WITHIN-HOST AND POPULATION-LEVEL
MODELING OF HUMAN IMMUNODEFICIENCY
VIRUS AND HEPATITIS C VIRUS DYNAMICS

RUTHIE BREINA BIRGER

A DISSERTATION
PRESENTED TO THE FACULTY
OF PRINCETON UNIVERSITY
IN CANDIDACY FOR THE DEGREE
OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE
BY THE DEPARTMENT OF
ECOLOGY AND EVOLUTIONARY BIOLOGY
ADVISER: BRYAN T. GRENFELL

NOVEMBER 2016
Abstract

The Human Immunodeficiency Virus (HIV) pandemic is one of the largest events to impact global human health in the past 30 years. Epidemic and within-host infection dynamics are affected by co-infections, especially by Hepatitis C Virus (HCV). In this dissertation, I present mathematical modeling analyses across scales to answer the question of which strategies have the greatest impact on controlling the co-epidemics of HIV and HCV. I address this question from the perspective of public health policy and intervention (Chapters 1-2), discussing the epidemiological dynamics for HIV in Newark, NJ, a city with one of the most severe epidemics in the US. This research shows that for there to be significant impact on incidence, care-continuum interventions must be bundled; and Pre-Exposure Prophylaxis (PrEP) is most effective when targeted at specific high-risk populations. These results underscore the need for addressing the “leaky” care pipeline.

I highlight the role of immune function in HCV clearance in a within-host model of HCV/HIV coinfection dynamics that incorporates treatment efficacy. Our analysis sheds light on the tradeoffs involved in choosing between treatment protocols, and how both duration and efficacy need to be taken into account carefully in coinfected patients, especially in light of new direct-acting antiviral treatments (DAAs) that are becoming available (Chapter 3).

Focusing in on HCV mono-infection, I build on the methodology and framework discussed in Chapter 3 to explore HCV’s unusual viral evolution dynamics. Testing various hypotheses including spatial-structure, latency, extra-hepatic replication and selective sweeps in a model of viral evolution can help elucidate HCV within-host dynamics, which can aid in effective treatment design (Chapter 4).

Coinfection and epidemiological modeling are combined in a nested approach that I use to explore relative impacts of antiretroviral and methadone maintenance treatment scale-up, and HCV treatment rollout on HIV/HCV disease burden in Ho Chi
Minh City, Vietnam (Chapter 5). These model results indicate that scale-up of antiretroviral therapy has a major impact on HIV, but a negligible impact on HCV. Methadone scale-up has an impact on both infections, and HCV treatment roll-out can increase multifold the reductions in death rates afforded by the other interventions.
Acknowledgements

My adviser, Bryan Grenfell, for so very cheerfully offering me a summer job back in July 2010. Who knew back then that you were opening the door for me to drag you into the muddy world of HIV modeling? Thank you for so many things: for always having an open door, for always having the two-second solution to a problem I had been puzzling over for weeks, for your unending patience and generosity with this haphazard dissertation that sent me scuttling all over the world looking for data and collaborators, for being such an efficient and no-nonsense proponent of social change within science, for keeping me topped up with my British humor fill when I was pining for the fjords of London, and of course for the EEID dance floor. Quite.

My committee, who have helped me immeasurably in so many ways:

Andrea Graham, for constant amazing insight on coinfection and the actual nitty-gritty biological processes, for chairing my committee meetings, for being such a powerful leader and ally for under-represented groups in science, and again, the EEID dance floor.

Jess Metcalf, for leading me through the workshop trenches, planting the seeds of inspiration for my postdoc work, challenging me to express my scientific ideas clearly, being a most reliable friend in chocolate and coffee, and for a third time, the EEID dance floor.

Andrea and Jess, I want to be both of you when I grow up.

Simon Levin, for always being there during my mathematical existential crises and for the amazing and inspiring range of science that your lab has produced even just during my time at Princeton.

Tim Hallett, for being a truly inspirational HIV modeler, mentor, and vacation buddy. How many people can point to one person and say that that person solved half their dissertation problems, and also their banana boat and karaoke problems?
I have been lucky enough to work with a host of amazing collaborators, who have each in their own way made this dissertation what it is:

Jonathan Dushoff, for being so unbelievably generous with your time and enthusiasm.

Roger Kouyos, for getting me into this whole HIV-HCV business and your calm, collected mentoring that has kept me sane through many chapters.

Sally Hodder and Anushua Sinha, for taking a chance on Bryan’s summer employee for your little Newark HIV modeling project all those years ago. Your clinical expertise, enthusiasm and mentorship have been shining lights these past 6 plus years.

Heidi Robbins, you’re such an inspiration. Thank you for doing all the dirty work, and such stellar clean work too.

Thuy Le, for welcoming me into the OUCRU community and all your patient insight on my project there.

Oli Pybus, Jose Lourenco, and Jayna Raghwani, for my summer at Oxford and bringing me into the wonderful world of phylogenetics.

The research carried out in this dissertation would not have been possible without generous funding from Princeton University, the Center for Health and Well-being, and the Princeton Department of Ecology and Evolutionary Biology (EEB).

Princeton EEB is a truly special place. I am convinced there is nowhere else in the world like it, and to some extent I have dreaded handing in this dissertation because it will mean my time there as a graduate student has come to an end. The people make the department, and I am grateful to each and every one of you for making my time here so special. Huge thanks to the administrative staff, especially Lolly O’Brien and Bernadette Penick for my sanity, among other things. In addition to my committee, I have received so much on-the-fly support and mentorship from the rest of the faculty, especially Corina Tarnita and Bridgett vonHoldt. I feel so lucky to have been a part of the Grenfell lab, with its illustrious members past and present.
from whom I have learned so much, especially Quentin Caudron for the Scotch and coffee and Alex Becker for the mezels, and the undergraduates I had the privilege of mentoring: Lisa Gong, Alex Jafari, and Sanjay Rao. From lab meetings to seminar lunches to beer hour to the EEB Holiday party, the EEB grad student and postdoc community has become like a family to me these past five years. I am grateful to everyone in it, and especially Molly Schumer, Emma Fuller, Anieke van Leeuwen, Christina Faust, Carla Staver, Albert Kao, Lisa McManus, Joe Bak-Coleman, Thom van Boeckel, Amy Wesolowski, Romain Garnier, Adrienne Tecza, my fellow grad-reps Sam Rabin and Josh Daskin, and of course Cleo Chou and Jennifer Schieltz (in your case literally) for making this place feel like home. I have also gotten so much support from outside the EEB bubble, special thanks to Stephanie Tung, Elena Krieger, Robin Dembroff, The Ladies’ High Tea Society, the Princeton LGBT Center, Hannah Babiker, who pulled me back up on my feet and kept me as a friend even when my dissertation crashed our holiday, and all my other friends out there who helped me make this happen.

Lastly, I want to thank my family. To my uncle Roger Glass, thank you for planting the spark of excitement about epidemiology in me on those family vacations all those years ago, and for continuing to help and open doors for me all along the way. To the Kaissars for all the love and noise, especially Molly —it was such a privilege to be in science school at the same time as you. To my parents, Ellen Glass Birger and Daniel Birger, thank you supporting me, feeding me, housing me, letting me run Hotel Ruthie out of your house —even though I didn’t go to medical school. Thank you for everything; I love you so much.
To my parents, Ellen Glass Birger and Daniel Birger; and to my grandmother,

Elinor “Tootsie” Werblin Glass (Dec 3, 1915 – Aug 14, 2016)
Contents

Abstract .......................................................... iii
Acknowledgements .................................................. v
List of Tables ..................................................... xi
List of Figures ..................................................... xiii

Introduction ................................................................ 1

1 Modeling the Impact of Interventions along the HIV Continuum of Care in Newark, New Jersey .............................................. 14

2 Impact of Pre-Exposure Prophylaxis with Tenofovir Disoproxil Fumarate/Emtricitabine on Human Immunodeficiency Virus Incidence in Newark, New Jersey .............................................. 36

3 Modeling the Effect of HIV Coinfection on Clearance and Sustained Virologic Response during Treatment for Hepatitis C Virus ........................................................................... 61

4 Modeling within-host HCV lineage patterns: a mechanistic approach .............................................................................. 83
5 The Impact of HCV Treatment roll-out in a high HIV-HCV Prevalence Population: A modeling study on People Who Inject Drugs in Ho Chi Minh City, Vietnam

Conclusion

Bibliography

Appendices

A Chapter 1 Supplementary Information

B Chapter 2 Supplementary Information

C Chapter 3 Supplementary Information

D Chapter 4 Supplementary Information

E Chapter 5 Supplementary Information
## List of Tables

1.1 Prior and Posterior parameter ranges ........................................... 21
1.2 Invariant Parameters ................................................................. 22
1.3 Impact of Interventions on Incident Cases ................................. 29
1.4 Impact of Interventions on Deaths ............................................... 30
2.1 Summary of PrEP Parameter Values in the Model ....................... 42
4.1 Model parameter estimates ......................................................... 93
4.2 Summary statistics for phylogenetic macro-signatures ................ 95
4.3 Data Estimates ............................................................................ 96
4.4 Top Two Statistics for Distinguishing from Null Non-Spatial Model 97
5.1 Parameter Values ......................................................................... 116
5.2 Maximum Likelihood Estimates of Fitted Parameters ................ 118
5.3 Cost in Dollars per Life-Year Saved for Different Intervention Combinations ........................................................................ 121
A.1 Partial Rank Correlation Coefficients ............................................. 170
B.1 Epidemiological Profile of the HIV/AIDS Epidemic in Newark in 2016 187
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>B.2 Number Needed to Treat (NNT) with PrEP to Avert One HIV Infection</td>
<td>188</td>
</tr>
<tr>
<td>in 2026 and Number of Cumulative PrEP Person-Years Needed to Avert</td>
<td></td>
</tr>
<tr>
<td>One HIV Infection 2016-2026</td>
<td></td>
</tr>
<tr>
<td>B.3 Literature Review of Standardized Metrics for Comparative Analysis</td>
<td>189</td>
</tr>
<tr>
<td>C.1 Parameter ranges from DebRoy et al 2011 [1]</td>
<td>191</td>
</tr>
<tr>
<td>C.2 Data Estimates</td>
<td>192</td>
</tr>
<tr>
<td>C.3 Parameter sets for Figs 2-6</td>
<td>192</td>
</tr>
</tbody>
</table>
List of Figures

1.1 **Graphical Representation of Model Flow and Risk Groups interactions.** (a) Susceptible individuals flow into the Acute Infection compartment via the force of infection. Acutely infected individuals flow into Asymptomatic Undiagnosed. Asymptomatic Individuals either get diagnosed before progressing to AIDS or not. Not pictured is the tested-ineligible category as this category no longer exists by the time interventions are implemented. Tested-eligible individuals can be lost to follow up or not linked to treatment or be linked to treatment and then either virologically suppress or not suppress. Unsuppressed individuals move into a dead-end AIDS post-ART compartment. Finally, Undiagnosed and Unlinked individuals progress to AIDS-never treated from which they either die or present for treatment and then either suppress or not. Flows are in blue arrows and excess deaths are in red. (b) Male and Female risk groups are shown with arrows indicating groups that most strongly assort.
1.2 **Testing and Care-continuum Interventions.** (a) Current, Modest
Intervention and Moderate Intervention levels are shown of Percentages
of individuals tested before reaching CD4<200, in Current, Modest In-
tervention and Moderate Intervention (b) Percentages of individuals
achieving each step in the care-continuum, in Current, Modest Inter-
vention and Moderate Intervention levels. . . . . . . . . . . . . . . . . 23

1.3 **Model Fits and Data.** Data from Project Impact (solid lines with
dots) are shown, overlaid with the range of model fits for each data
point. The twelve risk groups in the model are aggregated into six
shown here as follows: Het F comprises Low and High Risk non-IDU
Heterosexual Females, IDU F Low and High Risk IDU Heterosexual
females, Het M Low and High Risk non-IDU Heterosexual Males and
non-IDU Bisexual Males, MSM Low and High Risk non-IDU MSM,
IDU M Low and High Risk IDU Heterosexual Males and IDU Bisexual
Males, MSM IDU are just the IDU MSM group. . . . . . . . . . . . . . . . 25

1.4 **Incident Cases and Deaths Time Series, Percent Change Over
Time by Intervention (according to best-fit parameter set).**
(a) shows the projected change in incident cases over time under in-
creasing combinations of the top four most efficacious interventions.
(b) shows the projected change in deaths over time under increasing
combinations of the top four most efficacious interventions . . . . . 27

1.5 **Pressure Test of Interventions.** Impact on annual number of new
infection incidence of maximal implementation of all interventions. By
2023, there are fewer than 200 incident cases. . . . . . . . . . . . . . . . 28

1.6 **Pie-Charts of Incidence Sources.** Pie Charts designating prove-
nance of incident cases in 2023 for Baseline and Top Four most effica-
cious interventions at Moderate level. . . . . . . . . . . . . . . . . . . . . 29
2.1 PrEP compartmental model flow-chart and intervention parameters. Compartments for PrEP-users with high and low user-adherence ($P_H$ and $P_L$), as well as the compartments for the HIV care-continuum. Susceptible individuals flow into a PrEP compartment based on a population coverage scale-up equation, $H$. A proportion, $q$, of PrEP-users have high user-adherence, which is defined as 90% of sex-acts or intravenous injections protected by PrEP with 90% or greater efficacy. High-adherers take 3 or more daily doses of PrEP per week. $1-q$ is the proportion of PrEP-users with low user-adherence, defined as 30% of sex-acts or intravenous injections protected by PrEP with 90% or greater efficacy. Low-adherers take 2 or fewer doses of PrEP per week. $q$ is the population-adherence PrEP-users can still become infected and flow into HIV Acute Infection compartment, $I_0$. PrEP forces of infection, $\lambda_{p-low}, \lambda_{p-high}, \lambda_{p-drug-low}, \lambda_{p-drug-high}$, are influenced by PrEP efficacy and user-adherence. Interventions modeled assessed combinations of coverage levels (10%, 25%, and 50%) and population-adherence, $q$, levels (15%, 60%, and 95%) for 5 targeted populations: HRH, MSM, IDU, All high-risk groups, and the total population. Abbreviations: HRH= high-risk heterosexuals, MSM= men who have sex with men, IDU= injection drug users, TOT= total population.
2.2 Cumulative HIV incident cases averted 2016-2026 by PrEP interventions. Cumulative incident cases averted by each targeting strategy for intervention combinations of population coverage and population-adherence. Bars represent 95% confidence intervals. Abbreviations: HRH= high-risk heterosexuals, MSM= men who have sex with men, IDU= injection drug users, AHR= all high-risk groups, TOT= total population.

2.3 Herd effects of PrEP. The indirect effects of incident cases averted in other populations in addition to the targeted population by intervention combinations of population coverage and population-adherence. Abbreviations: HRH= high-risk heterosexuals, MSM= men who have sex with men, IDU= injection drug users, AHR= all high-risk groups, TOT= total population.

2.4 Percent reduction in HIV incidence by PrEP interventions. Impact of PrEP targeting strategies for intervention combinations of population coverage and population-adherence, with respect to the total population HIV incidence. Abbreviations: HRH= high-risk heterosexuals, MSM= men who have sex with men, IDU= injection drug users, AHR= all high-risk groups, TOT= total population.

2.5 Percent reduction in HIV incidence by PrEP interventions. Impact of PrEP targeting strategies for intervention combinations of population coverage and population-adherence, with respect to the HIV incidences of the specific population targeted. Abbreviations: HRH= high-risk heterosexuals, MSM= men who have sex with men, IDU= injection drug users, AHR= all high-risk groups, TOT= total population.
2.6 Median number of PrEP person-years required to prevent one new HIV infection over a 10-year period (2016-2026). Intervention combinations of population coverage and population adherence produced a range of person-years, which was used to calculate the median number. The range is shown in parentheses. Abbreviations: HRH = high-risk heterosexuals, MSM = men who have sex with men, IDU = injection drug users, AHR = all high-risk groups, TOT = total population.

3.1 Within-Host HIV-HCV Interactions. Dashed arrows represent mechanisms of HIV-HCV interaction described in the literature. Blue arrows represent population-biological effects taken into account in our model.

3.2 Model Comparison with Data. Panel a) shows patient data redrawn from (2), displaying the five types of long-term treatment outcome: SVR, Null Response, Partial Response, Breakthrough, and Relapse. Panel b) shows model scenarios from the model including free virus displaying these same five outcomes. Panel c) shows short-term treatment dynamics over the first two weeks of treatment; model outputs are paired with data redrawn from (3). Panel d) shows the same scenarios as b), using a version of the model wherein the quasi-steady state assumption has been implemented, demonstrating that the long-term dynamics can be recovered. Parameter values for each simulation are listed in Supp. Table C.3.
3.3 Bifurcation Diagram of Stable and Unstable Equilibrium

HCV Viral Loads with Varying CD4\(^+\) count. This figure shows the growing bistable region and corresponding shrinking region of spontaneous clearance for a parameter set with varying \(\delta\) and within-host \(R_0\). The x-axis represents what the within-host \(R_0\) would be in the HIV-free scenario the given clearance rate parameters used for the varying HIV scenarios. The y-axis represents equilibrium HCV viral load. To the left of the bifurcation points is the zone of spontaneous clearance. If viral load can be pushed below the dotted lines (unstable equilibria), with sufficient treatment, the patient will then be in the region of stable DFE and will be able to achieve SVR. \(\delta \times (1 + \alpha H)\) ranges from from \(d\) to \(3\) for varying CD4\(^+\) counts\(^*\). The other parameters are as follows: \(s = 4365, T_{\text{max}} = 4.016 \times 10^6, d = 1.06 \times 10^{-3}, \beta_c = 7.3 \times 10^{-8}, p = 13.48, c = 10.06, r_1 = 2.7, r_2 = 7.52, s_H = 9, \beta_H = 4.1 \times 10^{-6}, d_H = 9 \times 10^{-3}, \alpha = 5 \times 10^{-3}, \gamma = 2 \times 10^{-8}\). *Note: For lower CD4\(^+\) counts/higher HIV viral loads, the tips of the bifurcation trees could not be reached with clinically realistic values of \(\delta\). This fact implies that there would be no realistic zone of spontaneous clearance for this patient when CD4\(^+\) count falls below a certain point, though SVR might still be possible with sufficient treatment duration and efficacy.
3.4 Effect of depleted CD4$^+$ on Infection Clearance. This figure shows HCV viral load trajectories for two sample patients with different treatment efficacies and durations under initial conditions of normal CD4$^+$ count ($\sim 1000 \, \mu L$, HIV negative), and depleted CD4$^+$ count ($\sim 600/\mu L$, suppressed HIV with incomplete immunologic recovery), and very low CD4$^+$ count ($\sim 10/\mu L$, unsuppressed HIV) for treatment courses of 24 and 48 weeks. It can be seen that a treatment course of 24 weeks that is sufficient to achieve SVR when the patient has a normal CD4$^+$ count is no longer sufficient for either HIV positive scenario, even when CD4$^+$ count has partially recovered. However, when treatment is extended to 48 weeks, the patients are able to achieve SVR when CD4$^+$ count has partially recovered. Panel a) shows these scenarios for Patient I, a theoretical patient exhibiting classic bistability, while panel b) shows Patient II, a theoretical patient who requires the cure boundary condition to achieve SVR.

3.5 Treatment and Efficacy Combinations for a) Patient I and b) Patient II. Each point on each line denotes the minimum treatment duration necessary to achieve SVR for the corresponding efficacy, for varying initial CD4$^+$ counts. For any combination above a line, SVR is achieved, while for any combination below the line, SVR is not achieved. All scenarios were run for 96 weeks to assure that any potential relapse was captured.
3.6 **High-Efficacy Treatment with short-term cure.** Panels a) and b) demonstrate that high efficacy treatment, as with DAAs, can result in SVR over a short duration (12-24 weeks) for Patients I and II, respectively, even with incomplete CD4$^+$ recovery (though when each patient is severely immunocompromised, short term cure is still unlikely). However, panels c) and d) shows the same scenarios, but with treatment efficacy declining from 95% to 70-75% efficacy over the course of 12-16 weeks. When the CD4$^+$ count is depleted or low (HIV-positive), the patients respond initially, but relapse after treatment efficacy falls below a certain level.

4.1 **Model Flowchart.** The colored boxes represent different mechanisms tested. Infection occurs between free virions and target cells as well as between infected cells and target cells.

4.2 **Null Model Cell and viral dynamics.**

4.3 **Lattice representing Hepatocyte structure.** Note that dead infected cells may still contain HCV RNA, representing the transition to cirrhotic scar tissue that occurs in HCV patients there is evidence that HCV persists in cirrhotic tissue (4, 5). However, dead cell count is not included in viral load calculations.
4.4 **Null Non-spatial vs Non-Spatial Mechanistic Trees.** The left panel shows evolution trees generated by each non-spatial model with lowest and highest parameter values. For the latent model, the parameter is proportion of cells going long-term latent; for the Two-Patch model the parameter is the coefficient of inter-patch mixing; and for the Selective Sweep model, the parameters are selective fitness advantage and mutation rate (to fitter strain, distinct from evolutionary mutation rate). The right panel shows evolution trees for each version of the null model. 98

4.5 **Null Non-spatial vs Spatial Mechanistic Trees.** The left panel shows evolution trees for the null non-spatial model and blood-dominated spatial versions of each of the mechanistic model with lowest and highest parameter values (as described in caption for Figure 4.4). The right panel shows the same for cell-dominated spatial versions. 99

4.6 **Latent Model RFA Accuracy.** This figure shows the proportion of model runs versus the Null NonSpatial model that the RFA accurately categorizes by the proportion of cells going long-term latent for each type of model. The cell-dominated spatial model is most consistently identified correctly, even at very low latency proportion. 100

4.7 **Two-Patch Model RFA Accuracy.** This figure shows the proportion of model runs versus the Null NonSpatial model that the RFA accurately categorizes by the inter-patch mixing coefficient for each type of model. The cell-dominated spatial model is most consistently identified correctly. 100
4.8 Selective Sweep Model RFA Accuracy. This figure shows the proportion of model runs versus the Null NonSpatial model that the RFA accurately categorizes by selective fitness advantage for each type of model. The cell-dominated spatial model is most consistently identified correctly.

4.9 Nodetime value by model. This figure shows the variation in the value of the nodetime statistic with parameter value for each model.

4.10 Example Latent Model Cell and viral dynamics

5.1 Diagram of Model Flows. This figure shows a graphical representation of the model pathways. Full equations and descriptions can be found in the supplementary information.

5.2 Model Fit to HIV and HCV Prevalence. This figure shows the range of model estimates for HIV (blue) and HCV (green) prevalence among PWID in the shaded regions, with the estimate from the best-fit parameter set represented by the dashed line. Data estimates and corresponding confidence intervals to which the model was calibrated are represented by circles and error bars.

5.3 ART Scale-up, Incidence and Deaths changes over time. Each panel in this figure shows a plot of reductions in HIV and HCV incidence, prevalence or deaths with varying ART scale-up.

5.4 MMT Scale-up, Incidence and Deaths changes over time. Each panel in this figure shows a plot of reductions in HIV and HCV incidence, prevalence or deaths with varying MMT scale-up.

5.5 ART and MMT Scale-up, Reductions in deaths from disease over time. Each panel in this figure shows a plot of reductions in deaths from disease 10 years after interventionscale-up with varying ART and MMT scale-up.
5.6 HCV Treatment, Reductions in deaths from disease over time.
Each panel in this figure shows a plot of reductions in deaths from disease 10 years after roll-out of HCV treatment coverage, with maximum previous scale up of ART and MMT coverage (80% and 50%).

A.1 Best-fitting Parameter Set. Model output of best-fitting parameter set in solid lines, with data points on connected dots.

A.2 Contour Plots of Incidence and Deaths Reductions. (a) Shows contours of equivalent incidence reduction for combinations of the two most efficacious interventions. (b) Shows contours of equivalent reductions in deaths for combinations of the two most efficacious interventions.

B.1 Reported impact of PrEP in populations from different modeling studies. Impact, shown by the size of the circle, is the percent change in cumulative infections averted in the population of interest for the various interventions in each study. Panel A: PrEP interventions targeting MSM from different modeling studies on the basis of HIV prevalence in the MSM population, effectiveness, coverage, and resulting impact. Panel B: PrEP modeling studies of the total population. Abbreviations: HRH= high-risk heterosexuals, MSM= men who have sex with men, MMSM= men who mostly have sex with men, IDU= injection drug users, AHR= all high-risk groups, TOT= total population.

C.1 CD4+ counts over time. This figure, re-drawn from (6), demonstrates how CD4+ after ART initiation can be incomplete. It is strongly associated with CD4+ count at treatment initiation, p<.05 for all comparisons.
D.1 Example Non-Null Model Cell and viral dynamics

D.2 Latent Model RFA Accuracy. This figure shows the proportion of model runs versus the Null NonSpatial model that the RFA accurately categorizes by the proportion of cells going long-term latent for each type of spatial model with varying infinite sites mutation rates. As mutation rate increases, the ability of the RFA to distinguish models increases.

D.3 Null Non-spatial vs Non-Spatial Mechanistic Trees. The left panel shows pair-wise diversity, segregating sites, and Tajima’s D for the evolution trees generated by each non-spatial model with lowest and highest parameter values. For the latent model, the parameter is proportion of cells going long-term latent; for the Two-Patch model the parameter is the coefficient of inter-patch mixing; and for the Selective Sweep model, the parameters are selective fitness advantage and mutation rate (to fitter strain, distinct from evolutionary mutation rate). The right panel shows evolution trees for each version of the null model.

D.4 Null Non-spatial vs Spatial Mechanistic Trees. The left panel shows pair-wise diversity, segregating sites, and Tajima’s D for the evolution trees for the null non-spatial model and blood-dominated spatial versions of each of the mechanistic model with lowest and highest parameter values (as described in caption for Figure 4 in the main manuscript). The right panel shows the same for cell-dominated spatial versions.

E.1 ART Coverage in Vietnam. The first panel shows the total number of HIV patients on ART over time (both adults and children), while the second panel shows the proportion on ART who are eligible (i.e. CD4 count <200).
E.2 Incidence and Deaths changes over time. Each panel in this figure shows a plot of reductions in HIV and HCV incidence, prevalence or deaths with varying ART scale-up. ........................................... 208

E.3 Incidence and Deaths changes over time. Each panel in this figure shows a plot of reductions in HIV and HCV incidence, prevalence or deaths with varying MMT scale-up. ........................................... 209

E.4 Incidence and Deaths changes over time. Each panel in this figure shows a plot of reductions in HIV and HCV incidence, prevalence or deaths with varying ART scale-up. ........................................... 209

E.5 Incidence and Deaths changes over time. Each panel in this figure shows a plot of reductions in HIV and HCV incidence, prevalence or deaths with varying MMT scale-up. ........................................... 210

E.6 Model Fit to HIV and HCV Prevalence. This figure shows the range of model estimates for HIV (blue) and HCV (green) prevalence among PWID in the shaded regions, with the estimate from the best-fit parameter set represented by the dashed line. Data estimates and corresponding confidence intervals to which the model was calibrated are represented by circles and error bars. ........................................... 211

E.7 Incidence and Deaths changes over time. Each panel in this figure shows a contour plot of reductions in HIV and HCV incidence or deaths, or total deaths, with a combination of MMT and ART interventions. Each black line represents set of combinations of ART and MMT that yield the same reduction. ........................................... 211

E.8 PWID Population Dynamics over time. The model replicates the gradual increase in the PWID and ex-PWID populations over the past decades. ........................................... 212
E.9 **HCV Incidence changes over time with HCV treatment and MMT scale-up, High Treatment Efficacy.** This figure displays incident cases over time under three different MMT scale-up scenarios:

- No scale up (solid lines), 10% scale-up (dotted lines), 50% scale-up (x-lines), and three different HCV treatment scale-up: No scale-up (black), 10% scale-up (dark grey), 50% scale-up (light grey). Prior scale-up of MMT coverage lays the foundation for HCV treatment roll-out in a lower incidence and prevalence population, and without MMT scale-up, HCV treatment roll-out makes little difference to HCV incidence even at high coverage.
Introduction

Infectious disease dynamics are a major force shaping human ecology both at population and within-host levels. One of the most significant events in disease ecology over last 30 years was the arrival of the Human Immunodeficiency Virus/Acquired Immune Deficiency Syndrome (HIV/AIDS) pandemic. In turn, one of the most important coinfections of HIV is Hepatitis C Virus (HCV), which co-occurs frequently among people who inject drugs (PWID) and which can be profoundly impacted by the presence of HIV infection. Much work has been conducted in the field of biomedicine on HIV and HCV, and its effects scale from individual, within-host impacts to demographic perturbations and viral evolution. This is reflected in a long literature on parasite population dynamics and public health interventions, both rooted in dynamical models. In my thesis, I aim to deploy mathematical and computational models in tandem with insights from immunology, molecular biology and clinical work to address a series of novel questions in the dynamics of HIV and HCV at population and within-host levels. The thesis is presented as a series of self contained papers (Chapters 1-5). In this general introduction, and the concluding chapter, I review and discuss common themes and questions. In the rest of this chapter, I provide background on each pathogen and the role of models in its study, a discussion of open questions, and a dissertation prospectus.
Human Immunodeficiency Virus

HIV is a retrovirus that causes Acquired Immune Deficiency Syndrome (AIDS) by attacking the immune system mostly through infection of CD4+ cells (7). Since HIV was first reported in 1981 (8), it has spread to every continent; recent estimates indicate that approximately 35 million people are currently living with HIV-AIDS (9). HIV infection occurs in various stages with differing transmission probabilities. The early, acute stage is associated with high viral load and increased probability of transmission. It is followed by the asymptomatic stage, which begins with some CD4+ recovery and drop in viral load to a set point associated with lower transmission probability. After a median of roughly 8 years post-infection, viral load (and transmission probability) begin to increase rapidly and CD4+ count declines until death (10, 11). HIV can be transmitted through blood and mucous membranes as well as breast milk. Risk factors for HIV acquisition include unprotected sexual contact (vaginal and anal), injecting drug use, and contact with unsafe medical equipment or blood products (9). Initially fatal in almost all cases, effective chemotherapy has been available for HIV since the mid-late 1990’s and can vastly improve life expectancy as well as reducing transmissibility (12). However, treatment is expensive so often not available in low-income settings, and an additional challenge is that evolution of resistance is possible even when adherence to treatment is relatively high (13).

Hepatitis C Virus

Hepatitis C is a member of the Flaviviridae family and Hepacivirus genus that was isolated as a cause of non-A non-B viral hepatitis in 1989 (14). It is most efficiently transmitted through blood from contaminated needles or medical equipment, though it can also be transmitted percutaneously (15). The global prevalence of HCV is estimated to be on the order of 2.2-3%, though it varies widely between countries. In
many low-income countries, unsafe medical and blood products are the main causes of transmission, but in high-income countries and some middle- and low-income settings, the primary mode of transmission is via injecting drug use (16). The natural history of HCV infection is complex, with many potential outcomes. Infected individuals experience an acute stage of infection which is usually asymptomatic, after which some percentage of people will clear the virus. Upon clearance, there is some protective immunity though reinfection can be possible (17, 18). Both acute and chronic HCV have few symptoms, though some acutely infected individuals experience clinical symptoms including jaundice (14). Estimates vary, but the majority of individuals do not clear acute infection and go on to experience persistent viremia. Many of these individuals experience progressive fibrosis of the liver, which can ultimately lead to cirrhosis and/or hepatocellular carcinoma and death. The time span of this progression is extended (20-40 years), though certain factors such as male age, alcohol consumption, older age at infection and HIV coinfection are associated with increased progression rates of fibrosis (19). Treatment in the form of pegylated interferon-α and ribavirin is available for HCV, though efficacy is limited (15). New directly-acting drugs are starting to be made available that may have better efficacy, but also great risk of resistance evolution and high price (20, 21).

Role of Mathematical Models in clarifying disease dynamics and shaping control strategies

For infectious diseases in general and for HIV and HCV in particular, models have played a major role in understanding disease dynamics and sculpting interventions at both the epidemiological and within-host levels.
HIV at the Population Level

Early models of the HIV epidemic attempted to make short term predictions about HIV prevalence using simple, regression-based prediction methods (22, 23). In the early 2000’s, UNAIDS designed the Estimation and Projection Package (EPP) software to provide national estimates of prevalence trends in countries with generalized epidemics based on local sub-epidemic data. In addition to providing broad estimates for future trends in HIV prevalence, these types of models can be useful for baseline secular comparison with results from more mechanistic transmission models.

HIV is a sexually transmitted infection, but unlike some other STIs (e.g. *Neisseria gonorrhea, Chlamydia trachomatis*), it has a relatively low probability of transmission per sex act (24). The probability of transmission of infection can thus be quite sensitive to number of sex-acts within a partnership. Network models can be useful for exploring the role of multiple partnerships and concurrency in HIV models as well as general STI models (25, 26).

The population-level models used in this thesis (Chapters 1, 2, and 5) are primarily of a third type that will be discussed in more detail. Deterministic compartmental models of disease transmission have been used to explore the dynamics of epidemics since Daniel Bernoulli first proposed a set of equations for smallpox (reprinted in (27)). In the early 20th century, Kermack and McKendrick (28) published their seminal paper on SIR models, which has informed modern epidemiology. These models are extremely versatile and can be used to capture the dynamics infections ranging from seasonal influenza (e.g. (29, 30)) to hemorrhagic fever (e.g. (31)). They have informed and generated a wealth of literature, e.g. (32, 33, 34, 35, 36, 37, 38). In the early stages of the HIV epidemic, the bulk of recognized cases were in the developed world, and prevalence was particularly high among men who have sex with men (MSM). Modeling efforts focused on modeling sexual transmission of HIV in homogeneous groups could thus lend some insight into patterns of the early epidemic.
Developments followed through the years including expansion of models to include 2 sexes, differential transmission probability by disease stage, heterogeneity in transmission and transmission via other means, such as injection drug use \((11, 39, 40, 41)\).

A “second generation” of models began to appear after treatment and prevention methods had been identified for HIV. These models can be used either to assess retroactively the impact of interventions or secular changes or to predict the future impact of potential interventions \((42, 43)\). In particular, a paper published in 2009 by Granich et al has generated much discussion about the efficacy of interventions that involve expanded testing and treatment \((44)\). The study used data from ante-natal clinics in South Africa, representing a generalized, heterosexual epidemic, and made very favorable predictions about the impact of Test-and-Treat strategies, suggesting they could send the epidemic into elimination phase within 5 years. This optimistic result was met with skepticism by the HIV modeling community, and has spurred a field of inquiry into assessing more nuanced impacts of Test-and-Treat, and other interventions such as Pre-Exposure Prophylaxis (PrEP) using models. Chapters 1, 2, and 5 of this dissertation contribute to this field, and offer further review of the literature.

**HCV at the Population Level**

The modeling history of HCV at the population level has a smaller literature but a similar trajectory. As with HIV, several models have used trends in HCV prevalence over time to estimate progression rates and incidence of disease as well as to generate estimates of future disease burden \((45, 46)\). This approach is very useful for populating the current estimates of progression rates of various stages of fibrosis as well as generating estimates of the ranges of future prevalence in large-scale settings.

As with HIV, network models of HCV transmission are useful for exploring the effect of behavioral dynamics. Injecting Drug User (IDU) behavior rather than sexual
behavior, however, is the focus. These network models are able to lend useful insight into control methods such as behavior change and harm reduction.

Most relevant for this dissertation are deterministic compartmental models for HCV transmission that focus on transmission of HCV between injecting drug users. In high and in many low-income countries now, HCV is most prevalent among IDUs. It is notoriously difficult to collect data on injecting drug users—the most reliable data often comes from respondent-driven sampling—but deterministic models can help identify which parameters and data are most important to collect in terms of intervention planning, as well as making predictions on the efficacy of various interventions.

Treatment is difficult for HCV, and while it can be effective, it is usually only available in high income countries. Assessing HCV treatment in models is growing in importance with the advent of new, direct-acting antiviral drugs (DAAs). DAAs have good safety profiles, low risks of drug interactions, high tolerability, and can be administered via short courses, in addition to being very potent with >90% cure rates; this makes them better options than interferon-ribavirin (IFN-RBV) combination therapies that have been in use up until now and required long treatment courses with lower efficacy and severe side-effects. As with HIV, treatment can have an effect on onward transmission by reducing HCV viral load, especially if the treated patient ultimately achieves Sustained Virologic Response (SVR). A modeling literature similar to Test-and-Treat for HIV is starting to be formed for HCV e.g., but further work is necessary to assess treatment interventions, especially in the context of HIV coinfection.

**Within-host Dynamics**

Some of the most notable successes in within-host modeling have come from the study of HIV and HCV dynamics. In 1995, David Ho and Alan Perelson published a study
illuminating the dynamic nature of HIV viral load equilibrium, using a very simple mathematical model (56). Recognizing that at equilibrium, the viral production rate had to be balanced by the viral clearance rate, they were able to calculate the viral production rate by observing the decline in viral load in patients after initiation of treatment that blocked viral production. The insight that viral production rate is extremely rapid continues to be helpful in developing HIV therapies two decades after its discovery. Other models exploring latency, mutation, and other phenomena in HIV have similarly assisted with HIV control and therapy development (57, 58, 59, 60).

The first within-host models of HCV aimed to capture the dynamics of infection by pairing a mathematical model with patient data from trial of varying doses of interferon (IFN) (3, 61). The trial results showed a rapid initial decline in viral load after treatment, followed by a period of stasis or slow decline. The model of viral infection used was similar to a number of previous HIV models, and the results highlighted the likely mechanisms of how treatment with interferon (IFN) lowered viral load: the treatment patterns observed (i.e. rapid initial decline in viral load followed by extended slow decline) were consistent with the main mechanism of interferon being reduction in production of new virus by infected cells. While these models dealt with short-term dynamics of HCV immediately following treatment, later models were able to expand on this framework to address longer-term dynamics over time-frames in consistent with treatment duration (12-48 weeks), to make relevant clinical recommendations (62, 1).

As with population level models, within-host models of HCV are of growing current importance with the advent of DAAs. HCV undergoes rapid within-host evolution and exists in most patients as a quasi-species (63, 64). Pinning down the mechanisms of within-host transmission and treatment function will be crucial for preventing or slowing the acquisition of drug resistance to DAAs, especially in the

\[1\text{In the context of this thesis, quasi-species simply refers to a genetically variable within-host population and not classic the quasi-species model proposed by Eigen and Schuster (65).}\]
context of HIV coinfection. Chapter 3 provides further background detail on the mechanisms of HIV-HCV coinfection.

Open Questions

Several decades of research in modeling the dynamics of HIV, HCV and their interactions have delineated a range of key open questions, some of which I aim to address in this dissertation. Addressing these fundamentally cross-scale problems is a challenge, and here we utilized models that take advantage of the emergent simplicity in population-level interactions, while underpinning these models with an in-depth analysis of the within-host processes.

Since the controversial Granich paper in 2009 (44), many researchers have constructed similar models to evaluate the full range of impact of Test-and-Treat interventions. However, the care-continuum for HIV in the US and other developed countries is complex and involves many steps in addition to testing and treating, including retention in care and PrEP. Chapters 1 and 2 outline models that take into account these complexities. In Chapter 1, we evaluate the impact of each individual step along the HIV care-continuum, and make predictions about the efficacy of different interventions in the developed-world, impoverished urban setting of Newark, NJ. Chapter 2 adds PrEP into the slate of HIV interventions assessed, and there we find that when PrEP is made available to certain high-risk groups (e.g. MSM), it can have significant impact. The models used in both of these chapters are epidemiological, but they rely strongly on incorporation of the within-host dynamics of HIV that have been illuminated by the studies mentioned above. HIV viral load fluctuates over the course of infection in a characteristic pattern: there is a brief period during acute infection where viral load is very high; it then drops down to a set-point and remains there for a long duration until either the patient is treated and the viral load drops, or
the patient progresses to the AIDS stage of infection, and the viral load climbs again. Viral load is monotonically correlated with infection transmission probability, so patients have varying transmission probabilities during the course of infection especially if they are treated (transmission probability drops by several orders of magnitude). It is thus crucial to include this within-host insight into epidemiological models, and take a cross-scale approach to understanding the epidemic.

Cross-scale dynamics are an equally useful lens through which to study HCV, which is one of the most important coinfecting pathogens of HIV, particularly in developing world settings and among People who Inject Drugs (PWID). However, most likely due to the challenges of navigating two such complex, chronic infections and the complexity of their interactions, very few models thus far have looked at modeling HIV-HCV co-epidemics explicitly (66, 67), and the first model of within-host coinfection is described in Chapter 3 of this dissertation (68). With this type of within-host model, as well as data from clinical studies, we can begin to address the cross-scale dynamics of coinfection with these pathogens. With the advent of new treatments, it is becoming increasingly important to understand HCV within-host transmission and evolution. In Chapter 4 we simulate lineage evolution under different mechanistic models of within-host HCV transmission. Using these simulations, we explore methods for distinguishing between them in order to be able to analyze patient data to uncover mechanisms of transmission and evolution that may be crucial for designing treatment strategies.

Lastly, chapter 5 builds on the coinfection dynamics explored in Chapter 3 at a population level, modeling the HIV-HCV co-epidemics among people who inject drugs (PWID) in Ho Chi Minh City, Vietnam. Co-epidemic models such as this one have the capacity to make comparative predictions about the efficacy of interventions to address either or both epidemics, which is crucial for policy planning.
Prospectus

This dissertation comprises five main projects, as outlined below.

Chapter 1 focuses on modeling public health interventions for HIV in Newark, NJ. The HIV epidemic in Newark is among the most severe in the United States, with prevalence ranging up to 3.3% in some subpopulations. The aim of this project was to design a novel mathematical model of the epidemic in Newark that incorporated interventions along a continuum of care in order to assess which strategies could have the greatest impact on controlling the epidemic.

Interventions assessed were: increasing proportions tested, linked, and retained in care; linked, and adherent to treatment; and increasing testing frequency, high-risk-group testing, and adherence. The most effective interventions for reducing incidence were improving treatment adherence and increasing testing frequency and coverage. The most effective interventions for reducing deaths were increasing retention, linkage to care, testing coverage, and adherence.

We conclude that reducing HIV deaths in Newark over a 10-year period may be a realizable goal, but reducing incidence is less likely. Our results highlight the importance of addressing leaks across the care continuum, and reinforcing efforts in prevention of new HIV infections with additional interventions.

Chapter 2 extends the model proposed in Chapter 1 to include PrEP among the interventions assessed. Recent guidelines recommend PrEP for use by homogeneous sub-populations of high-risk individuals (e.g. men who have sex with men (MSM), high-risk heterosexuals (HRH), injection drug users (IDU)) to prevent HIV acquisition. Studies of tenofovir disoproxil fumarate (TDF) and emtricitabine (FTC) PrEP have demonstrated efficacy in specific populations. We model impact of TDF/FTC PrEP on HIV incidence in a heterogeneous population with multiple mechanisms of HIV transmission and complex contact patterns.
The compartmental model of HIV infection described in Chapter 1 was adapted to include PrEP interventions, and simulated PrEP delivery varying target populations, coverage, drug-efficacy, adherence, and scale-up rate based on ranges demonstrated in clinical trials. Incidence reduction and cumulative infections averted 10 years post-intervention and number of PrEP person-years required to avert one infection were assessed.

This model confirms that realizable PrEP delivery can achieve HIV incidence reductions in Newark, NJ. A comparison with other modeling studies further underlines that MSM-targeted PrEP could potentially decrease HIV incidence, with potential for substantial impact amongst MSM. Targeting IDU is the most efficient strategy to reduce HIV incidence, although population impact is less.

My third dissertation chapter drills down into the specifics of interactions between HIV and HCV. The two pathogens can exacerbate one another’s effects and transmission, and HIV coinfection can hinder HCV treatment, but until now there has been no modeling work that focuses specifically on within-host HCV/HIV coinfection. To explore these interactions, I have created a modeling framework for **within-host dynamics** of this coinfection that incorporates the role of the immune system in HCV clearance. Our model is able to replicate empirically-observed clinical patient outcomes: spontaneous clearance, clearance with treatment, relapse, and treatment non-response, and show how HIV coinfection can shift a patient between those outcomes by exploring **bifurcation dynamics** in the model. Incorporating treatment efficacy, our model sheds light on the tradeoffs involved in choosing between treatment protocols, and how both duration and efficacy need to be taken into account carefully in coinfected patients. Exciting breakthroughs are occurring in HCV treatment as new direct-acting antiviral agents (DAAs) are being found to outperform the traditional indirect treatments, but these new treatments have a high price tag.
and must be administered cautiously. Understanding the role of HIV coinfection in treatment is particularly pressing.

The breakthroughs in HCV treatment are remarkable not only for they way the may shift HCV epidemiology, but also because they were made with relatively little understanding of the within-host processes governing HCV dynamics. Chapter 4 focuses on HCV infection alone, and aims to use similar modeling techniques as described in the previous chapter to explore the unusual **viral evolution dynamics** that HCV displays and elucidate some of the potential mechanisms at work. Hepatitis C Virus (HCV) undergoes rapid within-host viral evolution, and genetic analysis indicates that multiple distinct lineages persist throughout infection. Not all lineages are detectable in the blood at all times, and there is evidence of viral genetic differences among groups of infected hepatocytes. Furthermore, in some patients, lineages observed early in infection appear to go extinct but reappear later, even after liver transplant, fueling the hypothesis that HCV replication outside the liver may occur. Incorporating these various hypotheses into a model of viral evolution during infection can help elucidate HCV within-host dynamics. In turn, this may help determine effective treatment design, especially with new (DAAs).

HIV-Hepatitis C coinfection is a **global epidemiological problem**, especially among **people who inject drugs**, as the pathogens are both blood-borne. Coinfected patients can have worse prognosis due to potentiating effects between HIV and HCV, as detailed in Chapter 3. Models have been used to a great extent for HIV care research, but work on modeling HIV-HCV coinfection has thus far been limited. In the final dissertation chapter, I use a mathematical model to explore relative impacts of **antiretroviral and methadone maintenance therapy scale-up**, and HCV treatment rollout on disease burden in the developing world setting of Ho Chi Minh City, VN. Drug users in Vietnam are burdened with an unusually high prevalence of HIV-HCV coinfection, exacerbated by poor linkage into health systems. HCV treat-
ment is currently unaffordable for most patients in Vietnam, but new medications have been developed that have the potential to be affordable within the next 10-15 years, so one of the aims of this project is to assess what can be done now to best prepare for a potential roll-out.

Our model results indicate that scale-up of antiretroviral therapy has a major impact on HIV incidence and deaths, but a negligible impact on HCV. Methadone scale-up has an impact on incidence of both infections and on total death rates, though a smaller impact on disease-specific death rates. **HCV treatment coverage can measurably reduce deaths in the population**, increasing multifold the reductions in death rates afforded by the other interventions. The one caveat is that without methadone scale-up, low levels of HCV treatment roll-out can actually yield more incident cases because treated individuals can get reinfected. With Vietnam’s transition to middle-income country status, many foreign sources that have previously funded HIV care are pulling out, so finding maximally efficient public health interventions that target both HIV and HCV is crucial.

HIV-HCV is a coinfection system with many facets, and it embodies many of the complexities of coinfection in general. Studying this system has sparked my broader interest in coinfection in particular in context of concomitant public health problems such drug resistance (69, 70), and how it can shape the way we think about disease ecology and public health.
Chapter 1

Modeling the Impact of Interventions along the HIV Continuum of Care in Newark, New Jersey

Ruthie B. Birger¹, Timothy Hallett², Anushua Sinha³, Bryan T. Grenfell¹,⁴, Sally L. Hodder⁵,⁶

¹ Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ
² Department of Infectious Disease Epidemiology, Imperial College London, London, UK
³ Department of Preventive Medicine and Community Health, New Jersey Medical School, Rutgers University, Newark, NJ
⁴ Fogarty International Center, National Institute of Health, Bethesda, MD
⁵ Department of Medicine, West Virginia University School of Medicine, Morgantown, WV
⁶ West Virginia Clinical and Translational Science Institute, Morgantown, WV
Abstract:

Background

The HIV epidemic in Newark, NJ is among the most severe in the US. Prevalence ranges up to 3.3% in some groups. The aim of this study is to use a mathematical model of the epidemic in Newark to assess the impact of interventions along the continuum of care, leading to virologic suppression.

Methods

A model was constructed of HIV infection including specific care-continuum steps. The model was calibrated to HIV/AIDS cases in Newark among different populations over a 10-year period. Interventions applied to model-fits were increasing proportions tested, linked and retained in care, linked and adherent to treatment, and increasing testing frequency, high-risk-group-testing and adherence. Impacts were assessed by measuring incidence and deaths reductions 10 years post-intervention.

Results

The most effective interventions for reducing incidence were improving treatment adherence and increasing testing frequency and coverage. No single intervention reduced incidence in 2023 by >5%, and the most effective combination of interventions reduced incidence by ~16%(2%-24%). The most efficacious interventions for reducing deaths were increasing retention, linkage to care, testing coverage and adherence. Increasing retention reduced deaths by approximately 27%(24%-29%); the most efficacious combination of interventions reduced deaths in 2023 by ~52%(46%-57%).
Conclusions

Reducing HIV deaths in Newark over a 10-year period may be a realizable goal, but reducing incidence is less likely. Our results highlight the importance of addressing leaks across the entire continuum of care and reinforcing efforts to prevention new HIV infections with additional interventions.\footnote{\textsuperscript{1}published in \textit{Clinical Infectious Diseases}, Vol. 12, pgs. 274–84 (2014); RB Birger, TB Hallett, A Sinha, BT Grenfell, SL Hodder \textsuperscript{41}. It has also been presented at CROI in Seattle, February 2012 and the EEID Conference at University of Michigan in May 2012}

Introduction

More than thirty years into the HIV/AIDS epidemic, Newark, New Jersey (NJ) continues to be one of the most severely HIV/AIDS-impacted communities within the United States. Newark residents have an HIV prevalence of 2.3%, nearly 7 times that of the rest of NJ \textsuperscript{(72)}. Moreover, the non-Hispanic Black community in Newark is disproportionately affected, with HIV prevalence of 3.3%, similar to the prevalence observed in some sub-Saharan African countries, e.g. Ghana (1.8%), Rwanda (2.9%) \textsuperscript{(73, 74, 75)}. The profile of risk behaviors in Newark is complex, with high prevalence and interaction of injecting drug-use (IDU) and risky sexual behaviors \textsuperscript{(72, 76, 77), .}

Prevention interventions over the last decade have not succeeded in substantially reducing HIV incidence in the US \textsuperscript{(78)}. Nevertheless, there is reason for new optimism as recent studies demonstrated that effective antiretroviral-therapy (ART) of HIV-infected persons results in 96% decrease in sexual HIV transmission \textsuperscript{(12)}. This has led to intense dialogue and new initiatives to increase HIV-testing coverage, improve connection/retention in care, and expand ART with maximal rates of virologic suppression. Mathematical modeling may help assess potential contributions of various interventions and identify combinations that may contribute most to overall epidemic control. Previous modeling work has suggested that universal Test-
and-Treat interventions may control local HIV epidemics (44). However, significant doubt remains, due to the difficulty of scaling up screening programs and the marked heterogeneity among HIV-infected persons with respect to disease transmission, access and acceptance of care and ART [11]. Additionally, impact of Test-and-Treat strategies may differ in areas with generalized epidemics compared with concentrated microepidemics.

The HIV epidemic in the US is a set of microepidemics— “hotspots” of infection (79). Results of modeling for generalized HIV epidemics may not accurately inform HIV control in the US. Modeling the impact of interventions to prevent HIV in Newark, NJ is an important case-study for several reasons: Newark is a well-described hotspot of HIV infection in the US, and its epidemic is driven by multiple factors including heterosexual transmission, intravenous drug-use and unreported sexual activity among men in prisons or other settings. Furthermore, as Newark is a small city, it may be possible to implement interventions and assess validity of modeling. Results of modeling may therefore, be useful for other US hotspots. Here, we use dynamic models, linked to epidemiological data, to assess probable impact of several interventions on future HIV incidence and mortality in Newark, NJ. Crucially, our model explicitly includes each step in the continuum of HIV care leading to virologic suppression (i.e., testing, linkage to care, retention in care, linkage to treatment, and adherence to treatment) (80).

Methods

Model

To capture the dynamics of the Newark epidemic, a compartmental model was used (Figure 1.1a). Following others (e.g. (44 81 82 83)), the model includes compartments for three infection stages: acute, high viral load but short duration (3-6
months); asymptomatic, lower viral load with duration of 8 years; and AIDS, intermediate viral load with duration \(\sim 2.5\) years. These stages are further stratified by awareness of infection, eligibility, and treatment status. Progression is as follows: Susceptible individuals move to acute infection, which progresses to asymptomatic --will be diagnosed before \(\text{CD}4 < 200\) cells/µL or asymptomatic --will be diagnosed after \(\text{CD}4 < 200\) cells/µL. Those who are tested and diagnosed move into either tested, eligible for ART or tested, ineligible for ART. Eligibility for ART depends on \(\text{CD}4\) count. Before 2010, those with \(\text{CD}4 < 350\) cells/µL are eligible. From 2011 onwards, all HIV+ are deemed eligible. Tested ineligible individuals become eligible as \(\text{CD}4\) count declines. Eligible individuals move into treated, suppressed compartment or treated, unsuppressed compartments. Current percentages of infected individuals passing each point in the continuum are included in the model: tested individuals are linked to care, individuals linked to care upon testing are retained in care, individuals retained in care are linked to ART, if eligible, with a mean delay of three months. Individuals linked to treatment achieve viral suppression, depending on ART adherence. Individuals who were never tested or dropped out from the care-continuum progress with infection, until presenting for care with \(\text{CD}4 < 200\) cells/µL. Individuals who progress to AIDS after failing on ART die after 1-3 years.

The equations are applied to males and females comprising high- and low-risk sexual activity classes, and IDU and non-IDU. The model takes into account heterosexual, homosexual, bisexual and IDU transmission, as well as combinations of sexual and IDU risk behaviors. Individuals move between IDU and non-IDU groups, and sexual mixing is assortative as per a study conducted using STI clinic data from Newark in the 1990’s (84) (Figure 1.1b). Additionally, the model includes a bounded decline in needle-sharing and drug-use-incidence as IDU has declined in Newark (85).

Ranges for model parameters were gathered from literature (Tables 1.1, 1.2). Uncertainty analysis was performed, altering parameters one at a time to see what
Figure 1.1: **Graphical Representation of Model Flow and Risk Groups Interactions.** (a) Susceptible individuals flow into the Acute Infection compartment via the force of infection. Acutely infected individuals flow into Asymptomatic Undiagnosed. Asymptomatic Individuals either get diagnosed before progressing to AIDS or not. Not pictured is the tested-ineligible category as this category no longer exists by the time interventions are implemented. Tested-eligible individuals can be lost to follow up or not linked to treatment or be linked to treatment and then either virologically suppress or not suppress. Unsuppressed individuals move into a dead-end AIDS post-ART compartment. Finally, Undiagnosed and Unlinked individuals progress to AIDS-never treated from which they either die or present for treatment and then either suppress or not. Flows are in blue arrows and excess deaths are in red. (b) Male and Female risk groups are shown with arrows indicating groups that most strongly assort.

Figure 1.1
influence ranging a given parameter had on results. In this way, a subset of influential parameters, those whose variation yielded >10% average shift over all model outputs, was selected for fitting. Next, Latin-hypercube-sampling was performed using the subset of influential parameters (86). A multi-dimensional grid describes the model’s parameter space. Latin-hypercube-sampling forces each grid-segment to be sampled exactly once, thus efficiently covering the whole parameter space. Model outputs from sampled parameter sets were then compared with observed counts stratified by risk-group of persons living with HIV/AIDS in Newark over time. Parameter sets were accepted if they fit within prior limits of the data (86). Sensitivity analysis was performed by calculating partial rank correlation coefficients (PRCCs) using standard methods (86). PRCCs provide a post-fitting breakdown of which parameters influence model outputs the most, thereby enabling validation of model assumptions.
Table 1.1: Prior and Posterior parameter ranges

<table>
<thead>
<tr>
<th>Parameter Name</th>
<th>Description</th>
<th>Prior Range</th>
<th>Posterior Range</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>death rate when CD4 $&lt;$ 200</td>
<td>0.2-2 yr$^{-1}$</td>
<td>0.20-1.99 yr$^{-1}$</td>
<td>(87, 11, 24)</td>
</tr>
<tr>
<td>$\zeta$</td>
<td>$1/\text{duration of drug use}$</td>
<td>0.025-1.667 yr$^{-1}$</td>
<td>0.03-0.9 yr$^{-1}$</td>
<td>(88)</td>
</tr>
<tr>
<td>$\theta$</td>
<td>$1/\text{duration of acute period}$</td>
<td>1.5-6 yr$^{-1}$</td>
<td>1.62-5.98 yr$^{-1}$</td>
<td>(11, 24)</td>
</tr>
<tr>
<td>$\epsilon$</td>
<td>Assortativeness coefficient</td>
<td>0.05-0.95</td>
<td>0.05-0.95</td>
<td>ME</td>
</tr>
<tr>
<td>$c(\text{m,h})$</td>
<td>rate of partner change in high-risk men (het)</td>
<td>3-10 partners/yr</td>
<td>3.02-9.79 partners/yr</td>
<td>(89)</td>
</tr>
<tr>
<td>$c(\text{m,l})$</td>
<td>rate of partner change in low-risk men (het)</td>
<td>0.01-2 partners/yr</td>
<td>0.04-1.98 partners/yr</td>
<td>(84, 89)</td>
</tr>
<tr>
<td>$c(\text{hrhm} : m)$</td>
<td>rate of male partner change in bisexual men</td>
<td>1-5 partners/yr</td>
<td>1.21-4.98 partners/yr</td>
<td>(89), ME</td>
</tr>
<tr>
<td>$c(\text{msm} : h)$</td>
<td>rate of partner change in high-risk MSM</td>
<td>3-10 partners/yr</td>
<td>3.01-9.97 partners/yr</td>
<td>(89), ME</td>
</tr>
<tr>
<td>$c(\text{msm} : l)$</td>
<td>rate of partner change in high-risk MSM</td>
<td>0.1-3 partners/yr</td>
<td>0.14-2.92 partners/yr</td>
<td>(89), ME</td>
</tr>
<tr>
<td>$m_x$</td>
<td>proportion of men exclusively MSM</td>
<td>0.03-0.12</td>
<td>0.04-0.12</td>
<td>(93)</td>
</tr>
<tr>
<td>$\eta_D$</td>
<td>needles shared per partner</td>
<td>50-350 needles/partner</td>
<td>52.64-349.65 needles/partner</td>
<td>(88)</td>
</tr>
<tr>
<td>$\psi_{\text{red}}$</td>
<td>coefficient of reduction in drug use incidence</td>
<td>0.001-1</td>
<td>0.0025-0.0951</td>
<td>ME</td>
</tr>
<tr>
<td>$\nu(l,l)$</td>
<td>number of sex acts in a low-low partnership</td>
<td>20-150 sex-acts/partner</td>
<td>20.09-149.95 sex-acts/partner</td>
<td>ME</td>
</tr>
<tr>
<td>$\nu(h,h)$</td>
<td>number of sex acts in a high-high partnership</td>
<td>1-30 sex-acts/yr</td>
<td>1.48-29.85 sex-acts/yr</td>
<td>ME</td>
</tr>
<tr>
<td>$\nu(l,h)$</td>
<td>number of sex acts in a low-high partnership</td>
<td>1-30 sex-acts/yr</td>
<td>1.02-28.79 sex-acts/yr</td>
<td>ME</td>
</tr>
<tr>
<td>$\beta_{\text{mm}}$</td>
<td>ratio of transmission probability male to female</td>
<td>1.5-6</td>
<td>1.54-5.8</td>
<td>(92)</td>
</tr>
<tr>
<td>$\beta_{\text{mf}}$</td>
<td>ratio of transmission probability male to male</td>
<td>1.5-6</td>
<td>1.52-5.92</td>
<td>(92)</td>
</tr>
<tr>
<td>$\omega_t_0$</td>
<td>ratio of infectiousness during acute phase</td>
<td>20-30</td>
<td>20.25-29.98</td>
<td>(11)</td>
</tr>
<tr>
<td>$\phi$</td>
<td>per-sex-act transmission probability</td>
<td>1e-5-0.01/sx-act</td>
<td>1e-4-0.0032/sx-act</td>
<td>(24)</td>
</tr>
<tr>
<td>$\sigma_d$</td>
<td>per-injection transmission probability</td>
<td>2e-5-0.05/injection</td>
<td>5e-4-0.049/injection</td>
<td>(93)</td>
</tr>
<tr>
<td>$p_{\text{du}}$</td>
<td>initial percent prevalence of drug use</td>
<td>0.5%-10%</td>
<td>4%-10%</td>
<td>(77)</td>
</tr>
<tr>
<td>$p_{\text{mdu}}$</td>
<td>initial percent prevalence of drug use among MSM</td>
<td>0. 5%-20%</td>
<td>1%-20%</td>
<td>(93)</td>
</tr>
<tr>
<td>$p_{\text{rm}}$</td>
<td>proportion of MSM high risk</td>
<td>0.1-.95</td>
<td>0.19-.95</td>
<td>(93)</td>
</tr>
<tr>
<td>$c_d$</td>
<td>rate of change of needle sharing partner</td>
<td>0.5-3 partners/yr</td>
<td>0.51-2.97 partners/yr</td>
<td>ME</td>
</tr>
<tr>
<td>$q_{\text{nm}}$</td>
<td>coefficient of reduction in unprotected sex acts when aware of status</td>
<td>.1-.9</td>
<td>0.11-.88</td>
<td>ME</td>
</tr>
</tbody>
</table>
## Table 1.2: Invariant Parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Value</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tau_{Test}$</td>
<td>Time between infection and testing if tested before CD4 count &lt;200/µL and eligible when tested</td>
<td>4 years</td>
<td>(95, 96, 97)</td>
</tr>
<tr>
<td>$z$</td>
<td>Percent of diagnosed individuals diagnosed before CD4&lt;200 µL</td>
<td>0.41</td>
<td>(97)</td>
</tr>
<tr>
<td>$z_{ratio1}$</td>
<td>Ratio of testing in other risk groups to heterosexual women</td>
<td>.68-1.1</td>
<td>(97)</td>
</tr>
<tr>
<td>$test_{link}$</td>
<td>Percent linked to care given tested</td>
<td>0.75</td>
<td>(98, 99, 100, 101)</td>
</tr>
<tr>
<td>$ret$</td>
<td>Percent retained in care given linked</td>
<td>0.51</td>
<td>(98, 99, 100, 101)</td>
</tr>
<tr>
<td>$tr_{link}$</td>
<td>Percent linked to treatment given retained in care</td>
<td>0.89</td>
<td>(98, 99, 100, 101)</td>
</tr>
<tr>
<td>$x$</td>
<td>Percent of treated who achieve viral suppression</td>
<td>0.77</td>
<td>(98, 99, 100, 101)</td>
</tr>
<tr>
<td>$x_{ratio1}$</td>
<td>Ratio of adherence/suppression in other risk groups to heterosexual women</td>
<td>.8-1</td>
<td>EO</td>
</tr>
<tr>
<td>$\tau_{Treatment}$</td>
<td>Time between eligibility and treatment</td>
<td>0.25</td>
<td>EO</td>
</tr>
<tr>
<td>$y$</td>
<td>Percent of Asymptomatics eligible for treatment when tested (200/µL &lt;CD4 count&lt;350/µL)</td>
<td>0.3</td>
<td>(97)</td>
</tr>
<tr>
<td>$\tau_{elig}$</td>
<td>Time to eligibility if tested and not eligible</td>
<td>2 years</td>
<td>(95, 96, 97)</td>
</tr>
<tr>
<td>$\omega_I^A, \omega_I^S, \omega_I^{NS}, \omega_A$</td>
<td>Weighting of transmission probability by stage of infection</td>
<td>1, .04, .9, 7</td>
<td>(11, 12)</td>
</tr>
<tr>
<td>$\gamma_{NT}$</td>
<td>Rate of progression for untreated individuals from HIV to AIDS</td>
<td>1/8 yr⁻¹</td>
<td>(87, 11, 24)</td>
</tr>
<tr>
<td>$\gamma_{NS}$</td>
<td>Reduced progression rate for treated, unsuppressed individuals</td>
<td>1/12 yr⁻¹</td>
<td>EO</td>
</tr>
<tr>
<td>$\mu, \mu_{IDU}$</td>
<td>Death rates/recruitment rates (non-IDU, IDU)</td>
<td>1/58 yr⁻¹, 1/36 yr⁻¹</td>
<td>(102)</td>
</tr>
<tr>
<td>$pdur, pdur_{IDU}$</td>
<td>Percent of IDU who engage in risky sexual behavior (m,f)</td>
<td>34%, 18%</td>
<td>(88)</td>
</tr>
<tr>
<td>$pmdur$</td>
<td>Percent of MSM IDU who engage in risky sexual behavior</td>
<td>0.8</td>
<td>ME</td>
</tr>
<tr>
<td>$pr$</td>
<td>Percent of Heterosexual men engaging in high risk behavior</td>
<td>0.3</td>
<td>(89)</td>
</tr>
<tr>
<td>$pr_f$</td>
<td>Percent of Heterosexual women engaging in high risk behavior</td>
<td>0.2</td>
<td>(89)</td>
</tr>
<tr>
<td>$fdu$</td>
<td>Percent of IDU who are female</td>
<td>0.36</td>
<td>(88)</td>
</tr>
<tr>
<td>$hrhm$</td>
<td>Percent of high risk heterosexual men engaging in bisexual behavior</td>
<td>0.15</td>
<td>(88), ME</td>
</tr>
<tr>
<td>$c(hrhm : f)$</td>
<td>Rate of female partner change in bisexual men</td>
<td>3 partners/yr</td>
<td>(88), ME</td>
</tr>
<tr>
<td>$\tau_{ART}$</td>
<td>Year of availability of ARTs</td>
<td>1995</td>
<td>(71)</td>
</tr>
<tr>
<td>$\psi_D$</td>
<td>Initial incidence of injecting drug use (approximated by prevalence/duration)</td>
<td>(71, 88)</td>
<td></td>
</tr>
<tr>
<td>$n_{red}$</td>
<td>Coefficient of reduction in needle sharing per 5 years</td>
<td>0.3</td>
<td>ME</td>
</tr>
</tbody>
</table>
Interventions

Increasing tested and adherent proportions overall (Figure 1.2a) and in high-risk groups by 5% and 25% and increasing testing frequency were implemented. For tested HIV-infected individuals, interventions at each point in the care-continuum were implemented at three different levels representing a modest, moderate or optimal level of intervention (Figure 1.2b).

![Figure 1.2: Testing and Care-continuum Interventions.](image)

(a) Percent Tested before CD4<200

(b) Care Cascade

Figure 1.2: Testing and Care-continuum Interventions. (a) Current, Modest Intervention and Moderate Intervention levels are shown of Percentages of individuals tested before reaching CD4<200, in Current, Modest Intervention and Moderate Intervention (b) Percentages of individuals achieving each step in the care-continuum, in Current, Modest Intervention and Moderate Intervention levels.
Data

Data used for the fitting came from Project IMPACT (Intensive Mobilization to Promote AIDS Awareness through Community-based Technologies), a NJ Department of Health initiative to address the HIV epidemic in NJ communities most affected. Project IMPACT data represent reported cases for Newark from 2002-2010 as ascertained by epidemiological surveillance.

Many parameter estimates came from National HIV Behavioral Surveillance (NHBS) data. Created in 2003 by the CDC to collect behavioral information on individuals at high risk for acquiring HIV, the NHBS is conducted in rotating cycles targeting MSM, IDUs and high-risk heterosexuals. A range of sampling techniques are used for each key-population, and standardized, anonymous questionnaires on behavior, testing and prevention are administered and HIV testing are offered.

Results

Model-fitting

The fitting process was run until 100 parameter sets were identified that produced model runs within the specified prior limits. Figure shows the range of fitted runs plotted with data points by risk group (Table shows prior and posterior intervals on fitted parameters).

PRCCs were calculated using the posterior range of each parameter against model outputs of prevalence and incidence. Different parameters showed peak influence for incidence and each risk group’s prevalence, though parameters that governed drug-use-incidence and duration and contact patterns were influential on all outputs (Supp. Table A.1). The amount of initial uncertainty that was reflected in the prior ranges of the influential parameters varied significantly; some parameters such
Figure 1.3: Model Fits and Data. Data from Project Impact (solid lines with dots) are shown, overlaid with the range of model fits for each data point. The twelve risk groups in the model are aggregated into six shown here as follows: Het F comprises Low and High Risk non-IDU Heterosexual Females, IDU F Low and High Risk IDU Heterosexual females, Het M Low and High Risk non-IDU Heterosexual Males and non-IDU Bisexual Males, MSM Low and High Risk non-IDU MSM, IDU M Low and High Risk IDU Heterosexual Males and IDU Bisexual Males, MSM IDU are just the IDU MSM group.

Interventions

To test the impact of interventions at each point in the care-continuum, parameters dictating the intensity at a given point were altered in isolation and conjunction, representing a range of intervention strength. Intervention scenarios were run on all
parameter sets to assess their efficacy. Parameters included: proportion of individuals tested before CD4<200 cells/µL, proportion connected to care given newly positive test, proportion retained in care given connection to care, proportion receiving ART given retention in care, proportion virally suppressing given ART (i.e. proportion adherent to ART), percent increase in testing and viral-suppression in high-risk groups, time between infection and test if tested before CD4<200 cells/µL, proportion eligible if tested before CD4<200 cells/µL.

The top four most efficacious interventions for averting infections were run in isolation and combination on all parameter sets. According to the best parameter-fit, these were, in descending order of efficacy: increasing ART-adherence in high-risk groups, increasing ART-adherence in all groups, reducing time between infection and testing, increasing proportion tested before CD4<200 cells/µL.

Combinations of these interventions resulted in larger reductions in incidence than single interventions (Table 1.3, Figure 1.4a).

The top four most efficacious interventions for reducing the number of deaths were similarly run in isolation and combination on all parameter sets. According to the best parameter-fit, these were, in descending order of efficacy: increasing proportion retained in care, increasing proportion linked to care, increasing testing coverage, increasing ART-adherence in all groups.

Again, combinations of interventions resulted in larger numbers of deaths averted. (Table 1.4, Figure 1.4b).

Of note, the order of efficacy of interventions varied slightly between parameter sets, but the four most efficacious remained the same across all parameter sets.
Figure 1.4: Incident Cases and Deaths Time Series, Percent Change Over Time by Intervention (according to best-fit parameter set). (a) shows the projected change in incident cases over time under increasing combinations of the top four most efficacious interventions. (b) shows the projected change in deaths over time under increasing combinations of the top four most efficacious interventions.

The ten-year impact on incidence of any individual intervention was small but when all interventions were run together at maximum levels, incidence declined steeply (Figure 1.5). This highlights the importance of closing gaps throughout the care-continuum. Each individual intervention can only have limited effect if the cascade is “leaky” elsewhere. Figure 5 also shows impact of secular declines in IDU over time. The red line shows a counterfactual scenario, i.e. one without reductions in
incidence of IDU and needles-shared. The epidemic in Newark had the potential to be worse, with $\sim 40\%$ higher annual incidence.

Figure 1.5: **Pressure Test of Interventions.** Impact on annual number of new infection incidence of maximal implementation of all interventions. By 2023, there are fewer than 200 incident cases.

Though no currently available individual intervention can effect a substantial reduction in incidence, combination interventions at intermediate levels can yield a reduction in incidence in 2023 of $\sim 16\%$. Interventions may have different levels of achievability and cost, and various combinations at different levels can yield similar results. Contour plots show the combination levels for pairs of interventions that can achieve similar results (Supp. Figure A.2). Changing the profile of combination interventions changes the distribution of incident infections sources (Figure 1.6).
Figure 1.6: **Pie-Charts of Incidence Sources.** Pie Charts designating provenance of incident cases in 2023 for Baseline and Top Four most efficacious interventions at Moderate level.

Table 1.3: Impact of Interventions on Incident Cases

<table>
<thead>
<tr>
<th>Intervention</th>
<th>Change from Baseline Level</th>
<th>% Incident Cases Change in 2023 (Range)</th>
<th>Absolute Change in Incident Cases 2023 (Range)</th>
<th>% Change in Cumulative Incident Cases 2013-2023 (Range)</th>
<th>Absolute Change in Cumulative Incident Cases 2013-2023 (Range)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Increase in High Risk ART adherence</td>
<td>25% increase</td>
<td>4.6 (0.8, 6.4)</td>
<td>22 (2, 43)</td>
<td>2.4 (0.5, 3.1)</td>
<td>13 (13, 260)</td>
</tr>
<tr>
<td>Increase All-group ART Adherence</td>
<td>77% to 95%</td>
<td>3.8 (0.6, 5.8)</td>
<td>18 (1, 38)</td>
<td>1.9 (0.3, 2.8)</td>
<td>108 (9, 225)</td>
</tr>
<tr>
<td>Increase in Testing Frequency</td>
<td>4 Years to 1.5 Years</td>
<td>2.6 (0.2, 6.2)</td>
<td>12 (0, 40)</td>
<td>2.3 (0.1, 4.4)</td>
<td>127 (7, 340)</td>
</tr>
<tr>
<td>Increase in HIV Testing Coverage</td>
<td>46% to 75%</td>
<td>2.3 (0.2, 4.3)</td>
<td>11 (1, 30)</td>
<td>1.2 (0.1, 2.3)</td>
<td>68 (4, 197)</td>
</tr>
<tr>
<td>Increase in ART Adherence; All-</td>
<td>as above</td>
<td>8.3 (1.2, 11.2)</td>
<td>38 (3, 72)</td>
<td>4.2 (0.8, 5.5)</td>
<td>235 (21, 422)</td>
</tr>
<tr>
<td>Increase in ART Adherence; All- and High Risk</td>
<td>as above</td>
<td>11.1 (1.6, 14.8)</td>
<td>52 (4, 98)</td>
<td>5.6 (1.1, 7.3)</td>
<td>314 (27, 565)</td>
</tr>
<tr>
<td>All top 4 interventions</td>
<td>as above</td>
<td>15.8 (2.1, 24.5)</td>
<td>73 (5, 158)</td>
<td>9.1 (1.6, 13.7)</td>
<td>598 (40, 1039)</td>
</tr>
</tbody>
</table>
Table 1.4: Impact of Interventions on Deaths

<table>
<thead>
<tr>
<th>Intervention</th>
<th>Change from Baseline Level</th>
<th>% Deaths Change in 2023 (Range)</th>
<th>Absolute Change in Deaths in 2023 (Range)</th>
<th>% Change in Cumulative Deaths 2013-2023 (Range)</th>
<th>Absolute Change in Cumulative Deaths 2013-2023 (Range)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Increase in Retention</td>
<td>51% to 67%</td>
<td>26.4 (24.2, 29.1)</td>
<td>89 (59, 150)</td>
<td>24.7 (23, 26.6)</td>
<td>1566 (184, 1659)</td>
</tr>
<tr>
<td>Increase in Linkage to care</td>
<td>75% to 83%</td>
<td>9.4 (8.4, 10.3)</td>
<td>32 (14, 59)</td>
<td>8.6 (8.3, 9.2)</td>
<td>370 (167, 516)</td>
</tr>
<tr>
<td>Increase in HIV Testing Coverage</td>
<td>68% to 75%</td>
<td>5 (3.3, 6.5)</td>
<td>17 (5, 31)</td>
<td>3.2 (1.7, 4.2)</td>
<td>135 (34, 275)</td>
</tr>
<tr>
<td>Increase All-group ART Adherence</td>
<td>25%</td>
<td>5.4 (3.4, 6.7)</td>
<td>18 (5, 31)</td>
<td>3.3 (1.1, 3)</td>
<td>97 (25, 176)</td>
</tr>
<tr>
<td>Increase in Retention and Linkage</td>
<td>as above</td>
<td>39.4 (36, 43.5)</td>
<td>132 (60, 224)</td>
<td>37.4 (34.9, 40.4)</td>
<td>1598 (741, 2496)</td>
</tr>
<tr>
<td>Increase in Retention, Linkage and Testing Coverage</td>
<td>as above</td>
<td>43.2 (39.1, 47.2)</td>
<td>145 (63, 247)</td>
<td>39.4 (36.6, 42.3)</td>
<td>1687 (762, 2656)</td>
</tr>
</tbody>
</table>

Discussion

HIV remains highly endemic in focal areas of the US. Despite availability of ART, HIV incidence in the US has been relatively stable with roughly 50,000 new infections per year (78). Recent empirical work has shown the efficacy of treatment as prevention in a clinical trial setting among stable serodiscordant couples and confirmed that maintenance of virologic suppression is key in HIV prevention (12). Given the recent national dialogue regarding the care-continuum, in which only 33-55% of HIV-infected persons are retained in care and 19-26% achieve virologic suppression (98, 99, 100, 101), our model addresses the key questions of which interventions are most likely to prove efficacious at decreasing HIV acquisition and mortality. Surprisingly, HIV incidence is estimated to decrease by just 16% when all interventions along the care-continuum of care are implemented at “achievable” levels, suggesting ART interventions cannot alone be relied upon to curb HIV transmission in hotspots and that additional prevention interventions are urgently required.

A major contributor to the modest impact of interventions is the “leaky” care-continuum. Large impacts on HIV incidence are only evident when all interventions approach 100%. In Newark, as in other settings, there is a subset of individuals who either cannot or will not access care. In a study of HIV acquisition among women living in areas of the US with high levels of poverty and HIV prevalence, 20% reported that they were unable to access needed medical care (104). The Affordable-Care Act

30
may improve access to care but other factors (e.g., distrust of the medical establishment and stigma associated with HIV infection) may continue to contribute to poor care access. Nonetheless, for individuals successfully linked/retained in care, this modeling exercise estimates mortality decreases of \(~35\%\). The challenges surrounding improved linkage and retention in care are particularly difficult, as improved retention requires on-going efforts rather than single-point change. Novel strategies to link and retain individuals in care need to be created and tested. Only by combining interventions that involve linking and retention in care can further reductions in mortality be realized.

Virologic suppression is a demonstrated powerful tool for averting HIV transmission. However, maximizing adherence to ART is challenging as it involves extensive follow-up and monitoring of individuals over long periods of time. Indeed, Test-and-Treat interventions may have the potential to increase numbers of patients initiating ART early, but without stabilizing the back-end of treatment continuation (i.e., care retention and ART adherence), Test-and-Treat strategies cannot achieve full potential.

The proportion of incident cases that arise from acute infections does not decline with successive interventions (Figure 1.6), so HIV transmission is maintained even in the four-intervention-scale-up scenario. This result suggests that unaddressed acute infection may play a critical role in the limited efficacy of Test-and-Treat interventions to decrease HIV incidence at least over the short time-scales examined here, however, further exploration of this hypothesis is needed. No intervention tested in this study addresses acute infection directly, as many commonly-used testing algorithms are unable to detect acute infection \(^{[105]}\). Implementation of newer generation HIV tests that simultaneously assess acute and chronic infection may be critical to further decreasing HIV incidence \(^{[105, 106]}\).
Previous modeling studies have assessed the efficacy of Test-and-Treat interventions in different settings. Granich et al predict transition of the epidemic into an elimination phase after introduction of Test-and-Treat in a setting of generalized HIV infection, but their assumptions differ from those in this study in several substantive ways: immediate treatment, higher adherence rates (i.e. elements of a perfect care-continuum), and only heterosexual transmission. The results of the model presented here are more consistent with the predictions of Dodd et al., who explore the efficacy of Test-and-Treat in a range of epidemiological settings. Though they also test scenarios that include perfect or near-perfect care-continua, they predict modest HIV incidence reductions in settings with greater variation in risky behavior, where key populations at high risk of infection and transmission drive the epidemic. Powers et al also predict modest reductions in HIV prevalence in a developing world setting (even with high coverage) unless individuals with early infection are targeted (106). Walensky et al, examining outcome measures in Washington, DC, draw similar conclusions on Test-and-Treat impact, predicting increases in life-expectancy but only a 15% reduction in HIV incidence over five years of overall population life-years spent with transmissible viral-load (107). Long et al, using a model based on 2007 US data, predict similarly modest effects of Test-and-Treat; incidence is reduced by 2.3% with enhanced testing, and ART averts only 10.3% of infections over 20 years (82). Lastly, Sorensen et al, modeling an urban population based on New York City MSM, report a greater impact on incidence, predicting roughly a 50% reduction over 20 years with intermediate levels of interventions implemented throughout the care-continuum (108). The present analysis offers a unique perspective on the Newark epidemic, but agrees broadly with some prior predictions in the literature about the impact of Test-and-Treat strategies in the US.

This study adds to the literature a bespoke model for a developed-world urban setting; it takes into account differential dynamics between heterosexual, homosexual
and bisexual HIV transmission in addition to IDU. While there is uncertainty around many of the model parameters, the fitting process used in this study helps quantify it. Additionally, this model incorporates all steps in the care-continuum from initial test through linkages and viral suppression.

There are inevitable simplifications in the model that may lend some uncertainty to the results. Only approximations are used for disease stage. There is no age-structure and the model does not take into account ethnic variation, which is likely an over-simplification of the transmission situation in Newark. While individuals can change drug-use status, no change in sexual risk-behavior over time is incorporated. Most parameters stay constant over time, and as there are multiple pathways to being treated, some individuals may suffer too short a survival penalty upon reaching this state, which may in turn slightly impact mortality estimates.

In addition, assumptions made about sexual mixing patterns were derived from a single study performed in Newark’s STD clinic roughly fifteen years ago. As can be seen in the PRCCs (Supp. Table A.1), the assortativity coefficient only comes up as influential for heterosexual prevalence measures. This result validates model assumptions: sexual-mixing diversity should be most important in heterosexual groups because they mix with the greatest number of other groups. Sexual assortativeness does not play a role in IDU transmission, as IDUs are far more likely to acquire infection through drug-use than sexual contact. Similarly, MSM are naturally more assortative as they only acquire infection from other MSM or the few bisexual men. Model prevalence estimates are also roughly validated by incidence predictions, though available data on incidence are limited. While these assumptions may compromise the specific quantitative accuracy of the model predictions, the general predictions of relative efficacy of various interventions are robust.

Our model of the Newark HIV epidemic suggests that realistic implementation of current interventions will have a modest impact on decreasing HIV incidence;
however, a more marked impact on mortality is possible. Our results emphasize the importance of reinforcing additional forms of HIV-prevention and reducing losses of patients throughout the entire care-continuum. We believe that the implications of these results are important, and may have policy applications for other US areas. It is imperative that innovative strategies to link and retain HIV-infected persons in care be developed and tested. Only by stopping care-continuum leakage will we be able to effectively decrease new HIV infections in the US.

See Appendix A for full supplementary information.

Acknowledgements

The authors would like to thank Paula Grant from the New Jersey State Department of Health, Division of HIV, STD and TB Services for assistance with data acquisition. Further thanks to Steven Saunders, MS, Director, Prevention and Education, New Jersey Department of Health, Division of HIV, STD and TB Services, Barbara Bolden, PhD, Acting Director, Epidemiologic Services, New Jersey Department of Health, Division of HIV, STD and TB Services, and Professors Andrea Graham and Simon Levin from the Princeton University Department of Ecology and Evolutionary Biology for their helpful comments.

Financial Support

The project was supported by a grant from the New Jersey Health Foundation, formerly University of Medicine and Dentistry Foundation. RBB was supported by a Princeton University Department of Ecology and Evolutionary Biology Fellowship. BTG was supported by the Bill and Melinda Gates Foundation and the RAPIDD program of the Science and Technology Directorate, U.S. Department of Homeland Security.
Security, and the Fogarty International Center, NIH. TBH was supported by the Bill and Melinda Gates Foundation and Imperial College.

Conflict of Interest

SLH has received research support from BMS, Gilead, Jansen Therapeutic and Viiv-GSK, and has served as a consultant, Jansen Therapeutic, BMS, Gilead, and Merck and Co. SLH’s spouse is a Becton-Dickinson Board of Director member, and has Merck and Co. stock options. AS has serves as a consultant for Merck Inc. All other authors: no conflicts.
Chapter 2

Impact of Pre-Exposure Prophylaxis with Tenofovir Disoproxil Fumarate/Emtricitabine on Human Immunodeficiency Virus Incidence in Newark, New Jersey

Heidi H. Robbins\textsuperscript{1,*}, Ruthie B. Birger\textsuperscript{1,*}, Anushua Sinha\textsuperscript{2}, Bryan T. Grenfell\textsuperscript{1,3}, Sally L. Hodder\textsuperscript{4,5}

\textbf{1} Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ

\textbf{2} Department of Preventive Medicine and Community Health, New Jersey Medical School, Rutgers University, Newark, NJ

\textbf{3} Fogarty International Center, National Institute of Health, Bethesda, MD

\textbf{4} Department of Medicine, West Virginia University School of Medicine, Morgantown, WV
West Virginia Clinical and Translational Science Institute, Morgantown, WV

* Authors contributed equally.

Abstract

Background

Recent guidelines recommend pre-exposure prophylaxis (PrEP) for use by homogeneous sub-populations of high-risk individuals (e.g. men who have sex with men (MSM), high-risk heterosexuals (HRH), injection drug users (IDU)) to prevent HIV acquisition. Studies of tenofovir disoproxil fumarate (TDF) and emtricitabine (FTC) PrEP have demonstrated efficacy in specific populations. We model impact of TDF/FTC PrEP on HIV incidence in a heterogeneous population with multiple mechanisms of HIV transmission and complex contact patterns.

Objective

Estimate effectiveness of tenofovir/emtricitabine PrEP on HIV incidence in a real-world population.

Design, Setting, and Participants

Compartmental model of HIV infection, fitted to Newark, NJ and developed to include PrEP interventions.

Interventions

We simulated PrEP delivery varying target populations, coverage, drug-efficacy, adherence, and scale-up rate based on ranges demonstrated in clinical trials.
Measurements

Incidence reduction and cumulative infections averted 10 years post-intervention, number of PrEP person-years required to avert one infection.

Results

Targeting MSM with 10-50% coverage generated a 6.3-20.1% reduction in population HIV incidence over a 10-year period, averting 159-688 infections. Targeting HRH reduced HIV incidence by 5.5-21.2%. Interventions targeting IDU were less impactful, reducing HIV incidence by 1.8-15.8%, but were the most efficient, averting one new HIV infection over a 10-year period for every 11-29 PrEP person-years. Herd effects were present, with PrEP averting secondary cases in populations not targeted. 10% coverage of the total population, with heterogeneous adherence, reduced HIV incidence by 21.3-46.2%.

Limitations

PrEP-users never discontinue PrEP. Targeting by age or during ‘seasons of risk’ was not assessed. MSM studies informed assumptions about dose-dependent adherence for all risk groups.

Conclusions

This model confirms that realizable PrEP delivery can achieve HIV incidence reductions in Newark, NJ. A comparison with other modeling studies further underlines that MSM-targeted PrEP could potentially decrease HIV incidence, with potential for substantial impact amongst MSM. Targeting IDU is the most efficient strategy to reduce HIV incidence, although population impact is less.\textsuperscript{1}

\footnote{\textsuperscript{1}This manuscript has been submitted to Clinical Infectious Diseases}
Introduction

Despite demonstration that antiretroviral therapy (ART) markedly decreases HIV transmission (99), HIV incidence remains high in particular communities in the U.S., with certain groups bearing a disproportionate burden of infection. Newark, NJ is one such hotspot for the HIV epidemic, with an overall HIV prevalence of 2.6%, ranging up to 3.1% for certain sub-populations (109, 110, 111). Newark’s risk profile includes men who have sex with men (MSM), high-risk heterosexual (HRH) individuals, and injection drug users (IDU), and there are complex interactions among HIV risk-groups. To plan effective prevention and curb HIV incidence in other settings, previous studies have employed mathematical models to evaluate Test-and-Treat interventions along the HIV care-continuum, which includes testing, linkage to care, retention in care, and viral suppression with ART (71, 83, 101, 112).

A novel public health strategy to control the HIV epidemic, pre-exposure prophylaxis (PrEP) using tenofovir disoproxil fumarate (TDF) and emtricitabine (FTC), has demonstrated efficacy in clinical trials (113, 114, 115, 116, 117). Previous modeling studies have demonstrated the potential of PrEP to decrease HIV incidence, resulting in significant clinical and policy interest (118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128). With recent FDA approval of co-formulated oral TDF/FTC PrEP and publication of clinical prescribing guidelines, it is now especially critical to model PrEP’s potential impact in the context of the realistic population dynamics in a U.S. hotspot (129). As previous studies have often focused on one route of HIV transmission in homogenous populations with high HIV prevalence, further study of PrEP’s effectiveness in a heterogeneous U.S. population, exploring the impact of targeted PrEP delivery to different high-risk groups, is warranted (130).

Here, we use a mathematical model to simulate different PrEP targeting scenarios amongst various populations in Newark, NJ. We extend a previously developed HIV care-continuum model (71), which successfully captures recent heterogeneous...
dynamics of Newark’s epidemic, to assess the impact of PrEP interventions on HIV incidence in Newark. We model a range of plausible and optimal interventions, assessing different combinations of PrEP drug efficacy, adherence, and coverage of various sub-populations.

In the discussion, we make a broad comparative analysis between our results and those of other studies modeling PrEP. This type of analysis is challenging since studies 1) model different populations; 2) assign different metrics for coverage, efficacy, and adherence; and 3) evaluate different health impacts at different population levels. Nevertheless, our model facilitates this analysis because interventions occur within the same population and have consistent metrics. We compare our simulations’ impacts to those of other PrEP modeling studies after standardizing other studies’ coverage, efficacy, adherence, and impact metrics.

Methods

Mathematical Model

Details regarding this compartmental model of HIV infection have previously been published (71). Briefly, it is of standard form (10), comprising compartments for three infection stages, further divided according to testing eligibility, known diagnosis, and viral suppression, and two compartments for PrEP-users: one for high-adherers (user-adherence level is 90%) and another for low-adherers (user-adherence is 30%) Figure 2.1 provides a flow chart of the compartmental model with PrEP, highlighting parameters specific to these PrEP compartments. The differential equations, describing the compartment inflows and outflows, and the force of infection equations, describing the infection risk, are provided in Appendix B with further explanation of model parameters.
The model was calibrated to HIV/AIDS cases in Newark among different populations over a 10-year period (71). It accounts for heterosexual, homosexual, bisexual, and injecting drug-use transmission, allowing for combinations of sexual and injecting drug-use risk behaviors to capture the epidemiological profile of the Newark epidemic: among persons living with HIV/AIDS in Newark in 2014, 17% of cases were from male-to-male sexual transmission, 42% from heterosexual transmission, 24% from injecting drug-use transmission, 2% from male-to-male sexual transmission or injecting drug-use transmission, and 15% from unreported exposure (110). The Newark population is thus stratified in the model into twelve mutually exclusive behavioral risk-groups distinguished by gender, high- or low-risk sexual activity, and injecting drug-use. High-risk sexual activity is defined by a high partner change rate: 3-10 partners per year for high-risk male heterosexuals, 0-3 partners per year for high-risk MSM, and 1-5 partners per year for bisexual men (71). The ten groups at high risk for acquiring HIV include female HRH, male HRH, female IDU, male IDU, female HRH/IDU, male HRH/IDU, MSM, MSM/IDU, male HRH/MSM, and male HRH/MSM/IDU. The remaining two risk groups, low-risk females and low-risk males, have a partner change rate of 0-2 partners per year (71). Further risk-group specifications are provided in Appendix B.

Model Parameters

A literature review informed values for model parameters: user-adherence, population-adherence, population-coverage, and population-specific drug efficacies to prevent sexual and/or parenteral transmission of HIV (Table 2.1) (113, 114, 115, 117, 121, 123, 127, 131, 132, 133).
### Table 2.1: Summary of PrEP Parameter Values in the Model

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Newark Model</th>
<th>Range Found in Literature</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Coverage Levels (PCOV) (%)</strong></td>
<td>10 (moderate-cov)</td>
<td>HRH: 4-12 [123]</td>
</tr>
<tr>
<td></td>
<td>25 (high-cov)</td>
<td>MSM: 0.5-85 [121, 131]</td>
</tr>
<tr>
<td></td>
<td>50 (maximal-cov)</td>
<td>TOT: 2.7-75 [119, 127]</td>
</tr>
<tr>
<td><strong>Population-Adherence (q) (%)</strong></td>
<td>15 (low-q)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>60 (medium-q)</td>
<td>33-95 [121, 131, 115]</td>
</tr>
<tr>
<td></td>
<td>95 (optimal-q)</td>
<td></td>
</tr>
<tr>
<td><strong>User-Adherence for High Adherers (pre$_{adh-high}$) (%)</strong></td>
<td>90</td>
<td>95 [131]</td>
</tr>
<tr>
<td><strong>User-Adherence for Low Adherers (pre$_{adh-low}$) (%)</strong></td>
<td>30</td>
<td>15-40 [131]</td>
</tr>
<tr>
<td><strong>PrEP Efficacy against Sexual Transmission (pre$_{eff}$) (%)</strong></td>
<td>HET: 91</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Male Bisexuals: 94</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MSM: 97</td>
<td></td>
</tr>
<tr>
<td></td>
<td>HET: 91 [133]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MSM: 99 [132]</td>
<td></td>
</tr>
<tr>
<td><strong>PrEP Efficacy against Parenteral/Injecting Drug-Use Transmission (pre$_{eff-IDU}$) (%)</strong></td>
<td>83</td>
<td>IDU: 83.5 [134]</td>
</tr>
<tr>
<td><em><em>Weighted Effectiveness against Sexual Transmission</em> (%)</em>*</td>
<td>TOT: 36-80</td>
<td></td>
</tr>
<tr>
<td></td>
<td>HRH: 35-79</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MSM: 37-84</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IDU: 35-49</td>
<td></td>
</tr>
<tr>
<td></td>
<td>HRH: 28-90 [123, 117]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>HRH: 28-90 [114]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MSM: 16.5-92 [115, 121]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MSM: 16.5-92 [131]</td>
<td></td>
</tr>
<tr>
<td><em><em>Weighted Effectiveness against Parenteral Transmission</em> (%)</em>*</td>
<td>IDU: 32-72</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IDU: 49 [113]</td>
<td></td>
</tr>
<tr>
<td><strong>Rate of Scale-up (r) years</strong></td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>

* Effectiveness was calculated as \((\text{efficacy} \times \text{population-adherence} \times \text{high user-adherence}) + (\text{efficacy} \times (1-\text{population-adherence}) \times \text{low user-adherence})\).

HRH=high risk heterosexuals; MSM=men who have sex with men; TOT=total population; HET=heterosexuals; IDU=intravenous drug users.

### User-Adherence

In our model, overall PrEP effectiveness is a function of drug efficacy and user-adherence, which we define as the proportion of daily sex-acts or injections protected by PrEP with 90% or greater efficacy. There are two PrEP-user adherence compartments: high adherers take at least three doses of PrEP each week, which translates to a user-adherence of 90%, while low adherers take at most two doses each week, translating to a user-adherence of 30%.
We arrived at these user-adherence values of 90% and 30% by using pharmacokinetic studies \(135, 136\) and studies pairing number of doses per week with drug concentrations achieved \(132, 137\). Our full assumptions are detailed in Appendix B.

### Population-Specific Drug Efficacies

Clinical trials report different drug effectiveness in populations with varying user-adherence levels and even find different drug efficacy despite optimal reported adherence \(114, 115, 132, 133, 134, 138, 139, 140\). Consequently, in our model of PrEP in a heterogeneous population, we chose population-specific drug efficacies based on studies correlating drug concentration, daily dosing patterns and/or user-adherence, and effectiveness estimates from clinical trials of PrEP in various populations \(114, 115, 132, 133, 137, 134, 139\). The drug efficacies assumed for each population are as follows: a reduction in risk of sexually-transmitted HIV of 97% in MSM, 91% in HET, and 94% in bisexual males, and a reduction in risk of intravenously-transmitted HIV of 83% in IDU (Table 2.1). We consider drug efficacy to be the risk reduction achieved with perfect or near-perfect user-adherence, which explains why the population-specific efficacy levels we calculate are higher compared to drug effectiveness values used by other models or reported by clinical trials in which partial adherence reduces the effectiveness. Our calculations and assumptions behind these population-specific efficacies are explained in Appendix B.

### Targeting and Coverage Levels

We simulated nine intervention scenarios for five targeted populations: 1) the entire population, 2) all high-risk groups, 3) HRH, 4) MSM, and 5) IDU (Figure 2.1). The nine interventions for each targeting strategy result from all possible combinations of three population coverage levels (10%-moderate, 25%-high, and 50%-maximal)
and three population-adherence levels (15%-low, 60%-medium, and 95%-optimal), where population-adherence is the percentage of the population that is high-adhering (user-adherence of 90%) versus low-adhering (user-adherence of 30%) (Figure 2.1). Population coverage levels of 10%, 25%, and 50% were chosen as round, identifiable coverage standards, which could translate easily into public health policy. The range includes more feasible or realistic coverage targets with 10% and 25% coverage, while 50% coverage provides a more intriguing pressure-test of PrEP’s impact, as 100% coverage and perfect adherence would avert 100% of all infection, drug resistance notwithstanding.

Outcomes and Confidence Bounds

All PrEP interventions begin in 2016 and assume PrEP-users continuously use PrEP initiation for the duration of the 10-year period that is modeled. Impact is measured as HIV incidence reductions in 2026, cumulative infections averted from 2016-2026, and number of PrEP person-years required to avert one new infection from 2016-2026. The 95% confidence interval for these health impacts is calculated from the upper- and lower- bounds of these health impacts produced by model runs with the initial model’s 100 parameter sets, whose outputs of HIV prevalence for Newark’s sub-populations fit the prior limits.

Results

Baseline Model HIV Prevalence and Incidence

We use the model, calibrated to Newark’s historic HIV epidemiologic profile, to forecast the natural history of HIV infection in the absence of PrEP or other expanded prevention strategies. We estimate that 526 (±3) new HIV infections will occur in 2016 among Newark’s population of 222,560 (SE ±148) sexually active individuals.
Figure 2.1: PrEP compartmental model flow-chart and intervention parameters. Compartments for PrEP-users with high and low user-adherence ($P_H$ and $P_L$), as well as the compartments for the HIV care-continuum. Susceptible individuals flow into a PrEP compartment based on a population coverage scale-up equation, $H$. A proportion, $q$, of PrEP-users have high user-adherence, which is defined as 90% of sex-acts or intravenous injections protected by PrEP with 90% or greater efficacy. High-adherers take 3 or more daily doses of PrEP per week. $1-q$ is the proportion of PrEP-users with low user-adherence, defined as 30% of sex-acts or intravenous injections protected by PrEP with 90% or greater efficacy. Low-adherers take 2 or fewer doses of PrEP per week. $q$ is the population-adherence PrEP-users can still become infected and flow into HIV Acute Infection compartment, $I_o$. PrEP forces of infection, $\lambda_{p\text{-low}}, \lambda_{p\text{-high}}, \lambda_{p\text{-drug\text{-low}}}, \lambda_{p\text{-drug\text{-high}}}$, are influenced by PrEP efficacy and user-adherence. Interventions modeled assessed combinations of coverage levels (10%, 25%, and 50%) and population-adherence, $q$, levels (15%, 60%, and 95%) for 5 targeted populations: HRH, MSM, IDU, All high-risk groups, and the total population. Abbreviations: HRH= high-risk heterosexuals, MSM= men who have sex with men, IDU= injection drug users, TOT= total population.
Model simulations of PrEP deployment in 2016 to the general population or targeted to MSM, IDU, or HRH risk-groups can significantly reduce Newark’s HIV incidence in the total population and in these specific sub-populations in subsequent years. Model estimates for the baseline epidemiological profile of Newark’s HIV/AIDS epidemic in 2016, stratified by high-risk sub-population are shown in Supp. Table B.1.

Newark’s population of high-risk MSM, which includes MSM and MSM who inject drugs, numbers 4,557 (SE ±287) and is 2.0% (SEp ±.13%) of the total population according to model calculations for 2016 which start with the baseline proportion of the population in each risk-group and project forward in time with a constant population size. The HRH population, with sexually high-risk women and men, injecting-drug-users of both sexes, and bisexuals, numbers 51,658 (±319) individuals and constitutes 23.2% (±.14%) of the population. The IDU population of 3,228 (±117) individuals is 1.4% (±.05%) of the total population and comprises sexually low-risk and high-risk heterosexuals, bisexuals, and MSM who inject drugs.

Impact of PrEP Interventions

Cumulative Infections Averted by PrEP Interventions (2016-2026)

Interventions targeting the entire (i.e., total) population averted the most infections over a 10-year period, with infections averted ranging from 582 for an intervention with the lowest population-adherence and coverage to 2665 for an intervention with the highest population-adherence and coverage. Other intervention combinations of intermediate population-adherence and coverage produced impacts within this range of 582-2665 infections averted (95% CI 198-4441), where the 95% confidence interval in parentheses is the lower bound of the least impactful intervention and the upper bound of the most impactful intervention (Figure 2.2). Targeting all high-risk groups was the next most impactful intervention, averting 344-1709 (67-3146) infections in the total population. HRH-targeting averted 148-762 (41-2576) infections, while MSM-
targeting averted 159-688 (7-1593) infections, and IDU-targeting averted 68-597 (2-1308) infections.

**Herd Effects**

A key issue in determining the population impact of PrEP is the extent of herd protection, the reduced risk of HIV transmission to susceptible sexual and needle-sharing
partners and others because of PrEP administration to at-risk groups. Because of herd protection, incident cases were averted in populations not targeted by a particular PrEP intervention (Figure 2.3). IDU-targeting averted 18-158 (95% CI 0-364) incident cases from 2016-2026 amongst HRH and 2-23 (0-69) cases amongst MSM. In this same time period, HRH-targeting averted 4-23 (0-28) cases amongst MSM and 30-204 (1-463) cases amongst IDU, and MSM-targeting averted 18-83 (0-112) cases amongst HRH and 6-40 (0-142) cases amongst IDU. These results signify complex dynamics in contact patterns between risk-groups and important indirect effects, with PrEP averting secondary cases.

Figure 2.3: **Herd effects of PrEP.** The indirect effects of incident cases averted in other populations in addition to the targeted population by intervention combinations of population coverage and population-adherence. Abbreviations: HRH= high-risk heterosexuals, MSM= men who have sex with men, IDU= injection drug users, AHR= all high-risk groups, TOT= total population.
Percent Reduction in HIV Annual Incidence in 2026 by PrEP Interventions

Depending on the population-adherence and coverage levels, reductions in HIV annual population incidence in 2026 achieved by targeting the total population were 21.3-68.8% (95% CI 6.7-83.9%) (Figure 2.4). Targeting all high-risk groups was the next most impactful targeting strategy and reduced population incidence by 12.5-46.4% (2.2-69.3%). HRH-targeting and MSM-targeting followed, with 5.5-21.2% (1.6-51.5%) and 6.3-20.1% (0.2-31.0%) reductions in population incidence, respectively. IDU-targeting reduced population incidence by 1.8-15.8% (0-36.9%).
Figure 2.4: Percent reduction in HIV incidence by PrEP interventions. Impact of PrEP targeting strategies for intervention combinations of population coverage and population-adherence, with respect to the total population HIV incidence. Abbreviations: HRH= high-risk heterosexuals, MSM= men who have sex with men, IDU= injection drug users, AHR= all high-risk groups, TOT= total population.
Targeting HRH, MSM, or IDU reduced within-population HIV incidence rates substantially. MSM-targeting was most impactful, reducing incidence amongst MSM by 29.3-83.2% (3.2-88.9%) (Figure 2.5). HRH-targeting followed in impact, reducing incidence amongst HRH by 17.6-61.8% (6.2-81.9%). Reductions in incidence amongst IDU for IDU-targeting were 3.5-30.1% (0-67.7%).
Per Capita Impact of PrEP

The efficiency or per capita impact of interventions was assessed as the number of person-years of PrEP coverage needed to avert one new HIV infection over a 10-year period. Targeting IDU was most efficient, averting one new HIV infection for every 11-29 person-years of PrEP coverage (95% CI 5-1044) (Figure 2.6, Supp. Table B.2). MSM-targeting was the next most efficient targeting method, requiring 27-62 (11-1359) person-years. HRH-targeting was less efficient and needed 414-1027 (106-3821) person-years. Targeting all high-risk groups and targeting the total population were both also relatively inefficient: targeting all high-risk groups required 197-471 (96-2418) person-years, while targeting the total population needed 477-1170 (274-3442) person-years.

Discussion

In this case study of a heterogeneous population with different HIV risk behaviors and transmission mechanisms, we find that PrEP can avert 4.7-10.8% of incident cases over a ten-year period in the total population with moderate (10%) coverage of either the HRH, MSM, or IDU sub-population and 60% population-adherence. This impact increases to 10.9-16.7% of incident cases averted with 95% population-adherence. Within high-risk sub-populations, substantial 10-year HIV incidence reductions; ranging from 47.2-70.2% among MSM, 30.5-48.1% among HRH, and 8.9-20.7% among IDU for 60% and 95% population-adherence levels; are achieved even with 10% coverage. Such targeted strategies benefit not only risk-group members but also their contacts by preventing secondary transmission (120). Targeting MSM and IDU realizes these incidence reductions with greater efficiency than other targeting strategies.

Providing PrEP to MSM is both an efficient and effective targeting strategy. HRH-targeting has an impact comparable to MSM-targeting, averting a similar number of
Figure 2.6: Median number of PrEP person-years required to prevent one new HIV infection over a 10-year period (2016-2026). Intervention combinations of population coverage and population adherence produced a range of person-years, which was used to calculate the median number. The range is shown in parentheses. Abbreviations: HRH= high-risk heterosexuals, MSM= men who have sex with men, IDU= injection drug users, AHR= all high-risk groups, TOT= total population.
infections and reducing population HIV-incidence by a similar amount. However, HRH-targeting is less efficient than targeting MSM or IDU in terms of per capita impact. Targeting IDU is the most efficient strategy, but the cumulative infections averted are fewer and the percent reduction in incidence is less than that of MSM-targeting and HRH-targeting because of the small number of IDUs. Population-wide targeting and targeting all high-risk groups have the largest health impact, with regard to cumulative infections averted and percent reduction in incidence, but are relatively inefficient.

If implementable, the results of a major deployment of PrEP, with high coverage and population adherence levels, would be compelling: IDU-targeting with high (25%) coverage and 60% population-adherence achieves a 5.9% HIV incidence reduction, and HRH- and MSM-targeting at this intervention level reduce HIV incidence by 13%. 10% coverage of the general population with PrEP reduces HIV incidence between 21.3-55.5%, depending on population-adherence.

The need for major deployment and high coverage levels aside, this model still demonstrates that PrEP can achieve significant HIV incidence reductions within Newark's heterogeneous population with moderate coverage (10%) and with drug efficacy, population-adherence, and user-adherence demonstrated in clinical trials (114, 115, 132, 133, 137, 134, 139).

**Comparative Analysis of Results**

Our results report similar health impacts to some published PrEP models (118, 120) and a smaller impact than others (119, 121, 123). To compare model results, we graphed PrEP intervention impacts for our and other modeling studies in the parameter space of assumed coverage, effectiveness, and population HIV prevalence after standardizing metrics for PrEP coverage, efficacy, adherence, and effectiveness (Supp. Figure B.1 and Supp. Table B.3). Despite a similar HIV prevalence among MSM in
Newark and New York City (16.4% compared to 14.6%), Desai et al’s 2008 model of MSM-targeted PrEP in New York City predicts a comparable impact with less coverage and drug effectiveness than our model and in half the amount of time [121]. Desai et al find a 5.4% cumulative incidence reduction over five years of 25% coverage of high-risk MSM (5.2% of the total MSM population) and assume low PrEP effectiveness (16.5%). MSM-targeting in Newark achieves a similar cumulative incidence reduction of 4.7-6% (95% CI 0-11.3%) with 10% coverage and 60%-95% population-adherence, calculating to 62-82% effectiveness. This difference may be explained by Desai et al’s faster rate of scale-up, calculation of HIV infection rate by age and transmission probability by partnership duration, and differences in parameters accounting for partner change rates and number of sex acts.

A model by Juusola et al calculated PrEP health impacts in a general and high-risk MSM population calibrated to U.S. national demographics where HIV prevalence and incidences were 12.3% and 0.8%, respectively and PrEP effectiveness 44%. 10% coverage of the general MSM population generated a 7% cumulative reduction in HIV incidence over the 20-year period [120]. In comparison, MSM-targeting in Newark with 15% population-adherence, calculating to 37% effectiveness, and 10% coverage resulted in a 10-year cumulative reduction of 2.9% (95% CI 0-5.5%). This result approximates that of Juusola et al when scaled over 20 years.

Chen and Dowdy’s 2014 model of MSM-targeted PrEP in the U.S. calculated the number needed to treat (NNT) to avert one HIV infection in a year and reported an NNT of 64 PrEP-users, assuming 19% HIV prevalence and 92% PrEP effectiveness [118]. In Newark with 16.4% HIV prevalence amongst MSM, we calculated needing 28-60 (95% CI 12-1646) PrEP-users in 2026 to avert one infection amongst MSM depending on the coverage and population-adherence combination (Supp. Table B.2). By this measurement, our model predicts a greater per capita impact of PrEP than that found by Chen and Dowdy. Possibly, secondary transmission in our model leads
to an increasing number of infections averted; this effect should increase after the
scale-up period of PrEP deployment (120).

More general population models of PrEP in endemic settings, such as those by
Cremin et al and Abbas et al in sub-Saharan Africa, assume a higher drug effec-
tiveness and, at times, higher coverage levels than our model. These models use
age-targeted and or high-risk-group-targeted PrEP delivery to populations in which
HIV prevalence is an order of magnitude greater than that of Newark (Supp. Figure
B.1) (71, 119, 123). Assuming high efficacy and adherence levels of 75% and 95%
which calculate to 71% PrEP effectiveness, Cremin et al find a 3.2% reduction in
cumulative infections with 7.3% coverage of susceptible 15-24-year-olds (119). Abbas
et al find a 3.3% reduction in cumulative infections with 25% coverage of the gen-
eral population and 30% PrEP effectiveness (123). Comparatively, population-wide
targeting in Newark achieves 3% cumulative incidence reductions with 25% coverage
and 36% PrEP effectiveness.

It is unsurprising that there is a variance in reported health impacts for PrEP
interventions and that discrepancies exist between our results and those of other
published PrEP models. It should be noted that the summary effectiveness parameter
for interventions in this model, influenced by our adherence assumptions, errs on the
conservative side, while there is a wide range of coverage levels. Whereas other
models have looked at particular sub-populations or a single type of transmission and
often model populations with much higher HIV prevalence, a strength of this model
is that it covers a very heterogeneous population with a variety of risk-groups and
transmission mechanisms and, for most interventions, assumes coverage, efficacy, and
adherence levels feasible for PrEP implementation, as opposed to optimal.
Limitations

The model included no option for PrEP-users to discontinue PrEP after initiating treatment or switch between high and low adherence compartments since the model assumes that risk behavior trajectories remain stable throughout a person’s sexual or injecting-drug-use career (i.e. high-risk individuals remains at high risk even later in life) (141). The model does not explore targeting individuals with PrEP during time-limited periods or ‘seasons’ of high-risk behavior (141), a strategy that could boost the efficiency of PrEP’s delivery. Likewise, if younger individuals are at higher-risk for acquiring HIV, an age-stratified model, allowing for age-targeted PrEP delivery, should enhance the efficiency of interventions while achieving the same impact. We also do not address the issue of behavioral risk compensation in this model, which could reduce PrEP’s effectiveness.

In our assumptions about user-adherence and drug efficacy, we assumed daily sex acts and drug injections. This is likely a significant overestimation, leading to underestimation of PrEP effectiveness and health impacts. Also, we assumed the same user-adherence values, informed by studies correlating drug concentration and dosing frequency with sexual transmission efficacy in MSM, for all types of transmission. If we assumed lower user-adherence values for IDU or HRH, we would expect PrEP’s impact in those sub-populations to be slightly less. In addition, higher drug levels may be required to afford similar protection to women engaging in vaginal sex with a man compared with MSM engaging in receptive anal intercourse.

The model assumes IDU are at high-risk for acquiring HIV, although we account for a reduction in the number of IDU over time. This is a reasonable assumption as HIV prevalence remains high among IDU, and there is a scarcity of data pertaining to the impact of Newark’s sterile syringe access programs (SAPs), which began in 2008, on HIV incidence among IDU (142). In the context of a strong needle exchange program, we might expect PrEP delivery to IDU to be less beneficial than in the
absence of needle exchange. However, our results for IDU-targeted PrEP predict a sizeable HIV incidence reduction amongst IDU and may underscore that SAPs are currently not reaching IDU most at-risk for acquiring and transmitting HIV.

Wide confidence intervals provide a buffer to over- or under-estimation of PrEP’s impact given these limitations. More importantly, the general trends regarding which targeting strategies are more effective and efficient in a heterogeneous population are robust and informative for public health policy, especially since there is a scarcity of models investigating PrEP delivery to IDU and HRH in the United States. Model comparisons of PrEP delivery to MSM qualitatively support the incidence reductions found in this model but also imply that more effort (i.e. identification and ensuring adherence) is needed to realize these impacts in a heterogeneous population.

Conclusions and Next Steps

From a public health standpoint for Newark, these results highlight the importance of identifying and linking high-risk persons to PrEP, especially IDU and MSM who bear a disproportionate burden of HIV infection. Further work should consider strategies to identify IDU and MSM at high risk, link them to PrEP treatment, and promote drug adherence. Additionally, effective PrEP implementation will require monitoring and testing PrEP-users to boost adherence and counter the risk of emerging drug resistance with HIV sero-conversion.

Identification might dovetail with current HIV prevention programs (i.e. testing centers and outreach programs) or SAPs. In the context of an earlier model’s results concerning care-continuum interventions in Newark (71), PrEP in combination with care-continuum strategies would likely present an even stronger arsenal against HIV infection, and strategies to increase HIV testing could operate in tandem with identification of high-risk persons for PrEP. However, at a certain point, successful
suppression of infected individuals’ HIV viral load and reduction in population incidence by care-continuum interventions should diminish the effectiveness of PrEP as a public health strategy.

In the case of IDU, methadone maintenance programs could foster adherence with directly observed therapy for both PrEP and methadone. Extensions of this modeling work already underway include modeling concurrent scale-up of methadone maintenance programs, HCV treatment, and the HIV care-continuum in Newark. Additional modeling of PrEP should focus on combined targeting of IDU and MSM and the impact of strengthening the care-continuum on PrEP’s effectiveness. Such information will be important for implementation and policy.

See Appendix B for full supplementary information.

Acknowledgements

The authors thank Professor Timothy Hallett from Imperial College London, School of Public Health for his contributions to the initial modeling study and his code expertise.

Financial support

HR and RB were supported by a research grant from the Program in U.S. Health Policy, Center for Health and Wellbeing, Princeton University. RB was supported by the Princeton University Department of Ecology and Evolutionary Biology. BG was supported by the Bill and Melinda Gates Foundation and the RAPIDD program of the Science and Technology Directorate, U.S. Department of Homeland Security, and Fogarty International Center, NIH. SH was supported by the National Institute of General Medical Sciences under Award Number U54GM104942.
Potential conflicts of interest

SH has disclosed that she has received consulting fees from Bristol-Myers Squibb, Gilead Sciences, Janssen, Merck, and ViiV; has received funds for research support from Bristol-Myers Squibb, Gilead Sciences, Janssen, and ViiV; and her spouse has ownership interest in Merck. All other authors report no potential conflicts.
Modeling the Effect of HIV Coinfection on Clearance and Sustained Virologic Response during Treatment for Hepatitis C Virus

Ruthie Birger\textsuperscript{1}, Roger Kouyos\textsuperscript{2}, Jonathan Dushoff\textsuperscript{3}, Bryan Grenfell\textsuperscript{1,4}

\textsuperscript{1} Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ, USA
\textsuperscript{2} Division of Infectious Diseases and Hospital Epidemiology, University Hospital Zürich, University of Zürich, Zürich, Switzerland
\textsuperscript{3} Department of Biology, McMaster University, Hamilton, Ontario, Canada
\textsuperscript{4} Fogarty International Center, National Institutes of Health, Bethesda, MD 20892, USA
Abstract

Background

HIV/Hepatitis C (HCV) coinfection is a major concern in global health today. Each pathogen can exacerbate the effects of the other and affect treatment outcomes. Understanding the within-host dynamics of these coinfecting pathogens is crucial, particularly in light of new, direct-acting antiviral agents (DAAs) for HCV treatment that are becoming available.

Methods and Findings

In this study, we construct a mathematical within-host model of HCV/HIV coinfection by adapting a previously-published model of HCV monoinfection to include an immune system component in infection clearance. We explore the effect of HIV-coinfection on spontaneous HCV clearance and sustained virologic response (SVR) by building in decreased immune function with increased HIV viral load. Treatment is modeled by modifying HCV burst-size, and we use clinically relevant parameter estimates.

Our model replicates real-world patient outcomes; it outputs infected and uninfected target cell counts, and HCV viral load for varying treatment and coinfection scenarios. Increased HIV viral load and reduced CD4$^+$ count correlate with decreased spontaneous clearance and SVR chances. Treatment efficacy/duration combinations resulting in SVR are calculated for HIV-positive and negative patients, and crucially, we replicate the new findings that highly efficacious DAAs reduce treatment differences between HIV-positive and negative patients. However, we also find that if drug efficacy decays sufficiently over treatment course, SVR differences between HIV-positive and negative patients reappear.
Conclusions

Our model shows theoretical evidence of the differing outcomes of HCV infection in cases where the immune system is compromised by HIV. Understanding what controls these outcomes is especially important with the advent of efficacious but often prohibitively expensive DAAs. Using a model to predict patient response can lend insight into optimal treatment design, both in helping to identify patients who might respond well to treatment and in helping to identify treatment pathways and pitfalls.\footnote{published in *Epidemics*, Vol. 58(2), pgs. 1–10 (2015); RB Birger, RD Kouyos, J Dushoff, BT Grenfell (71). It has also been presented at the Ecology and Evolution of Infectious Diseases Conference at Penn State University in May 2013 and the Epidemics Conference in Amsterdam in November 2013.}

Introduction and Background

HIV-HCV coinfection is a whole greater than the sum of its parts, due to the potentiating effect each virus can have on the other. It is an increasing concern in certain populations including people who inject drugs (PWID) and men who have sex with men (MSM). Understanding the within-host dynamics of coinfection is crucial for designing treatment strategies that will avoid complications such as hepatotoxicity and treatment failure, while minimizing the cost of treatment. There is a rich literature on monoinfection with each pathogen (e.g. (56, 3, 61, 143, 144, 145, 62, 1)) but to our knowledge, within-host HIV-HCV coinfection has not yet been modeled.

HIV and HCV are both viral infections that can be blood-borne. They are often transmitted together, especially among PWID (146, 147). The infections can interact synergistically. Broadly, HIV causes deterioration of the immune system, which can lead to poorer control and clearance rates of HCV as well as reduced probability of treatment success (148), while HCV may increase progression rates of HIV through chronic immune activation or increased CD4 apoptosis (149). Treatment success for
HCV is defined as achieving sustained virologic response (SVR), i.e. HCV RNA is undetectable during treatment and for 6 months beyond treatment end (150).

We approach clinical implications of coinfection with a dynamic-static mathematical model, based on a previously published dynamical model of HCV monoinfection (151, 152, 153) to include an explicit element of immune control. We then perturb the system by adding HIV infection as a static parameter, which erodes immune control and thus changes the HCV dynamics. Lastly, we explore how this loss of control impacts HCV treatment outcomes.

Understanding phenomena that may impact treatment outcomes is especially important with the advent of new direct-acting antiviral agents (DAAs) such as Sofosbuvir that are becoming available for HCV treatment. These new drugs are highly efficacious, but still very expensive, so thoughtful treatment design and administration is necessary (152, 153, 154).

**Immune/HCV Interactions**

The immune response to HCV is complex and involves both innate and adaptive components. The adaptive response to HCV is mostly T-cell dependent. Virus-specific CD4$^+$ and CD8$^+$ responses have been detected during acute infection, and it is thought that these cells may clear virus by lysing infected cells or by cytokine/chemokine-mediated effects (148).

HCV may be responsible for immune dysregulation itself; some studies indicate an inverse correlation between HCV viral load and CD4$^+$ count (155). HCV may also downregulate proliferation of T cells or increase apoptosis. Increased immune activation in coinfected versus monoinfected patients has been noted, which also may speed HIV progression (156). A schematic summary of HIV-HCV interaction effects is shown in Figure 3.1.
Empirical Evidence for Effect of HIV Infection on HCV Clearance and Treatment Response

Patients who are able to clear HCV spontaneously have been noted to mount intense multispecific CD4\(^+\) and CD8\(^+\) responses, in particular with HCV-specific CD4\(^+\) Th1 responses (157). Complementarily, some studies show that lower CD4\(^+\) counts are associated with reduced probability of clearing (158). However, HCV can avoid these responses through mutation and inhibition of dendritic cell activation and production of Th1 cytokines. This can prevent cytotoxic lymphocyte (CTL)-induced apoptosis of infected hepatocytes, which can in turn increase viral production (157).
Some HIV-positive patients do clear HCV, but in much lower proportions than HIV-negative individuals. The correlative evidence for mechanism is mixed. Higher rates of chronic HCV are inversely correlated with CD4+ count according to some studies reviewed by (148). A study in chimpanzees showed reduced endurance of HCV-specific CD8+ CTL response after depletion of CD4+ cells before reinfection; viremia upon reinfection was persistent despite the presence of functional CD8+ T-cells (158). Other studies have found more general negative correlation between HIV viral load and SVR (159, 157). The phenomenon of increased persistence of HCV among HIV-positive patients is related to and accompanied by an increased HCV viral load.

A number of empirical studies show elevated HCV viral load (.1-1 logs higher) among HIV-positive as compared to HIV-negative patients (160, 161, 155, 157). The cause is most likely reduced immune response: coinfection is associated with lack of critical CD4+ response to HCV (161, 162). Indeed, the study in chimpanzees mentioned above showed that CD8+ T-cells had impaired control of viral replication with insufficient CD4+ help (158), and there is evidence that broad CD4 responses play a major role in HCV clearance (163).

Role of HIV Treatment

In clinical settings where HCV treatment is available, most coinfected patients will have been treated with anti-retroviral therapy (ART) for HIV prior to receiving treatment for HCV despite some risks of hepatotoxicity from ART (164, 165). However, even when ART is successful and results in achievement of virological suppression, CD4+ recovery is often incomplete. Many studies have shown that CD4+ at similar durations after ART initiation varies widely and is correlated with a range of patient characteristics such as nadir CD4+ count (6, 166, 167) (see Supp. Fig. C.1). Corre-
spondingly, there is evidence that even coinfected patients who are treated for HIV have a lower probability of clearing HCV \((168)\) or achieving SVR \((169, 170, 171)\).

**Previous Models of HCV**

Some of the first within-host models of HCV aimed to capture the dynamics of infection by pairing a mathematical model with patient data from a trial of varying doses of interferon (IFN) treatment \((3, 61)\). The viral load patterns upon treatment observed (i.e. rapid initial decline in viral load followed by extended slow decline) were consistent with the main mechanism of IFN treatment being reduction in production of new virus by infected cells (burst size). Higher doses were more effective, and lower diversity of quasi-species were associated with better response to treatment.

Extensions to this model take into account complexities surrounding treatment response, such as non-response, rebound in viral load (relapse), or treatment with other therapies \((144, 143)\).

**Models of Extended Infection**

The models previously discussed deal with acute HCV infection and treatment dynamics over the course of several days or weeks. HCV can be a long course infection, however, and models can capture longer term dynamics as well.

The model created for this study is based on previous work \((62, 1, 151)\) and explores treatment dynamics past initial infection in detail. The conditions for achievement of SVR can be inconsistent. For example, the “End-of-Treatment” response, defined as undetectable viral load (below 2 logs) at the end of a 24 or 48 week course of treatment, is necessary but not sufficient for SVR. Debroy et al \((1)\) analyze medium- and long-term responses leading to either a viral endemic equilibrium or a disease-free equilibrium to establish mathematical criteria for each state. Depending on the initial conditions and biological parameter values, there exists a possibility of bistability.
for some patients (i.e. they have the capacity to clear infection, but only if treated adequately). The parameter values depend on patient immunological characteristics, as well as viral characteristics and interactions, which can vary by genotype (as explored by (172)). Our model aims to extend the analysis of these previous models in scenarios of HIV coinfection.

Methods

HCV Monoinfection Model

The model proposed in this study follows closely the above-referenced model (151, 62, 1). Following a standard ODE model proposed in earlier work (173), with some variations such as inclusion of infected hepatocytes in density dependence, the basic form of the equations (without the immune system component) are as follows

\[
\frac{dT}{dt} = s + r_1 T \left(1 - \frac{T + I}{T_{max}}\right) - dT - \beta TV
\]
\[
\frac{dI}{dt} = \beta TV + r_2 I \left(1 - \frac{T + I}{T_{max}}\right) - \delta I
\]
\[
\frac{dV}{dt} = pI - cV
\]

Here, \(T, I,\) and \(V\) are the state variables representing respectively uninfected Target cells (hepatocytes), Infected hepatocytes, and free HCV Virus, \(s\) is the recruitment rate for uninfected hepatocytes, \(r_1\) is the reproduction rate of uninfected hepatocytes, \(T_{max}\) is the maximum number/carrying capacity of Target cells, \(d\) is the death rate of uninfected Target cells, \(\beta\) is the mass-action infection parameter, \(r_2\) is the reproduction rate of infected hepatocytes, \(\delta\) is the clearance rate of infected
hepatocytes, $p$ is the number of virions an infected cell produces in its lifetime (which can also be interpreted as burst size), and $c$ is viral clearance rate. With this model of HCV monoinfection as a basis, we can build a new model of HCV that includes the role of the immune system – specifically, CD4$^+$ cells, denoted $H$ in the model equations. Work by (158) and others has illuminated to some extent the role of CD4$^+$ cells in HCV control, and activation of CD4$^+$ cells by HCV has been documented (174, 178). The model proposed here thus incorporates a dependence of the HCV clearance rate on CD4$^+$ count ($\alpha$) and a dependence of the activation rate of CD4$^+$ cells on HCV infected cell count ($\gamma$).

\[
\begin{align*}
\frac{dT}{dt} &= s + r_1 T \left(1 - \frac{T + I}{T_{\text{max}}}ight) - dT - \beta TV \\
\frac{dI}{dt} &= \beta TV + r_2 I \left(1 - \frac{T + I}{T_{\text{max}}}ight) - \delta (1 + \alpha H) I \\
\frac{dV}{dt} &= pI - cV \\
\frac{dH}{dt} &= s_H (1 + \gamma I) - d_H H
\end{align*}
\]

where $s_H$ is the recruitment rate of CD4$^+$ cells, and $d_H$ their death rate.

**Immunological Impact of HIV Coinfection on HCV**

Introducing HIV infection into the system will have an impact on its dynamics. Rather than including the full complexity of within-host HIV dynamics in this model, we have chosen to take advantage of the differing time scales of asymptomatic HIV infection and HCV treatment. Therefore, we use HIV set-point viral load as a constant parameter (rather than a state variable with its own dynamics) so it contributes as a mass-action depletor of the immune system compartment. Perturbing the system
by adding a component of HIV-infection changes the stability dynamics. To maximize simplicity in the model, HIV infection was therefore modeled as a single viral load, denoted by $V_H$, representing the set-point viral load. The equation for $H$ thus becomes

$$\frac{dH}{dt} = s_H(1 + \gamma I) - d_H H - \beta_H V_H H$$

where $\beta_H$ the is the mass-action infection parameter for HIV. Due to the static nature of the incorporation of HIV into this model, we are unable to capture dynamically the process of immune recovery after ART for HIV. However, as immune recovery is often incomplete, we can use lower values of $H$ (we can vary $s_H$ to generate these values when $V_H \approx 0$ ) to simulate a depleted immune system and thus capture the dynamics of HCV in HIV patients on ART.

**HCV Treatment**

Treatment efficacy was included in the model as a parameter $\varepsilon$ controlling viral production rate as in previous work (3), and following (2, 143), we have implemented a cure boundary such that virus stops being produced when infected cell count drops below 1. The equations become

$$\frac{dI}{dt} = \beta TV + r_2 I \left( 1 - \frac{T + I}{T_{max}} \right) - \delta(1 + \alpha H) I$$

$$\frac{dV}{dt} = (1 - \varepsilon)pI - cV$$

$$\frac{dH}{dt} = s_H(1 + \gamma I) - d_H H - \beta_H V_H H$$
This formulation allows for implementation of imperfect treatment for varying durations in this model. As explored in the next section, these combinations can reveal the uncertainty surrounding cure inherent in certain patients.

Bistability Analysis

This model allows for cure in two types patient systems. As explored by (2), including the cure boundary allows the model to replicate viral dynamics in patients in whom infection is effectively cleared (<1 infected hepatocyte in the modeled population). These patients do not exhibit true bistability; if infected hepatocytes are not completely cleared, the viral load will bounce back even from undetectable levels. In some patients, however, the system can by truly bistable and when infected cell count drops below a certain non-zero level, the patient will achieve SVR even if virus is not initially eradicated.

To explore the stability dynamics of the model, it can be useful to make the quasi-steady state approximation. Because the viral dynamics happen on a faster time scale than the cell dynamics, we can simplify the equations as

\[
\begin{align*}
\frac{dT}{dt} &= s + r_1 T \left( 1 - \frac{T + I}{T_{\text{max}}} \right) - dT - (1 - \varepsilon) \tilde{\beta} T I \\
\frac{dI}{dt} &= (1 - \varepsilon) \tilde{\beta} T I + r_2 I \left( 1 - \frac{T + I}{T_{\text{max}}} \right) - \delta (1 + \alpha H) I \\
\frac{dH}{dt} &= s_H (1 + \gamma I) - d_H H - \beta_H V_H H \\
V &= \frac{p}{c} I
\end{align*}
\]

where \( \tilde{\beta} = \frac{p}{c} \beta. \)

Following (1), we calculate the within-host \( R_0 \) by setting \( \frac{dI}{dt} > 0 \) at the start of the infection (before treatment, so the \( \varepsilon \) term disappears), when \( I \approx 0 \):
\[
\frac{dI}{dt} = \tilde{\beta}TI + r_2I \left( 1 - \frac{T + I}{T_{\text{max}}} \right) - \delta(1 + \alpha H)I > 0
\]

\[
\tilde{\beta}T + r_2 \left( 1 - \frac{T}{T_{\text{max}}} \right) > \delta(1 + \alpha H)
\]

\[
R_0 = \frac{\tilde{\beta}T_0 + r_2 \left( 1 - \frac{T_0}{T_{\text{max}}} \right)}{\delta(1 + \alpha H_0)} > 1
\]

where \(T_0, H_0\) are the initial values of uninfected hepatocytes and activated CD4\(^+\) cells in the absence of infection.

\[
T_0 = \frac{T_{\text{max}}}{2r_1} \left( (r_1 - d) + \sqrt{(r_1 - d)^2 + 4s \frac{r_1}{T_{\text{max}}}} \right)
\]

and

\[
H_0 = \frac{sH}{d_H}
\]

Bistability of a viral endemic equilibrium and a disease-free equilibrium can occur under certain circumstances \((1)\). In this monoinfection model, bistability occurs when the within-host \(R_0\) is close to or less than 1, and \(r_2 > \tilde{\beta}T_{\text{max}}, \ r_2 > \delta(1 + \alpha H_0)\). There is some evidence that HCV can increase the reproduction rate of infected cells to replace cells that were targeted successfully by immune response \((157, 148, 175, 176, 177)\), so stability of an endemic equilibrium given this condition has some clinical basis. The role of \(r_2\) in the bistability in this system has an intuitive basis as well: infected cell growth rate receives contributions both from infected replication and from virion infection of healthy cells. More basically, when \(r_2 > r_1\), i.e. the maximum proliferation rate of the infected cells is greater than the maximum proliferation rate of uninfected cells, the infection can invade even at a stable, disease-free equilibrium. Interestingly, HIV-coinfection impacts greatly the stability of the DFE, but does not
impact invasion probability at the stable DFE. Derivation of these conditions can be found in Appendix C.

The stability surrounding $R_0$ divides patients into four types: never achieving SVR ($R_0 > 1$), always achieving SVR ($R_0 < R_C$ where $R_C$ is the critical bifurcation value), or in the bistable region, either experiencing spontaneous cure or experiencing viremia but with the ability to be pushed into the clearance area by treatment. With this formulation, it is possible to analyze numerically which values of duration and efficacy can push a patient from one basin of attraction to the other (i.e. endemic equilibrium to clearance). SVR is thus dependent on a combination of critical efficacy and time. When the immunological impact of HIV-coinfection is included, however, the stability dynamics change, reducing the probability of achieving SVR (171).

**Results**

The model can qualitatively replicate viral dynamics observed in HCV patients after treatment and described in previous studies such as (3) and (2). Figure 3.2 demonstrates this replication, showing both the initial biphasic decline (model outputs paired with data redrawn from (3) in panel c)), and longer-term dynamics paired with data redrawn from (2) (panels a) and b)). The model outputs shown are not explicitly fitted to the data presented, but nonetheless replicate the viral trajectories. When the quasi-steady state assumption is made, these longer-term dynamics can similarly be qualitatively recovered, as seen in panel d) of Figure 3.2.

Using an example patient (Patient I) who exhibits true bistability of endemic- and disease-free equilibria allows us to visualize the impact of HIV infection. Simulating different levels of the depletion in CD4 count that might occur during the asymptomatic phase of HIV infection shows that the zone of bistability shrinks with decreasing CD4$^+$ count. The region of bistability is a region wherein the within-host
Figure 3.2: Model Comparison with Data. Panel a) shows patient data redrawn from (2), displaying the five types of long-term treatment outcome: SVR, Null Response, Partial Response, Breakthrough, and Relapse. Panel b) shows model scenarios from the model including free virus displaying these same five outcomes. Panel c) shows short-term treatment dynamics over the first two weeks of treatment; model outputs are paired with data redrawn from (3). Panel d) shows the same scenarios as b), using a version of the model wherein the quasi-steady state assumption has been implemented, demonstrating that the longterm dynamics can be recovered. Parameter values for each simulation are listed in Supp. Table C.3.

$R_0$ is less than or close to 1 but greater than a certain critical reproductive number, $R_C$ (see Appendix C for calculation of $R_C$). In this region, there are three distinct equilibria: the stable endemic equilibrium (EE), the disease-free equilibrium and a third, unstable EE. This unstable endemic equilibrium divides viral load measures into two basins of attraction: toward the disease-free equilibrium and toward the stable EE. Figure 3.3 shows how this region of bistability encroaches farther into the region of spontaneous clearance the lower the CD4$^+$ count, and eventually erases the possibility of spontaneous clearance for realistic values of $\delta$. The presence of HIV depletes the immune system component (even after successful treatment of HIV with ART) thus decreasing the probability of spontaneous clearance of HCV for this pa-
tient. The patient represented has values for parameters drawn from the realistic distributions described in (1) (and Supp. Table C.1). HIV parameters are adapted from (145); The set-point viral load values tested are derived from the estimates by (178) of the variability of HIV set-point viral load. The coinfection influence parameters are theoretical estimates chosen such that model reproduces viral load levels and relationships; while much empirical work has demonstrated qualitative relationships indicated by these parameters (e.g. (158, 174, 18)), this work does not allow for their precise determination in this context.

The bifurcation diagram demonstrates how spontaneous clearance can become non-clearance in a patient in whom the bistability criteria hold, but using this model we can also demonstrate differential treatment responses. Figure 3.4 shows two sample patients who are both able to achieve SVR in a non-immunosuppressed state. Panel a) shows Patient I, the same bistable patient as in the bifurcation diagram. Panel b) shows Patient II, who does not demonstrate true bistability, but is able to achieve SVR with the cure boundary (i.e. when the infected hepatocyte count drops below 1). For both patients, when HCV-monoinfected, SVR is achievable after 24 weeks of treatment at levels of \( \sim 80\% \) efficacy. However, for both patients, when CD4\(^+\) is depleted partially (as in a state where HIV viral load is controlled but CD4\(^+\) count has not fully recovered), 24 weeks of treatment is insufficient and relapse occurs: the patient must be treated for 48 weeks to achieve SVR. In the case when each patient is profoundly immunosuppressed (as in untreated HIV, or no immunologic response to treatment after very low nadir CD4\(^+\)), SVR cannot be achieved even with 48 weeks of treatment. This time series illustrates the sensitivity of treatment prognosis to HIV status, as well as the phenomenon that early treatment response is necessary but not sufficient to predict whether SVR will be achieved.

However, when treatment efficacy is high enough, as it can be with the new direct-acting drugs that have recently been approved (179, 180, 53), HIV coinfection with
Figure 3.3: Bifurcation Diagram of Stable and Unstable Equilibrium HCV Viral Loads with Varying CD4+ count. This figure shows the growing bistable region and corresponding shrinking region of spontaneous clearance for a parameter set with varying δ and within-host R0. The x-axis represents what the within-host R0 would be in the HIV-free scenario the given clearance rate parameters used for the varying HIV scenarios. The y-axis represents equilibrium HCV viral load. To the left of the bifurcation points is the zone of spontaneous clearance. If viral load can be pushed below the dotted lines (unstable equilibria), with sufficient treatment, the patient will then be in the region of stable DFE and will be able to achieve SVR. δ × (1 + αH) ranges from from d to 3 for varying CD4+ counts*. The other parameters are as follows: s = 4365, T_{max} = 4.016 × 10^6, d = 1.06 × 10^{-3}, \beta_c = 7.3 × 10^{-8}, p = 13.48, c = 10.06, r_1 = 2.7, r_2 = 7.52, s_H = 9, \beta_H = 4.1 × 10^{-6}, d_H = 9 × 10^{-3}, \alpha = 5 × 10^{-3}, \gamma = 2 × 10^{-8}. *Note: For lower CD4+ counts/higher HIV viral loads, the tips of the bifurcation trees could not be reached with clinically realistic values of δ. This fact implies that there would be no realistic zone of spontaneous clearance for this patient when CD4+ count falls below a certain point, though SVR might still be possible with sufficient treatment duration and efficacy.

incomplete CD4+ recovery no longer compromises SVR chances to the same extent. Sofosbuvir has been shown to effect SVR within 12-24 weeks in both HIV-positive and HIV-negative HCV patients \cite{52, 181, 182, 183}, and our model replicates this result for the above theoretical patients I and II (Figure 3.6 panels a) and b), (we still predict treatment failure for short-duration treatment when the patients are severely
Figure 3.4: **Effect of depleted CD4\(^+\) on Infection Clearance.** This figure shows HCV viral load trajectories for two sample patients with different treatment efficacies and durations under initial conditions of normal CD4\(^+\) count (\(\sim 1000\) \(\mu\)L, HIV negative), and depleted CD4\(^+\) count (\(\sim 600/\mu\)L, suppressed HIV with incomplete immunologic recovery), and very low CD4\(^+\) count (\(\sim 10/\mu\)L, unsuppressed HIV) for treatment courses of 24 and 48 weeks. It can be seen that a treatment course of 24 weeks that is sufficient to achieve SVR when the patient has a normal CD4\(^+\) count is no longer sufficient for either HIV positive scenario, even when CD4\(^+\) count has partially recovered. However, when treatment is extended to 48 weeks, the patients are able to achieve SVR when CD4\(^+\) count has partially recovered. Panel a) shows these scenarios for Patient I, a theoretical patient exhibiting classic bistability, while panel b) shows Patient II, a theoretical patient who requires the cure boundary condition to achieve SVR.

immunocompromised, but most trials of Sofosbuvir in coinfected patients have been among patients on ART with stable CD4\(^+\) e.g. ([181] [182] [183]). Figure 3.5 shows the pairs of treatment efficacies and durations that lead to sustained virologic response for Patients I and II with and without depleted CD4\(^+\) count from HIV coinfection. Each point on each line represents the minimum treatment duration necessary to achieve SVR for the corresponding treatment efficacy. In each of these scenarios, the model was run for 96 weeks (well past the maximum treatment duration tested) in order to capture any relapse that might occur. This figure demonstrates how the minimum acceptable treatment efficacies and durations depend strongly on HIV status. It can
be seen in this figure that duration/efficacy pairs that will result in SVR for the patients when they are HIV-negative fall in the “No SVR” region when the patient has a depleted CD$^+$ count. As CD4$^+$ drops, the minimum treatment efficacy necessary for SVR increases, as does minimum duration for a given efficacy in some cases. When the patients are severely immunocompromised, only very high efficacy treatment will result in SVR. Lastly, as Figure 3.5 suggests and Figure 3.6 shows in panels c) and d), in the likely event that treatment efficacy declines over time due to non-adherence, drug resistance, or concentration decay, the differences between monoinfected and coinfected patients may reappear, wherein coinfected patients relapse after treatment.

**Discussion and Conclusion**

In this study, we propose a model of within-host HCV infection that is able to capture broadly the impact of concurrent treated or untreated HIV infection on clearance and long-term cure of HCV in coinfected patients. To our knowledge, it is the first within-host model of HIV-HCV confection; it builds on previous models of HCV monoinfection (1) 62) and HIV monoinfection (145 173) and allows for an explicit role of the immune system in HCV disease course. Our model qualitatively replicates results of empirical research showing that HIV-coinfected HCV patients have reduced probability of spontaneous clearance of HCV as well as reduces rates of achieving sustained virologic response (148 155 160). Similarly, we replicate newer empirical findings indicating that highly efficacious direct-acting antiviral agents (DAAs) reduce treatment differences between HIV-positive and HIV-negative HCV patients (52).

This study adds to the literature a way of capturing HCV within-host dynamics while accounting for the role of the immune system under conditions of HIV infection. It contains the flexibility and tractability necessary for testing hypotheses about clearance and treatment. We provide a framework which may be useful for assessing a
patient’s chance of responding to treatment, given certain virologic, immunologic and therapeutic parameters. Our choice to model HIV statically allows our model capture how an immunocompromised patient can respond differently to HCV infection and subsequent treatment, while avoiding the problem of parameter proliferation.

There are several limitations of the model related to the complexity of the system. First, the bistability criterion that the proliferation rate of infected cells must be greater than the proliferation rate of uninfected cells only has partial support in
Figure 3.6: **High-Efficacy Treatment with short-term cure.** Panels a) and b) demonstrate that high efficacy treatment, as with DAAs, can result in SVR over a short duration (12-24 weeks) for Patients I and II, respectively, even with incomplete CD4+ recovery (though when each patient is severely immunocompromised, short term cure is still unlikely). However, panels c) and d) shows the same scenarios, but with treatment efficacy declining from 95% to 70-75% efficacy over the course of 12-16 weeks. When the CD4+ count is depleted or low (HIV-positive), the patients respond initially, but relapse after treatment efficacy falls below a certain level.

the literature. Some studies report that HCV core protein can induce proliferation in hepatocytes \cite{176,177,148,157} and inhibit immune-mediate cell killing \cite{175}, while others report slowed proliferation in HCV-infected cells \cite{184}. However, there may be other mechanism of bistability, such as interferon refractoriness of some cells as proposed by \cite{185} that would yield similar impacts under the HIV coinfection conditions proposed in this model. Second, the model assumes very simplified within-host HIV dynamics by only including HIV viral load as a static parameter, so may miss some of the subtleties of CD4 decay and viral load changes through the course of infection. However, this simplification is appropriate for the shorter relative time-courses of the HCV dynamics explored here. Similarly, while we cannot model the
explicit dynamics of treatment of HIV with ART, we are able to manually replicate them by using lower equilibrium values of CD4\(^+\) count. Third, the way we model HCV treatment does not take into account PK/PD complexities beyond simple, exponential decay of treatment efficacy over time, but it is useful for making baseline calculations. It has been used in the past for modeling IFN treatment \([11, 62, 61]\), but may in fact be a better representation of the mechanisms of DAAs \([153]\). IFN works by creating an antiviral environment inside susceptible cells, thereby decreasing their chances of getting infected \([150]\), while ribavirin can act as a mutagen and cause some proportion of virions to be non-infectious \([143]\). Empirical evidence suggests that response is not directly dependent on concentration over time, but maximum drug effectiveness is correlated with treatment response. DAAs, however, mainly work to inhibit HCV replication thereby having a more direct effect on viral burst-size and infectiousness \([154]\).

DAAs offer very promising prognoses for HCV patients, but are still prohibitively expensive for many \([186, 152]\). It is therefore crucial to assure that the drugs are being administered properly. Using a model to predict patient response may lend insight into pre-treatment estimation of treatment success, and could be helpful in monitoring effectiveness of therapy over the course of treatment. For example, if a patient requires or takes drug holidays, a model of this type may be helpful in predicting the maximum drug holiday allowable without compromising treatment. As Figure 3.6 shows, HIV has the potential alter treatment dynamics even in optimistic scenarios, underscoring further the importance of understanding this system.

The model results presented here give a theoretical demonstration of the effect that HIV co-infection can have on the course of HCV infection. While the system of HCV-HIV coinfection has many layers of complexity, we are able to use a model with relatively simple assumptions about pathogen and immune system interaction to qualitatively describe patient outcomes. Not only is an HIV-positive patient less
likely to clear HCV spontaneously, but also less likely to respond to HCV treatment when treatment efficacy is below a certain level \([148, 155, 160]\). Understanding what drives these treatment differences can help spare difficult treatment and side effects for patients who are unlikely to respond to treatment, as well as informing strategies to maximize treatment adherence. HIV-HCV coinfection is a growing issue not just among injecting drug users, but also among HIV-positive Men who have sex with Men (MSM) who may experience both increased transmissibility of HCV due to higher viral load and also increased susceptibility to sexually-transmitted HCV due to incomplete restoration of mucosal immunity \([148]\). It is thus vital to understand the within-host dynamics of these coinfecting pathogens in order to better assess treatment strategies and preempt shortfall and potential resistance acquisition.

See Appendix C for full supplementary information.

**Acknowledgments**

We thank Professors Andrea Graham and Simon Levin, Department of Ecology and Evolutionary Biology, Princeton University for their helpful comments and insights. RB was supported by the Princeton University Department of Ecology and Evolutionary Biology. RDK was supported by the Swiss National Science Foundation # PZ00P3-142411. BG was supported by the Bill and Melinda Gates Foundation and the RAPIDD program of the Science and Technology Directorate, U.S. Department of Homeland Security, and the Fogarty International Center, NIH.
Modeling within-host HCV lineage patterns: a mechanistic approach

Ruthie Birger\textsuperscript{1}, José Lourenço\textsuperscript{2}, Bryan Grenfell\textsuperscript{1,3}, Oliver Pybus\textsuperscript{2}

\textbf{1} Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ, USA
\textbf{2} Department of Zoology, University of Oxford, Oxford, UK
\textbf{3} Fogarty International Center, National Institutes of Health, Bethesda, MD 20892, USA

\textbf{Abstract}

\textbf{Introduction}

Hepatitis C Virus (HCV) undergoes rapid within-host viral evolution, and persists in patients with multiple distinct lineages persist throughout infection. Not all lineages are detectable in peripheral blood at all times, and there is evidence of viral genetic differences among groups of infected hepatocytes. Furthermore, in some patients, lineages observed early in infection appear to go extinct but reappear later, even
after liver transplant, fueling the hypothesis that extra-hepatic replication may occur. Incorporating these various hypotheses into a model of viral evolution during infection can help elucidate HCV within-host dynamics. In turn, this may help determine effective treatment design, especially with new direct-acting antivirals (DAAs).

**Methods**

We designed a mathematical framework to examine possible explanations of the lineage patterns observed in longitudinal studies of HCV patients. Proposed mechanisms include spatial structure in the liver, long-term latently-infected cells, extra-hepatic replication, and selective sweeps. The latter three mechanisms were tested using both non-spatial and spatial approaches. The non-spatial models are built using continuous-time Markov chains that monitor the state transitions of each cell. The spatial models are built on an individual-based lattice representing a patch of liver, and incorporating transmission between neighboring cells. Both types of model generate transmission chains of the infection process. These chains are combined with an infinite-sites model of molecular evolution to simulate phylogenies and summary population genetic measures from which lineage patterns are obtained. A Random Forest Algorithm (RFA) containing a series of tree statistics is run on the trees to test which trees are distinguishable from the non-spatial null model.

**Results**

Our results indicate that spatial trees generated by models with non-null mechanistic assumptions are largely distinguishable from the null non-spatial model by the RFA, even when they generate comparable cell and viral trajectories. This implies that if different mechanisms are involved in HCV within-host transmission, they may
detectable by analyzing genetic data even if they do not leave a signature in clinical data.

Conclusions

Modeling efforts such as this can lend insight into hypothesized physiological mechanisms of within-host HCV dynamics, the understanding of which is crucial for optimal deployment of DAAs\(^1\).

Introduction

Hepatitis C Virus (HCV) is a global public health concern, affecting nearly 200 million people worldwide\(^1\). The virus is most efficiently transmitted through blood from contaminated needles or medical equipment, though it can also be transmitted percutaneously and in transplanted organs\(^1\). In the United Kingdom (UK), HCV is now the leading cause of end-stage liver disease and transplantation failure\(^1\), remaining the world’s leading cause of cirrhosis of the liver and hepatocellular carcinoma (HCC)\(^1\). In the United States of America (USA), HCV is one of the fast-increasing causes of death, now outnumbering the total number of deaths caused by HIV\(^1\).

HCV is a member of the Flaviviridae family and Hepacivirus genus that was first isolated as a cause of non-A, non-B viral hepatitis in 1989\(^1\). The virus undergoes rapid within-host evolution, with one of the highest substitution rates among flaviviridae\(^1\), and exists within patients as a quasi-species\(^1\), with multiple distinct lineages developing and persisting throughout infection (Simmonds\(^1\) reviews the documentation of this evidence from many studies).

\(^1\)Earlier versions of this work were presented at the 2015 EEID conference in Athens, GA and the 2015 Epidemics Conference in Clearwater, FL.
Until recently, standard treatment for HCV has involved combination therapy of interferon and ribavirin, both of which function indirectly to control viral spread. These indirect therapies require long duration (24-48 weeks), and have limited efficacy (50-70%) (150). However, new direct-acting antivirals such as sofosbuvir and daclatasvir are becoming available in the developed world with much shorter treatment courses, and higher success rates (181, 180). Understanding the mechanisms behind the within-host dynamics is particularly important in light of these new treatments and their optimal design.

It is remarkable that the pharmacological breakthroughs in HCV treatment have occurred with relatively little understanding of the processes that govern HCV replication and within-host evolution. Most studies that measure HCV evolution derive sequences from peripheral blood, and calculate measures of sequence diversity and divergence. Basic models that assume a well-mixed viral population with identical evolution rates do not find associations between these measures and disease severity and progression (197). However, these models do not take into account spatial or other population structure, varying rates of replication or viral clearance between lineages, or other complexities of infection. Not all lineages are detectable throughout the course of infection, and there is evidence of differing distribution of variants between distinct groups of infected hepatocytes (liver cells). For example, Gray et al (198) report on a study assessing HCV genetic diversity in liver transplant patients pre- and post-transplant. In this study, sequences from serum were measured at multiple time points before and after transplant, so changes in diversity could be measured. It was previously hypothesized that the transplant would create a genetic bottleneck, and only certain lineages would survive transplantation, but this study showed that many patients had post-transplant sequences that came from multiple founder strains, and there was no between-patient evidence of a changing trend in viral diversity pre- and post-transplant. Additionally, rare variants persisted after the transplant, which
suggests occurrence of ongoing adaptive evolution and extra-hepatic viral latency. In another more recent study, Raghwani et al (199) analyze viral sequences sampled over time from a group of 15 HCV patients, and find extraordinary variation in molecular evolution within and between patients. Their results suggest that a structured viral population may explain some of the unexplained phenomena in HCV infection mentioned above (reappearance of long-unobserved lineages, speedy jumps in viral diversity), and highlight the fact that virus samples from peripheral blood do not capture the full genetic range of the virus at any one time. These phenomena are important to consider in the context of new therapies, especially with respect to full clearance and potential drug resistance, as well as for HCV transmission studies that involve molecular analysis.

Incorporating various hypotheses about within-host viral dynamics into a model of viral spread and evolution over the course of infection may be essential in understanding patterns that could lend insight into effective treatment design. Mathematical models have been used to capture and understand the dynamics of within-host disease progression of several infections including HCV. Some of the earliest within-host models were based on HIV infection and involved using model dynamics to estimate viral replication rate based on viral clearance after treatment (200, 56, 173). A similar approach has been used for HCV: early within-host models of HCV aimed to capture the dynamics of infection by pairing a mathematical model with patient data from trial of varying doses of interferon (IFN) (3, 61). These models captured simple viral dynamics such as the biphasic decline of viral load after treatment initiation, and augmenting these models by including replication of infected cells (151, 1) allowed them to explain more complex dynamics such as treatment failure. However, these models fall short at replicating other phenomena including cell-to-cell transmission of infection, clustering of infected hepatocytes within the liver, and the recent evidence
from liver biopsy samples that suggests the infection in the liver is seeded by viruses in the blood and then spreads locally \(201\, 202\, 203\).

In this study, we will use mathematical models to examine several hypotheses about mechanisms that might explain the lineage patterns observed in studies such as Gray et al \(197\), Raghwani et al \(199\) and others, and test a number of statistics to evaluate their potential to identify transmission trees generated by these mechanisms.

**Methods**

In order to test various hypotheses about within-host transmission mechanisms, we designed two families of stochastic models. The non-spatial family assumes a well-mixed hepatocyte (target cell) population, while the spatial family uses an explicit square lattice structure. Within each family, we test three further mechanisms: latency, extra-hepatic (two-patch) replication, and selective sweeps (described below). Figure 4.1 displays the model structures.

![Model Flowchart](image)

Figure 4.1: **Model Flowchart**. The colored boxes represent different mechanisms tested. Infection occurs between free virions and target cells as well as between infected cells and target cells.
Model Structures

We use ordinary differential equations (ODEs) to describe the infection process of the liver cells and blood during the course of chronic HCV infection.

Null Model

In the null model (i.e. no alternative mechanisms), we assume there are five types of cell:

- Uninfected Hepatocytes
- Short-term Latently Infected Hepatocytes
- Infected Hepatocytes
- Dead Infected Hepatocytes (Potential Scar Tissue)
- Dead Uninfected Hepatocytes

Uninfected Target cells are regenerated from Dead cells. Interacting with Infectious cells or free virus, they become Latently Infected (Exposed), and after the eclipse phase, they become Infectious. Both uninfected and infected hepatocytes die, with infected cells being eliminated by the immune response faster than natural death rates. Dead cells regenerate into uninfected target cells, but those dead after being infected with HCV have a lower probability of regenerating. Dead infected cells that do not regenerate represent the fibrotic scar tissue that accumulates over the course of infection and results in cirrhosis in some patients. Parameter descriptions and estimates are listed in Table 4.1.

Adapting a classic model of within-host HCV infection used in previous studies (61, 3), we have
\[
\begin{align*}
\frac{dT}{dt} &= \phi_{DT} D_T + (1 - \kappa)\phi_{DI} D_I - (\lambda_{\text{virions}} + \lambda_{\text{local}} + \nu_T)T \\
\frac{dE}{dt} &= (\lambda_{\text{virions}} + \lambda_{\text{local}})T - (\alpha + \nu_T)E \\
\frac{dI}{dt} &= \kappa\phi_{DI} D_I + \alpha E - \nu_I I \\
\frac{dDT}{dt} &= \nu_T(T + E + EX) - \phi_{DT} D_T \\
\frac{dDI}{dt} &= \nu_I I - \phi_{DI} D_I
\end{align*}
\]

where \( \lambda_{\text{local}} = \beta_L \times I \) and \( \lambda_{\text{virions}} = \beta_V \times V^2 \)

**Latent Model**

The Latent model builds on the Null model to include a second latent class, EX, that keeps cells latently infected for longer before becoming infectious.

\[
\begin{align*}
\frac{dT}{dt} &= \phi_{DT} D_T + (1 - \kappa)\phi_{DI} D_I - (\lambda_{\text{virions}} + \lambda_{\text{local}} + \nu_T)T \\
\frac{dE}{dt} &= (1 - \eta)(\lambda_{\text{virions}} + \lambda_{\text{local}})T - (\alpha + \nu_T)E \\
\frac{dEX}{dt} &= \eta(\lambda_{\text{virions}} + \lambda_{\text{local}})T - (\alpha_X + \nu_T)EX \\
\frac{dI}{dt} &= \kappa\phi_{DI} D_I + \alpha E + \alpha_X EX - \nu_I I \\
\frac{dDT}{dt} &= \nu_T(T + E + EX) - \phi_{DT} D_T \\
\frac{dDI}{dt} &= \nu_I I - \phi_{DI} D_I
\end{align*}
\]

\( ^2V = \frac{2}{p} I \) = viral load, and is calculated by applying the steady-state approximation to \( \frac{dV}{dt} = pI - cV \) where \( p = \) burst size and \( c = \) viral clearance rate.
Extra-hepatic replication (Two-Patch) Model

The Extra-hepatic replication model again builds on the Null model by adding a second patch with a simple SI model for the extra-hepatic compartment. There is scattered evidence that HCV can be found in non-liver tissues (e.g. nervous tissue [204, 205]). Variable mixing rates between the patches can be simulated.

\[
\begin{align*}
\frac{dT}{dt} &= \phi DT_D + (1 - \kappa) \phi DI_D - (\lambda_{\text{virions}1} + \lambda_{\text{local}} + m_{21} \lambda_{\text{virions}2} + \nu_T)T \\
\frac{dE}{dt} &= (\lambda_{\text{virions}1} + \lambda_{\text{local}} + m_{21} \lambda_{\text{virions}2})T - (\alpha + \nu_T)E \\
\frac{dI}{dt} &= \kappa \phi DI_D + \alpha E - \nu_I I \\
\frac{dDT}{dt} &= \nu_T (T + E + EX) - \phi DT DT_T \\
\frac{dDI}{dt} &= \nu_I I - \phi DI DI_I \\
\frac{dT_2}{dt} &= \nu_{T2} (T_2 + I_2) - (\lambda_{\text{virions}2} + \lambda_{\text{local}2} + m_{12} \lambda_{\text{virions}1} + \nu_{T2})T_2 \\
\frac{dI_2}{dt} &= (\lambda_{\text{virions}2} + \lambda_{\text{local}2} + m_{12} \lambda_{\text{virions}1})T_2 - \nu_{T2} I_2
\end{align*}
\]

The \( \lambda \) calculations follow the null model, e.g. \( \lambda_{\text{local}2} = \beta_L \times I_2 \).

Selective Sweep Model

Selective sweeps are modeled using another set of infected compartments with higher fitness \( S \), and mutation rate \( \mu \).
\[
\frac{dT}{dt} = \phi_{DT} DT - \phi_{DI} DI - (\lambda_{\text{virions}} + \lambda_{\text{local}} + \lambda_{\text{virions}}S + \lambda_{\text{local}}S + \nu_T)T
\]
\[
\frac{dE}{dt} = (1 - \mu)(1 - \eta)(\lambda_{\text{virions}} + \lambda_{\text{local}})T - (\alpha + \nu_T)E
\]
\[
\frac{dI}{dt} = \kappa \phi_{DI} DI + \alpha E - \nu_I I
\]
\[
\frac{dE_S}{dt} = (1 - \eta)(\mu \lambda_{\text{virions}} + \mu \lambda_{\text{local}} + \lambda_{\text{virions}}S + \lambda_{\text{local}}S)T - (\alpha + \nu_T)E
\]
\[
\frac{dI_S}{dt} = \kappa \phi_{DI} DI + \alpha E_S - \nu_I I_S
\]
\[
\frac{dD_T}{dt} = \nu_T (T + E + EX) - \phi_{DT} DT
\]
\[
\frac{dD_I}{dt} = \nu_I I - \phi_{DI} DI
\]

The \( \lambda \) calculations again follow the null model, and \( \lambda_{is} = S \times \lambda_I \), \( S \) = selective advantage.

The non-spatial version of each of these models is simulated using a continuous-time Markov Chain model with a tau-leap algorithm and transition probabilities calculated from the parameters above. See Supplementary information in Appendix D for full transition probabilities and matrix. The spatial version of each model uses a non-wrapping square lattice representing target hepatocytes, wherein each cell has a specified number of neighbors with which it can interact, and cell-to-cell transmission probability is adjusted for the fact that only neighboring cells can transmit. It is an Individual-Based Model (IBM) whose transitions follow the ODEs described above, and it is based on a model described in more detail by Lourenço et al (209). For both models, free virus load is calculated based on number of infectious cells, and contributes to the force of infection. In the spatial model, free virus can infect any cell and is not subject to neighbor constraints. In order to be able to compare the effect clustering in addition to the effect of spatial structure, we run two versions of the
Table 4.1: Model parameter estimates

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Value</th>
<th>Range</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>rate from short-term latency to infectiousness</td>
<td>1 day$^{-1}$</td>
<td>1/6h-1/24h</td>
<td>(203)</td>
</tr>
<tr>
<td>$\nu_T$</td>
<td>death rate of susceptible/latent cells</td>
<td>1/83 day$^{-1}$</td>
<td>$1 \times 10^{-2} - 1.4 \times 10^{-2}$ day$^{-1}$</td>
<td>(202)</td>
</tr>
<tr>
<td>$\nu_I$</td>
<td>death rate of infectious cells</td>
<td>1/8 day$^{-1}$</td>
<td>$1/5 - 1/7$ day$^{-1}$</td>
<td>(204)</td>
</tr>
<tr>
<td>$\phi_{DT}$</td>
<td>regeneration of $D_T$ patches</td>
<td>$10 \times \nu_T$</td>
<td>$1/8$ - $1/10$ day$^{-1}$</td>
<td>(206)</td>
</tr>
<tr>
<td>$\phi_{DI}$</td>
<td>regeneration of $D_I$ patches</td>
<td>$0.8 \times \phi_{DT}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_V$</td>
<td>free-virus transmission probability</td>
<td>$0.8 \times 10^{-8} - 1.1 \times 10^{-8}$</td>
<td>$1 \times 10^{-9} - 1 \times 10^{-6}$</td>
<td>(207)</td>
</tr>
<tr>
<td>$c$</td>
<td>viral clearance rate</td>
<td>8.18</td>
<td></td>
<td>(187)</td>
</tr>
<tr>
<td>$p$</td>
<td>viral burst size</td>
<td>22.3</td>
<td></td>
<td>(188)</td>
</tr>
<tr>
<td>$\beta_L$</td>
<td>infected cell transmission probability</td>
<td>$\beta_{CC} \times RR \times \frac{1440}{10000}$</td>
<td></td>
<td>(203)</td>
</tr>
<tr>
<td>$\beta_{CC}$</td>
<td>cell-bound RNA transmission probability</td>
<td>$1 \times 10^{-5}$/minute</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$RR$</td>
<td>Number of RNA per infected cell</td>
<td>4.1825</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Non-null model parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Value</th>
<th>Range</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\eta$</td>
<td>proportion of cells long-term latent</td>
<td>.01-.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha_X$</td>
<td>rate from long-term latency to infectiousness</td>
<td>1/76 days</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$N_2$</td>
<td>size of extra-hepatic compartment</td>
<td>$1 \times 10^3$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\nu_{T2}$</td>
<td>death rate of extra-hepatic target cells</td>
<td>$\nu_T$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_{L2}$</td>
<td>cell-cell transmission rate of extra-hepatic target cells</td>
<td>$\beta_L$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$m_{12,m_{21}}$</td>
<td>mixing percentages between patches</td>
<td>1%-50%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$S$</td>
<td>Fitness Advantage of mutated cells</td>
<td>1-1.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mu$</td>
<td>selective sweep mutation rate</td>
<td>$0.005 - 0.1 \times 10^{-3}$</td>
<td></td>
<td>(203)</td>
</tr>
</tbody>
</table>

Evolutionary model parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Value</th>
<th>Range</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu_{evo}$</td>
<td>evolutionary mutation rate for Infinite Sites Model</td>
<td>.005-.1</td>
<td></td>
<td>(195, 208)</td>
</tr>
</tbody>
</table>

Spatial model. In one, transmission is dominated by blood-borne free virus infection seeding events. In the other, transmission is dominated by cell-to-cell infection events.

Simulations and Statistics

Evolutionary Model

Independent of the model framework used, the full chain of transmission events is recorded, in essence storing the birth-death process of infected cells. In the context of population genetics, we take infected cells (latent and infectious) to be viral demes, with transmission into a new cell defined as a deme birth-event and cell death as a deme death-event. To model virion evolution, we simulate an infinite-sites model over the transmission chain with a fixed clock mutation rate. For simplicity, each viral
deme is assumed to contain a single consensus viral sequence, with possible mutation events taking place at time of transmission.

The output of the evolution model is an *multichotomous evolution tree*, representing the complete ancestry history of each transmission simulation - the root of the tree is the initial inoculating virus, assumed to have zero mutations; the tips represent observed viruses that have not generated progeny; and the internal nodes represent mutational events for unobserved viruses. The evolution tree is pruned uniformly in time to obtain a representative phylogenetic tree, for which multichotomies are solved by adding branching events with potential zero length.

We use a tree sample size of 200 tips, which is appropriate given the small empirical tree sizes found in the literature to carry out our analysis. The output of each phylogenetic genetic tree is then analyzed by calculating a range summary statistics described in Table 4.2.

**Random Forest Algorithm**

The Random Forest Algorithm (RFA) is an ensemble method that can be used for analyzing large datasets. Using information from multiple decision trees, it combines the outputs of multiple predictor variables for a response variable. Using the R-package *randomForest: Breiman and Cutler’s Random Forests for Classification and Regression* ([217]), we set the predictor variables to be the summary statistics described in Table 4.2 and the response variable to be the version of the model used to create the tree. The RFA then can assess how useful each summary statistic is in distinguishing between trees generated by different models.
Table 4.2: Summary statistics for phylogenetic macro-signatures.

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
<th>Reference(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Colless Index</td>
<td>Tree shape (balance): sum over all nodes of absolute value of difference in length between right and left daughter clades</td>
<td>(210, 211, 212)</td>
</tr>
<tr>
<td>Sackin Index</td>
<td>Tree shape (balance): sum of depths of the leaves</td>
<td>(213, 211, 212)</td>
</tr>
<tr>
<td>Cophenetic index</td>
<td>Tree shape (balance): sum of the depth of lowest common ancestor over all pairs of different leaves</td>
<td>(213)</td>
</tr>
<tr>
<td>Staircaseness</td>
<td>Tree shape (balance): portion of subtrees of a tree that are imbalanced</td>
<td>(214)</td>
</tr>
<tr>
<td>Cherries</td>
<td>Tree shape (branching): pairs of leaves adjacent to a common ancestor node</td>
<td>(215)</td>
</tr>
<tr>
<td>Nodetime</td>
<td>Lineages/nodes per unit time</td>
<td>-</td>
</tr>
<tr>
<td>CMT (mean)</td>
<td>Mean coalescent intervals</td>
<td>-</td>
</tr>
<tr>
<td>CMT (std. dev.)</td>
<td>Variance in coalescent intervals</td>
<td>-</td>
</tr>
<tr>
<td>Depth (mean)</td>
<td>Mean depth of internal nodes, where depth is number of edges between a given node and the root of the tree</td>
<td>-</td>
</tr>
<tr>
<td>Depth (std. dev.)</td>
<td>Variance depth of internal nodes</td>
<td>-</td>
</tr>
<tr>
<td>Delta (mean)</td>
<td>Maximum width/maximum depth, where width is number of nodes at a given depth</td>
<td>(216)</td>
</tr>
</tbody>
</table>

**Results**

All versions of the model output viral dynamics consistent with clinical observations (See Table 4.3) when run with proposed null parameters Figure 4.2 and non-null model parameters (Supplementary Figure D.1). In order to assess the impact of spatial clustering vs spatial structure alone, each version of the spatial models was run using viral transmission parameters and cell-to-cell transmission that generated similar viral loads but different spatial patterns of infection henceforth referred to as “blood-dominated” and “cell-dominated” spatial models (Figure 4.3).

Figures 4.4-4.5 show example trees generated by each mechanistic model, with the model parameters that yield the least and most difference from the null non-spatial model, e.g. the latent model trees are generated using .1% and and 20% of cells going long-term latent. The ability of the RFA to distinguish between trees generated
Table 4.3: Data Estimates

<table>
<thead>
<tr>
<th>Data type</th>
<th>Estimates</th>
<th>Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Viral Load</td>
<td>Initial: $10^6 - 10^8$, $10^6 - 10^7$ over time 30%, 25% ±16%, 6%-99%, 1%-30%, 21%-45% baseline ALT 50-250 IU/L: $\nu_l$ 0.05-0.4/day</td>
<td>(3) [1]</td>
</tr>
<tr>
<td>Proportion of cells infected</td>
<td>30%, 25% ±16%, 6%-99%, 1%-30%, 21%-45%</td>
<td>(203), (207), (187), (218), (202)</td>
</tr>
<tr>
<td>ALT (death rate of infected cells)</td>
<td>baseline ALT 50-250 IU/L: $\nu_l$ 0.05-0.4/day</td>
<td>(3)</td>
</tr>
</tbody>
</table>

Null Model

<table>
<thead>
<tr>
<th>Model</th>
<th>Non-spatial</th>
<th>Spatial Blood-Dominated</th>
<th>Spatial Cell-Dominated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Liver compartment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blood compartment</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 4.2: Null Model Cell and viral dynamics

by different models varied with model-type and parameter (Figures 4.6-4.8), which show the changes in proportion of simulations correctly identified by the RFA over changing parameter values. For each model, there are parameter values that allow it to be consistently distinguished from the null non-spatial model. Table 4.4 shows the top two statistics that distinguish between the models for each parameter set and mutation rate combination.

Overall, the nodetime statistic is the most useful for distinguishing between models. The model that can be most strongly, consistently distinguished from the null is the latent model. Figure 4.9 shows the variation in the nodetime statistic for each
Null Spatial Blood-Dominated

Null Spatial Cell-Dominated

Figure 4.3: Lattice representing Hepatocyte structure. Note that dead infected cells may still contain HCV RNA, representing the transition to cirrhotic scar tissue that occurs in HCV patients there is evidence that HCV persists in cirrhotic tissue \((4, 5)\). However, dead cell count is not included in viral load calculations.

Table 4.4: Top Two Statistics for Distinguishing from Null Non-Spatial Model

<table>
<thead>
<tr>
<th></th>
<th>Null</th>
<th>Latent</th>
<th>Two-Patch</th>
<th>Selective Sweep</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-Spatial</td>
<td>Nodetime, Cophenetic</td>
<td>Nodetime, Cophenetic</td>
<td>Nodetime, Cophenetic</td>
<td>Nodetime, Cophenetic</td>
</tr>
<tr>
<td>Blood-Dominated</td>
<td>Nodetime, Cophenetic</td>
<td>Nodetime, Sackin</td>
<td>Nodetime, Cophenetic</td>
<td>Nodetime, Sackin</td>
</tr>
<tr>
<td>Cell-Dominated</td>
<td>Nodetime, Cophenetic</td>
<td>Nodetime, Depth</td>
<td>Nodetime, Cophenetic</td>
<td>Nodetime, Cophenetic</td>
</tr>
</tbody>
</table>
Figure 4.4: Null Non-spatial vs Non-Spatial Mechanistic Trees. The left panel shows evolution trees generated by each non-spatial model with lowest and highest parameter values. For the latent model, the parameter is proportion of cells going long-term latent; for the Two-Patch model the parameter is the coefficient of inter-patch mixing; and for the Selective Sweep model, the parameters are selective fitness advantage and mutation rate (to fitter strain, distinct from evolutionary mutation rate). The right panel shows evolution trees for each version of the null model.
Figure 4.5: **Null Non-spatial vs Spatial Mechanistic Trees.** The left panel shows evolution trees for the null non-spatial model and blood-dominated spatial versions of each of the mechanistic model with lowest and highest parameter values (as described in caption for Figure 4.4). The right panel shows the same for cell-dominated spatial versions.
Figure 4.6: **Latent Model RFA Accuracy.** This figure shows the proportion of model runs versus the Null NonSpatial model that the RFA accurately categorizes by the proportion of cells going long-term latent for each type of model. The cell-dominated spatial model is most consistently identified correctly, even at very low latency proportion.

Figure 4.7: **Two-Patch Model RFA Accuracy.** This figure shows the proportion of model runs versus the Null NonSpatial model that the RFA accurately categorizes by the inter-patch mixing coefficient for each type of model. The cell-dominated spatial model is most consistently identified correctly.
Figure 4.8: **Selective Sweep Model RFA Accuracy.** This figure shows the proportion of model runs versus the Null NonSpatial model that the RFA accurately categorizes by selective fitness advantage for each type of model. The cell-dominated spatial model is most consistently identified correctly.

version of the latent model, and the versions of the two-patch and selective sweep models for which it works best. The latent model when run at its more extreme parameters, can alter the cell and viral dynamics, yielding a slightly lower viral load. In order to establish that the differences highlighted by the RFA are not simply due to lower viral load, we ran a version of the null model with lower viral load and generated a tree (right panel in Figure 4.4). The RFA was able to distinguish between this low viral load null model and the original, but the statistics that distinguished it (Sackin, Depth) were different than those used to distinguish the latent model. Lastly, we tested the latent spatial models against the null spatial models, and found that the RFA was still able to distinguish them using nodetime (and depth).
Figure 4.9: **Nodetime value by model.** This figure shows the variation in the value of the nodetime statistic with parameter value for each model.

**Discussion and Conclusion**

In this study, we have presented a framework for analyzing within-host transmission mechanisms of HCV. We have compared a range of dynamic models, comparing homogeneous-mixing frameworks with explicit spatial formulations, and further investigated the possibility of cell latency, extra-hepatic replication, and viral selective sweeps. Using a machine learning algorithm, we have been able to distinguish key phylogenetic signatures that are self-emergent between models assuming different mechanisms of transmission, selection and evolution.

Our results demonstrate that transmission models including cell latency present phylogenetic trees with an enrichment of internal nodes in time, when compared to a null model without latency. This can mean that heterogeneity in generation time of infectious virions gives rise to changing divergence over time.
The same pattern of enrichment of internal nodes in time is observable, though to a lesser extent, with the other mechanisms of extra-hepatic replication and selective sweeps. Other statistics involving tree balance are necessary to fully distinguish trees generated by these models from the null model.

Lastly, Figures 4.6-4.8 indicate that the trees generated by spatial models dominated by cell-to-cell transmission are most consistently distinguishable from the non-spatial null model, suggesting that clustering of infection in the liver impacts phylogenetic trees shape.

Our main conclusion, underscored most clearly by the models that include latency, are that varying mechanisms of infection can be responsible for differences between patients that are not observable using classical clinical measures (viral load, liver biopsy) but can be found using analysis of phylogenetic trees (See similar cell and viral population patterns in Figures 4.2 and 4.10). In other words, despite lack of heterogeneity in clinical measures, heterogeneity in phylogenetic signatures can lend insight into whether different mechanisms may be at work in different patients. Just using clinical measures, we can not rule out heterogeneity in generation time of infectious virions, as here we display different dynamics that can only be distinguished with genetic data, providing a counterpoint to previous studies that had eliminated latency as a mechanism present in within-host transmission (219).

There are several limitations in this study. We examine only chronic HCV infection and do not look explicitly at treatment dynamics. Our spatial models only look at a 10000 cell patch in the liver, as simulating the full liver size would be too computationally expensive. However, the chronic infection outputs that we generate would be most appropriate for comparison to patient data, and the 10000 cell size is commonly successfully used in other within-host models e.g. (1, 68) as it is the appropriate conversion factor to $1\mu L$ of blood. We assume an infinite sites evolutionary model and do not test alternative forms of mutation, though they have the potential
to impact trees \cite{220}. Before the variation built in by the alternative mechanisms explored in this study, there are three sources of variation present in the types of evolutionary trees presented here, and those that are available from patient data. Within trees generated using the same mechanistic assumptions, there is stochastic variation in the genealogy, i.e. between the true infection trees. There is further variation generated by the infinite sites model run on each of the true trees, depending on the mutation rate used. Supplementary Figure \ref{D.2} shows an example of how increasing the mutation rate increases the ability of the RFA to distinguish between models. We have built in this level of variation for appropriate future comparison with patient data, but it adds another level of noise in distinguishing models at this step. Lastly, there is variation generated by the sampling process which here is random. Future work will involve using these models for exploration of the latter two types of variation. The models can be used to test the impact of time-varying mutation rates
versus a perfect molecular clock. For comparison with patient data, we will implement a sampling scheme for the trees that mimics empirical sampling schemes used clinically, and test model trees sampled only from the peripheral blood compartment for consistency.

If there appears to be evidence for any of the non-null model versions upon comparison with data, it could be useful in explaining the unusual lineage patterns observed in HCV viral dynamics (See Supplementary Figures D.3 and D.4 for population genetic measures from each model). Additionally, as access to direct-acting antivirals grows, so does the risk of the evolution of drug resistance. Any of these mechanisms could aid drug resistance evolution; uncovering whether they are likely to be present can be crucial for opening new avenues of research into combating drug resistance.

This project lends itself well to several future extensions. Comparing model outputs to data from HCV patients will be most helpful for informing our understanding of viral dynamics and directions of treatment design, as well as for validating this framework. If successful, this framework can be easily extended to other pathogens, and may provide similar types of insights. The flexibility of the model structure means that it can be used for any pathogen that demonstrates SIR-like dynamics within host, giving it broad generalizability and utility.

See Appendix D for full supplementary information.
Chapter 5

The Impact of HCV Treatment roll-out in a high HIV-HCV Prevalence Population: A modeling study on People Who Inject Drugs in Ho Chi Minh City, Vietnam

Ruthie Birger\textsuperscript{1}, Thuy Le\textsuperscript{2}, Roger Kouyos\textsuperscript{3}, Bryan Grenfell\textsuperscript{1,4}, Timothy Hallett\textsuperscript{5}

\textbf{1} Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ, USA

\textbf{2} Oxford University Clinical Research Unit, Ho Chi Minh City, Vietnam

\textbf{3} Division of Infectious Diseases and Hospital Epidemiology, University Hospital Zürich, University of Zürich, Zürich, Switzerland

\textbf{4} Fogarty International Center, National Institutes of Health, Bethesda, MD 20892, USA.

\textbf{5} Department of Infectious Disease Epidemiology, Imperial College London, London, UK
Abstract

Background

Human Immunodeficiency Virus (HIV) and Hepatitis C Virus (HCV) coinfection is a major global health problem especially among people who inject drugs (PWID), with significant clinical implications. Mathematical models have been used to great effect to shape HIV care, but few have been proposed for HIV/HCV.

Methods

We constructed a deterministic compartmental ODE model that incorporated layers for HIV progression, HCV progression and PWID demography. Antiretroviral therapy (ART) and Methadone Maintenance Therapy (MMT) scale-ups were modeled as from 2016 and projected forward 10 years. HCV treatment roll-out was modeled beginning in 2026, after a variety of MMT scale-up scenarios, and projected forward 10 years.

Results

Our results indicate that scale-up of ART has a major impact on HIV though not on HCV burden. MMT scale-up has an impact on incidence of both infections. HCV treatment roll-out has a measurable impact on reductions of deaths, increasing multifold the mortality reductions afforded by just ART/MMT scale-ups, and may decrease cost-per-life-year saved if rolled out at subsidized cost.

Conclusion

HCV treatment roll-out can have major and long-lasting effects on averting PWID deaths on top of those averted by ART/MMT scale-up. Efficient intervention scale-up of HCV alongside HIV interventions is critical in Vietnam.

---

Footnote: This work is about to be submitted for publication. Earlier versions of this work have been previously presented at the 2014 Ecology and Evolution of Infectious Diseases Conference in Fort...
Introduction

Hepatitis C Virus (HCV) afflicts 150 million people globally, the majority of whom are people who inject drugs (PWID) living in Asia and Africa (221). Coinfection with human immunodeficiency virus (HIV) occurs in 5-10 million people (146, 222). Coinfection dynamics are complex at within- and between-host levels, as the two pathogens share a transmission route and can exacerbate each other (171, 68). Due to the challenges of navigating two such complex chronic infections and their interactions, very few models thus far have looked at modeling HIV-HCV co-epidemics explicitly (66, 67, 223). Past models of HIV and HCV mono-infection have been instrumental in gaining deeper understanding into infection dynamics, e.g. estimation of the basic reproduction number ($R_0$), and predicting the impact of interventions such as Pre-Exposure Prophylaxis (PrEP) and Test-and-Treat (e.g. (54, 119, 71)). HIV modeling studies that focus on PWID have explored the impact of harm reduction strategies such as needle-and-syringe programs (NSP) and methadone maintenance therapy (MMT) on the HIV epidemic among PWID, and found that gains in terms of infections and deaths averted can be substantial with adequate scale-up (224, 225).

Our study builds on previous work by incorporating a specific dynamical model in PWID to explore the potential of both HIV and HCV treatment interventions alongside harm-reduction interventions to assess gains in the HIV and HCV co-epidemic in a focused setting (67, 224, 54, 223, 225).

Assessing HCV treatment in such models is particularly timely and important with the advent of new, direct-acting antiviral drugs (DAAs). These new drugs have good safety profiles, low risks of drug interactions, high tolerability, and can be administered in 12-24 week courses, in addition to being very potent with $>90\%$ cure rates; this makes them better options than interferon-ribavirin (IFN-RBV) combi-
nation therapies that have been in use up until now (52, 53). Currently, even in high-income countries, these new drug regimens are often unaffordable, with prices ranging up to $168,000 for a 24-week course of therapy (152, 186). In low- and middle-income countries, where even the $15,000-$20,000 cost of IFN-RBV is too high for most patients, the DAAs as they are currently priced will not be an implementable solution. However, Gilead Sciences, the company that manufactures sofosbuvir and ledipasvir, has already implemented reduced-cost treatment in Egypt, offering the 12-week course for $900 (226). One aim in this study is to explore the feasibility of a roll-out of HCV treatment on a long-term time scale, dictated by these manufacturing constraints.

Ho Chi Minh City (HCMC), Vietnam is a setting for which modeling could be particularly helpful for long-term planning and policy. Estimates of HIV and HCV prevalence are available for model calibration, and infrastructure exists or is being designed for roll-out of various interventions. Efficient roll-out of disease interventions in Vietnam is crucial as the funding landscape is changing from primarily international donation to government funding (227). The HIV/AIDS epidemic in Vietnam is unevenly distributed; country-wide prevalence was estimated at 0.45% in 2011 (227), with the majority of cases among PWID, men who have sex with men (MSM), and female sex workers (FSW). Home to 7% of Vietnam’s population but 25% of Vietnam’s People Living with HIV/AIDS (PLWHA), HCMC is has an estimated HIV prevalence 46.1% among PWID (227, 224, 228).

Estimates of HCV prevalence among PWID in Vietnam have ranged up to 75% since the 1990’s (229, 230, 231, 232, 233). Coinfection with HIV and HCV is very common in PWID, with some estimates of the percentage of HIV-infected PWID coinfected with HCV at 100% (233, 232, 230). Coinfection introduces a series of complications, e.g. HIV speeds the progression rate of HCV, and HCV can complicate administration of antiretroviral therapy (ART) due to increased risk of hepatotoxicity.
Recent estimates suggest that nearly ∼90% of HIV-infected individuals in Vietnam have a CD4 count of <200 cells/µL upon ART initiation (234). This delayed access can mean that for coinfected individuals, chronic HCV may have progressed to chronic liver disease (CLD), which can compromise gains in life expectancy conferred by ART initiation to an even greater extent than late ART initiation does in mono-infected patients (235). In the pre-ART era, AIDS-related mortality was high enough to mask the effect of liver disease, as HCV-related cirrhosis and cancer deaths would come later than an HIV-related death. However, as studies in other populations have shown (e.g. (236)), the scale-up of ART and subsequent reduction in AIDS-related mortality have revealed the mortality impact of escalated liver disease progression, with liver disease accounting for 14-18% of non-AIDS deaths in the HIV-infected population around the globe (235). Currently, treatment for HCV is not widely available or publicly funded in Vietnam, though its impact could be substantial (54).

The HIV and HCV epidemics in Vietnam are fueled by the underlying problem of injection drug use. Since the 1990s, one of the main methods of injection drug use control has been detention of PWID in rehabilitation centers (known as 06 centers) (237), with an estimated 20% of known drug users being detained at any one time (238). Recognizing the low efficacy of this method, the government began supporting access to clean needles through NSPs in 2006-2007. However, NSPs have remained relatively small-scale and inaccessible (237, 238, 228, 224).

In 2008, a pilot program for MMT was rolled out in Hai Phong and HCMC, with the aim of reducing the rates of unsafe injecting, and thus HIV transmission and other adverse health outcomes among PWID. Reports from these pilot programs indicate that they may be quite efficacious in addressing the injecting drug use epidemic (239). The aim of this study is to use a mathematical model of HIV, HCV and injecting drug use in HCMC to predict the impact of proposed future MMT scale-up at the
city level, as well as the impacts of concurrent ART scale-up and potential future HCV treatment roll-out.

**Methods**

A compartmental deterministic model was created, incorporating levels for HIV and HCV dynamics, and PWID demography which are briefly described below (Figure 5.1). This model is a novel synthesis of existing standard HIV and HCV models, and harm-reduction interventions. Equations for all components of the model can be found in the supplementary information.

**PWID Demography**

The course of an injecting drug user’s life was modeled using three stages (only current and ex-PWID were included in the model). Persons enter the PWID community at a certain rate and remain there for the average duration that PWID inject in Vietnam. The rate of entry into the PWID community is time-dependent and constructed so as to reproduce the estimated change in the total number of PWID in this setting. During this time, PWID have an increased death rate due to overdose (240). Included in this group are PWID in detention centers who often still have access to drugs smuggled into the centers, and may have poor ART/MMT access (238). From the community, PWID can be recruited into MMT, or can cease injecting (and thus exit the community) spontaneously. PWID in MMT programs decrease their rates of injecting. However, retention is imperfect and relapse is common. Assumptions are that MMT increases life expectancy because excess risk of death due to drug-related complications is ameliorated, and also increases flow into ART programs. The latter assumption is based on the criterion for MMT initiation that PWIDs should be on ART (though not all are) or that ART should be started with MMT (227).
HIV Natural History Model

HIV progression and transmission were modeled using acute, asymptomatic and AIDS (CD4<200 cells/µL) stages of infection. Susceptible individuals flow into the acute infection state via a hazard function that is based on a weighted estimate of HIV prevalence. Asymptomatic infection is distributed over three stages of equal duration, reflecting an Erlang distribution \(241\) (see Appendix). Individuals initiate ART from each stage of infection at different rates but individuals who initiate ART when CD4<200 cells /µL have higher death rates on treatment. Parameter values are reported in Table 5.1.

ART scale-up in Vietnam began in 2005, so no treatment was included in the model before then. Treatment coverage rates as reported by National Committee for AIDS Drugs and Prostitution Prevention and Control were used (shown in Figure E.1) \(234, 227, 242\). A 2010 study conducted by Family Health International (FHI) reported that 60-70% of people on ART at two clinics in HCMC were current or former PWID, and with VAAC reporting that 60% of HIV infections were among PWID/ex-PWID, this indicates that the treatment rates are appropriate for PWID \(243, 234\).

HCV Natural History Model

HCV progression and transmission were similarly modeled using acute and chronic stages of infection. Susceptible individuals become acutely infected via a similar hazard function that is based on an estimate of HCV prevalence weighted for stage of infection and treatment status, and either clear infection (with or without protective immunity) \(244\) or move on to chronic infection. There are 4 stages of chronic infection, aligning with an Erlang distribution fit to data on HCV progression in PWID with and without HIV \(160\), and roughly corresponding to the four stages of fibrosis progression \(19\). As noted studies including in Di Martino et al \(160\).
HCV progression is faster in HIV coinfected individuals. The last stage is all-cause chronic liver disease, encompassing compensated and decompensated cirrhosis, and hepatocellular carcinoma.

Figure 5.1: **Diagram of Model Flows.** This figure shows a graphical representation of the model pathways. Full equations and descriptions can be found in the supplementary information.
Fitting

The model was calibrated to data on HIV and HCV prevalence using Maximum Likelihood Estimation (MLE) with a binomial likelihood function. We varied the parameters governing the hazard functions for HIV and HCV acquisition, initial HIV prevalence, and risk-reduction due to NSP so the model outputs reproduced the trends observed in the data. Parameters governing the hazard functions were allowed to vary between the pre- and post-ART availability time periods, with the impact of ART availability beginning between 2005 and 2007. Initial seeds were drawn from estimates for the range of each parameter, and the parameter space was explored using the Nelder-Mead algorithm with MATLAB 2015b’s fminsearch function (See appendix for further details).

Data

Estimates for HCV prevalence among PWID in Vietnam are uniformly high, approaching 90% in major cities (230, 231, 232, 233). In northern Vietnam, a prospective cohort study in 200 young male active heroin users showed that HCV prevalence increased linearly from $\sim 30\%$ to $\sim 70\%$ as duration of injection use increased from 10 months to 30 months (230), while another study in Bac Ninh province found that 229 out of 309 PWID tested positive for HCV (232), and a cross-sectional study in Northern Vietnam on PWID entering 06 drug treatment centers reported positive HCV tests among 350 of 455 PWID (231). In southern Vietnam, an earlier study reported 58 out of 67 PWID examined (87%) were positive for anti-HCV (229), and HCV prevalence in HCMC among PWID recruited into MMT pilot study in 2008 was 69.7% (n = 498) (239). Lastly, HCV prevalence estimates among men who actively inject drugs in HCMC according to sentinel surveillance reports of the Integrated Biological and Behavioral Surveillance (IBBS) from 2006 to 2009 were 71% (63.6-78.4%, n = 310) (245). HIV prevalence estimates are from sentinel surveillance reports com-
piled from the 2006 and 2009 IBBS (246, 247). IBBS is community-based systematic surveillance purposed for collecting information on health status and risk behaviors among high-HIV risk populations (MSM, commercial sex-workers and PWID). Recruitment was done by respondent-driven sampling, and information was obtained by one-on-one interviews and collection of biological samples.

**Interventions**

The initial interventions to be implemented in this model involve planned scale-up of MMT coverage and ART coverage. Each scale-up is implemented at varying coverage levels and rates, individually and in tandem. The ART and MMT percentages represent percentage of PWID initiated on the intervention each year. These initial interventions are projected out for a 10-year period, and then incidence of each infection and deaths are compared, to assess relative impact.

After the initial interventions are run for 10 years, HCV treatment coverage is implemented at varying levels and projected out a further 10 years on top of maximum previous scale up of ART and MMT coverage (80% and 50%). The coverage percentage indicates the proportion of chronic HCV patients who receive treatment by the end of each chronic stage, and move into an HCV-negative state. The same coverage of treatment is provided to all chronically HCV-infected patients (not restricted to those with chronic liver disease). It is assumed that treatment efficacy is high (around 95%) in line with estimates of the efficacy of the DAA combinations.

**Results**

**Model Fits**

Our model is able to reproduce the following observed trends in HIV and HCV prevalence: Figure 5.2 shows the range of model estimates of prevalence (light-colored
<table>
<thead>
<tr>
<th>Parameter Name</th>
<th>Description</th>
<th>Value</th>
<th>Range</th>
<th>Source</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PWID parameters</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mu I$</td>
<td>baseline death rate of PWID</td>
<td>$1/24.4 \text{ yr}^{-1}$</td>
<td>$[.026, .064]$</td>
<td>238</td>
<td>average age of PWID 34; average male life expectancy in VN 70.2 years</td>
</tr>
<tr>
<td>$\mu X$</td>
<td>baseline death rate of ex-PWID</td>
<td>$1/36 \text{ yr}^{-1}$</td>
<td></td>
<td>237</td>
<td></td>
</tr>
<tr>
<td>$\Lambda$</td>
<td>excess recruitment rate of new PWID</td>
<td>1.11</td>
<td></td>
<td>239</td>
<td></td>
</tr>
<tr>
<td>$\rho$</td>
<td>dropout rate/year of PWID in MMT</td>
<td>$1/7.69$</td>
<td>$[.05, .25]$</td>
<td>230</td>
<td>calculated by fitting to dropout data</td>
</tr>
<tr>
<td>$\gamma D$</td>
<td>1/duration of drug use career</td>
<td>$1/13 \text{ yr}^{-1}$</td>
<td></td>
<td>221</td>
<td></td>
</tr>
<tr>
<td>$\eta$</td>
<td>1/duration initial MMT phase</td>
<td>$1/2 \text{ yr}^{-1}$</td>
<td></td>
<td>221</td>
<td></td>
</tr>
<tr>
<td>$\nu$</td>
<td>recruitment rate into MMT clinic</td>
<td>$0.003 \text{ yr}^{-1}$</td>
<td></td>
<td>54</td>
<td></td>
</tr>
<tr>
<td>$mR$</td>
<td>reduction in risk among PWID entering MMT</td>
<td>0.8</td>
<td>$[.4, .9]$</td>
<td>230</td>
<td></td>
</tr>
<tr>
<td><strong>HIV parameters</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\theta$</td>
<td>Progression rate from acute stage (1/duration)</td>
<td>$4 \text{ yr}^{-1}$</td>
<td></td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>$c_P, c_I, c_A, c_T$</td>
<td>Transmission-weighting of prevalence term by stage of infection</td>
<td>$25, 1.7, .04$</td>
<td></td>
<td>11</td>
<td>Primary, Asymptomatic, AIDS, Treated</td>
</tr>
<tr>
<td>$\gamma_H$</td>
<td>1/duration of each stage of infection</td>
<td>$1/2.3 \text{ yr}^{-1}$</td>
<td></td>
<td>241</td>
<td></td>
</tr>
<tr>
<td>$\alpha_I$</td>
<td>treatment rate those with CD4$\geq$200</td>
<td>$0.021 \text{ yr}^{-1}$</td>
<td></td>
<td>221, 222</td>
<td>†</td>
</tr>
<tr>
<td>$\alpha_A$</td>
<td>treatment rate those with CD4$&lt;200$</td>
<td>see SI</td>
<td></td>
<td>221, 222</td>
<td></td>
</tr>
<tr>
<td>$\delta_I$</td>
<td>excess death-rate for treated, early initiation (including 10% LTF)</td>
<td>$1/21 \text{ yr}^{-1}$</td>
<td></td>
<td>250, 251</td>
<td></td>
</tr>
<tr>
<td>$\delta_A$</td>
<td>excess death-rate for Treated, late initiation (including 10% LTF)</td>
<td>$1/13.4 \text{ yr}^{-1}$</td>
<td></td>
<td>250, 251</td>
<td></td>
</tr>
<tr>
<td>$mT$</td>
<td>increase in treatment rates when linked into MMT</td>
<td>2</td>
<td>$[1, 5]$</td>
<td>230</td>
<td></td>
</tr>
<tr>
<td><strong>HCV parameters</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\kappa$</td>
<td>proportion of acutely infected individuals who clear infection</td>
<td>0.25</td>
<td>$[.15-.4]$</td>
<td>222</td>
<td></td>
</tr>
<tr>
<td>$\kappa_{HIV}$</td>
<td>proportion of acutely infected individuals who clear infection- HIV coinfected</td>
<td>0.1</td>
<td>$[.05-.15]$</td>
<td>222</td>
<td></td>
</tr>
<tr>
<td>$\phi$</td>
<td>proportion of cleared infections acquiring immunity</td>
<td>.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\phi_{HIV}$</td>
<td>proportion of cleared infections acquiring immunity- HIV coinfected</td>
<td>.01</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\gamma_A$</td>
<td>1/duration of Acute infection</td>
<td>$2 \text{ yr}^{-1}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\gamma_C$</td>
<td>progression rate through each stage of infection</td>
<td>$0.104 \text{ yr}^{-1}$</td>
<td>$[1/20, 1/8]$</td>
<td>160, 19</td>
<td></td>
</tr>
<tr>
<td>$\gamma_{C_{HIV}}$</td>
<td>acceleration of progression among HIV+</td>
<td>2</td>
<td>$[1.4, 3]$</td>
<td>160, 19</td>
<td></td>
</tr>
<tr>
<td>$\delta_L$</td>
<td>additional death rate due to chronic liver disease</td>
<td>$1/4 \text{ yr}^{-1}$</td>
<td>$[1/4, 1]$</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>$\epsilon$</td>
<td>HCV Treatment Efficacy</td>
<td>95%</td>
<td></td>
<td>180, 53</td>
<td></td>
</tr>
</tbody>
</table>

† In 2009 and 2010, roughly 12% of people initiating ART had CD4$\geq$200. Applying these percentages to number of people initiating ART in total for those years yields 1249 and 1297, out of an estimated 175510 and 176561 PLHIV with CD4$\geq$200. A rough estimate of the rate is this .7%/year or roughly 1/20th of the rate at which PLHIV with CD4 $<200$ initiate ART.
swathes) as well as the best-fitting parameter set output (dashed line) of the model to data. The parameters varied in the fitting process were coefficients in HIV and HCV hazard functions for incidence (coefficients changed after advent of ART), initial HIV prevalence, and impact of needle and syringe programs on the force of infection. Table 5.2 shows the maximum-likelihood estimates for each fitted parameter along with the 95% univariate bounds for each individual parameter estimate.

Figure 5.2: **Model Fit to HIV and HCV Prevalence.** This figure shows the range of model estimates for HIV (blue) and HCV (green) prevalence among PWID in the shaded regions, with the estimate from the best-fit parameter set represented by the dashed line. Data estimates and corresponding confidence intervals to which the model was calibrated are represented by circles and error bars.
Table 5.2: Maximum Likelihood Estimates of Fitted Parameters

<table>
<thead>
<tr>
<th>Name</th>
<th>Best Fit Mean</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_{HIV0}$</td>
<td>4.55E-05</td>
<td>8.84E-03</td>
<td>2.00E-02</td>
</tr>
<tr>
<td>$\beta_{HIV1}$</td>
<td>1.74E-01</td>
<td>1.72E-01</td>
<td>2.08E-01</td>
</tr>
<tr>
<td>$\beta_{HCV1}$</td>
<td>6.99E-01</td>
<td>7.92E-01</td>
<td>1.00E+00</td>
</tr>
<tr>
<td>Initial HIV prevalence</td>
<td>2.50E-02</td>
<td>1.29E-02</td>
<td>2.50E-02</td>
</tr>
<tr>
<td>$\omega_{ns}$</td>
<td>1.25E-01</td>
<td>2.23E-01</td>
<td>5.00E-01</td>
</tr>
</tbody>
</table>

Post-ART parameters

| $\omega_{HIV}$   | 5.46E-01      | 3.97E-01    | 8.39E-01    |
| $\beta_{HCV}$    | 6.45E-01      | 6.56E-01    | 1.00E+00    |

* The lower bound for $\beta_{HIV0}$ approaches zero because for some parameter sets, $\beta_{HIV1}$ is sufficient to explain HIV prevalence trends, while the lower bound for initial HIV prevalence approaches zero, because for some parameter sets, $\beta_{HIV0}$ is high enough to initiate the epidemic.

Interventions

Interventions were each scaled up to their final coverage levels over a five-year period and run at that level for a further five years. The plots in Figures 5.3 and 5.4 represent the reductions in HIV and HCV incidence, prevalence, and deaths under each type of intervention. HIV incidence and deaths are impacted by both ART scale-up and MMT scale-up. For HCV, MMT scale-up has a much more pronounced effect on incidence than ART scale-up. MMT has the strongest impact on incidence reduction because it stops exposure, targeting both infected and susceptible PWID. It has a greater impact even than ART on reducing HIV incidence, though ART has a more direct effect reducing HIV deaths. The impact of MMT on deaths from HIV and HCV is indirect: while MMT does have an impact on overall deaths from overdose, it only impacts deaths from infection by preventing future infections. With these prevention effects, reductions in deaths from HIV will follow within 10 to 20 years; but reductions in HCV deaths will not appear until longer after scale-up because of the longer duration of HCV infection. ART, however, will afford a small impact on HCV deaths due its reduction of the extent to which HIV speeds HCV progression. Combinations of ART and MMT scale-up have an additive effect on total deaths in
the population, with MMT scale-up contributing to reductions in deaths of otherwise-healthy PWID.

Figure 5.3: ART Scale-up, Incidence and Deaths changes over time. Each panel in this figure shows a plot of reductions in HIV and HCV incidence, prevalence or deaths with varying ART scale-up.

Figure 5.4: MMT Scale-up, Incidence and Deaths changes over time. Each panel in this figure shows a plot of reductions in HIV and HCV incidence, prevalence or deaths with varying MMT scale-up.

As can be seen in Figure 5.5, however, even with ART and MMT scale-up, the total number of deaths per year drops after 10 years of intervention, but the number of deaths due to HCV remains constant. This result suggests that including an HCV-specific intervention is necessary and may have significant marginal impacts.
Looking another 10 years forward to when high-efficacy DAA therapy for HCV is more affordable, it can be seen that implementing HCV treatment rollout in conjunction with ART and MMT scale-ups can provide substantial reductions in deaths in this population. Figure 5.6 demonstrates that including HCV treatment coverage on top of MMT and ART scale-ups can double or triple gains in reductions in deaths, indicating that the marginal benefit of rolling out HCV treatment is large. Even when ART and MMT have been scaled up at high levels, there are still many deaths not averted.

Figure 5.5: ART and MMT Scale-up, Reductions in deaths from disease over time. Each panel in this figure shows a plot of reductions in deaths from disease 10 years after interventionscale-up with varying ART and MMT scale-up.
Figure 5.6: **HCV Treatment, Reductions in deaths from disease over time.** Each panel in this figure shows a plot of reductions in deaths from disease 10 years after roll-out of HCV treatment coverage, with maximum previous scale up of ART and MMT coverage (80% and 50%).

### Cost per Life-Year Saved

Applying rough cost estimates of $200/year/person on ART and $900 for one-time treatment course for HCV, and integrating over the population size, we can calculate the cost per life-year saved for each intervention. The numbers in Table 5 indicate the cost in dollars per life-year saved projecting forward 30 years after implementing each combination of interventions under both low and high levels of MMT scale-up.

#### Table 5.3: Cost in Dollars per Life-Year Saved for Different Intervention Combinations

<table>
<thead>
<tr>
<th>Low MMT coverage</th>
<th>ART Scale-up Rate</th>
<th>HCV Coverage Rate</th>
<th>Mean (SD)</th>
<th>Mean (SD)</th>
<th>Mean (SD)</th>
<th>Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>55%</td>
<td>25%</td>
<td>504 (135)</td>
<td>498 (138)</td>
<td>487 (145)</td>
<td>477 (155)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50%</td>
<td>519 (115)</td>
<td>513 (117)</td>
<td>502 (124)</td>
<td>492 (133)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>75%</td>
<td>482 (91)</td>
<td>479 (94)</td>
<td>470 (101)</td>
<td>462 (109)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100%</td>
<td>456 (71)</td>
<td>453 (74)</td>
<td>446 (82)</td>
<td>439 (92)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>High MMT coverage</th>
<th>ART Scale-up Rate</th>
<th>HCV Coverage Rate</th>
<th>Mean (SD)</th>
<th>Mean (SD)</th>
<th>Mean (SD)</th>
<th>Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>55%</td>
<td>25%</td>
<td>606 (64)</td>
<td>603 (66)</td>
<td>599 (72)</td>
<td>596 (82)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50%</td>
<td>566 (53)</td>
<td>563 (54)</td>
<td>563 (57)</td>
<td>562 (64)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>75%</td>
<td>532 (36)</td>
<td>531 (37)</td>
<td>532 (42)</td>
<td>533 (50)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100%</td>
<td>519 (28)</td>
<td>519 (29)</td>
<td>520 (34)</td>
<td>521 (43)</td>
</tr>
</tbody>
</table>
Discussion and Conclusion

In this study, we use a mathematical model of the HIV and HCV co-epidemics in HCMC, Vietnam to analyze the effect of various interventions on future disease burden and cost-effectiveness of these infections. Our results indicate that scale-up of ART coverage can have an impact on HIV burden, though very little impact on HCV burden. MMT scale-up, however, can have an impact on both HIV and HCV incidence levels even at below-optimal coverage. We examine the impact of DAA therapy for HCV, and we conclude that HCV treatment roll-out in combination with MMT/ART scale-up has the potential to yield a multi-fold increase in the total number of deaths averted, compared to MMT/ART scale-up alone.

As with many modeling studies, this study has several limitations. Many parameters, such as transmission probability and impact of NSPs on transmission, are difficult to identify, and must be estimated using fitting techniques that cannot always account for co-linearity. The data that we use for calibration are from a variety of sources with varying sample sizes and confidence. However, we perform sensitivity analyses (see Appendix E) to demonstrate that our results are qualitatively robust across parameter sets, and our predictions conservative.

Our results confirm predictions by Durier et al. (54) that show very optimistic results for reductions in prevalence and incidence of HCV after roll-out of HCV treatment. In addition to extending healthy life expectancy for newly-cured HCV patients, HCV treatment coverage of coinfected patients insures the gains in life expectancy given by ART by protecting those patients from liver disease deaths. Our model indicates that ART and MMT scale-ups, while effective, still leave many deaths not averted unless HCV treatment is rolled out in tandem. This result holds across varying levels of ART and MMT scale-ups, suggesting that HCV treatment roll-out need not wait until ART and MMT hit optimal levels.
Currently, interferon and ribavirin treatment is prohibitively expensive in Vietnam, has limited efficacy, and is associated with side effects that can be severe (16). The new DAAs have shorter durations and much higher efficacy, but are likewise prohibitively expensive with costs ranging up to nearly $200,000 (253). However, the real costs of manufacturing these drugs are lower. In the future, it may be possible to produce 12-week regimens for $100-$250 (253). When costs for these treatments do drop in the next decade, it may be feasible to start incorporating HCV care into MMT programs. Alternatively, this population may be a good candidate population for a reduced-cost HCV treatment campaign such as in Egypt. A brief analysis indicates that if costs of HCV treatment are indeed dropped to $900 per treatment course, the cost per life-year saved decreases with increased HCV treatment coverage across varying levels of ART scale-up, and especially if MMT scale-up lags behind (Table 5).

In confronting the HIV epidemic in Vietnam, it is vital to take into account the high rates of HCV coinfection. Scale-up of ART can have significant impact on HIV incidence and deaths, but does very little for HCV control (though individual patients do benefit from ART (171)). Scale-up of MMT impacts both HIV and HCV incidence as well as deaths, though it is crucial for it to be tied to ART access. However, both of these interventions leave much to be desired in the way of averting HCV deaths, indicating that if HCV treatment programs can be rolled out in Vietnam, they will have the potential to greatly increase life expectancy gains beyond what is afforded by MMT and ART scale-up.

MMT scale-up is becoming a reality in Vietnam, with a proposed 20 new clinics opening in 2016, and a goal of reaching 80,000 PWID across the country by the end of 2016 (227). Studies thus far have indicated that MMT has good acceptability among PWID, and that retention rates are high (∼ 90% after 1 year) (239). MMT has the potential to offer longer life expectancy to PWID with and without HIV and HCV,
and as can be seen in the simulations, the potential to reduce incidence and deaths significantly.

The Vietnamese government has committed to ART and MMT scale-up efforts, and tremendous progress has been made: as of December 2011, 11 provinces have functioning MMT programs reaching nearly 7,000 PWID and benefits have been seen, with reduced family conflicts and crime reported in areas with these programs (227). These province-level programs receive the bulk of their funding from PEPFAR, DFID and other international assistance. However, these sources are expected to significantly reduce in the near future due to global financial concerns, and as Vietnam recently achieved middle-income status (227). The Vietnamese government is switching focus from relying on international funding to national funding and management, and it is crucial to ensure that gains already achieved in HIV care are not lost, and that future progress can be sustainable (227).

If HCV treatment programs are rolled out in the future, they could also be well incorporated into existing ART and MMT infrastructure. Targeting coinfected patients for co-treatment is effective for both treatment and prevention, especially if the patients are enrolled in MMT. Costs may be difficult to control, but building programs into existing clinics and satellite centers may help curb short-term costs for long-term gains.

See Appendix E for full supplementary information.
Conclusion

This dissertation grew out of an initial question of how to model and predict the potential impacts of HIV interventions on the epidemic in Newark, NJ. The project that we designed to address this question brought to light the importance of coinfecting pathogens in HIV epidemiology, and in turn the important cross-scale nature of coinfection dynamics.

The main findings of this dissertation both methodological and policy-oriented. Chapter 1 underscores the necessity of fixing the “leaky” care-continuum in order to effect a change in disease burden in Newark, NJ. In this chapter, and Chapter 2 we discuss the limitations of assessing the traditional Test-and-Treat interventions in favor of including a more comprehensive slate of interventions. In Chapter 2 we incorporate prevention interventions and show how targeting them toward specific high-risk groups can maximize their impact. Chapter 5 focuses exclusively disease burden in a high-risk group, and shows how even targeted interventions can miss achieving their full potential if coinfection is not taken into account, and we recommend rolling out treatments to this vulnerable population. The new DAAs for treatment of HCV are very promising on an epidemiological scale, and will likely change the face of the epidemic once they become widely available. With their distribution, it becomes increasingly important to interrogate the within-host dynamics of HCV. In Chapter 3 we describe a novel model of HCV-HIV within-host coinfect-
tion that can be used to make predictions about treatment failure in monoinfected vs coinfected patients. Getting out in front treatment failure with these pricy new drugs is crucial to prevent or delay the evolution of drug resistance. The modeling approach used in Chapter 4 delves further into the within-host evolutionary dynamics of HCV, and allows us to assess the quality and type of information we can infer from looking at the viral lineages existing in chronic HCV patients. Understanding the mechanisms of within-host transmission and how they impact viral evolution is similarly important in designing treatment strategies to maximize SVR chances and prevent drug resistance.

Tying all of these projects together are the themes of model utility and cross-scale dynamics. Each chapter delineates the power and limitations of a certain type of model, and together they underscore the importance of thinking across scales when addressing such profound public health problems and HIV and HCV. Without the within-host insight of models such as those proposed by Ho et al (56) and Neumann et al (3), both clinical and epidemiological research would be far behind. Conversely, epidemiological models can predict, and in some cases direct the direction of future therapies and intervention, and highlight the areas most pressing for evolutionary study. In this dissertation, as well as in parallel work on drug resistance (69, 70) I aim to highlight the complexities of within-host dynamics while using them to inform epidemiological work.

As with all modeling work, there are caveats with each of the types of model used in this dissertation. These caveats have been discussed in the individual chapters, but here I will present a brief general summary. Models of chronic infections can pose problems with proliferation of parameters and data scarcity. In contrast to acute infections such as measles or influenza, which can be described with relatively simple SIR-type models that lend themselves well to remarkably accurate inference (e.g. (254, 255, 256)), longer-term infections display heterogeneities and infection
durations that require more complex models to describe. The multiple stages of HIV infection, for example, each at a minimum require parameters to describe their duration and transmission probability, each of which are distinct between stages. There is a growing body of clinical work that can inform the ranges for these parameter estimates, but many still need to be estimated with models such as those proposed in Chapters 1, 2 and 5. As the number of parameters grows, as it will with models that describe complex populations and/or coinfection, the accuracy of these estimates will necessarily drop, especially if data availability remains limited. The prevalence and incidence data available for calibrating these chronic infections like HIV and HCV is often scarce, patchy, from multiple sources, or proxy populations, so there is a necessary trade off between model complexity and prediction accuracy. Previous models of HIV infection have also run up against these issues (see models discussed in introduction, (257, 81) and others); this is an open problem in chronic disease epidemiology.

However, cross-scale dynamics can inform such models, e.g. using different transmission probabilities for different stages of HIV infection based on breakthroughs in patient-level viral dynamics. Additionally, as new methods are being developed the data on which to base these models will improve. An ideal dataset for model calibration might include some combination of the following elements:

- A cohort follow-up study of HIV and HCV monoinfection and HIV-HCV coinfection patients over time with
  
  - regular measurements of CD4 count and HIV viral load before and during treatment
  
  - regular HCV viral load measures
  
  - regular genetic analysis of HCV viral load
  
  - bedside data collection including deep sequencing
• an accurate record of contact tracing

• detailed data on order of infection

- HIV and HCV prevalence data at the scale and accuracy of Project IMPACT (the data source described in Chapter [1]) across all US cities.

These data would have the ability to reconcile parameter estimates as well as models across scales, and help delineate where the cross-scale interactions are most relevant. However, even without this ideal dataset, these models have great utility and it would be a missed opportunity not to use them for broad-level public health estimates. While they may not be able to generate estimates of exact \( R_0 \)'s, or exact number of incident cases averted with a certain intervention, they are crucial for comparative analysis of public health interventions both prospectively and retrospectively (where they can be validated). Their flexibility allows them also to sync well not only with within-host data, but also with with economic data in cost-effectiveness studies that have high utility for informing policy.

When uncertainty is present in model parameters, model structure can help quantify this uncertainty through Bayesian comparison with data. There is also uncertainty inherent in model structure and specification; we cannot know a priori if the assumptions used to construct the model are valid, though we can have good guesses from empirical work. If model assumptions are laid out clearly, the model creates framework for testing and feedback across scales. Crucially, the results about interventions presented in Chapters [1] [2] and [5] are robust across model uncertainties and stand up to sensitivity analysis, thus creating testable predictions which can accommodate feedback from future data generate by implementing these interventions.

These model limitations and potential solutions further highlight the importance of cross-scale approaches to understanding and modeling pathogen dynamics. Many challenges remain in reconciling the data available with models, and future work should focus on designing models and collecting data that can work together syner-
gistically. Future modeling efforts will benefit from taking advantage of the increasing ranges of data available (e.g. genetic, real-time population location and density data, electronic medical records, other sources of digital epidemiology), and should be constructed in a flexible manner that scales (e.g. designing network models that become mass-action models within a certain range of parameter values). Statistical methods such as the Random Forest Algorithm discussed in Chapter 4 will help bridge these gaps.

Future Work

There are many directions for future work; salient ones include the following. The model of HIV in Newark described in Chapter 1 is being extended in several directions. Cost-effectiveness analysis is being done on the results to price out and compare the various care-continuum interventions. Additionally, we are incorporating HCV into the model and exploring opiate-substitution therapy and HCV treatment interventions using methods similar to those described in Chapter 5. Potential future work stemming from Chapter 4 will involve a comparative analysis with the data presented by Raghwani et al. (199), and may also involve expanding the framework described to analyze within-host patterns of other viral infections. Lastly, the cross-scale nature of the research described in this dissertation acts as the first step in my future research program, which will focus on analyzing coinfection interactions across population and geographic scales and interventions.

Across many epidemiological settings, coinfection is the rule rather than the exception, yet many public health strategies do not take this phenomenon into account, due to the complexity of designing individual and public-health-level treatment approaches. One set of systems (in addition to HIV-HCV) where these questions are particularly pressing are viral-bacterial coinfections. A general looming problem
infectious disease is the growing specter of drug resistance. Improper use of antimicrobial chemotherapy has furthered the spread of drug resistance and hindered the sustainability of the shrinking number of antimicrobial drugs available. Viral-bacterial interactions on the within-host level have been described extensively in the literature and their importance is starting to be noted in the context of drug resistance (see (70) and others). These interactions are complex and myriad: pathogens can interact through immune-mediated pathways (258, 259), as well as directly (260, 259) with the presence of a coinfecting virus often increasing bacterial load multifold. Within-host modeling work is starting to elucidate coupled interactions and mechanisms in these systems (e.g. (261)), and can be used to inform epidemiological studies, population-level predictions, and policy.

One example of the complex cross-scale feedback in this system is that a number of viral infections that are associated with bacterial infections (e.g. influenza, respiratory syncytial virus (RSV) with S. pneumoniae or pneumococcus) can have strong seasonal and environmental patterns. The impact of this seasonality on bacterial infection patterns has been investigated recently with mechanistic models coupled with incidence data to suggest that viral infections can be a main driver of bacterial infection seasonality (262, 263, 264). Crucially, seasonality also plays a role in antibiotic consumption. Several studies have found strong associations with increases in prescriptions during winter months and with flu incidence (265, 266). There is evidence that seasonal trends in prescribing and consumption translate into lagged seasonal increases in incidence of antibiotic-resistant E. coli and Methicillin-Resistant Staphylococcus Aureus (MRSA) (267). Drug resistance is a growing problem in bacterial infection, and given these strong seasonal drivers from viral and prescribing patterns, it seems likely that there may be a connection between viral dynamics and drug resistance. The growth of antimicrobial drug resistance is phenomenologically underpinned by pathogen-host evolution, so understanding how pathogens evolve and
spread in the human body and in populations is crucial for designing effective interventions, from drug-design to implementation of public health policy. Sustainability of the full complement of antimicrobials is a fraught challenge, and forecasting models will be key to planning successful public health campaigns.

At the broader population scale, factors related to seasonality are important in the context of urbanization: seasonal effects will be different in urban and non-urban settings (268, 269) contributing to the complexity of this system. With over half of the world’s population currently living in cities, and a projected global urban population of 6.3 billion by 2050 (270), urban population dynamics will be one of the key determinants of public health and socioeconomic well-being over the next decades. In light of these population shifts, there is a growing need to focus on how changing urban exposures might impact human health, and in particular infectious disease transmission (271). Indeed, urbanization can either aid infectious disease spread, e.g. via overcrowding and increased contact rates, or help control it, e.g. via improved healthcare access relative to rural areas. It is thus extremely important to better understand the drivers and outcomes of infectious disease in an urban context, as accurate prediction of how urbanization might impact human, pathogen and drug interactions will be necessary for planning and sustaining 21st-century urban health systems.

Because urbanization, infectious disease, and drug resistance can interact in such a complex manner, it is a challenge to even frame the open questions that arise from this system. Designing modeling frameworks can help delineate these questions in this emerging field and move toward valuable predictions. In undertaking this research program, I will build on the types of methods I have used in my PhD work to connect the across the various levels at play, from genetic to geographic. Working with both data and theory, I aim to use models to untangle cross-scale feedback inherent in phenomena such as seasonality and treatment optimization. For example,
I will combine my previous review work (69, 70) into a broad theoretical model that encompasses both coinfection and treatment strength on within-host and epidemiological levels. With recent advances in empirical understanding of the mechanisms that govern drug resistance evolution and spread (e.g. (272)), it will be possible to parameterize this model in a realistic manner. Models of this type are already being used to understand pneumococcal strain dynamics, e.g. work by Mitchell et al (273) showing antibiotic pressure as the strongest driver of emergence of resistance in pneumococcus. Expanding on this type of model and building in the flexibility to include the influence of coinfecting pathogens as well as ranges in treatment intensity can help make population level predictions for a range of coinfection systems with the potential to inform clinical and public health decisions.

Overall, my research has informed my thinking about two useful and complementary aspects of modeling: the first relates to how models can be utilized to understand the dynamics of complex systems. The second, which can often rely strongly on the first, is how models can be used to design policies or strategies that change systems. These two issues will drive the modeling work I intend to pursue in my career as an academic.
Bibliography


[7] Dalgleish AG, Beverley PCL, Clapham PR, Crawford DH, Greaves MF, Weiss RA. The CD4 (T4) antigen is an essential component of the receptor for the


[40] Boily MC, Lowndes C, Alary M. The impact of HIV epidemic phases on the effectiveness of core group interventions: insights from mathematical models. Sexually Transmitted Infections. 2002 apr;78(Supplement 1):i78–i90. Available from: http://sti.bmj.com/content/78/suppl{1}/i78.abstracthttp://sti.bmj.com/content/78/suppl{1}/i78.short.


Ho DD, Neumann AU, Perelson AS, Chen W, Leonard JM, Markowitz M. Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. Nature. 1995 jan;373(6510):123–6. Available from: [http://dx.doi.org/10.1038/373123a0](http://dx.doi.org/10.1038/373123a0).


Ribeiro RM, Mohri H, Ho DD, Perelson AS. In vivo dynamics of T cell activation, proliferation, and death in HIV-1 infection: why are CD4+ but not


Examining the promise of HIV elimination by test.


147


Takuva S, Maskew M, Brennan AT, Long L, Sanne I, Fox MP. Poor CD4 recovery and risk of subsequent progression to AIDS or...


31-Jul-2012{&}view=c{&}originContentFamily=serial{&}wchp=dGLbV1S-zSkWb{&}md5=d00f1800d42c78d60ca9c7bae62e449e{&}pid=1-s2.0-S0955395911002337-main.pdf


[243] Family Health International. Results of the program evaluation of patients initiating antiretroviral therapy in two health facilities in Ho Chi Minh City, Viet Nam. Ha Noi, Viet Nam. Family Health International; 2010.


[265] Suda KJ, Hicks LA, Roberts RM, Hunkler RJ, Taylor TH. Trends and seasonal variation in outpatient antibiotic prescription rates in the United States,


Appendices
Figure A.1: **Best-fitting Parameter Set.** Model output of best-fitting parameter set in solid lines, with data points on connected dots.
Figure A.2: Contour Plots of Incidence and Deaths Reductions. (a) Shows contours of equivalent incidence reduction for combinations of the two most efficacious interventions. (b) Shows contours of equivalent reductions in deaths for combinations of the two most efficacious interventions.
Table A.1: Partial Rank Correlation Coefficients

<table>
<thead>
<tr>
<th>Parameter</th>
<th>PRCC</th>
<th>Parameter</th>
<th>PRCC</th>
<th>Parameter</th>
<th>PRCC</th>
<th>Parameter</th>
<th>PRCC</th>
<th>Parameter</th>
<th>PRCC</th>
<th>Parameter</th>
<th>PRCC</th>
<th>Parameter</th>
<th>PRCC</th>
<th>Incidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>ζ</td>
<td>0.42 ***</td>
<td>Ψref -0.75 ***</td>
<td>η</td>
<td>0.70 ***</td>
<td>(msm,b) 0.48 ***</td>
<td>Ψref -0.71 ***</td>
<td>pdmdu 0.77 ***</td>
<td>Ψref -0.62 ***</td>
<td>a  -0.28 ***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mx</td>
<td>-0.23 *</td>
<td>pdu 0.60 ***</td>
<td>η</td>
<td>-0.54 ***</td>
<td>(h,h) 0.45 ***</td>
<td>pdu 0.54 ***</td>
<td>Ψref -0.62 ***</td>
<td>a  -0.28 ***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cd</td>
<td>0.15</td>
<td>cd -0.34 **</td>
<td>η</td>
<td>0.31 **</td>
<td>(msm,b) 0.42 ***</td>
<td>cd -0.37 ***</td>
<td>prm 0.57 ***</td>
<td>ζ  0.15 **</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ε</td>
<td>-0.15</td>
<td>ζ  -0.32 **</td>
<td>η</td>
<td>-0.31 **</td>
<td>(msm,l) -0.41 ***</td>
<td>a  -0.34 **</td>
<td>mx 0.52 ***</td>
<td>η  -0.03 **</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ν(l)</td>
<td>0.15</td>
<td>ν(l) -0.27 *</td>
<td>η</td>
<td>(msm,l) 0.31 **</td>
<td>hrc -0.30 **</td>
<td>σd -0.33 **</td>
<td>(msm,b) -0.44 ***</td>
<td>βmm -0.02 **</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ψref</td>
<td>-0.14</td>
<td>σd -0.24 *</td>
<td>η</td>
<td>(m,l) -0.30 **</td>
<td>(h,h) -0.29 *</td>
<td>ν(l) -0.29 *</td>
<td>ν(h,h) -0.41 ***</td>
<td>(msm,b) -0.44 *</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>α</td>
<td>0.14</td>
<td>c(hrm:m) -0.20</td>
<td>prn -0.22</td>
<td>prn 0.24 *</td>
<td>prdμ 0.21</td>
<td>c(msm,l) 0.30 **</td>
<td>c(m,l) -0.29 *</td>
<td>c(l) -0.28</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>β</td>
<td>-0.14</td>
<td>ωl 0.16</td>
<td>ν(l) -0.20</td>
<td>ω 0.20</td>
<td>c(hrm:m) -0.19</td>
<td>cd -0.29 *</td>
<td>φ  -0.28</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>prn</td>
<td>-0.13</td>
<td>mx 0.16</td>
<td>Ψref -0.18</td>
<td>mx 0.19</td>
<td>φ 0.18</td>
<td>a  -0.28 *</td>
<td>ν(h,h) -0.41</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*** = p<.001, ** = p <.01, * p<.05
Equations for Newark Model

Main Equation

\[
\begin{align*}
\frac{dS(i)}{dt} &= \Lambda(i) \times N(i) - (\lambda_S(i) + \lambda_D(i) + \mu(i) + \psi(i, t)) \times S(i) + \zeta \times S(i') \\
\frac{dI_0(i)}{dt} &= (\lambda_S(i) + \lambda_D(i)) \times S(i) - (\mu(i) + \theta + \psi(i, t)) \times I_0(i) + \zeta \times I_0(i') \\
\frac{dI_T(i)}{dt} &= z(i) \times \theta \times I_0(i) - \left(\mu(i) + \frac{1}{T_{test}} + \psi(i, t)\right) \times I_T(i) + \zeta \times I_T(i') \\
\frac{dI_{TNE}(i)}{dt} &= (1 - y(i)) \times \frac{I_T(i)}{T_{test}} - \left(\mu(i) + \frac{1}{T_{elig}} + \psi(i, t)\right) \times I_{TNE}(i) + \zeta \times I_{TNE}(i') \\
\frac{dI_{TE}(i)}{dt} &= y(i) \times \frac{I_T(i)}{T_{test}} + (test_{ink} \times ret) \frac{I_{TNE}(i)}{T_{elig}} - \left(\mu(i) + \frac{1}{T_{treat}} + \psi(i, t)\right) \times I_{TE}(i) + \zeta \times I_{TE}(i') \\
\frac{dI_U(i)}{dt} &= (1 - test_{ink} \times ret) \frac{I_{TNE}(i)}{T_{elig}} + (1 - test_{ink} \times ret \times trt_{ink}) \times \frac{I_{TE}(i)}{T_{treat}} - \left(\mu(i) + \frac{1}{\gamma_{NM} - T_{test}} + \psi(i, t)\right) \times I_U(i) + \zeta \times I_U(i') \\
\frac{dI_N(i)}{dt} &= x(i) \times test_{ink} \times ret \times trt_{ink} \times \left(\frac{I_{TE}(i)}{T_{treat}} + \alpha\Lambda_{NT}(i)\right) - (\mu(i) + \psi(i, t)) \times I_N(i) + \zeta \times I_N(i') \\
\frac{dI_{NS}(i)}{dt} &= (1 - x(i)) \times test_{ink} \times ret \times trt_{ink} \times \left(\frac{I_{TE}(i)}{T_{treat}} + \alpha\Lambda_{NT}(i)\right) - (\mu(i) + \psi(i, t) + \gamma_{NS}) \times I_{NS}(i) + \zeta \times I_{NS}(i') \\
\frac{dA_T(i)}{dt} &= \gamma_{NS} \times I_N(i) - (\mu(i) + \alpha + \psi(i, t)) \times A_T + \zeta \times A_T(i') \\
\frac{dI_{NT}(i)}{dt} &= (1 - z(i)) \times \theta \times I_0(i) - (\mu(i) + \gamma_{NT} + \psi(i, t)) \times I_{NT}(i) + \zeta \times I_{NT}(i') \\
\frac{dA_{NT}(i)}{dt} &= \gamma_{NT} \times I_{NT}(i) + \left(\frac{1}{\gamma_{NM} - T_{test}}\right) \times I_U(i) - (\mu(i) + \alpha + \psi(i, t)) \times A_{NT} + \zeta \times A_{NT}(i') 
\end{align*}
\]

Where \(S\) represents Susceptible Individuals, \(I_0\) Acute infection, \(I_T\) asymptomatic will be tested before CD4+<200, \(I_{TNE}\) Tested - ineligible , \(I_{TE}\) Tested eligible, \(I_U\) Tested but unlinked/not on treatment, \(I_N\) Treated and Suppressed, \(I_{NS}\) Treated not Suppressed, \(A_T\) AIDS post-ART, \(I_{NT}\) will not be tested before CD4+<200, and \(A_{NT}\) AIDS not tested. \(i\) indicates risk group, and \(i'\) corresponding non-IDU risk group. 

Note: The equations above represent non-IDU classes. Equations for IDU classes are identical except that \(\psi(i, t)\) and \(\zeta\) switch places.
Inflow/Outflow guide to equations

**Change in susceptible population:** Inflow is recruitment rate, outflow is force of infection from sexual and IDU transmission

**Change in Acute Infecteds:** Inflow is force of infection from sexual and IDU transmission, outflow is the reciprocal of the average duration in the Acute stage of infection.

**Change in population who will be tested before CD4<200:** Inflow is based on progression rate out of Acute Infected times proportion of infected individuals who get tested before CD4<200. Outflow a rate based on the average waiting time to get tested.

**Change in population who are NOT eligible for ART when tested:** Inflow is the rate based on the waiting time to get tested times the proportion who are not eligible when tested. Outflow is a rate based on the average waiting time to ART eligibility adjusted for having already waited to get tested.

**Change in population who ARE eligible for ART when tested:** Inflow is the rate based on the waiting time to get tested times the proportion who are eligible when tested. Outflow is a rate based on the average waiting time to ART linkage.

**Change in unlinked/linked but untreated population:** Inflows are based on 1) proportion of individuals not eligible for ART when tested who are not retained in care 2) proportion of individuals eligible for ART when tested who are not linked to ART. Outflow is based on average time until CD4 drops below 200 adjusted for the previous waiting times.
**Change in treated, suppressed population:** Inflow is based on proportion of tested, eligible, retained, linked individuals who are adherent to treatment and so achieve virologic suppression.

**Change in treated, unsuppressed population:** Inflow is based on proportion of tested, eligible, retained, linked individuals who are not adherent to treatment and so do not achieve virologic suppression. Outflow is a rate based on the average waiting time for unsuccessfully treated individuals to progress to CD4<200.

**Change in CD4<200 post-ART population:** Inflow is based on the average waiting time for unsuccessfully treated individuals to progress to CD4<200. Outflow is based on the average time until death from CD4<200.

**Change in asymptomatic, not tested before CD4<200 population:** Inflow is based on progression rate out of Acute Infected times proportion of infected individuals who do not get tested before CD4<200. Outflow is based on progression rate to CD4<200 in the absence of ART.

**Change in CD4<200, ART-naive population:** Inflow is based on progression rate to CD4<200 in the absence of ART. Outflow is based on the average time until death from CD4<200.

**Parameter Definitions**

Λ is the rate of population growth

μ(i) is the rate of leaving sexual activity (differs between IDU and non-IDU, and treated individuals have a slightly increased rate compared to susceptible individuals)

λ_S, λ_D are forces of infection (equations below)
\( \psi(t) = \psi_0 \times \exp(-\psi_{red} \times (t - T_0)) \) is the incidence of injecting drug use (differs between heterosexual men, women and MSM and decays over time up until minimum value is reached, in accordance with secular trends in Newark)

\( \zeta \) is the rate of stopping drug use (1/average duration of using drugs)

\( \theta \) is the rate of progression from Acute Infection to asymptomatic infection

\( z(i) \) is the proportion tested before CD4\(^+\) < 200 in each risk group

\( T_{test} \) is average time between progression to asymptomatic phase and testing

\( T_{elig} \) is the average time to become eligible if initially ineligible when tested

\( T_{treat} \) is the average time between becoming eligible and initiating treatment (if tested)

\( test_{\text{link}}, ret, trt_{\text{link}} \) are the proportion of those tested who are linked to care, proportion of those linked to care retained in care, and proportion of those retained in care linked to treatment, respectively

\( x(i) \) is proportion of patients linked to treatment who successfully achieve viral suppression

\( \alpha \) is the accelerated death rate for individuals with AIDS

\( \gamma_{NT}, \gamma_{NS} \) are the rate of progression from the asymptomatic stage to AIDS stage in untreated and non-adherent treated classes, respectively

There are two components to the force of infection: force of infection from sexual transmission, \( \lambda_S \), and force of infection from shared needles among injecting drug users, \( \lambda_D \). The formulae for differ between the sexes because of the MSM component, while the \( \lambda_D \) are the same. They are as follows:
Force of infection from Sexual Transmission

Men

\[
\lambda_S(m, i, t) = \sum_j p_s(m, i, j, t) c(m, i) p_w_f(j) \times \frac{1}{N(f, j, t)} \times \sum_{\text{stage}} \Omega_{\text{stage}fm} + \\
\sum_j p_s(m, i, j, t) c(m, i) p_w_m(j) \times \frac{1}{N(m, j, t)} \times \sum_{\text{stage}} \Omega_{\text{stagemm}}
\]

Women

\[
\lambda_S(f, i, t) = \sum_j p_s(f, i, j, t) c(f, i) \times \frac{1}{N(m, j, t)} \times \sum_{\text{stage}} \Omega_{\text{stagemf}}
\]

With

\[
\Omega_{\text{stage}kk'} = (1 - (1 - \beta_{kk'} \omega_{\text{stage}} \phi)^{\nu(kk', i, j, t)}) \times N_{\text{stage}}
\]

Where \( p_s \) is the term in the sexual contact mixing matrix indicating the probability of an individual of risk group \( i \) interacting with an individual of risk group \( j \) (of opposite gender for females and heterosexual males, and of same gender for homosexual males or bisexual men interacting with male partners), determined as follows:

\[
p_s(k, i, j, t) = \varepsilon \delta_{ij} + (1 - \varepsilon) \frac{c(k', j) N(k', j, t)}{\sum_s c(k', s) N(k', s, t)}
\]

where \( \varepsilon \) is the assortativeness coefficient and \( \delta_{ij} \) is the Kronecker delta.

\( c(k, i) \) is partner change rate in gender \( k \), risk group \( i \)

\( pw_k \) is the proportion of men in a risk group that interact with gender \( k \) (e.g. for homosexual men is 0, while for low risk men is 1-proportion exclusively homosexual)
Ω terms represent weighted prevalence by stage of disease and risk groups interacting, and comprise the terms

\[ \beta_{kk'} \] is the weighting of per-sex-act transmission probability depending on which sexes are interacting and which partner is infected.

\[ \omega_{stage} \] is the weighting of transmission probability depending on stages of disease (i.e. viral load weighting).

\[ \phi \] is per-sex-act transmission probability during the asymptomatic stage (baseline).

\[ \nu_{kk',ij} \] is the number of unprotected sex acts in a partnership of a (k,i) individual and a (k,j) individual.

\[ N_{stage} \] is the number of persons in disease stage.

**Force of Infection from Injecting Drug Use**

\[ \lambda_d(k,i,t) = \sum_j c_d p_d(k,l,i,j) \frac{1}{N(l,j,t)} \sum_{stage} \Upsilon_{stage} \]

where

\[ \Upsilon_{stage} = (1 - (1 - \omega_{stage}\sigma)^\nu) \times N_{stage} \]

Where \( c_d \) is the number of other IDUs with whom needles are shared in one year, \( p_d \) is the term in the needle-sharing mixing matrix indicating the probability of an individual of gender \( k \), risk group \( i \) sharing a needle with an individual of gender \( l \), risk group \( j \). The IDU mixing matrix is assumed to be non-assortative (proportionate mixing) only weighted by the higher number of male IDU than female. The elements \( p_d(k,l,i,j) \) are calculated by assuming 4 male risk groups and 2 female risk groups with contact with the male risk groups being twice as likely. Thus

\[ 4p_m + 4p_f = 4(2 \times p_f) + 2p_f = 10p_f = 1 \rightarrow p_f = .1, p_m = .2 \]
The $\Upsilon$ terms, like the $\Omega$ terms, represent weighted prevalence by stage of disease and comprise the terms

$\omega_{\text{stage}}$ as defined above

$\sigma$ the per-injection transmission probability,

$\eta = \eta_0 \times \exp(-\eta_{\text{red}}(t - T_0))$ the number of shared needles per year which, as drug use incidence, decays over time in accordance with secular trends in Newark

$N_{\text{stage}}$ as defined above

**Initial Conditions**

We used the initial proportions of each risk group as defined in the parameter tables applied to an initial population of Newark from the 1980 census, multiplied by the proportion of the population over 15 years of age (.72*329248), and we seeded the epidemic in 1980 with a very low prevalence (3 cases in each risk group).
Appendix B

Chapter 2 Supplementary Information

Supplementary Methods

Model Equations

Differential Equations for HIV Care-Continuum Compartmental Model with PrEP

Equations 1-11 are those for the HIV care-continuum, were previously published in Birger et al [71] and can also be found in Appendix A. Equations 12 and 13 are for the two PrEP compartments.

1. \( \frac{dS}{dt} = \Lambda (i) \times N(i) - [\lambda S(i) + \lambda D(i) + \mu(i) + \psi(i,t)] \times S(i) + \zeta \times S'(i) \)

2. \( \frac{dI_0}{dt} = [\lambda s(i) + \lambda d(i)] \times S(i) - [\mu(i) + \theta + \psi(i,t)] \times I_0(i) + \zeta \times I_0'(i) \)

3. \( \frac{dI_T}{dt} = \zeta(i) \times \theta \times I_0(i) - [\mu(i) + \frac{1}{T_{treat}} + \psi(i,t)] \times I_T(i) + \zeta \times I_T'(i) \)

4. \( \frac{dT_{TNE}}{dt} = (1 - y(i)) \times \frac{I_T(i)}{T_{test}} - \left[ \mu(i) + \frac{1}{T_{elig}} + \psi(i,t) \right] \times I_{TNE}(i) + \zeta \times T_{TNE}(i') \)

5. \( \frac{dI_{U}}{dt} = y(i) \times \frac{I_T(i)}{T_{test}} + (test_{link} \times ret) \frac{I_T(i)}{I_{elig}} \times I_{TNE}(i) - \left[ \mu(i) + \psi(i,t) + \frac{1}{T_{treat}} \right] \times I_{TE}(i) + \zeta \times I_{TNE}(i') \)

6. \( \frac{dI_N}{dt} = (1 - test_{link} \times ret) \frac{I_{TNE}(i)}{I_{elig}} + (1 - test_{link} \times ret \times trt_{link}) \times \frac{I_{TNE}(i)}{I_{treat}} - \left[ \mu(i) + \left( \frac{1}{T_{NM}} - \frac{1}{T_{test}} \right) + \psi(i) \right] \times I_U(i) - \zeta \times I_U(i') \)

7. \( \frac{dI_{S}}{dt} = x(i) \times test_{link} \times ret \times trt_{link} \times \frac{I_{TNE}(i)}{I_{treat}} + \alpha \times A_{NT}(i) - \left[ \mu(i) + \psi(i,t) \right] \times I_S(i) + \zeta \times I_S(i') \)

8. \( \frac{dI_{NS}}{dt} = (1 - x(i)) \times test_{link} \times ret \times trt_{link} \times \frac{I_{TNE}(i)}{I_{treat}} + \alpha \times A_{NT}(i) - \left[ \mu(i) + \psi(i,t) + \gamma_{NS} \right] \times I_{NS}(i) + \zeta \times I_{NS}(i') \)
\[ \frac{dA}{dt} = \gamma_{NS} I_{NS} (i) - [\mu (i) + \alpha + \psi (i, t)] \times A_T (i) + \zeta \times A_T (i') \]

\[ \frac{dI_{NT}}{dt} = (1 - z (i)) \times \theta \times I_0 (i) - [\mu (i) + \gamma_{NT} + \psi (i, t)] \times I_{NT} (i) + \zeta \times I_{NT} (i') \]

\[ \frac{dA_{NT}}{dt} = \gamma_{NT} I_{NT} - [\mu (i) + \alpha + \psi (i, t)] \times A_{NT} (i) + \zeta \times A_{NT} (i') \]

\[ \frac{dP_L}{dt} = (1 - q (i)) \times H (i) \times S (i) - (\lambda_{P-low} (i) + \lambda_{P-drug-low} (i) + \mu (i) + \varphi (i, t)) \times P_L (i) + \zeta \times P_L (i') \]

\[ \frac{dP_H}{dt} = q (i) \times H (i) \times S (i) - (\lambda_{P-high} (i) + \lambda_{P-drug-high} (i) + \mu (i) + \varphi (i, t)) \times P_H (i) + \zeta \times P_H (i') \]

Where \( S \) represents Susceptible Individuals, \( I_0 \) Acute infection, \( I_T \) asymptomatic – will be tested before CD4<200, \( I_{TNE} \) Tested ineligible, \( I_{TE} \) Tested eligible, \( I_U \) Tested but unlinked/not on treatment, \( I_S \) Treated and Suppressed, \( I_{NS} \) Treated not Suppressed, \( A_T \) AIDS post-ART, \( I_{NT} \) will not be tested before CD4<200, \( A_{NT} \) AIDS not tested, \( P_L \) and \( P_H \) PrEP-users with low user-adherence (30%) and high user-adherence (90%), respectively.

\( i \) indicates risk group, and \( i' \) corresponding non-IDU risk group.

Note: The equations above represent non-IDU classes. Equations for IDU classes are identical except that \( \psi (i, t) \) and \( \zeta \) switch places.

\( \Lambda \) is the rate of population growth,
\( \mu \) is the rate of leaving sexual activity (differs between IDU and non-IDU, and treated individuals have a slightly increased rate compared to susceptible individuals),
\( \lambda_S, \lambda_D \) are forces of infection (equations below),
\( \psi (t) = \psi_0 \times \exp (-\psi_{red} (t - T_0)) \) is the incidence of injecting drug use (differs between heterosexual men, women and MSM and decays over time in accordance with secular trends in Newark),
\( \zeta \) is the rate of stopping drug use (1/average duration of using drugs), \( i' \) indicates corresponding non-drug-using group
\( \theta \) is the rate of progression from Acute Infection to asymptomatic infection,
\( z(i) \) is the proportion tested before CD4<200 in each risk group,
\( T_{test} \) is average time between progression to asymptomatic phase and testing,
test\text{\_link}, \text{ret}, \text{trt\_link} are the proportion of those tested who are linked to care, proportion of those linked to care retained in care, and proportion of those retained in care linked to treatment, respectively.

\( x_k \) is proportion of patients linked to treatment who successfully achieve viral suppression,

\( \alpha \) is the accelerated rate at which individuals with AIDS leave sexual activity,

\( \gamma_{NM}, \gamma_{NS} \) are the rate of progression from the asymptomatic stage to AIDS stage in untreated and non-adherent treated, respectively,

\( q \) is population-adherence, the proportion of the population with high user-adherence,

\( H \) is the rate of initiation on PrEP,

\( \lambda_{p\text{-low}} \) and \( \lambda_{p\text{-high}} \) are the forces of infection for sexual transmission among PrEP-users with low or high user-adherence to PrEP, and

\( \lambda_{p\text{-drug-low}} \) and \( \lambda_{p\text{-drug-high}} \) are the forces of infection for parenteral transmission due to shared needles among persons who inject drugs (IDU) with low or high user-adherence to PrEP.

**Non-PrEP-User Force of Infection Equations for Sexual and Injection Drug-Use Transmission, \( \lambda_S, \lambda_D \)**

*Force of infection from Sexual Transmission, \( \lambda_S, \)*

\[ \lambda_s (m, i, t) = \sum_j p_s (m, i, j, t) c (m, i) p w_f (j) \times \frac{1}{N (f, j, t)} \times \sum_{\text{stage}} \Omega_{\text{stage } fm} + \]

\[ \sum_j p_s (m, i, j, t) c (m, i) p w_m (j) \times \frac{1}{N (m, j, t)} \times \sum_{\text{stage}} \Omega_{\text{stage } mm} \]

Women
\[ \lambda_s(f, i, t) = \sum_j p_s(f, i, j, t) c(f, i) \times \frac{1}{N(f, j, t)} \times \sum_{\text{stage}} \Omega_{\text{stage mf}} \]

With

\[ \Omega_{\text{stage } kk'} = \left( 1 - (1 - \beta_{kk'} \omega_{\text{stage}} \phi)^{\nu(kk', i, j, t)} \right) \times N_{\text{stage}} \]

Where \( p(i, j, k, t) \) is the term in the sexual contact mixing matrix indicating the probability of an individual of risk group \( i \) interacting with an individual of risk group \( j \) (of opposite gender for females and heterosexual males, and of same gender for homosexual males or men on the down low interacting with male partners), determined as follows:

\[ p_s(k, i, j, t) = \epsilon \delta_{ij} + (1 - \epsilon) \frac{c(k', j) N(k', j, t)}{\sum_s c(k', s) N(k', s, t)} \]

where \( \epsilon \) is the assortativeness coefficient and \( \delta_{ij} \) is the Kronecker delta,

\( c(k, i) \) is partner change rate in gender \( k \), risk group \( i \),

\( pw_k \) is the proportion of men in a risk group that interact with gender \( k \) (e.g. \( pw_f \) for homosexual men is 0, while \( pw_f \) for low risk men is 1-proportion exclusively homosexual).

The \( \Omega \) terms represent weighted prevalence by stage of disease and risk groups interacting, and comprise the terms

\( \beta_{kk'} \) is the weighting of per-sex-act transmission probability depending on which sexes are interacting and which partner is infected,

\( \omega_{\text{stage}} \) is the weighting of transmission probability depending on stages of disease (i.e. viral load weighting),

\( \phi \) is per-sex-act transmission probability,

\( \nu(kk', i, j) \) is the number of unprotected sex acts in a partnership of a \((k,i)\) individual and a \((k',j)\) individual,
$N_{\text{stage}}$ is the number of persons in disease stage.

**Force of Infection from Injecting Drug Use, $\lambda_D$**

\[
\lambda_d(k, i, t) = \sum_j c_d p_d(k, l, i, j) \frac{1}{N(l, j, t)} \sum_{\text{stage}} \Upsilon_{\text{stage}}
\]

\[
\Upsilon_{\text{stage}} = (1 - (1 - \omega_{\text{stage}}\sigma)^\eta) \times N_{\text{stage}}
\]

Where $c$ is the number of other IDUs with whom needles are shared in one year, $p_d(k, l, i, j)$ is the term in the needle-sharing mixing matrix indicating the probability of an individual of gender $k$, risk group $i$ sharing a needle with an individual of gender $l$, risk group $j$.

The IDU mixing matrix is assumed to be non-assortative (random mixing) only weighted by the higher number of male IDU than female. The elements $p_d(k, l, i, j)$ are calculated by assuming 4 male risk groups and 2 female risk groups with contact with the male risk groups being twice as likely. Thus

\[
4p_m + 4p_f = 4(2 \times p_f) + 2p_f = 10p_f = 1 \rightarrow p_f = .1, \quad p_m = .2
\]

The $\Upsilon$ terms, like the $\Omega$ terms, represent weighted prevalence by stage of disease and comprise the terms

- $\omega_{\text{stage}}$ as defined above,
- $\sigma$, the per-injection transmission probability,
- $\eta = \eta_0 \times \exp(\eta_{\text{red}}(t - T_0))$, the number of shared needles per year which, as drug use incidence, decays over time in accordance with secular trends in Newark,
- $N_{\text{stage}}$, as defined above.
PrEP-User Force of Infection Equations for Sexual and Injecting Drug Use Transmission, \( \lambda_{p\text{-low}}, \lambda_{p\text{-high}}, \lambda_{p\text{-drug-low}}, \) and \( \lambda_{p\text{-drug-high}} \)

In the force of infection equations, PrEP is included in the calculation of weighted prevalence, \( \Omega \) for sexual transmission of HIV, and \( \Upsilon \) for injecting drug use transmission according to HIV disease stage, risk-groups interacting, and adherence to PrEP:

\[
\Omega_{stagekk'} = (1 - (1 - \beta_{kk'} \times \omega_{stage} \times \phi) \nu(kk',i,j,t) \times (1 - \text{prep}_{adh\text{-high}}(i)) \times
(1 - (1 - \text{prep}_{eff}(i)) \times \beta_{kk'} \times \omega_{stage} \times \phi) \nu(kk',i,j,t) \times \text{prep}_{adh\text{-high}}(i) \times N_{stage}
\]

\[
\Omega_{stagekk'} = (1 - (1 - \beta_{kk'} \times \omega_{stage} \times \phi) \nu(kk',i,j,t) \times (1 - \text{prep}_{adh\text{-low}}(i)) \times
(1 - (1 - \text{prep}_{eff}(i)) \times \beta_{kk'} \times \omega_{stage} \times \phi) \nu(kk',i,j,t) \times \text{prep}_{adh\text{-low}}(i) \times N_{stage}
\]

\[
\Upsilon_{stage} = (1 - (1 - \omega_{stage} \times \sigma) \eta(j) \times (1 - \text{prep}_{adh\text{-high}}(i)) \times
(1 - (1 - \text{prep}_{eff-IDU}(i)) \times \omega_{stage} \times \sigma) \eta(j) \times \text{prep}_{adh\text{-high}}(i) \times N_{stage}
\]

\[
\Upsilon_{stage} = (1 - (1 - \omega_{stage} \times \sigma) \eta(j) \times (1 - \text{prep}_{adh\text{-low}}(i)) \times
(1 - (1 - \text{prep}_{eff-IDU}(i)) \times \omega_{stage} \times \sigma) \eta(j) \times \text{prep}_{adh\text{-low}}(i) \times N_{stage}
\]

Where \( \text{prep}_{eff} \) is the efficacy of PrEP for sexual transmission, \( \text{prep}_{eff-IDU} \) is the efficacy of PrEP for parenteral (IDU) transmission, and \( \text{prep}_{adh\text{-high}} \) and \( \text{prep}_{adh\text{-low}} \) are the user-adherence levels for PrEP-users in the high and low user-adherence compartments, respectively.
**PrEP Scale-up Equations**

The scale-up equations for PrEP use a linear rate of scale-up \(1/r\) to achieve the desired population coverage level:

for \(t \geq T_{prep}\) and \(t < T_{prep} + r\)

\[
H(i) = \left( \left( \frac{t - T_{prep}}{r} \right) \times PCOV(i) \right)
\]

for \(t \geq T_{prep} + r\)

\[
H(i) = PCOV(i)
\]

Where \(T_{prep}\) is the year PrEP interventions begin (2016 in this model) and PCOV is the population coverage level.

**Population Specifications**

The breakdown of the twelve behavioral risk groups was previously published (1). Briefly, 19% of heterosexual (HET) women are in the female high-risk-heterosexuals (HRH) group; 24% of HET men are in male HRH; 20% of female HET/IDU are in female HRH/IDU; 32% of male IDU are in male HRH/IDU; 4-12% of all males are exclusive men who have sex with men (MSM), with 19-95% engaging in high-risk sexual activity; 16% of HRH males are bisexual (HRH/MSM); MSM/IDU are 0.6% of high-risk MSM, and 0.4% of bisexual males inject drugs (HRH/MSM/IDU).

**User-Adherence Assumptions**

We assumed the half-life of TDF-FTC was three days (135, 136, 274), a full dose of TDF-FTC achieved a TFV-DP (tenofovir is phosphorylated within cells to TFV-DP) concentration in peripheral blood mononuclear cells (PBMCs) of 36-42 fmol/M (132, 137), and a TFV-DP concentration of 16 fmol/M corresponded to a 90% risk reduction (132); resulting in at least a 90% risk reduction by the third day when
the concentration drops to 18-21 fmol/M. Therefore, for a high adherer taking at least three doses per week, five out of seven days (71%) or more of daily sex acts are protected with greater than 90% efficacy; four doses per week covers six out of seven days (86%); five, six, or seven doses per week provides 100% coverage. We averaged these values to get the high adherence value, which rounded to 90%. We followed the same process to calculate the low adherence value, which rounded to 30%. For parenteral transmission in IDU, we assume daily injections and the same adherence values.

**Population-Specific Drug Efficacies**

For MSM, a TFV-DP drug concentration in PBMCs of 42 fmol/M corresponding to daily dosing (i.e. perfect user-adherence) was predicted to have a 99% risk reduction in the iPrEx study population (132). We assumed a slightly lower efficacy of 97% since a separate pharmacokinetic study found the mean steady-state concentration of TFV-DP in PBMCs for daily dosing to be 36 fmol/M in a heterogeneous population (137). For HET, drug efficacy is 91% to match the reported risk reduction of 91% in a daily dosing cohort of trial participants whose TFV plasma concentrations were greater than 40 ng/mL (133). We assumed this efficacy and understood it to be conservative since TFV plasma concentrations of 40 ng/mL likely do not represent perfect adherence when other pharmacokinetic studies have found the mean steady-state concentration to be between 52 and 54 ng/mL in pharmacokinetic studies (135, 137). For bisexual males, efficacy is 94%, the average of the efficacy values assumed above for heterosexual transmission (91%) and MSM transmission (97%). For parenteral transmission in IDU, the 83% efficacy is the Cox regression estimate by Martin et al of risk reduction when daily adherence is greater than or equal to 97.5% from adherence-effectiveness data from the Bangkok Tenofovir Study (113, 134).
Our assumptions and estimates here are conservative, allowing for refinements later. Even in the absolute best-case scenario modeled where 95% of the population is highly adherent (i.e. user-adherence is 90%, corresponding to three or more doses per week), 5% of the population still has a low user-adherence of 30%. The average user-adherence for the population is then 87%, which corresponds roughly to four doses per week and is in keeping with the Ipergay study in which participants took a median of 16 pills per month (275). The effectiveness of PrEP in this population is calculated to be 84% (average user-adherence x drug efficacy of 97% for MSM), which is consistent with the 86% risk reduction found in the recent Ipergay study (275).

Supplementary Discussion

Standardization for Comparative Analysis of Results

We reviewed the literature for studies modeling PrEP in populations of MSM and the general population. In order to better compare and contrast results, we strove to standardize the metrics of effectiveness, coverage, HIV prevalence, and percent reduction in cumulative infections by PrEP so that these metrics were with respect to the same population level (i.e. impact was reported as percent reduction in HIV incidence with respect to the MSM population as opposed to with respect to the high-risk MSM population or the general population) (Table B.3). For modeling studies that did not report a functional effectiveness, we calculated overall effectiveness for comparison’s sake as a summary metric of efficacy and adherence: 

\[(\text{high-adherence efficacy} \times \text{population adherence}) + (\text{low-adherence efficacy} \times (1-\text{population adherence}))\].

Modeling studies looking at the impact of PrEP in MSM communities included Desai et al’s study (121) of high-risk MSM in New York City, Juusola et al’s model (120) of PrEP delivered to MSM in the United States, and Gomez et al’s study (131) of PrEP delivered to men who mostly have sex with men (MMSM) in Lima, Peru. From
Desai et al’s interventions, we highlighted two of their targeting strategies: basic, where efficacy is either 50% or 70% if adherent and 0% otherwise, and adherence-dependent, where efficacy is higher if fully adherent and reduced if partially adherent either at 50%/30% or 70%/50%. Desai et al add another layer to this scheme with three levels of population-adherence: 95% (high), 50% (medium), and 33% (low). Coverage is 25% of high-risk MSM. In Gomez et al’s study, approximately 2% of MMSM are covered, and efficacy and adherence are assumed to reflect the effectiveness profile found in the iPrEx study. Juusola et al’s model of delivering PrEP to 10% of MSM in the United States also assumed the iPrEx effectiveness.

Modeling studies of PrEP in the general population included Abbas et al’s model (123) of PrEP among a population of sub-Saharan African heterosexuals where PrEP delivery was either non-targeted or targeted to the two highest sexual activity classes. Optimistic, neutral, and pessimistic levels of intervention corresponded to combinations of assumed coverage and effectiveness. In two scenarios from Cremin et al’s model (119) of PrEP in KwaZulu-Natal, South Africa, PrEP is delivered either to 7.3% of 15-24 year olds or 4.4% of 15-54 year olds, and efficacy and adherence is 75% and 95%, respectively.

### Table B.1: Epidemiological Profile of the HIV/AIDS Epidemic in Newark in 2016

<table>
<thead>
<tr>
<th>Measurement</th>
<th>HRH (± SE)</th>
<th>MSM (± SE)</th>
<th>IDU (± SE)</th>
<th>All High-Risk Groups (AHR) (± SE)</th>
<th>Total Population (TOT) (± SE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incidence (%)</td>
<td>0.26 (±0.00)</td>
<td>2.22 (±0.06)</td>
<td>23.95 (±1.36)</td>
<td>0.66 (±0.01)</td>
<td>0.25 (±0.00)</td>
</tr>
<tr>
<td>Incident Cases (N)</td>
<td>128 (±1)</td>
<td>75 (±0)</td>
<td>206 (±3)</td>
<td>353 (±3)</td>
<td>526 (±3)</td>
</tr>
<tr>
<td>Prevalence (%)</td>
<td>3.13 (±0.01)</td>
<td>16.37 (±0.35)</td>
<td>39.13 (±0.35)</td>
<td>5.58 (±0.02)</td>
<td>2.77 (±0.01)</td>
</tr>
<tr>
<td>Diagnosed Prevalent Cases (N)</td>
<td>1,619 (±3)</td>
<td>746 (±1)</td>
<td>1,263 (±5)</td>
<td>3,268 (±6)</td>
<td>6,166 (±9)</td>
</tr>
<tr>
<td>Susceptible Population (N)</td>
<td>49,099 (±41)</td>
<td>3,319 (±30)</td>
<td>857 (±5)</td>
<td>53,053 (±75)</td>
<td>212,660 (±170)</td>
</tr>
<tr>
<td>Total Population (N)</td>
<td>51,658 (±32)</td>
<td>4,557 (±34)</td>
<td>3,228 (±11)</td>
<td>58,551 (±57)</td>
<td>222,560 (±148)</td>
</tr>
</tbody>
</table>

1These values are predicted by the model for 2016 prior to any PrEP intervention.
2Prevalent cases + susceptibles do not equal total population size since undiagnosed prevalent cases are unaccounted for.
Table B.2: Number Needed to Treat (NNT) with PrEP to Avert One HIV Infection in 2026 and Number of Cumulative PrEP Person-Years Needed to Avert One HIV Infection 2016-2026

<table>
<thead>
<tr>
<th>PrEP Intervention</th>
<th>Targeted Population: Coverage (%) / Population-Adherence (%)</th>
<th>NNT for 2026 (N)</th>
<th>95% Confidence Interval (N)</th>
<th>Number of Cumulative PrEP Person-Years Needed for 2016-2026 (Person-Years)</th>
<th>95% Confidence Interval (Person-Years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOT: 10/15</td>
<td></td>
<td>1185</td>
<td>635-3730</td>
<td>1072</td>
<td>597-3152</td>
</tr>
<tr>
<td>TOT: 25/15</td>
<td></td>
<td>1278</td>
<td>691-4090</td>
<td>1131</td>
<td>635-3306</td>
</tr>
<tr>
<td>TOT: 50/15</td>
<td></td>
<td>1307</td>
<td>706-4280</td>
<td>1170</td>
<td>657-4442</td>
</tr>
<tr>
<td>TOT: 10/60</td>
<td></td>
<td>702</td>
<td>389-2159</td>
<td>621</td>
<td>354-1791</td>
</tr>
<tr>
<td>TOT: 25/60</td>
<td></td>
<td>783</td>
<td>437-2449</td>
<td>672</td>
<td>388-1906</td>
</tr>
<tr>
<td>TOT: 50/60</td>
<td></td>
<td>812</td>
<td>454-2561</td>
<td>705</td>
<td>409-1997</td>
</tr>
<tr>
<td>TOT: 10/95</td>
<td></td>
<td>549</td>
<td>310-1669</td>
<td>477</td>
<td>274-1355</td>
</tr>
<tr>
<td>TOT: 25/95</td>
<td></td>
<td>626</td>
<td>362-1907</td>
<td>535</td>
<td>309-1457</td>
</tr>
<tr>
<td>TOT: 50/95</td>
<td></td>
<td>657</td>
<td>384-2005</td>
<td>556</td>
<td>335-1539</td>
</tr>
<tr>
<td>AHR: 10/15</td>
<td></td>
<td>496</td>
<td>210-3410</td>
<td>449</td>
<td>204-2319</td>
</tr>
<tr>
<td>AHR: 25/15</td>
<td></td>
<td>517</td>
<td>227-3628</td>
<td>465</td>
<td>218-2418</td>
</tr>
<tr>
<td>AHR: 50/15</td>
<td></td>
<td>512</td>
<td>230-3495</td>
<td>471</td>
<td>226-2378</td>
</tr>
<tr>
<td>AHR: 10/60</td>
<td></td>
<td>290</td>
<td>132-1703</td>
<td>257</td>
<td>123-1209</td>
</tr>
<tr>
<td>AHR: 25/60</td>
<td></td>
<td>309</td>
<td>156-1771</td>
<td>272</td>
<td>136-1170</td>
</tr>
<tr>
<td>AHR: 50/60</td>
<td></td>
<td>307</td>
<td>156-1763</td>
<td>277</td>
<td>144-1157</td>
</tr>
<tr>
<td>AHR: 10/95</td>
<td></td>
<td>225</td>
<td>107-1223</td>
<td>197</td>
<td>96-861</td>
</tr>
<tr>
<td>AHR: 25/95</td>
<td></td>
<td>244</td>
<td>125-1264</td>
<td>210</td>
<td>110-849</td>
</tr>
<tr>
<td>AHR: 50/95</td>
<td></td>
<td>243</td>
<td>132-1236</td>
<td>216</td>
<td>118-857</td>
</tr>
<tr>
<td>HRH: 10/15</td>
<td></td>
<td>1058</td>
<td>236-4424</td>
<td>973</td>
<td>222-3509</td>
</tr>
<tr>
<td>HRH: 25/15</td>
<td></td>
<td>1108</td>
<td>258-4617</td>
<td>1005</td>
<td>238-3672</td>
</tr>
<tr>
<td>HRH: 50/15</td>
<td></td>
<td>1121</td>
<td>265-4844</td>
<td>1027</td>
<td>250-3821</td>
</tr>
<tr>
<td>HRH: 10/60</td>
<td></td>
<td>599</td>
<td>149-3116</td>
<td>549</td>
<td>135-1859</td>
</tr>
<tr>
<td>HRH: 25/60</td>
<td></td>
<td>633</td>
<td>173-2560</td>
<td>573</td>
<td>151-1957</td>
</tr>
<tr>
<td>HRH: 50/60</td>
<td></td>
<td>642</td>
<td>182-2717</td>
<td>588</td>
<td>162-2046</td>
</tr>
<tr>
<td>HRH: 10/95</td>
<td></td>
<td>453</td>
<td>121-1718</td>
<td>414</td>
<td>106-1365</td>
</tr>
<tr>
<td>HRH: 25/95</td>
<td></td>
<td>484</td>
<td>145-1912</td>
<td>435</td>
<td>122-1441</td>
</tr>
<tr>
<td>HRH: 50/95</td>
<td></td>
<td>492</td>
<td>156-2041</td>
<td>449</td>
<td>133-1509</td>
</tr>
<tr>
<td>MSM: 10/15</td>
<td></td>
<td>56</td>
<td>25-1401</td>
<td>56</td>
<td>22-1208</td>
</tr>
<tr>
<td>MSM: 25/15</td>
<td></td>
<td>59</td>
<td>26-1551</td>
<td>60</td>
<td>23-1289</td>
</tr>
<tr>
<td>MSM: 50/15</td>
<td></td>
<td>60</td>
<td>26-1646</td>
<td>62</td>
<td>24-1359</td>
</tr>
<tr>
<td>MSM: 10/60</td>
<td></td>
<td>35</td>
<td>15-811</td>
<td>34</td>
<td>13-659</td>
</tr>
<tr>
<td>MSM: 25/60</td>
<td></td>
<td>39</td>
<td>16-931</td>
<td>38</td>
<td>15-716</td>
</tr>
<tr>
<td>MSM: 50/60</td>
<td></td>
<td>40</td>
<td>17-992</td>
<td>40</td>
<td>17-766</td>
</tr>
<tr>
<td>MSM: 10/95</td>
<td></td>
<td>28</td>
<td>12-624</td>
<td>27</td>
<td>11-494</td>
</tr>
<tr>
<td>MSM: 25/95</td>
<td></td>
<td>33</td>
<td>14-728</td>
<td>30</td>
<td>13-541</td>
</tr>
<tr>
<td>MSM: 50/95</td>
<td></td>
<td>34</td>
<td>15-790</td>
<td>33</td>
<td>14-583</td>
</tr>
<tr>
<td>IDU: 10/15</td>
<td></td>
<td>37</td>
<td>10-4169</td>
<td>27</td>
<td>9-1044</td>
</tr>
<tr>
<td>IDU: 25/15</td>
<td></td>
<td>39</td>
<td>11-3619</td>
<td>28</td>
<td>9-951</td>
</tr>
<tr>
<td>IDU: 50/15</td>
<td></td>
<td>40</td>
<td>12-3071</td>
<td>29</td>
<td>10-850</td>
</tr>
<tr>
<td>IDU: 10/60</td>
<td></td>
<td>19</td>
<td>6-1956</td>
<td>14</td>
<td>5-464</td>
</tr>
<tr>
<td>IDU: 25/60</td>
<td></td>
<td>20</td>
<td>7-1679</td>
<td>14</td>
<td>6-424</td>
</tr>
<tr>
<td>IDU: 50/60</td>
<td></td>
<td>20</td>
<td>8-1435</td>
<td>15</td>
<td>6-380</td>
</tr>
<tr>
<td>IDU: 10/95</td>
<td></td>
<td>14</td>
<td>5-1444</td>
<td>11</td>
<td>4-338</td>
</tr>
<tr>
<td>IDU: 25/95</td>
<td></td>
<td>15</td>
<td>6-1253</td>
<td>11</td>
<td>5-310</td>
</tr>
<tr>
<td>IDU: 50/95</td>
<td></td>
<td>16</td>
<td>7-1078</td>
<td>11</td>
<td>5-279</td>
</tr>
</tbody>
</table>

3 The NNT for 2026 is the number of current PrEP-users required to prevent one incident HIV infections in 2026. For cumulative PrEP person-years needed to avert 1 HIV infection over a 10-year period (2016-2026), the total number of persons on PrEP during that time period was divided by the cumulative infections averted for 2016-2026. We use the term “person-years” since the total number of persons on PrEP in that 10-year period does not distinguish between one individual starting PrEP in 2016 and continuing for 10 years and two individuals starting PrEP in 2021 and continuing for 5 years.
Table B.3: Literature Review of Standardized Metrics for Comparative Analysis of Results

<table>
<thead>
<tr>
<th>Population-Adherence (Proportion of the Population Adherent/Highly Adherent) (%)</th>
<th>33</th>
<th>50</th>
<th>95</th>
<th>33</th>
<th>50</th>
<th>95</th>
<th>33</th>
<th>50</th>
<th>95</th>
<th>95</th>
<th>95</th>
</tr>
</thead>
<tbody>
<tr>
<td>PrEP Efficacy for Low/Partial Adherence (%)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>30</td>
<td>30</td>
<td>50</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>PrEP Efficacy for High/Full Adherence (%)</td>
<td>50</td>
<td>50</td>
<td>70</td>
<td>70</td>
<td>50</td>
<td>50</td>
<td>70</td>
<td>70</td>
<td>92</td>
<td>75</td>
<td>75</td>
</tr>
<tr>
<td>PrEP Overall Effectiveness (%)</td>
<td>17</td>
<td>25</td>
<td>48</td>
<td>23</td>
<td>35</td>
<td>67</td>
<td>37</td>
<td>49</td>
<td>57</td>
<td>60</td>
<td>69</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Targeted Population</th>
<th>Very High-Risk MSM (30% of total MSM population)</th>
<th>MMSM</th>
<th>MSM</th>
<th>15-24 year-olds</th>
<th>15-54 year-olds</th>
<th>15-49 year-olds</th>
<th>2 highest-risk sexual activity groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>Location</td>
<td>New York City, USA</td>
<td>Lima, Peru, USA</td>
<td>KwaZulu-Natal, Sub-Saharan Africa</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>PrEP Coverage of Targeted Population (%)</td>
<td>25</td>
<td>75</td>
<td>50</td>
</tr>
<tr>
<td>PrEP Coverage of Reference Population (%)</td>
<td>5.2</td>
<td>2.1</td>
<td>10.0</td>
</tr>
</tbody>
</table>

| Impact Reported as 10-Year Cumulative Incidence Reduction (%) | 5.4 | 8.7 | 14.3 | 7.0 | 11.3 | 20.0 | 14.8 | 15.8 | 16.6 | 21.9 | 22.5 | 23.1 | 1.2 | 74.4 |

| HIV Prevalence in Reference Population (%) | 14.6 | 14.6 | 14.6 | 14.6 | 14.6 | 14.6 | 14.6 | 14.6 | 14.6 | 17.9 | 12.3 | 22 | 22 | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 |

Note: The highlighted value indicates an anomaly or an atypical result.
Figure B.1: **Reported impact of PrEP in populations from different modeling studies.** Impact, shown by the size of the circle, is the percent change in cumulative infections averted in the population of interest for the various interventions in each study. Panel A: PrEP interventions targeting MSM from different modeling studies on the basis of HIV prevalence in the MSM population, effectiveness, coverage, and resulting impact. Panel B: PrEP modeling studies of the total population. Abbreviations: HRH= high-risk heterosexuals, MSM= men who have sex with men, MMSM= men who mostly have sex with men, IDU= injection drug users, AHR= all high-risk groups, TOT= total population.
Appendix C

Chapter 3 Supplementary Information

Parameter and Data Estimates

Table C.1: Parameter ranges from DebRoy et al 2011 (1)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Units</th>
<th>Min</th>
<th>Max</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_{max}$</td>
<td>Maximum Hepatocyte count (cells/mL) *</td>
<td></td>
<td>$4 \times 10^7$</td>
<td>$1.3 \times 10^7$</td>
<td>(187)</td>
</tr>
<tr>
<td>$d$</td>
<td>Natural Hepatocyte death rate (day$^{-1}$)</td>
<td></td>
<td>$1 \times 10^{-3}$</td>
<td>$1.4 \times 10^{-3}$</td>
<td>(187)</td>
</tr>
<tr>
<td>$s$</td>
<td>Natural Hepatocyte production rate (cells/mL/day)</td>
<td></td>
<td>$1 \times d \times T_{max}$</td>
<td>$1 \times T_{max}$</td>
<td>(187)</td>
</tr>
<tr>
<td>$r_1$</td>
<td>Natural Hepatocyte proliferation rate (day$^{-1}$)</td>
<td></td>
<td>.1</td>
<td>3</td>
<td>(187)</td>
</tr>
<tr>
<td>$\beta_c$</td>
<td>Infection rate (mL/virion/day)</td>
<td></td>
<td>$1 \times 10^{-8}$</td>
<td>$1 \times 10^{-6}$</td>
<td>(187)</td>
</tr>
<tr>
<td>$\delta$</td>
<td>Clearance rate of Infected Hepatocytes (day$^{-1}$)</td>
<td></td>
<td>$3/(1 + \alpha \times H_0)$</td>
<td>$3/(1 + \alpha \times H_0)$</td>
<td>(187)</td>
</tr>
<tr>
<td>$r_2$</td>
<td>Infected Hepatocyte proliferation rate (day$^{-1}$)</td>
<td></td>
<td>1</td>
<td>10</td>
<td>(188)</td>
</tr>
<tr>
<td>$p$</td>
<td>Viral burst size (virions/day)</td>
<td></td>
<td>.1</td>
<td>45</td>
<td>(188)</td>
</tr>
<tr>
<td>$c$</td>
<td>Viral clearance rate (day$^{-1}$)</td>
<td></td>
<td>.8</td>
<td>22</td>
<td>(187)</td>
</tr>
</tbody>
</table>

*This is a calculation made in order for hepatocyte count to correspond to HCV RNA load, which is measured per mL of blood. There is an estimated 15L of blood in an average 70kg human, and the liver is estimated to contain $10^{-11} \log_{10}$ hepatocytes by (276, 277) and others, so the corresponding number hepatocytes to 1 mL of blood is on the order of $10^6 - 10^7$. 191
Table C.2: Data Estimates

<table>
<thead>
<tr>
<th>Parameter</th>
<th>HIV viral load</th>
<th>Impact of HIV coinfection HCV</th>
<th>Impact of HIV treatment on HCV Viral Load</th>
</tr>
</thead>
<tbody>
<tr>
<td>HCV viral load</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HIV neg: 7.0-7.2 log\text{10} RNA copies/mL</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HIV pos: 6.8-7.2 log\text{10} RNA copies/mL</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HIV neg: 6.6 log\text{10} eq/mL, HIV pos: 7.0 log\text{10} eq/mL</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HIV neg: 6.1 log\text{10} eq/mL, HIV pos: 6.4 log\text{10} eq/mL</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Compiled from (155), (178), (279):
HCV VL decreased 0.2-0.3 logs over long-term ART
Several studies showed no difference w.r.t HIV status, or only showed difference with CD4<sup>+</sup> < 200

Table C.3: Parameter sets for Figs 2-6

<table>
<thead>
<tr>
<th>Parameter</th>
<th>SVR (also Patient II)</th>
<th>Null Responder</th>
<th>Partial Responder</th>
<th>Breakthrough</th>
<th>Relapse</th>
<th>Biphasic Decline 1</th>
<th>Biphasic Decline 2</th>
<th>Patient 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s$</td>
<td>9199</td>
<td>1.16×10&lt;sup&gt;4&lt;/sup&gt;</td>
<td>903</td>
<td>8096</td>
<td>8096</td>
<td>1.43×10&lt;sup&gt;1&lt;/sup&gt;</td>
<td>6.38×10&lt;sup&gt;2&lt;/sup&gt;</td>
<td>4635</td>
</tr>
<tr>
<td>$r_1$</td>
<td>2.6</td>
<td>0.34</td>
<td>1.7</td>
<td>2.88</td>
<td>2.88</td>
<td>2.09</td>
<td>2.07</td>
<td>2.7</td>
</tr>
<tr>
<td>$c$</td>
<td>7.27</td>
<td>16.9</td>
<td>16.1</td>
<td>19.52</td>
<td>19.52</td>
<td>23.5</td>
<td>9.75</td>
<td>10.06</td>
</tr>
<tr>
<td>$\beta_c$</td>
<td>6.09×10&lt;sup&gt;-7&lt;/sup&gt;</td>
<td>9.60×10&lt;sup&gt;-7&lt;/sup&gt;</td>
<td>1.25×10&lt;sup&gt;-3&lt;/sup&gt;</td>
<td>1.01×10&lt;sup&gt;-3&lt;/sup&gt;</td>
<td>1.01×10&lt;sup&gt;-3&lt;/sup&gt;</td>
<td>1.11×10&lt;sup&gt;-3&lt;/sup&gt;</td>
<td>1.37×10&lt;sup&gt;-3&lt;/sup&gt;</td>
<td>1.06×10&lt;sup&gt;-3&lt;/sup&gt;</td>
</tr>
<tr>
<td>$\delta$</td>
<td>4.35</td>
<td>1.80</td>
<td>1.77</td>
<td>4.06</td>
<td>4.06</td>
<td>2.09</td>
<td>1.30</td>
<td>7.52</td>
</tr>
<tr>
<td>$\epsilon$</td>
<td>77%</td>
<td>50%</td>
<td>63%</td>
<td>59%</td>
<td>60%</td>
<td>78%</td>
<td>63%</td>
<td></td>
</tr>
</tbody>
</table>

Incomplete CD4<sup>+</sup> Recovery

Figure C.1: CD4<sup>+</sup> counts over time. This figure, re-drawn from (6), demonstrates how CD4<sup>+</sup> after ART initiation can be incomplete. It is strongly associated with CD4<sup>+</sup> count at treatment initiation, p<.05 for all comparisons.
Backward bifurcation and bistability

This bistability phenomenon is an example of backward bifurcation. Following a model of HTLV proposed by \cite{190}, and closely adapting their methods to this study we can calculate the conditions for bistability.

\( H \) acts as a constant in this system at equilibrium (and its equilibrium value can be determined by HIV viral load or previous HIV infection). At equilibrium, \( T, I \) satisfy the equations:

\[
0 = s + r_1 T \left(1 - \frac{T + I}{T_{\text{max}}}\right) - dT - \tilde{\beta}TI \quad (C.1)
\]

\[
0 = \tilde{\beta}TI + r_2 I \left(1 - \frac{T + I}{T_{\text{max}}}\right) - \delta(1 + \alpha H)I \quad (C.2)
\]

At the disease-free equilibrium, \( I = 0 \) so equation (1) becomes

\[
0 = s + r_1 T_0 \left(1 - \frac{T_0}{T_{\text{max}}}\right) - dT_0
\]

which implies that \( T_0 \) is a solution to the equation

\[
f_1(T) = 0
\]

where

\[
f_1(T) = s + (r_1 - d)T - \frac{r_1}{T_{\text{max}}} T^2
\]

At the endemic equilibrium (EE), \( I > 0 \) so can be canceled from equation (2), yielding

\[
0 = \tilde{\beta}T + r_2 \left(1 - \frac{T + I}{T_{\text{max}}}\right) - \delta(1 + \alpha H)
\]
which implies that

\[ I^* = \frac{T_{\text{max}}}{r_2} \left[ (\tilde{\beta} - \frac{r_2}{T_{\text{max}}})T + r_2 - \delta(1 + \alpha H) \right] \]

Plugging this value back into equation (1) yields

\[ 0 = s + r_1 T \left( 1 - \frac{T + \frac{T_{\text{max}}}{r_2} \left[ (\tilde{\beta} - \frac{r_2}{T_{\text{max}}})T + r_2 - \delta(1 + \alpha H) \right]}{T_{\text{max}}} \right) - dT \]

\[ -\tilde{\beta} T \frac{T_{\text{max}}}{r_2} \left[ (\tilde{\beta} - \frac{r_2}{T_{\text{max}}})T + r_2 - \delta(1 + \alpha H) \right] \]

Isolating \( f_1 \) from the above equation gives the equality

\[ f_1(T^*) = f_2(T^*) \]

where

\[ f_2(T) = \left( \tilde{\beta} T_{\text{max}} + r_1 \right) \left( \frac{1}{r_2} \right) \left[ (\tilde{\beta} - \frac{r_2}{T_{\text{max}}})T + r_2 - \delta(1 + \alpha H) \right] T \]

Where those two functions intersect, there will be endemic equilibria. If they intersect twice, there will be two equilibria - one stable and one unstable. It can be shown that \( f_1(T) \) is necessarily a concave-down parabola with exactly one positive root, because \( -\frac{r_1}{T_{\text{max}}} < 0 \), and \( \frac{r_1\delta}{T_{\text{max}}} > 0 \). \( f_2(T) \) is a parabola with one root at 0, so in order for it to intersect with \( f_1 \) twice, it must be concave down with its other root being positive.

For \( f_2 \) to be concave down,

\[ \tilde{\beta} - \frac{r_2}{T_{\text{max}}} < 0 \implies r_2 > \tilde{\beta} T_{\text{max}} \]
For the non-zero root to be positive the following must also be true:

\[ r_2 - \delta(1 + \alpha H) > 0 \rightarrow r_2 > \delta(1 + \alpha H) \]

It can be seen here that if \( H \) shrinks (due to HIV infection), the value which \( r_2 \) must exceed also shrinks thereby allowing values of \( r_2 \) that previously would have allowed spontaneous clearance to allow instead for bistability.

Calculation of \( R_C \):
\( R_C \) is reached when the two functions are tangent, which occurs when the vertex of of \( f_2 \) falls on \( f_1 \). The vertex of \( f_2 \) occurs at

\[ T_C = \left( \beta T_{\text{max}} + r_1 \right) \left( \frac{1}{r_2} \right) \times \frac{\delta(1 + \alpha H) - r_2}{2 \times \left( \beta - \frac{r_2}{T_{\text{max}}} \right)} \]

so \( R_C \) occurs when \( T^* = T_C \).

**Cycles**

Equations of this format can have areas of parameter space that generate periodic behavior. This can have interesting consequences in systems with natural noise. In simulations not shown here, we have tested various parameter sets that demonstrate periodic behavior by adding noise or pushing the system far from the equilibrium, and found that noise does not alter the longterm dynamics qualitatively. However, this may be in interesting area for further exploration in this model.
Bifurcation calculation from simplified equations

By transforming the equations to be dimensionless, and removing the recruitment rate of healthy hepatocytes, we can delve into the bifurcation criteria at a more basic level.

Starting with the original equations:

\[
\frac{dT}{dt} = s + r_1 T \left( 1 - \frac{T + I}{T_{\text{max}}} \right) - dT - \tilde{\beta}TI \\
\frac{dI}{dt} = \tilde{\beta}TI + r_2 I \left( 1 - \frac{T + I}{T_{\text{max}}} \right) - \delta(1 + \alpha H)I
\]

we remove \(s\), transform time into a dimensionless quantity using \(t' = r_1 t\), and solve for a new set of state variables \(x = \frac{T}{T_{\text{max}}} \), \(y = \frac{I}{T_{\text{max}}} \). We thus have:

\[
\frac{dx}{dt'} = \frac{dT/T_{\text{max}}}{dr_1 t} = x(1 - x - y) - d_1 x - bxy
\]

and

\[
\frac{dy}{dt'} = \frac{dI/T_{\text{max}}}{dr_1 t} = bxy + ry(1 - x - y) - cy
\]

where

\[
d_1 = \frac{d}{r_1}, b = \frac{\tilde{\beta} \times T_{\text{max}}}{r_1}, r = \frac{r_2}{r_1}, c = \frac{\delta(1 + \alpha H)}{r_1}
\]

This system has four equilibria, one of which does not have meaning for this infection system, as it represents total hepatocyte count of zero, \([0,0]\). The others represent total infection: \([0,1 - \frac{c}{r}]\), Disease-Free equilibrium: \([1 - d,0]\), and endemic equilibrium:

\[
\left[ \frac{c + bc - br - dr \cdot (b - c - bd + dr)}{(b + 1)(b + r)} \right], \frac{(b + 1)(b + r)}{(b + 1)(b + r)}
\]
In order to calculate the conditions for invasion at the disease-free equilibrium, it is necessary to calculate the Jacobian and solve for both the left and right eigenvectors. Following (189), we want to explore the conditions under which the infection can invade when the within-host $R_0 = 1$. We do this by assuming the infection encroaches a small amount along the dominant eigenvector, calculating the Jacobian vector field along the dominant eigenvector close to the DFE, and multiplying by the dominant left eigenvector to assess if the directionality is positive or negative. If positive, then the infection can invade at the DFE, if negative, it cannot.

The first step is to assess the conditions for stability of the DFE. The Jacobian is as follows:

$$J = \begin{bmatrix} (1 - 2x - y) - by - d_1 & -x(1 + b) \\ 0 & r(1 - x - 2y) + bx - c \end{bmatrix}$$

At the disease-free equilibrium,

$$J_{DFE} = \begin{bmatrix} (1 - 2(1 - d_1)) - d_1 & -(1 - d_1)(1 + b) \\ 0 & r(1 - (1 - d_1)) + b(1 - d_1) - c \end{bmatrix} = \begin{bmatrix} -(1 - d_1) & -(1 - d_1)(1 + b) \\ 0 & rd_1 + b(1 - d_1) - c \end{bmatrix}$$

Bifurcation occurs when the bottom left corner of the matrix is equal to zero, so

$$c = b(1 - d_1) + rd_1$$

The DFE is thus stable (within-host $R_0 = 1$ when

$$c \geq b(1 - d_1) + rd_1$$
(because the upper lefthand term in the matrix is $< 0$). HIV coinfection will change this stability: recall that

$$c = \frac{\delta(1 + \alpha H)}{r_1}$$

So in order for

$$c \geq b(1 - d_1) + rd_1$$

we must have

$$\frac{\delta(1 + \alpha H)}{r_1} \geq b(1 - d_1) + rd_1$$

plugging in the original values, we get

$$\frac{\delta(1 + \alpha H)}{r_1} \geq \tilde{\beta} \times T_{\text{max}} \left(1 - \frac{d}{r_1}\right) + r_2 \frac{d}{r_1}$$

$$\delta(1 + \alpha H) \geq \tilde{\beta} \times T_{\text{max}}(1 - \frac{d}{r_1}) + r_2 \frac{d}{r_1}$$

$$H \geq \frac{\tilde{\beta} \times T_{\text{max}}(1 - \frac{d}{r_1}) + r_2 \frac{d}{r_1}}{\delta \alpha} - \delta$$

where $H$ is the immune system contribution. If $H$ drops below this critical value, then the DFE is no longer stable. HIV will cause $H$ to drop, changing the stability of the system.

Assuming the stability criterion is fulfilled, $c = b(1 - d_1) + rd_1$, the Jacobian evaluated at the DFE is equal to

$$\begin{bmatrix}
-\frac{(1 - d_1)}{r_1} & -\frac{(1 - d_1)(1 + b)}{r_1} \\
0 & 0
\end{bmatrix}$$
At the bifurcation point, the dominant eigenvalue of this matrix will be zero, so the dominant right eigenvector $\vec{v}$ must fulfill the equation $J\vec{v} = 0$. Therefore

$$\vec{v} = \begin{pmatrix} -(1 + b) \\ 1 \end{pmatrix}$$

The dominant left eigenvector, $\vec{w}$, similarly, must fulfill the equation $\vec{w}J = 0$. Therefore

$$\vec{w} = \begin{pmatrix} 0 & 1 \end{pmatrix}$$

To obtain the Jacobian vector field along the direction of the dominant right eigenvector, we must obtain the $x$ and $y$ directional components of the Jacobian by taking the second derivatives, and multiply them by $\vec{v}$.

$$J_x = \begin{pmatrix} -2 & -(1 + b) \\ 0 & b - r \end{pmatrix}$$

$$J_y = \begin{pmatrix} -(1 + b) & 0 \\ b - r & -2r \end{pmatrix}$$

so

$$J_v = \begin{pmatrix} (1 + b) & (1 + b)^2 \\ b - r & (r - b)(1 + b) - 2r \end{pmatrix}$$

The next step is to calculate the vector field displacement along the dominant eigenvector:

$$J_v\vec{v} = \begin{pmatrix} (1 + b) & (1 + b)^2 \\ b - r & (r - b)(1 + b) - 2r \end{pmatrix} \times \begin{pmatrix} -(1 + b) \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 2((r - b)(1 + b) - r) \end{pmatrix}$$
To assess if the directionality is positive or negative, we multiply by the left eigenvector, \( \vec{w} = \begin{pmatrix} 0 & 1 \end{pmatrix} \), obtaining \( 2((r-b)(1+b)-r) \). The sign of this expression is what will determine invasibility, so we can simplify.

\[
(r - b)(1 + b) - r = r + br - b - b^2 - r = b(r - 1 - b)
\]

So if \( r > 1 + b \), the infection can invade close to the DFE. \( b > 1 \) by definition, so \( r > 1 \) implies that \( \frac{r}{r_1} > 1 \) which means that, for invasion to occur, \( r_2 \) must be greater than \( r_1 \), and in fact greater than \( (1 + b) \times r_1 \).
Appendix D

Chapter 4 Supplementary Information

Example CTMC Transition Probabilities and Generator Matrix for Non-Spatial Neutral Model

Let $\vec{X}(t) = [T(t), E(t), EX(t)I(t), D_T(t), D_I(t)]$, so the probability of state change is defined as $\text{Prob}\{\Delta \vec{X}(t) = (a, b, c, d, e, f)|\vec{X}(t)\}$, where $a$ represents the change in state $T$, $b$ in state $E$, etc. We assume that the time step is small enough that each change is only in one cell, so $a - f$ can only take the values 0 or ±1. The transition probabilities are as follows.
The generator matrix $Q$ derived from these transition probabilities is thus as follows

$$Q = \begin{bmatrix}
0 & (1 - \eta)(\lambda_{\text{virions}} + \lambda_{\text{local}})T & \eta(\lambda_{\text{virions}} + \lambda_{\text{local}})T & 0 & \nu_T T & 0 \\
0 & 0 & \alpha E & 0 & \nu_T E & 0 \\
0 & 0 & \alpha_X EX & 0 & \nu_T E & 0 \\
0 & 0 & 0 & 0 & 0 & \nu_I I \\
\phi_{DT} DT & 0 & 0 & 0 & 0 & 0 \\
(1 - \kappa)\phi_{DI} DI & 0 & 0 & \kappa\phi_{DI} & 0 & 0 \\
\end{bmatrix}$$
Cell Population Trajectory Examples for Non-null Models

Latent Non-Spatial, 20%

Two Patch Spatial, Cell-Dominated, 25%

Selective Sweep Spatial, Cell-Dominated, $S = 1.5, \mu = 1e^{-3}$

Figure D.1: Example Non-Null Model Cell and viral dynamics
Sensitivity of RFA to Infinite Sites Model Mutation Rate

Figure D.2: **Latent Model RFA Accuracy.** This figure shows the proportion of model runs versus the Null NonSpatial model that the RFA accurately categorizes by the proportion of cells going long-term latent for each type of spatial model with varying infinite sites mutation rates. As mutation rate increases, the ability of the RFA to distinguish models increases.
Figure D.3: Null Non-spatial vs Non-Spatial Mechanistic Trees. The left panel shows pair-wise diversity, segregating sites, and Tajima’s D for the evolution trees generated by each non-spatial model with lowest and highest parameter values. For the latent model, the parameter is proportion of cells going long-term latent; for the Two-Patch model the parameter is the coefficient of inter-patch mixing; and for the Selective Sweep model, the parameters are selective fitness advantage and mutation rate (to fitter strain, distinct from evolutionary mutation rate). The right panel shows evolution trees for each version of the null model.
Figure D.4: Null Non-spatial vs Spatial Mechanistic Trees. The left panel shows pair-wise diversity, segregating sites, and Tajima’s D for the evolution trees for the null non-spatial model and blood-dominated spatial versions of each of the mechanistic model with lowest and highest parameter values (as described in caption for Figure 4 in the main manuscript). The right panel shows the same for cell-dominated spatial versions.
Chapter 5 Supplementary Information

ART initiation

We use ART coverage as reported by the National Committee for AIDS Drugs and Prostitution Prevention and Control. Note: eligibility criteria changed to CD4<350 from 200 OR WHO stage 3 or 4 in 2009, so for post-2009, we extrapolated a linear increase in coverage \[221\ 227\ 242\ 283\].

Figure E.1: ART Coverage in Vietnam. The first panel shows the total number of HIV patients on ART over time (both adults and children), while the second panel shows the proportion on ART who are eligible (i.e. CD4 count <200).
Sensitivity Analysis Figures

Our results are qualitatively consistent across parameter estimates that produce model fits within realistic ranges. HIV incidence rate among PWID in Ho Chi Minh City, VN was estimated at 7.7% (2.9%, 12%) in 2011-2012 (284). The best-fit parameter set generates an out-of-sample 2011 incidence rate estimate of 8.8%, so we have run the intervention scenarios on parameter sets that generate out-of-sample 2011 incidence rate estimates at the low and high ends of the reported confidence interval (3.6%, 11.4%). Figures E.2, E.3, E.4, and E.5 demonstrate the impacts of ART and MMT scale-up (corresponding to Figures 5.3 and 5.4).

Figure E.2: Incidence and Deaths changes over time. Each panel in this figure shows a plot of reductions in HIV and HCV incidence, prevalence or deaths with varying ART scale-up.
Figure E.3: Incidence and Deaths changes over time. Each panel in this figure shows a plot of reductions in HIV and HCV incidence, prevalence or deaths with varying MMT scale-up.

Figure E.4: Incidence and Deaths changes over time. Each panel in this figure shows a plot of reductions in HIV and HCV incidence, prevalence or deaths with varying ART scale-up.
Figure E.5: **Incidence and Deaths changes over time.** Each panel in this figure shows a plot of reductions in HIV and HCV incidence, prevalence or deaths with varying MMT scale-up.

In this analysis, we assume an interaction between HIV and HCV – namely that HIV speeds HCV progression and makes spontaneous cure less likely. This interaction is likely responsible for some increase in HCV prevalence. Figure E.6 shows model predictions of HCV and HIV prevalence without this interaction, demonstrating a slightly lower estimate of HCV prevalence. Figure E.7 shows ART scale-up without this interaction, demonstrating no change in HCV prevalence, incidence or deaths with ART scale-up.
Figure E.6: Model Fit to HIV and HCV Prevalence. This figure shows the range of model estimates for HIV (blue) and HCV (green) prevalence among PWID in the shaded regions, with the estimate from the best-fit parameter set represented by the dashed line. Data estimates and corresponding confidence intervals to which the model was calibrated are represented by circles and error bars.

Figure E.7: Incidence and Deaths changes over time. Each panel in this figure shows a contour plot of reductions in HIV and HCV incidence or deaths, or total deaths, with a combination of MMT and ART interventions. Each black line represents set of combinations of ART and MMT that yield the same reduction.
Figure E.8: **PWID Population Dynamics over time.** The model replicates the gradual increase in the PWID and ex-PWID populations over the past decades.

Figure E.9: **HCV Incidence changes over time with HCV treatment and MMT scale-up, High Treatment Efficacy.** This figure displays incident cases over time under three different MMT scale-up scenarios: No scale up (solid lines), 10% scale-up (dotted lines), 50% scale-up (x-lines), and three different HCV treatment scale-up: No scale-up (black), 10% scale-up (dark grey), 50% scale-up (light grey). Prior scale-up of MMT coverage lays the foundation for HCV treatment roll-out in a lower incidence and prevalence population, and without MMT scale-up, HCV treatment roll-out makes little difference to HCV incidence even at high coverage.
Low levels of HCV treatment roll-out combined with low-levels of MMT scale-up can have a seemingly paradoxical effect on incidence. Figure E.9 shows the changes in HCV incidence after introduction and scale up of HCV treatment programs with high efficacy therapy. Scale-up of treatment can have an impact on incidence when implemented at high coverage levels. However, at low coverage levels, treatment scale-up can actually increase incidence (though prevalence is still reduced). This phenomenon occurs because prevalence is high enough that reinfection is extremely likely. With higher treatment coverage, prevalence is pushed low enough after 10 years to avoid this phenomenon. Similarly, when MMT has been scaled up sufficiently before HCV treatment is rolled out, the HCV prevalence is effectively diluted down to levels that are not associated with high risks of reinfection.

To demonstrate the theoretical conditions for the increase in incidence observed after scale-up of HCV treatment, it is helpful to used a pared-down version of the model. Scaling back to a simple SIS model retains the important characteristics. We thus have

\[
\begin{align*}
\frac{dS}{dt} &= \mu - \beta SI - \mu S + \alpha I \\
\frac{dI}{dt} &= \beta SI - \mu I - \alpha I
\end{align*}
\]

As this is a closed population, \( I = 1 - S \). The force of infection is \( \beta SI = \beta (S - S^2) \), and it is thus maximized when \( S = .5, I = .5 \) (thus when \( R_{\text{effective}} = 2 \)). In the case of HCV prevalence among PWID in HCMC, the prevalence is around 70%. At equilibrium before treatment intervention, i.e. when \( \alpha = 0 \), \( S^* = \frac{\mu}{\beta}, I^* = \frac{\beta - \mu}{\beta} \). When \( \alpha \) increases, after treatment scale-up is implemented, \( S^* \) becomes \( \frac{\mu + \alpha}{\beta} \). As \( S \) approaches the new \( S^* \), \( \beta SI \) will increase even as \( I \) decreases as \( S \) approaches .5. Note: after even the highest levels of treatment intervention, there is an initial increase in incidence, but at high levels, the increase quickly reverses.
Model Equations

HIV Progression Equations

Where

\( i \) is PWID status

\( j \) is HCV status

\( \mu_{in} \) is the PWID demographic in-flow term for each HIV stage

\( \mu_{out} \) is the PWID demographic out-flow term for each HIV stage

\( hcv_{in} \) is the HCV in-flow term for each HIV stage

\( hcv_{out} \) is the HCV out-flow term for each HIV stage

\[
\begin{align*}
\frac{dS(i,j)}{dt} &= -V_{HIV}(i,j) + \mu_{in}(S,i,j) + hcv_{in}(S,i,j) - \mu_{out}(S,i,j) - hcv_{out}(S,i,j) \\
\frac{dP(i,j)}{dt} &= V_{HIV}(i,j) - \theta P(i,j) + \mu_{in}(P,i,j) + hcv_{in}(P,i,j) - \mu_{out}(P,i,j) - hcv_{out}(P,i,j) \\
\frac{dI_1(i,j)}{dt} &= \theta P(i,j) - \gamma H I_1(i,j) - \alpha I_1(i,j) + \mu_{in}(I_1,i,j) + hcv_{in}(I_1,i,j) - \mu_{out}(I_1,i,j) - hcv_{out}(I_1,i,j) \\
\frac{dI_2(i,j)}{dt} &= \gamma H I_1(i,j) - \gamma H I_2(i,j) - \alpha I_2(i,j) + \mu_{in}(I_2,i,j) + hcv_{in}(I_2,i,j) - \mu_{out}(I_2,i,j) - hcv_{out}(I_2,i,j) \\
\frac{dI_3(i,j)}{dt} &= \gamma H I_2(i,j) - \gamma H I_3(i,j) - \alpha I_3(i,j) + \mu_{in}(I_3,i,j) + hcv_{in}(I_3,i,j) - \mu_{out}(I_3,i,j) - hcv_{out}(I_3,i,j) \\
\frac{dT_E(i,j)}{dt} &= \alpha I \sum_s I_s(i,j) - \delta E T_E(i,j) + \mu_{in}(T_E,i,j) + hcv_{in}(T_E,i,j) - \mu_{out}(T_E,i,j) - hcv_{out}(T_E,i,j) \\
\frac{dA(i,j)}{dt} &= \gamma H I_3(i,j) - \gamma H A(i,j) - \alpha A(i,j) + \mu_{in}(A,i,j) + hcv_{in}(A,i,j) - \mu_{out}(A,i,j) - hcv_{out}(A,i,j) \\
\frac{dT_L(i,j)}{dt} &= \alpha A(i,j) - \delta L T_L(i,j) + \mu_{in}(T_L,i,j) + hcv_{in}(T_L,i,j) - \mu_{out}(T_L,i,j) - hcv_{out}(T_L,i,j)
\end{align*}
\]
Where $S$ is susceptible, $P$ is Primary (Acute) infection, $I$ is Asymptomatic Infection (3 different Erlang stages), $T_E$ is early ART initiation, $A$ is symptomatic AIDS, and $T_L$ is late ART initiation.

**Incidence hazard function**

$$V_{HIV}(i, j) = F_{HIV}(I_{HIV}) \times S(i, j) \times (1 - \omega_{ns}(t))(1 - \omega_{mmt}(i))$$

$F_{HIV}$, the force of infection is formatted using a transmission coefficient ($\beta_{HIV1}$) multiplying model prevalence, $I_{HIV}$, and an extrinsic force of infection component ($\beta_{HIV0}$) as well weighting terms $\omega_{ns}$ for the impact of needle and syringe programs on the force of infection, and $\omega_{mmt}$ for the reduction in risk for PWID in Methadone Maintenance Therapy, and was fitted such that the model would reproduce prevalence estimates as described in the Data section:

$$F_{HIV} = \omega_{HIV} \beta_{HIV1} I_{HIV} + \omega_{HIV} \beta_{HIV0}$$

where $I_{HIV}$ is weighted model prevalence

$$I_{HIV} = \sum_i \sum_j \omega_{mmt}(i) \left( c_P P(i, j) + c_I I(i, j) + c_A A(i, j) + c_T (T_E(i, j) + T_L(i, j)) \right) / N(i, j)$$

where the $c$ terms are weights assigned for different transmission probabilities at the different stages of infection. $\beta_{HIV1}$, $\omega_{HIV}$ and $\omega_{ns}$ are estimated by using maximum likelihood; the $\beta$ terms are transmission coefficients, while $\omega_{HIV}$ term is a weighting factor that accounts for a change in hazard after the advent of ART (i.e. $\omega_{HIV1} = \omega_{HIV0} = 1$ prior to ART availability). Note: $\beta_0$ is included because there is likely
a component of the force of HIV infection on PWID that comes from sources other than injecting drug use. It can be noted that the model trajectory misses two of the data points. There may have been an inconsistency in data collection during these years.

Parameters

\( \mu_s \) inflow/outflow (death) rate (dependent on PWID class)*

\( \theta \) 1/duration of Acute stage

\( \gamma_H \) 1/duration of each phase of infection (the scale parameter that corresponds to the Erlang fit with shape = 3, Erlang being the special case of a Gamma distribution with integer shape parameter)

\( \alpha_s \) treatment rate from asymptomatic or symptomatic stages

\( \delta_s \) death rate on treatment (early or late)

\( c_s \) weighting for HIV stage

\( \omega_{mnt} \) weighting for MMT status

\( \omega_{ns} \) weighting for needle and syringe program effect

*Death rate on MMT, \( \mu_{MT} \) is calculated as a weighted average of PWID and ex-PWID death rates, weighted by \( mR \), the proportion of PWID on MMT who cease injecting.
PWID demographics

PWID demographics are layered on the HIV equations as birth in-flows and death out-flows. The stand-alone PWID demographics are described by the following equations.

\[
\frac{dC}{dt} = \sum_i \mu_i N_i + \lambda \sum_i N_i (1 - \frac{\sum_i N_i}{K}) + \rho_1 M_1 + \rho_2 M_2 - \gamma_D C - \nu C - \mu_D C \\
\frac{dM_1}{dt} = \nu C - \rho_1 M_1 - \eta M_1 - \mu_M T M_1 \\
\frac{dM_2}{dt} = \eta M_1 - \rho_2 M_2 - \mu_M T M_2 \\
\frac{dX}{dt} = \gamma_D C - \mu_X X
\]

Where \( C \) are PWID in the community, \( M \) are PWID in Methadone Maintenance, \( X \) are ex-PWID.

The in- and out-flows take the following format, where \( j \) is HCV status and \( k \) is HIV status.

\[
\mu_{in}(1, j, k) = (1 + \Lambda) \sum_i \mu_i N_i^* + \rho_1 M_1(j, k) + \rho_2 M_2(j, k) \\
\mu_{in}(2, j, k) = \nu C \\
\mu_{in}(3, j, k) = \eta M_1 \\
\mu_{in}(4, j, k) = \gamma_D C
\]

*This term is only present for the HIV and HCV uninfected class, i.e. new individuals entering the population are assumed to always be initially uninfected.
\[
\begin{align*}
\mu_{out}(1, j, k) &= \gamma_D C + \nu C + \mu_M TC \\
\mu_{out}(2, j, k) &= \rho_1 M_1 + \eta_1 M_1 + \mu_M TM_1 \\
\mu_{out}(3, j, k) &= \rho_2 M_2 + \mu_M TM_2 \\
\mu_{out}(4, j, k) &= \mu_X X 
\end{align*}
\]

**Parameters**

\(\Lambda\) excess recruitment rate of new PWID

\(\rho_i\) dropout rate of PWID in MMT

\(\gamma_D\) 1/duration of community drug use

\(\eta\) 1/duration of early-stage MMT

\(\nu\) recruitment rate into MMT (different for HIV positive and negative in scale-up scenarios)
HCV progression

HCV progression is layered on the HIV equations as in-flows and out-flows due to HCV. The stand-alone HCV dynamics are described by the following equations.

\[
\begin{align*}
\frac{dS}{dt} &= (1 + \Lambda)\left(\sum_i \mu_i N_i\right) + \kappa \times (1 - \phi)\gamma_A A - V_C(t) + \\
&\quad \alpha_T \epsilon (1 - \phi) \left[ C_1 + C_2 + C_3 + C_4 \right] - \mu_i S \\
\frac{dA}{dt} &= V_C(t) - \gamma_A A - \mu_i A \\
\frac{dR}{dt} &= \kappa \phi \gamma_A A + \alpha_T \epsilon \phi \left[ C_1 + C_2 + C_3 + C_4 \right] - \mu_i R \\
\frac{dC_1}{dt} &= (1 - \kappa)\gamma_A A - (\gamma_C + \mu_i) \times C_1 \\
\frac{dC_2}{dt} &= (1 - \alpha_T \epsilon)\gamma_C C_1 - (\gamma_C + \mu_i) \times C_2 \\
\frac{dC_3}{dt} &= (1 - \alpha_T \epsilon)\gamma_C C_2 - (\gamma_C + \mu_i) \times C_3 \\
\frac{dC_4}{dt} &= (1 - \alpha_T \epsilon)\gamma_C C_3 - (\gamma_C + \mu_i) \times C_4 \\
\frac{dL}{dt} &= (1 - \alpha_T \epsilon)\gamma_C C_4 - \delta_L L - \mu_i L
\end{align*}
\]

Translated into in-flows and out-flows, the equations become the following, where \(i\) is PWID group and \(k\) is HIV stage. For parameters that vary with HIV status, such as \(\kappa\), \(\kappa(k) = \kappa\) when \(k\) indicates a susceptible or treated stage, and \(\kappa(k) = \kappa_{HIV}\) when \(k\) indicates an HIV-positive and untreated stage.
\[ hcv_{in}(i, S, k) = \delta_L L(i, k) + \kappa(k)(1 - \phi(k))\gamma_A A(i, k) + (1 - \phi(k))\alpha_T(k)\epsilon(k)\gamma_C(k) \sum_s C_s(i, k) + \alpha_T(k)\epsilon(k)(1 - \kappa(k))\gamma_A A hcv(i, k) \]

\[ hcv_{in}(i, A, k) = V_{HCV}(i, k) \]

\[ hcv_{in}(i, P, k) = \kappa(k)\phi(k)\gamma_A A(i, k) + \phi(k)\alpha_T(k)\epsilon(k)\gamma_C(k) \sum_s C_s(i, k) \]

\[ hcv_{in}(i, C_1, k) = (1 - \alpha_T(k)\epsilon(k))(1 - \kappaappa(k))\gamma_A A(i, k) \]

\[ hcv_{in}(i, C_2, k) = (1 - \alpha_T(k)\epsilon(k))\gamma_C C_1(i, k) \]

\[ hcv_{in}(i, C_3, k) = (1 - \alpha_T(k)\epsilon(k))\gamma_C C_2(i, k) \]

\[ hcv_{in}(i, C_4, k) = (1 - \alpha_T(k)\epsilon(k))\gamma_C C_3(i, k) \]

\[ hcv_{in}(i, L, k) = (1 - \alpha_T(k)\epsilon(k))\gamma_C C_4(i, k) \]

\[ hcv_{out}(i, S, k) = V_{HCV}(i, k) \]

\[ hcv_{out}(i, A, k) = \gamma_A A(i, k) \]

\[ hcv_{out}(i, C_1, k) = \gamma_C C_1(i, k) \]

\[ hcv_{out}(i, C_2, k) = \gamma_C C_2(i, k) \]

\[ hcv_{out}(i, C_3, k) = \gamma_C C_3(i, k) \]

\[ hcv_{out}(i, C_4, k) = \gamma_C C_4(i, k) \]

\[ hcv_{out}(i, L, k) = \delta_L L \]
Incidence hazard function

\[ V_{HCV}(i,k) = F_{HCV}(I_{HCV}) \times S(i,k) \times (1 - \omega_{ns}(t))(1 - \omega_{mmt}(i)) \]

As with \( F_{HIV} \), \( F_{HCV} \) is formatted as a transmission coefficient (\( \beta_{HCV1} \)) times the model prevalence estimate \( I_{HCV} \) (with \( \omega_{ns}, \omega_{mmt} \) as described in the HIV Incidence Hazard section, though with no extrinsic term), and was fitted so the model would reproduce prevalence estimates as detailed in the Data section:

\[ F_{HCV} = \beta_{HCV1}I_{HCV} \]

where \( I_{HCV} \) is model prevalence:

\[ I_{HCV} = \sum_i \sum_k \omega_{mmt}(i) \left( A(i,k) + \sum_s C_s(i,k) + L(i,k) \right) / N(i,k) \]

and the \( \beta \) term is estimated by MLE. As with HIV, the \( \beta \) term is allowed to vary to account for a change in hazard after the advent of ART. Parameters

\( \Lambda(t) \) recruitment rate of new PWID

\( \kappa \) proportion of Acute HCV infections that clear spontaneously (differs with HIV status)

\( \phi \) proportion of cleared infections who develop protective immunity (differs with HIV status)

\( \gamma_A \) 1/duration of Acute HCV infection

\( \gamma_C \) duration of each stage of chronic infection (differs with HIV status)

\( \delta \) additional death rate from chronic liver disease
Maximum Likelihood Estimation

The maximum likelihood estimator functions used were binomial. The model output prevalence, $p$, for each infection at each time point, while the data was formatted as $k$ cases in sample size $N$ at each time point, so the log-likelihood was calculated as

$$\sum_{t} k(t) \times \log(p(t)) + \left( N(t) - k(t) \right) \times \log(1 - p(t))$$

for each infection, and then the log-likelihoods for both infections were summed and maximized.

Scale-up

Scale-up of interventions (MMT access, ART coverage and HCV treatment) is calculated as follows\(^\text{[119]}\). $x_{int}$ is the flow rate into the intervention compartment over time. The scale-up takes place over a duration, $r_{int}$ starting at time $t_{int}$, so the average rate of linear scale-up is $\frac{1}{r_{int}}$. Coverage increases until it reaches $Int_{cov}$, and then remains constant.

Thus for $t \leq t_{int}$

$$x_{int} = 0$$

for $t_{int} < t \leq t_{int} + r_{int}$

$$x_{int} = \left( \frac{t - t_{int}}{r_{int}} \right) \times Int_{cov} - \frac{Total_{int}}{Total_{need}}$$

and for $t > t_{int} + r_{int}$

$$x_{int} = Int_{cov} - \frac{Total_{int}}{Total_{need}}$$