard Mukenge-Tshibaka, France Bernier, Nassirou Geraldo, Catherine M. Lowndes, Honorè Meda, Cyriaque A. B. Gnintoungbè, Severin Anagonou, and Jean R. Joly. Decline in the prevalence of HIV and sexually transmitted diseases among female sex workers in Cotonou, Benin, 1993–1999. *AIDS*, 16(3):463–470, 2002. doi: 10.1097/00002030-200202150-00019.

5. Roy M. Andersen and Robert M. May. Epidemiological parameters of {HI} V transmission. *Nature*, 333(6173):514–519, jun 1988. doi: 10.1038/333514a0. URL `http://dx.doi.org/10.1038/333514a0`.

6. Jared M. Baeten, Deborah Donnell, Patrick Ndase, Nelly R. Mugo, James D. Campbell, Jonathan Wangisi, Jordan W. Tappero, Elizabeth A. Bukusi, Craig R. Cohen, and Elly et al. Katabira. Antiretroviral prophylaxis for HIV prevention in heterosexual

men and women. *New England Journal of Medicine*, 367(5):399–410, 2012. doi: 10.1056/nejmoa1108524.

7. Richard E. Berger. Re: Effectiveness and Safety of Tenofovir Gel an Antiretroviral Microbicide, for the Prevention of {HIV} Infection in Women. *The Journal of Urology*, 185(5):1729, may 2011. doi: 10.1016/s0022-5347(11)60197-3. URL `http://dx.doi.org/10.1016/S0022-5347(11)60197-3`.

8. Md Haider Ali Biswas, Luís Tiago Paiva, and Maria Do Rosário De Pinho. A seir model for control of infectious diseases with constraints. *Mathematical Biosciences and Engineering*, 2013.

9. SM Blower, D Hartel, H Dowlatabadi, RM Anderson, and RM May. Drugs, sex and hiv: a mathematical model for new york city. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 331(1260):171–187, 1991.

10. Norbert Bráu, Mirella Salvatore, Carlos F. Ríos-Bedoya, Alberto Fernández-Carbia, Fiorenzo Paronetto, José F. Rodríguez-Orengo, and Maribel Rodríguez-Torres. Slower fibrosis progression in HIV/HCV-coinfected patients with successful HIV suppression using antiretroviral therapy. *Journal of Hepatology*, 44(1):47–55, 2006. doi: 10.1016/j.jhep.2005.07.006.

11. Center for Disease Control and Prevention. Cdc trial and another major study find PrEP can reduce risk of HIV infection among heterosexuals., Accessed 20 Oct. 2015. URL `http://www.cdc.gov/nchhstp/newsroom/PrEPHeterosexuals.html`.

12. Centers for Disease Control and Prevention. Preexposure prophylaxis for the prevention of HIV infection in The United States - 2014, Accessed 25 Sept 2015. URL `http://www.cdc.gov/HIV/pdf/guidelines/PrEPguidelines2014.pdf`.

13. Centers for Disease Control and Prevention. Preexposure prophylaxis for the prevention of HIV infection in The United States - 2014: Clinical providers' supplement, Accessed 25 Sept 2015. URL `http://www.cdc.gov/hiv/pdf/prepprovidersupplement2014.pdf`.

14. Tim Clayton, Scott Duke-Sylvester, Louis J Gross, Suzanne Lenhart, and Leslie A Real. Optimal control of a rabies epidemic model with a birth pulse. *Journal of biological dynamics*, 4(1):43–58, 2010.

15. Susan M Cleary, Di McIntyre, and Andrew M Boulle. The cost-effectiveness of antiretroviral treatment in Khayelitsha, South Africa–a primary data analysis. *Cost Effectiveness and Resource Allocation*, 4(1):20, 2006. doi: 10.1186/1478-7547-4-20.

16. Thomas J Coates, Peter Aggleton, Felix Gutzwiller, Don Des Jarlais, Masahiro Kihara, Susan Kippax, Martin Schechter, and J Anneke R van den Hoek. Hiv prevention in developed countries. *The Lancet*, 348(9035):1143–1148, 1996.

17. J. Cohen. HIV treatment as prevention. *Science*, 334(6063):1628–1628, 2011. doi: 10.1126/science.334.6063.1628.

18. Myron S. Cohen, Ying Q. Chen, Marybeth McCauley, Theresa Gamble, Mina C. Hosseinipour, Nagalingeswaran Kumarasamy, James G. Hakim, Johnstone Kumwenda, Beatriz Grinsztejn, and Jose H.S. et al. Pilotto. Prevention of HIV-1 infection with early antiretroviral therapy. *New England Journal of Medicine*, 365(6):493–505, 2011. doi: 10.1056/nejmoa1105243.

19. Edward M Connor, Rhoda S Sperling, Richard Gelber, Pavel Kiselev, Gwendolyn Scott, Mary Jo O'Sullivan, Russell VanDyke, Mohammed Bey, William Shearer, Robert L Jacobson, et al. Reduction of maternal-infant transmission of human immunodeficiency

virus type 1 with zidovudine treatment. *New England Journal of Medicine*, 331(18): 1173–1180, 1994.

20. Siobhan Crowley. Preventing HIV transmission with antiretrovirals. *Bulletin of the World Health Organization*, 87(7):488–488, 2009. doi: 10.2471/blt.09.067330.

21. Kevin M De Cock, Charles F Gilks, Ying-Ru Lo, and Teguest Guerma. Can antiretroviral therapy eliminate HIV transmission? *The Lancet*, 373(9657):7–9, 2009. doi: 10.1016/s0140-6736(08)61732-8.

22. Maria Do Rosário De Pinho, Igor Kornienko, and Helmut Maurer. Optimal control of a seir model with mixed constraints and l 1 cost. In *CONTROLO'2014–Proceedings of the 11th Portuguese Conference on Automatic Control*, pages 135–145. Springer, 2015.

23. Mahesh Devnani, Anil K Gupta, and Yan Guo. Global hiv funding and local contexts. *Health Affairs*, 34(2):359–359, 2015.

24. Avinash K Dixit and Robert S Pindyck. *Investment Under Uncertainty*. Princeton University Press, Princeton, NJ, 1994. ISBN 978-0691034102.

25. David T Dunn, Marie-Louise Newell, AE Ades, and Catherine S Peckham. Risk of human immunodeficiency virus type 1 transmission through breastfeeding. *The Lancet*, 340(8819):585–588, 1992.

26. Wendell Fleming and Raymond Rishel. *Deterministic and Stochastic Optimal Control*. Springer New York, 1975. doi: 10.1007/978-1-4612-6380-7. URL `http://dx.doi.org/10.1007/978-1-4612-6380-7`.

27. Geoffrey P Garnett and Roy M Anderson. Balancing sexual partnership in an age and

activity stratified model of hiv transmission in heterosexual populations. *Mathematical Medicine and Biology*, 11(3):161–192, 1994.

28. MR Gold, JE Siegel, LB Russell, and MC Weinstein. *Cost-Effectiveness in Health and Medicine*. Oxford University Press, New York, 1996. ISBN 978-0195108248.

29. Robert M. Grant, Javier R. Lama, Peter L. Anderson, Vanessa McMahan, Albert Y. Liu, Lorena Vargas, Pedro Goicochea, Martín Casapía, Juan Vicente Guanira-Carranza, and Maria E. et al. Ramirez-Cardich. Preexposure chemoprophylaxis for HIV prevention in men who have sex with men. *New England Journal of Medicine*, 363(27):2587–2599, 2010. doi: 10.1056/nejmoa1011205.

30. Ronald H Gray, Maria J Wawer, Ron Brookmeyer, Nelson K Sewankambo, David Serwadda, Fred Wabwire-Mangen, Tom Lutalo, Xianbin Li, Thomas vanCott, and Thomas C Quinn. Probability of HIV-1 transmission per coital act in monogamous, heterosexual, HIV-1-discordant couples in Rakai, Uganda. *The Lancet*, 357(9263): 1149–1153, 2001. doi: 10.1016/s0140-6736(00)04331-2.

31. Daniel T Halperin and Helen Epstein. Concurrent sexual partnerships help to explain Africa's high HIV prevalence: implications for prevention. *The Lancet*, 364(9428):4–6, 2004. doi: 10.1016/s0140-6736(04)16606-3.

32. Khalid Hattaf and Noura Yousfi. Two optimal treatments of hiv infection model. *World Journal of Modelling and Simulation*, 8(1):27–35, 2012.

33. Herbert W. Hethcote. The mathematics of infectious diseases. *SIAM Review*, 42(4): 599–653, 2000. doi: 10.1137/s0036144500371907.

34. Alan C. Hindmarsh. Brief description of ODEPACK - a systematized collection of ODE

solvers double precision version, Accessed 25 Sept 2015. URL http://www.netlib.org/odepack/opkd-sum.

35. Leigh F Johnson, Thomas M Rehle, Sean Jooste, and Linda-Gail Bekker. Rates of hiv testing and diagnosis in south africa: successes and challenges. *AIDS*, 29(11): 1401–1409, 2015.

36. Eric Jones, Travis Oliphant, Pearu Peterson, and other. SciPy: Open source scientific tools for Python, Accessed 25 Sept 2015. URL http://www.scipy.org.

37. Eunok Jung, Shingo Iwami, Yasuhiro Takeuchi, and Tae-Chang Jo. Optimal control strategy for prevention of avian influenza pandemic. *Journal of theoretical biology*, 260 (2):220–229, 2009.

38. S C Kalichman. HIV testing attitudes, AIDS stigma, and voluntary HIV counselling and testing in a black township in Cape Town, South Africa. *Sexually Transmitted Infections*, 79(6):442–447, 2003. doi: 10.1136/sti.79.6.442.

39. Matt J Keeling and Pejman Rohani. *Modeling infectious diseases in humans and animals*. Princeton University Press, 2008.

40. Denise Kirschner, Suzanne Lenhart, and Steve Serbin. Optimal control of the chemotherapy of HIV. *Journal of Mathematical Biology*, 35(7):775–792, aug 1997. doi: 10.1007/s002850050076. URL http://dx.doi.org/10.1007/s002850050076.

41. J Kreiss, M Carael, and A Meheus. Role of sexually transmitted diseases in transmitting human immunodeficiency virus. *Genitourinary medicine*, 64(1):1, 1988.

42. Suzanne Lenhart, E. Jung, and Z. Feng. Optimal control of treatments in a two-strain

tuberculosis model. *DCDS-B*, 2(4):473–482, aug 2002. doi: 10.3934/dcdsb.2002.2.473. URL `http://dx.doi.org/10.3934/dcdsb.2002.2.473`.

43. Suzanne Lenhart, Vladimir Protopopescu, Eunok Jung, and Charles Babbs. Optimal control for a standard CPR model. *Nonlinear Analysis: Theory Methods & Applications*, 63(5-7):e1391–e1397, nov 2005. doi: 10.1016/j.na.2005.02.023. URL `http://dx.doi.org/10.1016/j.na.2005.02.023`.

44. Jay A. Levy. Mysteries of {HIV}: challenges for therapy and prevention. *Nature*, 333(6173):519–522, jun 1988. doi: 10.1038/333519a0. URL `http://dx.doi.org/10.1038/333519a0`.

45. Mary Lou Lindegren, Robert H Byers Jr, Pauline Thomas, Susan F Davis, Blake Caldwell, Martha Rogers, Marta Gwinn, John W Ward, and Patricia L Fleming. Trends in perinatal transmission of hiv/aids in the united states. *Jama*, 282(6):531–538, 1999.

46. Ira M. Longini, W. Scott Clark, Robert H. Byers, John W. Ward, William W. Darrow, George F. Lemp, and Herbert W. Hethcote. Statistical analysis of the stages of {HIV} infection using a Markov model. *Statist. Med.*, 8(7):831–843, jul 1989. doi: 10.1002/sim.4780080708. URL `http://dx.doi.org/10.1002/sim.4780080708`.

47. James H McMahon, Julian H Elliott, Silvia Bertagnolio, Rachel Kubiak, and Michael R Jordan. Viral suppression after 12 months of antiretroviral therapy in low-and middle-income countries: a systematic review. *Bulletin of the World Health Organization*, 91 (5):377–385, 2013.

48. Margaret L McNairy, Myron Cohen, and Wafaa M El-Sadr. Antiretroviral therapy for prevention is a combination strategy. *Current HIV/AIDS Reports*, 10(2):152–158, 2013.

49. Médecins Sans Frontières. Untangling the web of antiretroviral price reductions: 14th edition—July 2011, Accessed 21 Oct 2015. URL http://d2pd3b5abq75bb.cloudfront.net/2012/07/16/14/42/23/52/UTW_14_ENG_July2011.pdf.

50. Julio SG Montaner, Robert Hogg, Evan Wood, Thomas Kerr, Mark Tyndall, Adrian R Levy, and P Richard Harrigan. The case for expanding access to highly active antiretroviral therapy to curb the growth of the HIV epidemic. *The Lancet*, 368(9534): 531–536, 2006. doi: 10.1016/s0140-6736(06)69162-9.

51. Rachael Miller Neilan and Suzanne Lenhart. An introduction to optimal control with an application in disease modeling. *DIMACS Series in Discrete Mathematics*, 75: 67–81, 2010.

52. KO Okosun, OD Makinde, and I Takaidza. Impact of optimal control on the treatment of hiv/aids and screening of unaware infectives. *Applied Mathematical Modelling*, 37 (6):3802–3820, 2013.

53. World Health Organization et al. Global update on hiv treatment 2013: results, impact and opportunities. *World Health Organization*, 2013.

54. World Health Organization et al. *March 2014 supplement to the 2013 consolidated guidelines on the use of antiretroviral drugs for treating and preventing HIV infection: recommendations for a public health approach*, 2014. URL http://apps.who.int/iris/bitstream/10665/104264/1/9789241506830_eng.pdf.

55. World Health Organization et al. Consolidated guidelines on general hiv care and the use of antiretroviral drugs for treating and preventing hiv infection: recommendations for a public health approach. 2013. *World Health Organization*, 20, 2015.

56. Thomas A Peterman, D Peter Drotman, and James W Curran. Epidemiology of the acquired immunodeficiency syndrome (aids). *Epidemiologic reviews*, 7:1–21, 1985.

57. Lev Semenovich Pontryagin. *Mathematical theory of optimal processes*. CRC Press, 1987.

58. Travis C Porco, Jeffrey N Martin, Kimberly A Page-Shafer, Amber Cheng, Edwin Charlebois, Robert M Grant, and Dennis H Osmond. Decline in HIV infectivity following the introduction of highly active antiretroviral therapy. *AIDS*, 18(1):81–88, 2004. doi: 10.1097/00002030-200401020-00010.

59. Michael JD Powell. A direct search optimization method that models the objective and constraint functions by linear interpolation. In *Advances in optimization and numerical analysis*, pages 51–67. Springer, 1994.

60. Thomas C. Quinn, Maria J. Wawer, Nelson Sewankambo, David Serwadda, Chuanjun Li, Fred Wabwire-Mangen, Mary O. Meehan, Thomas Lutalo, and Ronald H. Gray. Viral load and heterosexual transmission of human immunodeficiency virus type 1. *New England Journal of Medicine*, 342(13):921–929, 2000. doi: 10.1056/nejm200003303421303.

61. M L Richter, M Chersich, M Temmerman, and S Luchters. Characteristics, sexual behaviour and risk factors of female, male and transgender sex workers in South Africa. *South African Medical Journal*, 103(4), 2013. doi: 10.7196/samj.6170.

62. Bernhard Schwartländer, John Stover, Timothy Hallett, Rifat Atun, Carlos Avila, Eleanor Gouws, Michael Bartos, Peter D Ghys, Marjorie Opuni, David Barr, et al. Towards an improved investment approach for an effective response to hiv/aids. *The Lancet*, 377(9782):2031–2041, 2011.

63. Nathan Shaffer, Rutt Chuachoowong, Philip A Mock, Chaiporn Bhadrakom, Wimol Siriwasin, Nancy L Young, Tawee Chotpitayasunondh, Sanay Chearskul, Anuvat Roongpisuthipong, Pratharn Chinayon, et al. Short-course zidovudine for perinatal hiv-1 transmission in bangkok, thailand: a randomised controlled trial. *The Lancet*, 353(9155):773–780, 1999.

64. James D Shelton. Ten myths and one truth about generalised HIV epidemics. *The Lancet*, 370(9602):1809–1811, 2007. doi: 10.1016/s0140-6736(07)61755-3.

65. LC Simbayi, O Shisana, T Rehle, D Onoya, S Jooste, N Zungu, and K Zuma. *South African National HIV Prevalence, Incidence and Behaviour Survey, 2012*. HSRC Press, Cape Town, South Africa, 2014. ISBN 978-0-7969-2483-4. URL `http://www.hsrc.ac.za/en/research-data/ktree-doc/15031`.

66. Stephen M Smith. Pre-exposure chemoprophylaxis for hiv: it is time. *Retrovirology*, 1 (1):16, 2004.

67. Statistics South Africa. Mid-year population estimates 2014, Accessed 21 Oct 2015. URL `http://www.statssa.gov.za/publications/P0302/P03022014.pdf`.

68. Jonathan AC Sterne, Miguel A Hernán, Bruno Ledergerber, Kate Tilling, Rainer Weber, Pedram Sendi, Martin Rickenbach, James M Robins, Matthias Egger, Swiss HIV Cohort Study, et al. Long-term effectiveness of potent antiretroviral therapy in preventing aids and death: a prospective cohort study. *The Lancet*, 366(9483):378–384, 2005.

69. Amy S Sturt, Emily Kainne Dokubo, and Tin Tin Sint. Antiretroviral therapy (art) for treating hiv infection in art-eligible pregnant women. *The Cochrane Library*, 2010.

70. UNAIDS.org. The GAP report | UNAIDS, Accessed 25 Sept 2015. URL http://www.unaids.org/sites/default/files/en/media/unaids/contentassets/documents/unaidspublication/2014/UNAIDS_Gap_report_en.pdf.

71. UNAIDS.org. AIDSinfo | UNAIDS, Accessed 25 Sept 2015. URL http://www.unAIDS.org/en/dataanalysis/datatools/AIDSinfo.

72. UNAIDS.org. 90–90–90: An ambitious treatment target to help end the AIDS epidemic, Accessed 25 Sept 2015. URL http://www.unAIDS.org/en/resources/documents/2014/90-90-90.

73. Ard I van Sighem, Mark A van de Wiel, Azra C Ghani, Mariélle Jambroes, Peter Reiss, Inge C Gyssens, Kees Brinkman, Joep MA Lange, and Frank de Wolf. Mortality and progression to AIDS after starting highly active antiretroviral therapy. *AIDS*, 17 (15):2227–2236, 2003. doi: 10.1097/00002030-200310170-00011.

74. Nancy A Wade, Guthrie S Birkhead, Barbara L Warren, Tina T Charbonneau, P Tyler French, Ling Wang, Jeanne B Baum, James M Tesoriero, and Robert Savicki. Abbreviated regimens of zidovudine prophylaxis and perinatal transmission of the human immunodeficiency virus. *New England Journal of Medicine*, 339(20):1409–1414, 1998.

75. Xuezhong Wang. Solving optimal control problems with matlab-indirect methods: Indirect methods, 2009.

76. Maria J. Wawer, Ronald H. Gray, Nelson K. Sewankambo, David Serwadda, Xianbin Li, Oliver Laeyendecker, Noah Kiwanuka, Godfrey Kigozi, Mohammed Kiddugavu, Thomas Lutalo, Fred Nalugoda, Fred Wabwire-Mangen, Mary P. Meehan, and Thomas C. Quinn. Rates of {HIV}-1 Transmission per Coital Act by Stage of {HIV}-1

Infection, in Rakai, Uganda. *The Journal of Infectious Diseases*, 191(9):1403–1409, may 2005. doi: 10.1086/429411. URL http://dx.doi.org/10.1086/429411.

77. Robin A Weiss. How does hiv cause aids? *Science*, 260(5112):1273–1279, 1993.

78. WHO. *Interim Who Clinical Staging of HIV/AIDS and HIV/AIDS Case Definitions For Surveillance African Region*, 2005 (accessed September 6, 2014). URL http://www.who.int/hiv/pub/guidelines/clinicalstaging.pdf.

79. WHO. *Rapid advice: antiretroviral therapy for HIV infection in adults and adolescents*, 2009 (accessed September 6, 2014). URL http://www.who.int/hiv/pub/arv/rapid_advice_art.pdf.

80. WHO. *WHO: People Most at Risk of HIV are not Getting the Health Services they Need*, 2014 (accessed September 6, 2014). URL http://www.who.int/mediacentre/news/releases/2014/key-populations-to-hiv/en/.

81. WHO, 2013. Adult hiv prevalence (15-49 years)., 2013. URL http://www.unAIDS.org/en/dataanalysis/datatools/AIDSinfo.

82. James A Wiley, Stephen J Herschkorn, and Nancy S Padian. Heterogeneity in the probability of HIV transmission per sexual contact: The case of male-to-female transmission in penile–vaginal intercourse. *Statistics in Medicine*, 8(1):93–102, 1989. doi: 10.1002/sim.4780080110.

83. David P. Wilson, Alexander Hoare, David G. Regan, and Matthew G. Law. Importance of promoting {HIV} testing for preventing secondary transmissions: modelling the Australian {HIV} epidemic among men who have sex with men. *Sexual Health*, 6(1): 19, 2009. doi: 10.1071/sh08081. URL http://dx.doi.org/10.1071/SH08081.

84. World Health Organization. AIDS in africa: Three scenarios to 2025 - 2014, Accessed 20 Oct. 2015. URL http://data.unaids.org/Publications/IRC-pub07/jc1058-AIDSinafrica_en.pdf.

85. Mike Youle and Mark A Wainberg. Could chemoprophylaxis be used as an hiv prevention strategy while we wait for an effective vaccine? *Aids*, 17(6):937–938, 2003.

86. Tunde T Yusuf and Francis Benyah. Optimal strategy for controlling the spread of hiv/aids disease: a case study of south africa. *Journal of biological dynamics*, 6(2): 475–494, 2012.

APPENDICES

A APPENDIX Supplementary Material for Chapter 2

We constructed a mathematical model that incorporates risk, gender and treatment. We parameterized the model from from a variety of data sources from South Africa and, more broadly, Sub-Saharan Africa.

A1 Model

Our model stratifies the population by gender, level of sexual activity, HIV status and treatment status (Figure 2.1). The female population is divided into low, medium and high risk, while the male population is divided into low and medium risk. Low-risk people are in monogamous marriages, single or otherwise have zero or one sexual partner. Medium-risk people have more than one sexual partner. High-risk females are commercial sex workers, sometimes called transactional sex workers. We neglected the high-risk class for males because male and transgender commercial sex workers in Sub-Saharan Africa are rare [61]. Each of the risk groups is further stratified by HIV infection status (i.e. HIV negative and HIV positive) and by treatment status (treated and untreated).

The model only considers the adult population. People enter the model at per capita rate b, when they attain the age of 15, as low risk and leave either through natural death, with per capita rate μ, or death due to AIDS, with per capita rate γ. Movement between risk groups is due to changes in sexual behavior and movement within risk group is due to either infection or people electing to start or stop treatment. HIV-negative people can only become infected by heterosexual sexual contact; men who have sex with men, injecting drug use and vertical transmission are not included in the model.

We assumed perfect adherence to treatment, i.e. individuals who start using treatment will strictly adhere the prescribed regime until they decide to stop treatment. All other HIV

prevention measures such as male circumcision, education and condom use are aggregated in the HIV risk per sex act and are assumed constant for the span of the model.

The forces of infection for females and males are

$$\begin{aligned}\lambda_{\mathrm{F}ij}(t) &= \sum_{\substack{k\in\{\mathrm{L,M}\}\\ \ell\in\{\mathrm{U,T}\}}} \psi c_{ijk\ell}\frac{I_{\mathrm{M}k\ell}(t)}{N(t)}, \qquad \text{for } i\in\{\mathrm{L,M,H}\} \text{ and } j\in\{\mathrm{U,T}\},\\ \lambda_{\mathrm{M}ij}(t) &= \sum_{\substack{k\in\{\mathrm{L,M,H}\}\\ \ell\in\{\mathrm{U,T}\}}} c_{ijk\ell}\frac{I_{\mathrm{F}kl}(t)}{N(t)}, \qquad \text{for } i\in\{\mathrm{L,M}\} \text{ and } j\in\{\mathrm{U,T}\},\end{aligned} \tag{A1}$$

where the total population size is

$$N(t) = \sum_{\substack{i\in\{\mathrm{F,M}\}\\ j\in\{\mathrm{L,M,H}\}\\ k\in\{\mathrm{U,T}\}}} \left[S_{ijk}(t) + I_{ijk}(t)\right], \tag{A2}$$

with S_{ijk} and I_{ijk} defined to be 0 for the groups not in the model (FLT, MLT, MHU, and MHT), and the transmission rate is

$$c_{ijk\ell} = d_{ik}\cdot e_j \cdot f_\ell, \tag{A3}$$

with the infectivity set by the riskiest partner,

$$d_{ik} = \begin{cases} \beta_{\mathrm{L}} & \text{if } i = k = \mathrm{L},\\ \beta_{\mathrm{M}} & \text{if } i = \mathrm{M} \text{ or } (k = \mathrm{M} \text{ and } i \neq \mathrm{H}),\\ \beta_{\mathrm{H}} & \text{if } i = \mathrm{H},\end{cases} \tag{A4}$$

the transmissibility reduced if the uninfected partner is on PrEP,

$$e_j = \begin{cases} 1 & \text{if } j = \mathrm{U},\\ 1-\theta_{\mathrm{PrEP}} & \text{if } j = \mathrm{T},\end{cases} \tag{A5}$$

and the transmissibility is reduced if the infected partner is on ART

$$f_\ell = \begin{cases} 1 & \text{if } \ell = \mathrm{U},\\ 1-\theta_{\mathrm{ART}} & \text{if } \ell = \mathrm{T}.\end{cases} \tag{A6}$$

The model is specified as the system of differential equations

$$\frac{\mathrm{d}S_{\mathrm{FLU}}}{\mathrm{d}t} = \frac{bN}{2} + y_{\mathrm{SFMU}}F_{\mathrm{NMU}} - (x_{\mathrm{SFLU}} + \lambda_{\mathrm{FLU}} + \mu)S_{\mathrm{FLU}}, \tag{A7a}$$

$$\frac{\mathrm{d}I_{\mathrm{FLU}}}{\mathrm{d}t} = \lambda_{\mathrm{FLU}}S_{\mathrm{FLU}} + y_{\mathrm{IFMU}}I_{\mathrm{FMU}} + \sigma I_{\mathrm{FLT}} - (x_{\mathrm{IFLU}} + \pi_{\mathrm{IML}} + \gamma + \mu)I_{\mathrm{FLU}}, \tag{A7b}$$

$$\frac{\mathrm{d}I_{\mathrm{FLT}}}{\mathrm{d}t} = y_{\mathrm{IFMT}}I_{\mathrm{FMT}} + \pi_{\mathrm{IFL}}F_{\mathrm{PLU}} - (x_{\mathrm{IFLT}} + \sigma + \mu)I_{\mathrm{FLT}}, \tag{A7c}$$

$$\frac{\mathrm{d}S_{\mathrm{MLU}}}{\mathrm{d}t} = \frac{bN}{2} + y_{\mathrm{SMMU}}S_{\mathrm{MMU}} - (x_{\mathrm{SMLU}} + \lambda_{\mathrm{MLU}} + \mu)S_{\mathrm{MLU}}, \tag{A7d}$$

$$\frac{\mathrm{d}I_{\mathrm{MLU}}}{\mathrm{d}t} = y_{\mathrm{IMMU}}I_{\mathrm{MMU}} + \lambda_{\mathrm{MLU}}S_{\mathrm{MLU}} + \sigma I_{\mathrm{MLT}} - (x_{\mathrm{IMLU}} + \pi_{\mathrm{IML}} + \gamma + \mu)I_{\mathrm{MLU}}, \tag{A7e}$$

$$\frac{\mathrm{d}I_{\mathrm{MLT}}}{\mathrm{d}t} = y_{IMMT}I_{\mathrm{MMT}} + \pi_{\mathrm{IML}}I_{\mathrm{MLU}} - (x_{\mathrm{IMLT}} + \sigma + \mu)I_{\mathrm{MLT}}, \tag{A7f}$$

$$\begin{aligned}\frac{\mathrm{d}S_{\mathrm{FMU}}}{\mathrm{d}t} &= x_{\mathrm{SFLU}}S_{\mathrm{FLU}} + y_{\mathrm{SFHU}}S_{\mathrm{FHU}} + \sigma S_{\mathrm{FMT}} \\ &\quad - (y_{\mathrm{SFMU}} + x_{\mathrm{SFMU}} + \pi_{\mathrm{SFM}} + \lambda_{\mathrm{FMU}} + \mu)S_{\mathrm{FMU}},\end{aligned} \tag{A7g}$$

$$\frac{\mathrm{d}S_{\mathrm{FMT}}}{\mathrm{d}t} = y_{\mathrm{SFHT}}S_{\mathrm{FHT}} + \pi_{\mathrm{SFM}}S_{\mathrm{FMU}} - (x_{\mathrm{SFMT}} + \sigma + \lambda_{\mathrm{FMT}} + \mu)S_{\mathrm{FMT}}, \tag{A7h}$$

$$\begin{aligned}\frac{\mathrm{d}I_{\mathrm{FMU}}}{\mathrm{d}t} &= x_{\mathrm{IFLU}}I_{\mathrm{FLU}} + y_{\mathrm{IFHU}}I_{\mathrm{FHU}} + \sigma I_{\mathrm{FMT}} + \lambda_{\mathrm{FMU}}S_{\mathrm{FMU}} \\ &\quad - (y_{\mathrm{IFMU}} + x_{\mathrm{IFMU}} + \pi_{\mathrm{IFM}} + \gamma + \mu)I_{\mathrm{FMU}},\end{aligned} \tag{A7i}$$

$$\begin{aligned}\frac{\mathrm{d}I_{\mathrm{FMT}}}{\mathrm{d}t} &= x_{\mathrm{IFLT}}I_{\mathrm{FLT}} + y_{\mathrm{IFHT}}I_{\mathrm{FHT}} + \pi_{\mathrm{IFM}}I_{\mathrm{FMU}} + \lambda_{\mathrm{FMT}}S_{\mathrm{FMT}} \\ &\quad - (y_{\mathrm{IFMT}} + x_{\mathrm{IFMT}} + \sigma + \mu)I_{\mathrm{FMT}},\end{aligned} \tag{A7j}$$

$$\frac{\mathrm{d}S_{\mathrm{MMU}}}{\mathrm{d}t} = x_{\mathrm{SMLU}}S_{\mathrm{MLU}} + \sigma S_{\mathrm{MMT}} - (y_{\mathrm{SMMU}} + \pi_{\mathrm{SMM}} + \lambda_{\mathrm{MMU}} + \mu)S_{\mathrm{MMU}}, \tag{A7k}$$

$$\frac{\mathrm{d}S_{\mathrm{MMT}}}{\mathrm{d}t} = \pi_{\mathrm{SMM}}S_{\mathrm{MMU}} - (\sigma + \lambda_{\mathrm{MMT}} + \mu)S_{\mathrm{MMT}}, \tag{A7l}$$

$$\frac{\mathrm{d}I_{\mathrm{MMU}}}{\mathrm{d}t} = x_{\mathrm{IMLU}}I_{\mathrm{MLU}} + \sigma I_{\mathrm{MMT}} + \lambda_{\mathrm{MMU}}S_{\mathrm{MMU}} - (y_{\mathrm{IMMU}} + \pi_{\mathrm{IMM}} + \gamma + \mu)I_{\mathrm{MMU}}, \tag{A7m}$$

$$\frac{\mathrm{d}I_{\mathrm{MMT}}}{\mathrm{d}t} = x_{\mathrm{IMLT}}I_{\mathrm{MLT}} + \pi_{\mathrm{IMM}}I_{\mathrm{MMU}} + \lambda_{\mathrm{MMT}}S_{\mathrm{MMT}} - (y_{\mathrm{IMMT}} + \sigma + \mu)I_{\mathrm{MMT}} \tag{A7n}$$

$$\frac{\mathrm{d}S_{\mathrm{FHU}}}{\mathrm{d}t} = x_{\mathrm{SFMU}}S_{\mathrm{FMU}} + \sigma S_{\mathrm{FHT}} - (y_{\mathrm{SFHU}} + \pi_{\mathrm{SFH}} + \lambda_{\mathrm{FHU}} + \mu)S_{\mathrm{FHU}}, \tag{A7o}$$

$$\frac{\mathrm{d}S_{\mathrm{FHT}}}{\mathrm{d}t} = x_{\mathrm{SFMT}}S_{\mathrm{FMT}} + \pi_{\mathrm{SFH}}S_{\mathrm{FHU}} - (y_{\mathrm{SFHT}} + \sigma + \lambda_{\mathrm{FHT}} + \mu)S_{\mathrm{FHT}}, \tag{A7p}$$

$$\frac{\mathrm{d}I_{\mathrm{FHU}}}{\mathrm{d}t} = x_{\mathrm{IFMU}} I_{\mathrm{FMU}} + \sigma I_{\mathrm{FHT}} + \lambda_{\mathrm{FHU}} S_{\mathrm{FHU}} - (y_{\mathrm{IFHU}} + \pi_{\mathrm{IFH}} + \gamma + \mu) I_{\mathrm{FHU}}, \tag{A7q}$$

$$\frac{\mathrm{d}I_{\mathrm{FHT}}}{\mathrm{d}t} = x_{\mathrm{IFMT}} I_{\mathrm{FMT}} + \pi_{\mathrm{IFH}} I_{\mathrm{FHU}} + \lambda_{\mathrm{FHT}} S_{\mathrm{FHT}} - (y_{\mathrm{IFHT}} + \sigma + \mu) I_{\mathrm{FHT}}. \tag{A7r}$$

A2 Parameterization

We parameterized our model with demographic data from South Africa because of the availability of excellent demographic statistics (Tables A1–A3). Our model is initialized for the beginning of the year 2014 ($t = 0$), although we used demographic data from 2012, the date of the last South African National HIV Prevalence, Incidence and Behavior Survey [65].

The initial numbers of people in each compartment (Table A2) were derived from separate South African data on the current prevalence by gender and risk and on the proportion of infected people currently receiving ART. From the 2012 South African National HIV Prevalence, Incidence and Behavior Survey [65], adults who reported having one sexual partner in last 12 months were considered low-risk. Multiple sexual partners increases the likelihood of exposure to HIV through expanding sexual networks, so people who reported having more than one sexual partners in the last 12 months were classified as medium risk. Female sex workers were considered as the high-risk population. The number of people in each gender and risk group, combined with the current prevalence by gender and risk from the same survey, were used to determine the initial number of people infected and susceptible:

$$I_{ij}(0) = \phi_{ij} N_{ij}(0), \qquad S_{ij}(0) = (1 - \phi_{ij})\, N_{ij}(0), \tag{A8}$$

for $i \in \{\mathrm{F}, \mathrm{M}\}$ and $j \in \{\mathrm{L}, \mathrm{M}, \mathrm{H}\}$, where ϕ_{ij} is the prevalence and $N_{ij}(0)$ is the initial number of people, both susceptible and infected, in that gender and risk group. We then initialized the number of infected people on ART in each gender and risk group to 42%, the

Parameter	Description	Value	Source
b	Birth rate	0.0226 y^{-1}	[67]
μ	Natural death rate	0.0106 y^{-1}	[67]
γ	Death rate due to AIDS	0.0909 y^{-1}	[71]
β_{L}	Low-risk contact rate	0.0050 y^{-1}	[30]
β_{M}	Medium-risk contact rate	0.0075 y^{-1}	—
β_{H}	High-risk contact rate	0.0250 y^{-1}	—
ψ	Male-to-female relative transmission rate	2	[82]
θ_{ART}	Efficacy of ART at preventing transmission	96%	[18]
θ_{PrEP}	Efficacy of PrEP at preventing transmission	71%	[6]
$r_{\max}$	Rate of enrolling people on treatment	20 y^{-1} $\mathrm{treatment}^{-1}$	—
c_{I}	Cost of an infection	\$1,000 y^{-1}	[15]
c_{D}	Cost of a death	\$100,000	—
c_{T}	Cost of treatment	\$120 y^{-1}	[49]
r	Discount rate for costs	3% y^{-1}	[28]
σ	Rate of treatment stoppage	0.2 y^{-1}	[65]
ϕ_{FL}	Initial prevalence amongst low-risk females	14.4%	[65]
ϕ_{FM}	Initial prevalence amongst medium-risk females	23.2%	[65]
ϕ_{FH}	Initial prevalence amongst high-risk females	59.6%	[65]
ϕ_{ML}	Initial prevalence amongst low-risk males	17.3%	[65]
ϕ_{MM}	Initial prevalence amongst medium-risk males	14.5%	[65]

Table A1: Model parameters. See also Tables A3 & A2.

current proportion of infected people on ART in South Africa [71]:

$$I_{ij\mathrm{T}}(0) = 0.42 I_{ij}(0), \qquad I_{ij\mathrm{U}}(0) = (1 - 0.42) I_{ij}(0), \tag{A9}$$

for $i \in \{\mathrm{F}, \mathrm{M}\}$ and $j \in \{\mathrm{L}, \mathrm{M}, \mathrm{H}\}$. We further assumed no widespread use of PrEP, i.e. $S_{ij\mathrm{T}}(0) = 0$ for $i \in \{\mathrm{F}, \mathrm{M}\}$ and $j \in \{\mathrm{L}, \mathrm{M}, \mathrm{H}\}$.

We took the low-risk contact rate, β_{L}, to be annual risk for serodiscordant couples [30], and then the medium- and high-risk contact rates, β_{M} and β_{H}, to be respectively 50% and 400% greater than the low-risk contact rate.

Treatment is generally constrained by number of drugs available, the manpower available to distribute the drugs and other resources. The rate of enrolling people on treatment per drug available per untreated person ($r_{\max}$) was assumed to be 20 $\mathrm{y}^{-1}\mathrm{treatment}^{-1}$, i.e. it takes about 1/20 y or about 18 days for an eligible person to get on treatment.

The cost of death (c_D) is the financial benefit that an otherwise healthy person would contribute to society on average over his or her remaining lifetime was assumed to be \$100,000.

Due to the absence of behavioral data, values for the rates of transition between risk groups ($x_{ijk\ell}$ and $y_{ijk\ell}$) were assumptions (Table A3).

A3 Objectives

We determined the proportion of each group to be treated according to four different outcome measures: the number of new infections, the number of total infection-years, the number of deaths and total cost. These were all evaluated over a time horizon of $t_{\mathrm{end}} = 10$ y.

New infections The total number of new infections is the sum of the rates of new infections

Variable	**Value**
$S_{\mathrm{FLU}}(0)$	14,767,171
$I_{\mathrm{FLU}}(0)$	2,017,210
$I_{\mathrm{FLT}}(0)$	1,071,932
$S_{\mathrm{MLU}}(0)$	11,865,435
$I_{\mathrm{MLU}}(0)$	1,764,793
$I_{\mathrm{MLT}}(0)$	717,335
$S_{\mathrm{FMU}}(0)$	719,587
$S_{\mathrm{FMT}}(0)$	0
$I_{\mathrm{FMU}}(0)$	98,296
$I_{\mathrm{FMT}}(0)$	52,234
$S_{\mathrm{MMU}}(0)$	2,984,922
$S_{\mathrm{MMT}}(0)$	0
$I_{\mathrm{MMU}}(0)$	443,959
$I_{\mathrm{MMT}}(0)$	180,456
$S_{\mathrm{FHU}}(0)$	76,419
$S_{\mathrm{FHT}}(0)$	0
$I_{\mathrm{FHU}}(0)$	73,617
$I_{\mathrm{FHT}}(0)$	39,120

Table A2: Model initial conditions.

Parameter	Risk Transition Rate	Value
x_{SFLU}	From low to medium risk for untreated susceptible females	0.05 y^{-1}
x_{SFMU}	From medium to high risk for untreated susceptible females	0.06 y^{-1}
y_{SFMU}	From medium to low risk for untreated susceptible females	0.02 y^{-1}
x_{SFMT}	From medium to high risk for treated susceptible females	0.05 y^{-1}
y_{SFHU}	From high to medium risk for untreated susceptible females	0.02 y^{-1}
y_{SFHT}	From high to medium risk for treated susceptible females	0.02 y^{-1}
x_{SMLU}	From low to medium risk for untreated susceptible males	0.05 y^{-1}
y_{SMMU}	From medium to low risk for untreated susceptible males	0.03 y^{-1}
x_{IFLU}	From low to medium risk for untreated infected females	0.02 y^{-1}
x_{IFLT}	From low to medium risk for treated infected females	0.06 y^{-1}
x_{IFMU}	From medium to high risk for untreated infected females	0.06 y^{-1}
y_{IFMU}	From medium to low risk for untreated infected females	0.04 y^{-1}
x_{IFMT}	From medium to high risk for treated infected females	0.07 y^{-1}
y_{IFMT}	From medium to low risk for treated infected females	0.02 y^{-1}
y_{IFHU}	From high to medium risk for untreated infected females	0.07 y^{-1}
y_{IFHT}	From high to medium risk for treated infected females	0.01 y^{-1}
x_{IMLU}	From low to medium risk for untreated infected males	0.06 y^{-1}
x_{IMLT}	From low to medium risk for treated infected males	0.07 y^{-1}
y_{IMMU}	From medium to low risk for untreated infected males	0.03 y^{-1}
y_{IMMT}	From medium to low risk for treated infected males	0.04 y^{-1}

Table A3: Model risk transition rates.

arising in each of the 8 uninfected classes integrated over the time period:

$$\int_0^{t_{\text{end}}} \sum_{\substack{i\in\{\text{F,M}\}\\ j\in\{\text{L,M,H}\}\\ k\in\{\text{U,T}\}}} \lambda_{ijk}(t) S_{ijk}(t)\,\mathrm{d}t, \tag{A10}$$

with λ_{ijk} defined to be 0 for the groups not in the model (FLT, MLT, MHU, and MHT).

Total infection-years The total infection-years is the sum of the number of all the infected people (both treated and untreated) in the 10 infected classes at each time, integrated over the time period, i.e.

$$\int_0^{t_{\text{end}}} I(t)\,\mathrm{d}t, \tag{A11}$$

where the number of infected people is

$$I(t) = \sum_{\substack{i\in\{\text{F,M}\}\\ j\in\{\text{L,M,H}\}\\ k\in\{\text{U,T}\}}} I_{ijk}(t). \tag{A12}$$

Deaths due to AIDS Treatment prevents progression from HIV to AIDS, thus reducing deaths. We assumed that AIDS-related deaths only occur in infected, untreated people. Thus, the total number of deaths due to AIDS is the number of untreated infected people at each time,

$$I_{\text{U}}(t) = \sum_{\substack{i\in\{\text{F,M}\}\\ j\in\{\text{L,M,H}\}}} I_{ij\text{U}}(t), \tag{A13}$$

multiplied by the rate of death due to AIDS:

$$\int_0^{t_{\text{end}}} \gamma I_{\text{U}}(t)\mathrm{d}t. \tag{A14}$$

Total cost The total cost consists of cost of infections, cost of deaths, and cost of treatment. The disease cost per year, which includes monetary equivalent loss of the infected

people, like lost productivity etc., is average cost of disease per person per year times the total number of people infected,

$$C_{\mathrm{I}}(t) = c_{\mathrm{I}} I(t). \tag{A15}$$

The cost per year of deaths is the the cost per death times the number of deaths per year,

$$C_{\mathrm{D}}(t) = c_{\mathrm{D}} \gamma I_{\mathrm{U}}(t). \tag{A16}$$

The treatment cost per year is the cost per person per year times the number of people treated:

$$C_{\mathrm{T}}(t) = c_{\mathrm{T}} T(t), \tag{A17}$$

where the number of people treated is

$$T(t) = \sum_{\substack{i \in \{\mathrm{F,M}\} \\ j \in \{\mathrm{L,M,H}\}}} \left[S_{ij\mathrm{T}}(t) + I_{ij\mathrm{T}}(t) \right]. \tag{A18}$$

The cost objective is discounted sum of these costs, integrated over the time period:

$$\int_0^{t_{\mathrm{end}}} \left[C_{\mathrm{I}}(t) + C_{\mathrm{D}}(t) + C_{\mathrm{T}}(t) \right] \mathrm{e}^{-rt} \mathrm{d}t. \tag{A19}$$

The total cost is discounted at rate r, representing the rate a policymaker is willing to pay as trade-off for the cost today versus future cost [24].

A4 Controls and Constraints

We considered two different ways to model government-level intervention programs that use antiretroviral drugs to control HIV. In both intervention models, the antiretroviral drugs can either be used as treatment of infected people (ART) or as prophylaxis for uninfected people (PrEP). The first intervention model, global ART & PrEP, infected people are given ART and uninfected people are given PrEP at coverage levels that are

independent of their risk of transmission. In the second intervention model, ART & high-risk PrEP, there are coverage levels of ART and PrEP for high-risk people and another coverage level of ART for low- and medium-risk people, with no PrEP for low- and medium-risk people. The latter intervention model was chosen to evaluate the CDC's current recommendation that ART be given to all infected people and PrEP only be given to high-risk people [13].

The control variables are the total number of people targeted to be on treatment in each of the designated groups (V_G) at any one time. From these treatment targets, the flows of people into treatment per unit time were taken to be

$$\pi_G(t) = r_{\max} \max\left(V_G - T_G(t), 0\right), \tag{A20}$$

where $T_G(t)$ is the total number of people in group G currently on treatment. The quantity $\max(V_G - T_G(t), 0)$ is the number of treatment slots available at time t and $r_{\max}$ is the rate of enrolling people on treatment per drug available per untreated person.

Global ART & PrEP This treatment strategy is modeled with two control groups, one for ART for all infected people,

$$G_{\text{ART}} = \{I_{\text{FL}}, I_{\text{FM}}, I_{\text{FH}}, I_{\text{ML}}, I_{\text{MM}}\}, \tag{A21}$$

and the other for PrEP for all eligible (i.e. non-low-risk) susceptible people,

$$G_{\text{PrEP}} = \{S_{\text{FM}}, S_{\text{FH}}, S_{\text{MM}}\}. \tag{A22}$$

The control variables are the total number of people targeted to be on ART (V_{ART}) and PrEP (V_{PrEP}) at any one time. The numbers of people on ART and PrEP at

time t are

$$\begin{aligned} T_{\mathrm{ART}}(t) &= \sum_{\substack{i\in\{\mathrm{F,M}\}\\ j\in\{\mathrm{L,M,H}\}}} I_{ij\mathrm{T}}(t), \\ T_{\mathrm{PrEP}}(t) &= \sum_{\substack{i\in\{\mathrm{F,M}\}\\ j\in\{\mathrm{M,H}\}}} S_{ij\mathrm{T}}(t), \end{aligned} \tag{A23}$$

and the flows into treatment for all infected and susceptible people, respectively, are

$$\begin{aligned} \pi_{\mathrm{IFL}} = \pi_{\mathrm{IFM}} = \pi_{\mathrm{IFH}} = \pi_{\mathrm{IML}} = \pi_{\mathrm{IMM}} = \pi_{\mathrm{ART}}, \\ \pi_{\mathrm{SFM}} = \pi_{\mathrm{SFH}} = \pi_{\mathrm{SMM}} = \pi_{\mathrm{PrEP}}. \end{aligned} \tag{A24}$$

The constraints on the controls are that both are positive and that their sum is less than the total amount of treatment available, $T_{\max}$:

$$\begin{aligned} V_{\mathrm{ART}} \geq 0, \quad V_{\mathrm{PrEP}} \geq 0, \\ V_{\mathrm{ART}} + V_{\mathrm{PrEP}} \leq T_{\max}. \end{aligned} \tag{A25}$$

Using this model for the control effort, the optimization problems we solved were to find the control variables $(V_{\mathrm{ART}}, V_{\mathrm{PrEP}})$ that minimize one of the objectives—new infections (A10), total infection-years (A11), deaths due to HIV (A14), or total cost (A19)—subject to the constraints (A25). For each proposed value of the control variables $(V_{\mathrm{ART}}, V_{\mathrm{PrEP}})$, the value of the objective was calculated by numerically solving the model differential equations (A7). These problems, for each of the four objectives, were solved for varying levels of total treatment $T_{\max}$.

ART & high-risk PrEP We modeled this treatment strategy with three control groups, high-risk people on ART,

$$G_{\mathrm{ART,H}} = \{I_{\mathrm{FH}}\}, \tag{A26}$$

non-high-risk people on ART,

$$G_{\mathrm{ART,LM}} = \{I_{\mathrm{FL}}, I_{\mathrm{FM}}, I_{\mathrm{ML}}, I_{\mathrm{MM}}\}, \tag{A27}$$

and high-risk people on PrEP

$$G_{\mathrm{PrEP,H}} = \{S_{\mathrm{FH}}\}, \tag{A28}$$

with no PrEP for non-high-risk people. The control variables ($V_{\mathrm{ART,H}}$, $V_{\mathrm{ART,LM}}$, $V_{\mathrm{PrEP,H}}$) are the total number of people targeted to be on treatment in each of the control groups at any one time. The total numbers of people in the treatment groups at time t are

$$\begin{aligned} T_{\mathrm{ART,H}}(t) &= I_{\mathrm{FHT}}, \\ T_{\mathrm{ART,LM}}(t) &= \sum_{\substack{i\in\{\mathrm{F,M}\} \\ j\in\{\mathrm{L,M}\}}} I_{ij\mathrm{T}}(t), \\ T_{\mathrm{PrEP,H}}(t) &= S_{\mathrm{FHT}}, \end{aligned} \tag{A29}$$

and the flows into treatment for the control groups are

$$\begin{aligned} \pi_{\mathrm{IFH}} &= \pi_{\mathrm{ART,H}}, \\ \pi_{\mathrm{IFL}} = \pi_{\mathrm{IFM}} = \pi_{\mathrm{IML}} = \pi_{\mathrm{IMM}} &= \pi_{\mathrm{ART,LM}}, \\ \pi_{\mathrm{SFH}} &= \pi_{\mathrm{PrEP,H}}, \\ \pi_{\mathrm{SFM}} = \pi_{\mathrm{SMM}} &= 0. \end{aligned} \tag{A30}$$

The constraints on the controls are that each of them are positive and that their sum is less than the total amount of treatment available, $T_{\max}$:

$$\begin{aligned} V_{\mathrm{ART,H}} \geq 0, \quad V_{\mathrm{ART,LM}} \geq 0, \quad V_{\mathrm{PrEP,H}} \geq 0, \\ V_{\mathrm{ART,H}} + V_{\mathrm{ART,LM}} + V_{\mathrm{PrEP,H}} \leq T_{\max}. \end{aligned} \tag{A31}$$

Using this model for the control effort, the optimization problems we solved were to find the control variables $(V_{\mathrm{ART,H}}, V_{\mathrm{ART,LM}}, V_{\mathrm{PrEP,H}})$ that minimize one of the objectives subject to the constraints (A31). As before, for each proposed value of the control variables, the value of the objective was calculated by numerically solving the model differential equations (A7). These problems, for each of the four objective, were solved for varying levels of total treatment $T_{\max}$.

The system of ordinary differential equations (A7) was solved in Python with `scipy.integrate.odeint` [36], which uses LSODA (Livermore Solver for Ordinary Differential Equations) [34]. The constrained optimization problems were solved numerically in Python using `scipy.optimize.fmin` with `method = 'COBYLA'` [36], which uses the COBYLA (Constrained Optimization by Linear Approximation) algorithm [59]. To avoid finding minima which are locally but not globally optimal, the best result was kept from running the optimization routine started with several different initial guesses for the control variables. The initial guesses used were: all control variables 0, all control variables equal to $T_{\max}/N$, one control variable equal to $T_{\max}$ and the others 0, and 10 uniform random vectors in $[0, T_{\max}]^N$, where N is the number of control variables.

B APPENDIX Supplementary Material for Chapter 4

Given optimal controls (U_2^*, U_3^*, U_4^*) and the corresponding state solutions $X^* = (I_{\mathrm{Ua}}^*, I_{\mathrm{U1}}^*, I_{\mathrm{U2}}^*, I_{\mathrm{U3}}^*, I_{\mathrm{U4}}^*, I_{\mathrm{Da}}^*, I_{\mathrm{D1}}^*, I_{\mathrm{D2}}^*, I_{\mathrm{D3}}^*, I_{\mathrm{D4}}^*, I_{\mathrm{T2}}^*, I_{\mathrm{T3}}^*, I_{\mathrm{T4}}^*)$ from solving the state system (4.4), there exist adjoint variables θ that satisfy the adjoint equations

$$\dot{\theta}_i = \left(-\frac{\partial \mathcal{H}}{\partial I_i}\right), \qquad i = 1, 2, \ldots, 14, \tag{A1}$$

with the transversality condition

$$\theta_i(T) = 0, \qquad i = 1, 2, ..., 14. \tag{A2}$$

The optimal control characterization holds,

$$\left(\frac{\partial \mathcal{H}}{\partial U_i}\right)\bigg|_{U_i = U_i^*} = 0, \tag{A3}$$

and

$$\mathcal{H}(t, X^*, U^*, \theta^*) \leq \mathcal{H}(t, X, U, \theta). \tag{A4}$$

To analyze all four objective functions at once, we can rewrite (A1) as

$$\dot{\theta}_i = \left(-\frac{\partial \mathcal{H}}{\partial X_i}\right) = -\left(\frac{\partial f_k}{\partial X_i}(t, X) + \theta^T \frac{\partial \ell}{\partial X_i}(t, X) + \theta^T \frac{\partial \psi}{\partial X_i}(t, U)\right) \tag{A5}$$

for $i = 1, 2, 3, \ldots, 14$ and $k \in \{\mathrm{I}, \mathrm{NI}, \mathrm{D}, \mathrm{C}\}$. The partial derivative of the objective functions

$\frac{\partial f_j}{\partial X_i}(t, S, I, U)$ are

$$\frac{\partial f_{\mathrm{I}}}{\partial X} = [0,\ 1,\ 1,\ 1,\ 1,\ 1,\ 1,\ 1,\ 1,\ 1,\ 1,\ 1,\ 1,\ 1]\,, \tag{A6}$$

$$\begin{aligned}\frac{\partial f_{\mathrm{NI}}}{\partial X} = \Big[&\frac{(I_{\mathrm{Ua}} + I_{\mathrm{Da}})\beta_a + (I_{\mathrm{U1}} + I_{\mathrm{D1}})\beta_1 + (I_{\mathrm{U2}} + I_{\mathrm{D2}} + \alpha I_{\mathrm{T2}})\beta_2 + (I_{\mathrm{U3}} + I_{\mathrm{D3}} + \alpha I_{\mathrm{T3}})\beta_3}{N},\\ &\frac{\beta_4 S}{N}, \frac{\beta_1 S}{N}, \frac{\beta_2 S}{N}, \frac{\beta_3 S}{N}, 0, \frac{\beta_4 S}{N}, \frac{\beta_1 S}{N}, \frac{\beta_2 S}{N}, \frac{\beta_3 S}{N}, 0, \frac{\alpha\beta_2 S}{N}, \frac{\alpha\beta_3 S}{N}, 0\Big],\end{aligned} \tag{A7}$$

$$\frac{\partial f_{\mathrm{D}}}{\partial X} = [0,\ 0,\ 0,\ 0,\ \gamma_4,\ 0,\ 0,\ 0,\ 0,\ \gamma_4,\ 0,\ 0,\ \gamma_4]\,, \tag{A8}$$

$$\frac{\partial f_{\mathrm{C}}}{\partial X} = [0, C_{\mathrm{I}}, C_{\mathrm{I}}, C_{\mathrm{I}}, C_{\mathrm{I}}, C_{\mathrm{I}} + \gamma_4 C_{\mathrm{D}}, C_{\mathrm{I}}, C_{\mathrm{I}}, C_{\mathrm{I}}, C_{\mathrm{I}}, C_{\mathrm{I}} + \gamma_4 C_{\mathrm{D}}, C_{\mathrm{I}} + C_{\mathrm{T}}, C_{\mathrm{I}} + C_{\mathrm{T}}, C_{\mathrm{I}} + C_{\mathrm{T}}]\,. \tag{A9}$$

$\psi(t, U)$ collects all the terms with the control variables U_i in the Hamiltonian,

$$\begin{aligned}\psi(t, U) = {} & r_{\max}\max\left(U_2(t) - I_{T2},\ 0\right) I_{\mathrm{D2}}(\theta_{12} - \theta_9)\\ & + r_{\max}\max\left(U_3(t) - I_{T3},\ 0\right) I_{\mathrm{D3}}(\theta_{13} - \theta_{10})\\ & + r_{\max}\max\left(U_4(t) - I_{T4},\ 0\right) I_{\mathrm{D4}}(\theta_{14} - \theta_{11}).\end{aligned} \tag{A10}$$

This has derivatives

$$\begin{aligned}\Psi_9 &= \frac{\partial \psi}{\partial X_9}(t, U) = r_{\max}\max\left(U_2 - I_{T2},\ 0\right)(\theta_{12} - \theta_9),\\ \Psi_{10} &= \frac{\partial \psi}{\partial X_{10}}(t, U) = r_{\max}\max\left(U_3 - I_{T3},\ 0\right)(\theta_{13} - \theta_{10}),\\ \Psi_{11} &= \frac{\partial \psi}{\partial X_{11}}(t, U) = r_{\max}\max\left(U_4 - I_{T4},\ 0\right)(\theta_{14} - \theta_{11}),\\ \Psi_{12} &= \frac{\partial \psi}{\partial X_{12}}(t, U) = -r_{\max}(\theta_{12} - \theta_9) I_{\mathrm{D2}} H(U_2 - I_{\mathrm{T2}}),\\ \Psi_{13} &= \frac{\partial \psi}{\partial X_{13}}(t, U) = -r_{\max}(\theta_{13} - \theta_{10}) I_{\mathrm{D3}} H(U_3 - I_{\mathrm{T3}}),\\ \Psi_{14} &= \frac{\partial \psi}{\partial X_{14}}(t, U) = -r_{\max}(\theta_{14} - \theta_{11}) I_{\mathrm{D4}} H(U_4 - I_{\mathrm{T4}}),\end{aligned} \tag{A11}$$

so that

$$\frac{\partial \psi}{\partial X} = [0, 0, 0, 0, 0, 0, 0, 0, \Psi_9, \Psi_{10}, \Psi_{11}, \Psi_{12}, \Psi_{13}, \Psi_{14}]\,. \tag{A12}$$

The parts of the partial derivative of the all terms adjoint equation arising from the right hand side of the system of differential equations minus all the controls terms $\theta^T \frac{\partial \ell}{\partial I_i}(t, X)$ is

$$\theta^T \frac{\partial \ell}{\partial X_i}(t, X) = h_i \qquad i = 1, 2, 3, \ldots, 14, \tag{A13}$$

where

$$\begin{aligned}
h_1 &= \frac{(N\mu\theta_1 + ((I_{\mathrm{Ua}} + I_{\mathrm{Da}})\beta_a + (I_{\mathrm{U1}} + I_{\mathrm{D1}})\beta_1 + (I_{\mathrm{U2}} + I_{\mathrm{D2}} + \alpha I_{\mathrm{T2}})\beta_2}{N} \\
&\quad + \frac{(I_{\mathrm{U3}} + I_{\mathrm{D3}} + \alpha I_{\mathrm{T3}})\beta_3(\theta_1 - \theta_2)}{N}, \\
h_2 &= \frac{\beta_4 S(\theta_1 - \theta_2) + N[(\mu + d + r_a)\theta_2 - r_a\theta_3 - d\theta_7]}{N}, \\
h_3 &= \frac{\beta_1 S(\theta_1 - \theta_2) + N[(\mu + d + r_1)\theta_3 - r_1\theta_4 - d\theta_8]}{N}, \\
h_4 &= \frac{\beta_2 S(\theta_1 - \theta_2) + N[(\mu + d + r_2)\theta_4 - r_2\theta_5 - d\theta_9]}{N}, \\
h_5 &= \frac{\beta_3 S(\theta_1 - \theta_2) + N[(\mu + r_3 + d)\theta_5 - r_3\theta_6 - d\theta_{10}]}{N}, \\
h_6 &= (\mu + d + \gamma_4)\theta_6 - d\theta_{11}, \\
h_7 &= \frac{\beta_4 S(\theta_1 - \theta_2) + N(\mu + r_a)\theta_7 - N r_a\theta_8}{N}, \\
h_8 &= \frac{\beta_1 S(\theta_1 - \theta_2) + N(\mu + r_1)\theta_8 - N r_1\theta_9}{N}, \\
h_9 &= \frac{\beta_2 S(\theta_1 - \theta_2) + N(\mu + r_2)\theta_9 - N r_2\theta_{10}}{N}, \\
h_{10} &= \frac{\beta_3 S(\theta_1 - \theta_2) + N(\mu + r_3)\theta_{10} - N r_3\theta_{11}}{N}, \\
h_{11} &= (\mu + \gamma_4)\theta_{11}, \\
h_{12} &= \frac{\alpha\beta_2 S(\theta_1 - \theta_2) - N\tau\theta_9 + N(\mu + \tau)\theta_{12}}{N}, \\
h_{13} &= \frac{\alpha\beta_3 S(\theta_1 - \theta_2) - N(\tau\theta_{10} + y_3\theta_{12}) + N(\mu + y_3 + \tau)\theta_{13}}{N}, \\
h_{14} &= -\tau\theta_{11} - y_4\theta_{13} + (\mu + y_4 + \tau + \gamma_4)\theta_{14}.
\end{aligned} \tag{A14}$$