Changing Risk Behaviours and the HIV Epidemic: A Mathematical Analysis in the Context of Treatment as Prevention

Bojan Ramadanovic¹, Krisztina Vasarhelyi²,³, Ali Nadaf¹,⁴, Ralf W. Wittenberg⁴, Julio S. G. Montaner²,⁵, Evan Wood²,⁵, Alexander R. Rutherford¹,⁴*

¹ The IRMACS Centre, Simon Fraser University, Burnaby, British Columbia, Canada, ² British Columbia Centre for Excellence in HIV/AIDS, Vancouver, British Columbia, Canada, ³ Faculty of Health Sciences, Simon Fraser University, Burnaby, British Columbia, Canada, ⁴ Department of Mathematics, Simon Fraser University, Burnaby, British Columbia, Canada, ⁵ Division of AIDS, Department of Medicine, Faculty of Medicine, University of British Columbia, Vancouver, British Columbia, Canada

Abstract

Background: Expanding access to highly active antiretroviral therapy (HAART) has become an important approach to HIV prevention in recent years. Previous studies suggest that concomitant changes in risk behaviours may either help or hinder programs that use Treatment as Prevention strategy.

Analysis: We consider HIV-related risk behaviour as a social contagion in a deterministic compartmental model, which treats risk behaviour and HIV infection as linked processes, where acquiring risk behaviour is a prerequisite for contracting HIV. The equilibrium behaviour of the model is analysed to determine epidemic outcomes under conditions of expanding HAART coverage along with risk behaviours that change with HAART coverage. We determined the potential impact of changes in risk behaviour on the outcomes of Treatment as Prevention strategies. Model results show that HIV incidence and prevalence decrease only above threshold levels of HAART coverage, which depends strongly on risk behaviour parameter values. Expanding HAART coverage with simultaneous reduction in risk behaviour act synergistically to accelerate the drop in HIV incidence and prevalence. Above the thresholds, additional HAART coverage is always sufficient to reverse the impact of HAART optimism on incidence and prevalence. Applying the model to an HIV epidemic in Vancouver, Canada, showed no evidence of HAART optimism in that setting.

Conclusions: Our results suggest that Treatment as Prevention has significant potential for controlling the HIV epidemic once HAART coverage reaches a threshold. Furthermore, expanding HAART coverage combined with interventions targeting risk behaviours amplify the preventive impact, potentially driving the HIV epidemic to elimination.

Introduction

Highly active antiretroviral therapy (HAART) suppresses HIV viral replication, which not only reduces morbidity and mortality [1,2], but also the transmission of HIV [3–5]. As a result, HAART has emerged as a potentially high-impact global prevention strategy [6,7]. Ecological and cohort studies have documented significant associations between increasing treatment coverage and declines in new HIV diagnoses [8] or incidence [9]. More recently, a randomized control trial found that HIV transmission in serodiscordant couples decreased by 96% when the infected partner received immediate HAART [10]. These findings are fuelling accelerated efforts to implement Treatment as Prevention programs worldwide, through expanding testing and offering earlier treatment to those infected with HIV. Implementation studies are either planned or currently underway throughout the world to evaluate the preventive effectiveness of Treatment as Prevention under field conditions [11].

Potential negative consequences of a large-scale expansion of treatment have been debated in the literature [6,12–16]. One concern is the possible increase in risk behaviours over time [17,18]. The argument is that expanding awareness of the beneficial effects of HAART can reduce fears of acquiring or transmitting HIV infection, which can lead to behavioural disinhibition commonly referred to as HAART optimism or risk compensation [4].

A number of empirical studies have investigated HAART optimism in heterosexual men and women, men who have sex with men (MSM), and injection drug users (IDU) [19]. The picture emerging from these studies is complex. Increases in risk behaviour since HAART was introduced in 1996 have been most often reported for MSM [20–33], but examples of no change [34] and also of decreasing risk behaviour have also been reported [35].
A number of studies of heterosexual individuals find that sexual risk behaviour drops or does not change [18,21,36,37] after the initiation of HAART, but some report increasing risk behaviours [39]. For IDU, the evidence for HAART optimism affecting sexual and injection risk behaviour is also mixed, with studies reporting increased risk behaviour [39], decreased risk behaviour [40], and no change in risk behaviour [41].

The potential impact of HAART optimism on the population-level preventive effects of HAART have also been studied using mathematical models. For example, a compartmental model of the San Francisco MSM community incorporates the evolution of drug-resistant strains and assumes that sexual risk behaviour increases with time [42]. In this model, an only 10% increase in risk behaviour was sufficient to overwhelm the gains achieved in preventing new infections through expanded HAART coverage. However, this model assumes that drug resistance plays a significant role in curtailing benefits from expanded HAART coverage.

Several studies have concluded that HAART optimism can substantially limit or completely overwhelm the effectiveness of HAART in preventing new infections [43-46]. Other studies predicted benefits despite behavioural disinhibition [12,47]. No clear pattern can be discerned from these studies. This may be due in part to variation in model assumptions. For example, studies which assume lower preventive efficacy for HAART, also tend to find greater negative impact due to HAART optimism. However, additional sources of variability may include the inherent heterogeneity in risk behaviour specific to geographical setting or risk group, as well as methodological differences between modelling studies.

One challenge is that the concept of HAART optimism is poorly defined [18,48]. Changing social norms around HIV risk due to treatment can be envisaged to influence both HIV-positive and HIV-negative individuals because diffusion of opinions and attitudes affects the population as a whole. However, empirical studies of HAART optimism tend to focus on subgroups of those diagnosed with HIV or specifically those on treatment. Behavioural disinhibition in the undiagnosed or susceptible subpopulations may have equal or even greater influence on population-level preventive HAART effects [48], and at least one study of MSM in the Netherlands provides supporting evidence [32].

A methodological limitation common to many models of HAART optimism involves the representation of risk behaviour. As people respond to changing social norms by changing their attitudes and behaviours, they are also changing the social norms themselves. This is an example of a dynamic interaction that can have a profound systematic impact on the HIV epidemic. Mathematical models that treat HIV-related risk behaviour dynamically are rare [50]. In most models, risk behaviour is represented as an exogenous parameter, with a predefined value. Furthermore, these models typically make assumptions specific to one setting and, therefore, it may be difficult to translate their results to different situations and locations.

In this analysis, we examine in a general sense the role that either increasing or decreasing risk behaviours may play in influencing the population-level impact of Treatment as Prevention. Our approach is to represent the spread of risk behaviour as a social contagion [51,52] and to use a two-disease model [53,54] in which we treat both HIV and risk behaviour as infectious processes. In the context of HIV, two-disease models have been developed previously to study co-infections by HIV and TB [55-64] or HIV and gonorrhea [65,66]. We are not aware of other studies that applied this approach to HIV and risk behaviour.

In our model, both the acquisition of risk behaviour and infection with HIV are by contact with others. The acquisition of risk behaviour is a precondition for possible subsequent infection with HIV. We perform a mathematical analysis of the epidemiologically relevant equilibria for arbitrary parameter values. Detailed calculations and proofs are provided in the Mathematical Supplement SI. One benefit of this is that our results are quite general, not specific to any particular locale.

The model is used to investigate the impact on HIV incidence and prevalence of simultaneous changes in risk behaviour and HAART coverage. Risk behaviour may be influenced independently through targeted interventions such as harm reduction, or be coupled to HAART coverage through HAART optimism. Data on the HIV epidemic in Vancouver’s Downtown Eastside inner-city neighbourhood are used to demonstrate a specific application of the model.

Analysis

Model Description and Analysis of Model Behaviour

Model structure. Our deterministic compartmental model of an HIV epidemic driven by the spread of risk behaviour is illustrated in Figure 1. The model has five states or compartments consisting of the general subpopulation G not engaging in risk behaviour, the susceptible subpopulation S who engage in risk behaviour but are not infected by HIV, the HIV-positive subpopulation A in the early acute phase of infection, the untreated HIV-positive subpopulation U in the post-acute chronic phase of infection, and the HIV-positive subpopulation H in the post-acute chronic phase who are receiving treatment with HAART. The chronic phase combines both the latent phase and AIDS. The mortality rate and infectiousness in the chronic phase are calculated as a time average of these values in each of the latent and AIDS phases. It is assumed that patients in the acute phase do not receive treatment.

The system of differential equations governing the time evolution of subpopulations in the model is

\[ \frac{dG}{dt} = r(G) - \beta_{SAC}G(S + A + U + H) - \delta_{G}G, \]
\[ \frac{dS}{dt} = \beta_{SAC}G(S + A + U + H) - (\lambda_{A}A + \lambda_{U}U + \lambda_{H}H) - \delta_{S}S, \]
\[ \frac{dA}{dt} = S(\lambda_{A}A + \lambda_{U}U + \lambda_{H}H) - \rho_{A}A - \delta_{A}A, \]
\[ \frac{dU}{dt} = \rho_{A}A - \delta_{U}U - \phi_{U}U, \]
\[ \frac{dH}{dt} = \phi_{U}U - \delta_{H}H. \]

Individuals in the subpopulations S and A, as well as in the two chronic phase subpopulations U and H, are assumed to equally influence individuals in the general subpopulation to engage in risk behaviour. This influence is modelled by a contact term with coefficient \( \beta_{SAC} \). This coefficient may be interpreted as the probability per unit time that social influence between a single individual engaged in risk behaviour and one not engaged in risk behaviour will result in the latter individual becoming involved in risk behaviour. The infectivities of the subpopulations in the HIV acute phase, untreated HIV chronic phase, and treated HIV chronic phase are \( \lambda_{A} \), \( \lambda_{U} \), and \( \lambda_{H} \), respectively. Specifically, \( \lambda_{A} \) is the rate of infection per unit time between a single HIV-negative individual engaged in risk behaviour and an HIV-positive individual in the acute phase of infection. The infectivities \( \lambda_{U} \) and \( \lambda_{H} \) have analogous interpretations. Note that \( \lambda_{A} > \lambda_{U} > \lambda_{H} \).
because HIV viral load is much higher in the acute phase than in the chronic phase of HIV infection and treatment further reduces viral load. The parameter $\rho_A$ is the rate at which HIV-positive individuals transition from the acute phase to the chronic phase. In other words, $\frac{1}{\rho_A}$ is the average duration of the acute phase. The parameter $\phi_U$ is the rate at which HIV chronic phase individuals are diagnosed and initiate treatment with HAART. The death rates in the $G$, $S$, $A$, $U$, and $H$ subpopulations are $\delta_G$, $\delta_S$, $\delta_A$, $\delta_U$, and $\delta_H$, respectively. The term $r(t)$ is the replenishment rate at which new individuals enter the model.

The infectivity parameters $\lambda_A$, $\lambda_U$, and $\lambda_H$ are influenced by factors such as the number of partners, length of partnerships, number of risk acts within a partnership, and the average probability of infection per single risk act between serodiscordant partners. This probability is in turn influenced by viral load. In this analysis, we primarily focus on the impact of risk behaviour on the infectivity parameters.

We assume that the death rates in the general subpopulation, susceptible subpopulation, acute phase HIV-positive subpopulation, and treated chronic phase population are equal and let $\delta_{GSAH} = \delta_G = \delta_S = \delta_A = \delta_H$. The death rate $\delta_U$ in the untreated chronic phase HIV-positive subpopulation takes into account transitions from the HIV latent phase to AIDS and subsequent death due to AIDS-related causes. As a result, we assume that $\delta_U > \delta_{GSAH}$. The death rate in the acute HIV phase is not elevated, because patients do not die of AIDS-related causes directly from the acute phase. The assumption that $\delta_H$ is equal to the death rate in the general population corresponds to assuming that patients on HAART who are virally suppressed have a normal life expectancy.

We make the simplifying assumption that the model population has constant size $N$, which entails setting $r(t) = \delta_{GSAH} (G(t) + S(t) + A(t) + H(t)) + \delta_U U(t)$. Substituting this expression for $r(t)$ into the system of equations (1), dividing all equations by $N$, and rescaling the parameters gives the following system of equations:

\begin{align*}
\frac{dg}{dt} &= \delta_{GSAH} (s + a + h) + \delta_A a - \beta_{SGH} g(s + a + u + h), \\
\frac{ds}{dt} &= \beta_{SGH} g(s + a + u + h) - s(\lambda_A a + \lambda_A u + \lambda_A h) - \delta_{GSAH} s, \\
\frac{da}{dt} &= s(\lambda_A a + \lambda_A u + \lambda_A h) - \rho_A a - \delta_{GSAH} a, \\
\frac{du}{dt} &= \rho_A a - \delta_A a - \phi_H u, \\
\frac{dh}{dt} &= \phi_H u - \delta_H h.
\end{align*}

In these equations, $g$, $s$, $a$, $u$, and $h$ are the fractions of the population in each of the states $G$, $S$, $A$, $U$, and $H$, respectively.
These functions satisfy the constraint that \( g(t) + s(t) + a(t) + \beta_u u(t) + \delta u(t) + h(t) = 1. \) Therefore, the system of equations (2) consists of only four independent equations. In order to understand the dynamics of the model defined by the system of equations (2), we study the equilibria of this system.

**Model equilibria.** The task of determining the model equilibria of the system of equations (2) can be simplified by combining the last two equations, which are linear equations with an inflow of \( \rho_a a. \) The total fraction of the population that is in the HIV chronic phase is \( c - u + h \) and the fraction of the HIV chronic phase population that is on treatment is \( \tau = \frac{h}{u} \). This gives the following system of equations for \( g, s, a, \) and \( c: \)

\[
\begin{align*}
\frac{dg}{dt} &= \delta g(s + a) + \left(1 - \tau\right)\delta u(s + a + c) - \beta_u u g(s + a + c), \\
\frac{ds}{dt} &= \beta_u u g(s + a + c) - s(\lambda_a a + \lambda_c c) - \delta g s, \\
\frac{da}{dt} &= s(\lambda_a a + \lambda_c c) - \rho_a a - \delta g a, \\
\frac{dc}{dt} &= \rho_a a - \delta c.
\end{align*}
\]

The last equation in the system (2) implies that at any equilibrium of this system, the equilibrium values of \( u \) or \( h \) must satisfy

\[
h_* = \frac{c* - u_*}{\delta g_*},
\]

where the subscript * is used to denote a generic equilibrium value. Therefore, the value of \( \tau \) at any equilibrium for which \( u_* + h_* \neq 0 \) is

\[
\tau_* = \frac{h_*}{u_* + h_*} = \frac{c*}{\delta g_* + \delta g_*}.
\]

This equation gives \( \phi_* = \frac{\tau_* \delta g_*}{\delta g_* (1 - \tau_*)} \). Therefore, we can replace the role of \( \phi_* \) in the model by \( \tau_* \), which is defined as the fraction of the HIV chronic phase subpopulation which is on treatment at equilibrium. This equilibrium HAART coverage parameter is treated as an exogenous parameter in the model.

To find the equilibria of the system of equations (3), it is sufficient to study the following system, which has \( \tau \) replaced by \( \tau_*: \)

\[
\begin{align*}
\frac{dg}{dt} &= \delta g(s + a) + \left(1 - \tau_*\right)\delta u(s + a + c) - \beta_u u g(s + a + c), \\
\frac{ds}{dt} &= \beta_u u g(s + a + c) - s(\lambda_a a + \lambda_c c) - \delta g s, \\
\frac{da}{dt} &= s(\lambda_a a + \lambda_c c) - \rho_a a - \delta g a, \\
\frac{dc}{dt} &= \rho_a a - \delta c.
\end{align*}
\]

This system of equations has the same equilibria as (3). We have defined

\[
\lambda_c = \left(1 - \tau_*\right)\lambda_u + \tau_* \lambda_h \quad \text{and} \quad \delta_c = \left(1 - \tau_*\right)\delta_u + \tau_* \delta_g h, \quad (7)
\]

which are the effective HIV chronic phase infectivity and death rate at equilibrium. These are obtained by using \( \tau_* \) to calculate weighted averages from the infectivities and death rates in the undiagnosed and treated subpopulations. It follows from our assumptions \( \lambda_u > \lambda_H > \lambda_L \) and \( \delta u > \delta_c g h H \) that the effective chronic phase infectivity and death rate in the equations (7) satisfy the inequalities \( \lambda_c < \lambda_L \) and \( \delta_c > \delta_g h. \)

The system of equations (6) is analyzed rigorously in the Mathematical Supplement SI to this paper. It is found that if \( \delta_c > \delta_g h \) and \( \lambda_c < \lambda_L \), then the model exhibits three types of equilibrium states: the risk-free state in which there is neither endemic risk behaviour nor endemic HIV, the risk-endemic state in which risk behaviour is endemic but there is no endemic HIV, and the HIV-endemic state in which both risk behaviour and HIV are endemic.

Each of the equilibrium states may in turn be stable, depending on the values of the parameters in the model. The importance of stable equilibria is that they represent the state approached by the epidemic as time evolves. In the Mathematical Supplement SI, conditions for the stability of each of these three equilibria are determined and we shall simply quote the results here.

**Risk-free equilibrium.** The risk-free equilibrium state is stable if

\[
0 < \beta_u u < \delta_g h. \quad (8)
\]

The population proportions at this equilibrium are

\[
g_0 = 1, \quad s_0 = 0, \quad a_0 = 0, \quad c_0 = 0. \quad (9)
\]

**Risk-endemic equilibrium.** The risk-endemic equilibrium is stable if

\[
\frac{\beta_u u}{\delta_g h} + \frac{\lambda_a + \rho_a}{\delta_c} \lambda_c < \frac{\beta_u u (\rho_a + \delta_g h)}{\beta_u u - \delta_g h}. \quad (10)
\]

The population proportions at this equilibrium are

\[
g_1 = \frac{\delta_g h}{\beta_u u}, \quad s_1 = 1 - \frac{\delta_g h}{\beta_u u}, \quad a_1 = 0, \quad c_1 = 0. \quad (11)
\]

**HIV-Endemic equilibrium.** The HIV-endemic equilibrium is stable if

\[
\frac{\beta_u u}{\delta_g h} > \frac{\lambda_a + \rho_a}{\delta_c} \lambda_c > \frac{\beta_u u (\rho_a + \delta_g h)}{\beta_u u - \delta_g h}. \quad (12)
\]
The population proportions at this equilibrium are

\[
g_2 = 1 - s_2 - a_2 - c_2, \\
c_2 = \frac{\delta_c (\delta_{g2} + \rho_u)}{\lambda_u \delta_c + \lambda_c \rho_u}, \\
a_2 = \frac{\delta_c}{2(\delta_c + \rho_u)} - \frac{\delta_c \delta_{g2} + \rho_u(\delta_{g2} + 2\delta_{p2}) + \rho_u(\delta_{g2} + 2\delta_{p2}) + \rho_u}{2\delta_p(\delta_c + \rho_u)} + \frac{\delta_c}{(\delta_{g2} - \delta_{g2})\delta_c + (\delta_{g2} - \delta_{g2})\rho_u} + \delta_c, \\
(13)
\]

\[
(\delta_{g2} - \delta_{g2})\delta_c + (\delta_{g2} - \delta_{g2})\rho_u
\]

\[
c_2 = \frac{\rho_u}{\delta_c} a_2.
\]

The interpretation of these results is that a minimum value for \(\beta_{vac}\), the rate of spread for risk behaviour, is required for risk behaviour to become endemic within the model. Given endemic risk behaviour, minimum values for the HIV infectivities \(\lambda_u\) and \(\lambda_c\) are required for HIV to become endemic in the model. Note that this minimum rate for the spread of HIV depends on the rate at which risk behaviour is spreading. The model describes two epidemics, with one driving the other.

Observe that for any given set of epidemiologically possible values of the parameters, exactly one of the three equilibria is stable. The behaviour of the system is determined by which equilibrium is stable. We can formulate public health goals in terms of guiding the system towards one of the equilibria by altering key parameters through interventions. For example, if our goal is the eradication of HIV, we need to determine what changes in parameter values would drive our system away from the HIV-endemic equilibrium by making it unstable. Alternatively, we may define a less ambitious goal of driving down HIV incidence and prevalence by changing the value of the HIV-endemic equilibrium in solution (13), through interventions designed to impact the parameters in the model.

**Implementation of treatment.** Treatment is implemented in the model by assuming that a specified proportion of the HIV-positive subpopulation in the chronic phase at equilibrium is receiving HAART, as described in the previous subsection. Aside from the mathematical advantages of implementing treatment this way, using HAART coverage rather than treatment initiation rate \(\phi\) as a model parameter aligns the model better with standard public health metrics.

It is important to note that we define HAART coverage as the proportion of the entire chronically infected subpopulation – regardless of diagnostic status – that is on treatment with fully suppressed viral load. This definition does not address issues of low adherence, treatment failure and other factors that reduce effective coverage. In practice, HAART coverage is usually measured as a proportion of the diagnosed subpopulation prescribed HAART. This interpretation will lead to a higher HAART coverage value than the value in the model, because it uses the smaller reference population of only the diagnosed rather than all chronically infected individuals. Including treated patients with unsuppressed viral load would further increase the measured coverage above the HAART coverage value in the model. The analyses in this paper refer to the model definition of HAART coverage. Thus our estimates of the HAART coverage required to produce an effect on the course of the epidemic are likely to be lower than the actual coverage needed to achieve the same effect in practice.

We assume in this model that HIV-positive individuals in the acute phase are never on treatment, because it is unlikely that diagnosis and referral for treatment would be completed within the first two months of infection. This assumption is supported by data from British Columbia, Canada, where between 2006 and 2009 approximately 5% of newly diagnosed patients were in the acute phase [67]. Even fewer of these individuals would still be acutely infected when starting treatment.

We assume that individuals on HAART are not infectious and are not participating in disease dynamics [68], which implies that \(\delta_a = 0\). The fraction \((1 - \tau_u)\) of the chronically infected subpopulation remains infectious in the model. This implies that HIV infectivity and death rate in the chronic phase from equations (7) simplify to

\[
\lambda_c(\tau_u) = (1 - \tau_u) \lambda_u \\
\delta_c(\tau_u) = (1 - \tau_u) \delta_u + \tau_u \delta_{g2}.
\]

Both the stability condition of the HIV-endemic equilibrium in equations (12) and the population state at equilibrium in equations (13) depend on \(\lambda_c(\tau_u)\) and \(\delta_c(\tau_u)\). This implies that by changing HAART coverage, we can change both the HIV prevalence at equilibrium and the threshold for the switch from the state of endemic HIV to a state of endemic risk behaviour without HIV.

**Estimation of model parameters.** Model parameters were either taken directly from or were estimated using published sources. The ratio between infectivity in the acute phase and infectivity of untreated individuals in the chronic phase depends on the viral loads in the two disease phases. Therefore, it is convenient to write

\[
\lambda_u = z \lambda_u,
\]

where the ratio \(z\) is a parameter related to the viral load in the two phases. We use \(z \approx 40\) based on data from the Rakai, Uganda seroconversion study [69].

The length of the acute phase of an HIV infection is approximately 2 months. Therefore, we take \(\rho_u \approx 0.50\) months\(^{-1}\) \(\approx 6.0\) years\(^{-1}\). The life expectancy is defined as the average length of time from when individuals become susceptible to risk behaviour until death. The life expectancy for the uninfected subpopulation is taken to be 32 years, based on demographic information for Vancouver's Downtown Eastside neighbourhood [68], which means that \(\delta_{g2} \approx 0.031\) year\(^{-1}\). The life expectancy for an untreated HIV-positive individual is assumed to be 11 years [68], implying that \(\delta_u \approx 0.091\) year\(^{-1}\). The values of these parameters are summarized in Table 1.

The parameters \(\tau_u\), \(\lambda_u\), and \(\beta_{vac}\) are treated as free parameters. The equilibrium HAART coverage \(\tau_u\) depends on the effectiveness of HIV testing programs and the guidelines for initiating treatment. The infectivity \(\lambda_u\) of untreated HIV-positive individuals depends on the level and nature of risk behaviour among HIV-positive individuals. Harm reduction programs would be expected to primarily impact \(\lambda_u\). In this model, any impact on \(\lambda_u\) is also reflected in the acute phase infectivity, because \(\lambda_a = z \lambda_u\). The risk behaviour propagation coefficient \(\beta_{vac}\) reflects the propensity for individuals to become engaged in risk behaviour and is modified by social influences such as HAART optimism.

**Model Results**

**Control and eradication of the HIV epidemic in the presence of risk behaviour.** Risk behaviour is captured in the model through the parameters \(\beta_{vac}\) and \(\lambda_u\). The parameter \(\beta_{vac}\)
measures the rate of risk behaviour propagation and incorporates both the receptivity of susceptible individuals to becoming engaged in risk behaviour, and peer pressure that risk-engaged individuals exert, without making a distinction between the two effects. The infectivity parameter for untreated individuals in the chronic phase, \( \lambda_u \), incorporates the frequency of risk behaviour, through the effect that risk behaviour has on infectivity. By equations (14) and (15), \( \lambda_u \) is related to both \( \lambda_u \) and \( \lambda_c \). This determines the impact of risk behaviour on infectivity in both the acute and chronic phases in the model.

Analyses of equilibrium states, described in the "Model Equilibria" subsection, enable us to determine the necessary conditions for the persistence of the HIV epidemic, which occurs when the HIV-endemic equilibrium in equations (13) is stable. The stability condition for the HIV-endemic equilibrium in equation (12) can be rewritten using equations (14) and (15) as a condition on the undiagnosed chronic phase infectivity:

\[
\lambda_u > \frac{\beta^u u ((1 - \tau_c)\delta_u + \tau_c \delta^u u (\rho_u + \delta^u u))}{(\beta^u u - \delta^u u ((1 - \tau_c)(\delta_u + \rho_u) + \tau_c \delta^u u))}. \tag{16}
\]

This condition for stability of the HIV-endemic equilibrium can be rewritten as the condition:

\[
\beta^u u > \frac{\lambda_u \delta^u u ((1 - \tau_c)(\delta_u + \rho_u) + \tau_c \delta^u u)}{(1 - \tau_c)(\delta_u + \rho_u) + \tau_c \delta^u u)}, \tag{17}
\]

Substitution of \( \tau_c = 0 \) into equation (16) shows that if:

\[
\lambda_u > \frac{\beta^u u \delta^u u (\rho_u + \delta^u u)}{(\beta^u u - \delta^u u) (\delta_u + \rho_u)} \tag{18}
\]

then the HIV epidemic is not sustainable, even without HAART. Furthermore, if the risk-endemic equilibrium in equations (11) is stable, then risk behaviour persists even in the absence of an HIV epidemic. The maximum HAART coverage possible is \( \tau_c = 1 \). Substituting this into the inequality (16) shows that if:

\[
\lambda_u > \frac{\beta^u u (\rho_u + \delta^u u)}{\alpha (\beta^u u - \delta^u u)} \tag{19}
\]

then HAART cannot eliminate the HIV epidemic in the model. This occurs because individuals in the acute phase of HIV infection are never treated with HAART in the model and the inequality (19) is the condition for the epidemic to be sustainable with only transmissions from individuals in the acute phase. Provided that the HIV-endemic equilibrium is unstable when \( \tau_c = 1 \) and stable when \( \tau_c = 0 \), the condition on \( \tau_c \) for the HIV-endemic equilibrium to be stable is:

\[
\tau_c < \frac{\lambda_u (\beta^u u - \delta^u u) (\delta_u + \rho_u) - \beta^u u \delta^u u (\delta^u u + \rho_u)}{\lambda_u (\beta^u u - \delta^u u) (\delta_u + \rho_u) - \beta^u u \delta^u u (\delta^u u + \rho_u) + \delta^u u (\delta_u + \rho_u) (\delta^u u + \rho_u) - \lambda_u (\beta^u u - \delta^u u)} \tag{20}
\]

The surface representing the boundary between regions where HIV is endemic and where it is extinct within the three-dimensional parameter space defined by \( \beta^u u, \lambda_u \), and \( \tau_c \) is obtained by setting \( \lambda_u \) equal to the expression in (16). This surface is plotted in Figure 2, with the parameters \( x, \beta^u u, \delta^u u \), and \( \delta^u u \) set equal to the values in Table 1. To make visualisation of these regions easier, we alternately fix one of the three parameters and plot the boundary curve in two dimensions.

First, we set chronic phase HAART coverage \( \tau_c \) to 0 and use the inequalities (10) and (18) to plot regions for \( \lambda_u \) and \( \beta^u u \) in the two-dimensional parameter space, where each of the risk-free, risk-endemic, and HIV-endemic equilibria are stable. The resulting diagram, known as a two-parameter bifurcation diagram, is shown in Figure 3. We find that the HIV epidemic can be eliminated by reducing either \( \lambda_u \) or \( \beta^u u \) to a small enough value. The concave shape of the boundary curve implies that, at least theoretically, there exists an optimal strategy for driving the HIV epidemic to elimination. Starting with an HIV-endemic state, elimination of the HIV epidemic corresponds to driving the system along a path in the \( \lambda_u \) and \( \beta^u u \) parameter space, which leaves the HIV-endemic region. The shortest such path is one possible definition of an optimal intervention. The concave curve implies that the shortest path to the boundary is neither vertical nor horizontal, regardless of where it starts within the HIV-endemic region. Therefore, a combination intervention, which impacts both \( \lambda_u \) and \( \beta^u u \), is optimal. Whether a theoretically optimal strategy is also optimal in practice depends on the cost or feasibility of the intervention. The model can be applied to real-world situations by modifying the path-length definition of optimality through weighting the \( \lambda_u \) and \( \beta^u u \) components of the path length by the cost associated with interventions that modify these parameters.

Next, we show in Figure 4 the two-parameter bifurcation diagrams obtained by setting the value of \( \lambda_u \) to 0.33 year\(^{-1} \) and to approximately half that value, 0.17 year\(^{-1} \). The stability regions in these bifurcation diagrams for \( \beta^u u \) and \( \tau_c \) are obtained from the inequalities (10) and (17). The value of 0.33 year\(^{-1} \) for \( \lambda_u \) is the estimate obtained from our analysis of the Vancouver Downtown Eastside epidemic in the following section. As HAART coverage increases, the threshold value of \( \beta^u u \) required for HIV to become endemic also increases. For \( \lambda_u = 0.17 \) year\(^{-1} \), above 80\% HAART coverage, the boundary curve rises rapidly, implying that by expanding HAART coverage, the epidemic can be eliminated even with rapid spread of risk behaviour. Similar to Figure 3, the boundary of the HIV-endemic region in Figure 4 is concave, implying that an optimal intervention involving \( \beta^u u \) and \( \tau_c \) is a combination intervention which impacts both parameters.

We now fix the risk propagation coefficient \( \beta^u u \) at 0.059 year\(^{-1} \), corresponding to the estimate used for the Vancouver Downtown Eastside analysis below and determine conditions on the infectivity \( \lambda_u \) of untreated HIV-positive individuals and HAART coverage \( \tau_c \) for endemic HIV. The stability regions in the two-parameter bifurcation diagram of \( \lambda_u \) and \( \tau_c \) are shown in Figure 5. The graph in this figure is obtained from the inequality (20), with the parameter values taken from Table 1. It follows from equation (18) that if \( \lambda_u \) is less than approximately 0.12 year\(^{-1} \), then an HIV epidemic is not sustainable, even without any HAART coverage. Equation (19) implies that if \( \lambda_u \) is greater than approximately 0.32 year\(^{-1} \), then endemic HIV would persist, even with complete HAART coverage. In our model, we are assuming that treatment with HAART only occurs in the chronic phase of the disease and in the last scenario the epidemic is being driven entirely by acute phase infections. From the graph in Figure 5, we can see that if \( \lambda_u \) is less than approximately 0.2 year\(^{-1} \), then eradication of the epidemic can be achieved with less than 80\% HAART coverage.
Table 1. Model parameters.

<table>
<thead>
<tr>
<th>Parameter Description</th>
<th>Symbol</th>
<th>Model Analysis</th>
<th>DTES Value</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>progression rate from acute to chronic infection</td>
<td>$\mu_a$</td>
<td>6.0 year$^{-1}$</td>
<td>6.0 year$^{-1}$</td>
<td>[78]</td>
</tr>
<tr>
<td>death rate for HIV-negative and acute phase HIV</td>
<td>$\delta_{out}$</td>
<td>0.031 year$^{-1}$</td>
<td>0.031 year$^{-1}$</td>
<td>[68]</td>
</tr>
<tr>
<td>death rate for untreated chronic phase HIV</td>
<td>$\delta_a$</td>
<td>0.091 year$^{-1}$</td>
<td>0.091 year$^{-1}$</td>
<td>[68]</td>
</tr>
<tr>
<td>ratio of acute HIV infectivity to chronic infectivity</td>
<td>$\alpha$</td>
<td>40</td>
<td>40</td>
<td>[69], [78]</td>
</tr>
<tr>
<td>risk behaviour propagation coefficient</td>
<td>$\beta_{bac}$</td>
<td>free</td>
<td>0.09 year$^{-1}$</td>
<td>this paper</td>
</tr>
<tr>
<td>untreated chronic phase HIV infectivity</td>
<td>$\lambda_a$</td>
<td>free</td>
<td>0.33 year$^{-1}$</td>
<td>this paper</td>
</tr>
<tr>
<td>equilibrium HAART coverage for HIV chronic phase subpopulation</td>
<td>free</td>
<td>0.20</td>
<td></td>
<td>[68]</td>
</tr>
</tbody>
</table>

Values for the general model analysis and for the specific application to Vancouver’s Downtown Eastside are shown. doi:10.1371/journal.pone.0062321.t001

Furthermore, for $\lambda_a$ less than 0.2 year$^{-1}$, the level of HAART coverage required to eradicate the epidemic drops rapidly as $\lambda_a$ decreases.

Instead of eradicating the epidemic altogether, reducing HIV incidence and prevalence may be a more realistic short-term goal. Therefore, our next objective is to examine how equilibrium incidence and prevalence depend on the free parameters in the model. HIV prevalence, expressed as a proportion of the total population, is

$$p = a + c.$$  \hspace{1cm} \text{(21)}

The equilibrium prevalence when HIV is endemic is

$$p_2 = \frac{\delta_c (\tau_s a)}{\delta (\tau_s)} a_2,$$  \hspace{1cm} \text{(22)}

where $a_2$ is given in equations (13) and $\delta_c$ is given by equation (14). HIV incidence is defined as the number of new infections per unit time and given by

$$I = \lambda_a SA + \lambda_c SC,$$  \hspace{1cm} \text{(23)}

where $C = U + H$ and $\lambda_c = (1 - \tau_s) \lambda_U$. It is convenient to also express incidence relative to the size of the total population. Incidence per unit population is

$$i = s (\lambda_a a + \lambda_c c),$$  \hspace{1cm} \text{(24)}

and hence the incidence per unit population at the HIV-endemic equilibrium is

$$i_2 = s_2 (\lambda_a a_2 + \lambda_c c_2) = (\delta_{bac} + \rho_a) a_2.$$  \hspace{1cm} \text{(25)}

Figure 6 shows level curves for multiple values of equilibrium HIV prevalence, given by equation (21), as a function of $\tau_s$ and $\beta_{bac}$, plotted for values of $\lambda_a$ fixed at 0.33 year$^{-1}$ and 0.17 year$^{-1}$. The value of 0.33 year$^{-1}$ was chosen, because this is the estimate obtained for Vancouver Downtown Eastside below. The prevalence level curves for $\lambda_a = 0.17$ year$^{-1}$ show the impact of reducing $\lambda_a$ by approximately half. Each level curve represents the combination of the $\tau_s$ and $\beta_{bac}$ parameters which corresponds to constant HIV prevalence at the indicated value. Figure 6 shows that once HAART coverage increases beyond a threshold, the spread of risk behaviour has little further impact on HIV prevalence. Furthermore, this effect depends on $\lambda_a$, becoming more pronounced as infectivity drops. Reducing $\lambda_a$ from 0.33 year$^{-1}$ to 0.17 year$^{-1}$ shifts the range in which HAART is highly effective at reducing prevalence from approximately 95% coverage to approximately 80% coverage.

The graphs in Figure 6 show that there is a maximum value of the risk propagation coefficient $\beta_{bac}$ below which the HIV epidemic is not sustainable, because equilibrium prevalence is zero. This value can be computed using the inequality (17). When $\tau_s = 1$, this inequality gives the following maximum value of $\beta_{bac}$ for which the epidemic can be eliminated through solely an increase in HAART coverage:

$$\beta_{bac} \leq \frac{2 \lambda_a \delta_{bac}}{2 \lambda_a - \rho_a - \delta_{bac}}.$$  \hspace{1cm} \text{(26)}

When $\lambda_a \approx 0.33$ year$^{-1}$, the maximum value of the risk propagation coefficient for which the HIV epidemic can be eliminated solely through HAART expansion is approximately 0.057 year$^{-1}$. If $\lambda_a$ decreases to 0.17 year$^{-1}$, then this maximum value increases to 0.27 year$^{-1}$.

These results show that expanding HAART coverage can be highly effective in containing and perhaps eradicating the HIV epidemic, provided that the infectivity of undiagnosed HIV-positive individuals is sufficiently low. The model further suggests that the benefit of using Treatment as Prevention increases significantly when combined with measures such as harm reduction that decrease untreated infectivity.

Impact of HAART optimisation. HAART optimisation is defined in this study as an increase in the social propagation of risk behaviour in an environment of expanded HAART coverage. It is incorporated in the model by assuming that the risk behaviour propagation coefficient $\beta_{bac}$ depends on the level of HAART coverage $\tau_s$. HAART optimisation implies that $\beta_{bac}$ is an increasing function of $\tau_s$, on the interval 0 to 1. Furthermore, $\beta_{bac}$ must be positive when $\tau_s = 0$, because people engage in risk behaviour even in the absence of HAART. For simplicity, we postulate that $\beta_{bac}$ depends linearly on $\tau_s$ and write

$$\beta_{bac}(\tau_s) = \omega \tau_s + \beta_0,$$  \hspace{1cm} \text{(27)}
where $\omega > 0$ and $\beta_0 > 0$. We refer to $\omega$ as the intensity of HAART optimism. The parameter $\beta_0$ is the value of $\beta_{\text{arc}}$ in the absence of HAART. We examine the impact of varying both $\omega$ and $\beta_0$ on HIV prevalence and incidence.

Figure 7 shows equilibrium HIV prevalence $p_2$ plotted as a function of HAART coverage $\tau_\star$ for three different values of $\omega$ and $\beta_0$. These plots are generated using equations (22) and (27), with the parameters $\rho_{\text{d}}, \delta_{\text{gph}}, \delta_{\text{a}},$ and $\alpha$ set to values given in Table 1 and $\lambda_u$ set to 0.33 year$^{-1}$. HAART optimism initially causes prevalence to increase. However, there is a threshold level of HAART coverage above which prevalence decreases, even in the presence of HAART optimism. As expected, larger values of the intensity of HAART optimism $\omega$ cause greater increases in prevalence for levels of HAART coverage below this threshold. This threshold value of $\tau_\star$ can be computed by numerically calculating the value of $\tau_\star$ between 0 and 1 for which the derivative of equilibrium HIV prevalence $p_2$ with respect to HAART coverage $\tau_\star$, is zero. We denote this threshold level of HAART coverage for prevalence by $\tau_p$. First we examine how $\tau_p$ varies with $\beta_0$. Setting the infectivity $\lambda_u$ to 0.33 year$^{-1}$ and $\omega$ to 0.10 year$^{-1}$, the threshold level of HAART coverage $\tau_p$ is approximately 0.84 when $\beta_0 = 0.05$ year$^{-1}$ and it decreases to 0.69 when $\beta_0$ is increased to 0.20 year$^{-1}$. Now consider how $\tau_p$ varies with $\omega$.

Setting $\lambda_u$ to 0.33 year$^{-1}$ and $\beta_0$ to 0.10 year$^{-1}$, the value of $\tau_p$ is 0.80 when $\omega$ is 0.05 year$^{-1}$ and it decreases to 0.75 when $\omega$ increases to 0.20 year$^{-1}$. Not only does the threshold level of HAART coverage $\tau_p$ not increase with $\omega$, but it decreases slowly for this range of values for $\beta_0$ and $\omega$. The dependence of $\tau_p$ on $\beta_0$ and $\omega$ is shown in the surface plot in Figure 8.

Figure 9 shows equilibrium HIV incidence $i_2$ plotted as a function of HAART coverage $\tau_\star$ for three different values of $\omega$ and $\beta_0$. These plots are generated using equations (25) and (27), with the parameters $\rho_{\text{d}}, \delta_{\text{gph}}, \delta_{\text{a}},$ and $\alpha$ set to values given in Table 1 and $\lambda_u$ set to 0.33 year$^{-1}$. The plots of incidence in the presence of HAART optimism in Figure 8 show that for low levels of HAART coverage, HAART optimism causes incidence to increase. As with prevalence, there is a threshold level of HAART coverage, above which HAART causes incidence to decrease. The threshold level of HAART coverage for incidence is denoted by $\tau_i$. The value of $\tau_i$ is computed by numerically solving for the value of $\tau_\star$ where.

$$\frac{dp_2}{d\tau_\star} = \left( \frac{\rho_{\text{d}}\delta_2}{(\delta_{\text{a}}(\tau_\star))^2} - \frac{\delta_{\text{a}}(\tau_\star) + \rho_{\text{d}}\delta_2}{\delta_{\text{a}}(\tau_\star)} \right) (\delta_u - \delta_{\text{gph}})$$

$$+ \frac{\delta_{\text{a}}(\tau_\star) + \rho_{\text{d}}}{\delta_{\text{a}}(\tau_\star)} \left( \frac{\omega}{\delta_{\text{gph}}} \frac{\partial a_2}{\partial \alpha} - \frac{\lambda_u}{\delta_{\text{a}}} \frac{\partial a_2}{\partial \lambda_u} \right)$$

is equal to zero. We denote this threshold level of HAART coverage for prevalence by $\tau_p$. First we examine how $\tau_p$ varies with $\beta_0$. Setting the infectivity $\lambda_u$ to 0.33 year$^{-1}$ and $\omega$ to 0.10 year$^{-1}$, the threshold level of HAART coverage $\tau_p$ is approximately 0.84 when $\beta_0 = 0.05$ year$^{-1}$ and it decreases to 0.69 when $\beta_0$ is increased to 0.20 year$^{-1}$. Now consider how $\tau_p$ varies with $\omega$.

$$\frac{di_2}{d\tau_\star} = \left( \delta_{\text{gph}} + \rho_{\text{d}} \right) \left( \frac{\omega}{\delta_{\text{gph}}} \frac{\partial a_2}{\partial \alpha} - \frac{\lambda_u}{\delta_{\text{a}}} \frac{\partial a_2}{\partial \lambda_u} \right)$$

$$- \left( \delta_2 - \delta_{\text{gph}} \right) \frac{\partial a_2}{\partial \delta_{\text{a}}}$$

is equal to zero. If $\frac{dp_2}{d\tau_\star}$ is never zero for $0 \leq \tau_\star \leq 1$, then it must be either strictly negative or strictly positive on this interval. If $\frac{dp_2}{d\tau_\star}$ is strictly negative, then the threshold HAART coverage $\tau_t$ is 0 and if it is strictly positive, then $\tau_t$ is 1. To illustrate how $\tau_i$ varies with $\beta_0$ in Figures 9 and 10, we set $\lambda_u$ to 0.33 year$^{-1}$ and $\omega$ to 0.10 year$^{-1}$. The threshold HAART coverage $\tau_t$ is then approximately 0.65 when $\beta_0 = 0.05$ year$^{-1}$; however, it decreases substantially to 0.047 when $\beta_0 = 0.20$ year$^{-1}$. To examine how $\tau_t$
Figure 3. Equilibrium stability regions with no HAART coverage. Regions represent values of the risk behaviour propagation coefficient $\beta_{\text{sec}}$ and the infectivity in the undiagnosed chronic HIV phase $\lambda_\mu$ for which both HIV and risk behaviour are endemic, only risk behaviour is endemic, and neither are endemic. The red curve separates the HIV-endemic and HIV-free regions. The blue line separates risk behaviour-endemic and risk behaviour-free regions. The equilibrium HAART coverage parameter $\tau_\tau$ is set to 0 for this plot.

doi:10.1371/journal.pone.0062321.g003

Figure 4. Equilibrium stability regions with fixed untreated chronic phase infectivity. The untreated chronic phase infectivity is fixed at (a) $\lambda_\mu = 0.33$ year$^{-1}$ and (b) $\lambda_\mu = 0.17$ year$^{-1}$. Regions represent values of the risk behaviour initiation rate $\beta_{\text{sec}}$ and equilibrium HAART coverage $\tau_\tau$ for which both HIV and risk behaviour are endemic, only risk behaviour is endemic, and neither HIV nor risk behaviour are endemic. The red curve separates the HIV-endemic and HIV-free regions. The blue line separates risk behaviour-endemic and risk behaviour-free regions.

doi:10.1371/journal.pone.0062321.g004
varies with \( \omega \), we set \( \lambda_u \) to 0.33 year\(^{-1} \) and \( \beta_0 \) to 0.10 year\(^{-1} \). In this case, \( \tau_{c} \) is approximately 0.46 and 0.42, when \( \omega \) is 0.05 year\(^{-1} \) and 0.20 year\(^{-1} \), respectively. More generally, the dependence of the threshold HAART coverage \( \tau_{c} \) on \( \beta_0 \) and \( \omega \) is shown as a surface plot in Figure 10. In this plot we can see that as \( \beta_0 \) increases, eventually \( \tau_{c} \) becomes zero. This occurs when \( \frac{d\tau_{c}}{d\beta_0} \) is strictly negative or in other words, \( \tau_{c} \) decreases with \( \beta_0 \) over the entire interval 0 to 1.

In summary, we find that in the presence of HAART optimism, HIV incidence and prevalence may either increase or decrease with increasing HAART coverage. For HIV prevalence, we demonstrated the existence of a threshold level of HAART coverage, below which prevalence increases with expanded HAART coverage, but above which HAART expansion overcomes the impact of HAART optimism and prevalence decreases with increasing HAART coverage. Likewise, we also demonstrated that there exists a threshold level of HAART coverage for HIV incidence. The threshold values for HAART coverage are different for prevalence and incidence. This difference stems from the fact that increasing HAART coverage reduces mortality among the HIV-positive population, which exerts upward pressure on prevalence, but not on incidence.

The relationship between the threshold levels for HAART coverage and the HAART optimism parameters is nonlinear and exhibits some counterintuitive properties. For example, we find that for some values of \( \beta_0 \) and \( \omega \), one or both of the threshold levels of HAART coverage may decrease slightly with increasing HAART optimism intensity \( \omega \). Of more significance, we find that the HAART coverage thresholds only increase significantly with HAART optimism intensity \( \omega \) when both \( \omega \) and the risk propagation coefficient \( \beta_0 \) are very small. This scenario is unlikely to be epidemiologically relevant because at these levels of the risk behaviour parameters the HIV epidemic is barely sustainable.

The HIV Epidemic in Vancouver’s Downtown Eastside
In the 1990s, the inner-city Downtown Eastside neighbourhood of Vancouver experienced the most severe HIV epidemic in North America. HIV transmission was primarily driven by syringe sharing among injection drug users. Intensive harm reduction and efforts to expand HAART coverage were followed by a reduction in HIV incidence [9,70]. However, HIV prevalence continues to be relatively high in the neighbourhood and HIV transmission has not been fully contained. We examine the combined HIV prevention impact of HAART and harm reduction, as well as the potential influence of HAART optimism in this setting.

**Parameter values for the downtown eastside.** The parameters \( \beta_{sac} \), \( \lambda_u \), and \( \tau_{c} \), which were previously treated as free, are now set to values specific to the Downtown Eastside. The parameters \( \beta_{sac} \) and \( \lambda_u \) are influenced by complex behavioural factors that are difficult to measure. However, equations (13) link parameters to measurable demographic properties of the epidemic, such as the fraction of the population that is infected with HIV and the fraction of the population that is engaged in risk behaviour. Demographic data for the Downtown Eastside are

![Figure 5. Equilibrium stability regions with fixed risk behaviour propagation rate.](image)

The red curve, which represents the equilibrium HAART coverage needed for extinction of the HIV epidemic as given by equation (20), separates the HIV-endemic and HIV-free regions. For \( \lambda_u \) less than approximately 0.12 year\(^{-1} \), the HIV epidemic is not sustained. For \( \lambda_u \) greater than approximately 0.32 year\(^{-1} \), the epidemic cannot be eliminated by HAART in the chronic HIV phase alone. The risk behaviour propagation rate is fixed at \( \beta_{sac} = 0.059 \) year\(^{-1} \).

doi:10.1371/journal.pone.0062321.g005
available for 1999 [68]. This study reports that the size of the Downtown Eastside community was 19,015 individuals and that there was no significant net migration of HIV-positive and HIV-negative individuals. Furthermore, HIV prevalence was approximately 7%, approximately 20% of HIV-positive individuals were receiving HAART, and an estimated 5000 injection drug users were engaged in risk behaviour [68]. By assuming that the HIV epidemic in the Downtown Eastside is reasonably close to equilibrium, then we can use results from the study [68] to solve for $\beta_{inc}$, $\lambda_u$, and $\tau_e$.

First we solve for $\tau_e$. At the HIV endemic equilibrium in equations (13), the fraction of the total HIV-positive population that is on HAART is

$$\sigma_2 = \frac{\tau_e \sigma_2}{a_2 + c_2} = \frac{\tau_e \rho_u}{\delta_u + \rho_u}.
$$

(30)

Substituting for $\delta_e$ from equation (7) and solving for $\tau_e$ gives

$$\tau_e = \frac{(\delta_u + \rho_u)\sigma_2}{\rho_u + (\delta_u - \delta_{gut})\sigma_2}.
$$

(31)

Therefore, using 0.20 as the estimate for $\sigma_2$ and substituting values for the other parameters from Table 1 gives approximately 0.20 for $\tau_e$.

To solve for $\lambda_u$, we use the equation for $s_2$ in (13), along with the equations (14) and (15). The result is that

$$\lambda_u = \frac{(1 - \tau_e)\rho_u + \tau_e \delta_{gut}}{s_2 [(1 - \tau_e)(2\delta_u + \rho_u) + \tau_e \delta_{gut})].
$$

(32)

The study in reference [68] estimates $s_2$, the fraction of the population that is engaged in injection drug use risk behaviour, but is HIV-negative, as $\frac{3000}{19815} \approx 0.15 \approx 0.18$. Substituting this value for $s_2$ into equation (32), along with $\tau_e \approx 0.20$ and the values of the other parameters taken from Table 1 gives $\lambda_u \approx 0.33$ year$^{-1}$. Harm reduction efforts in the Downtown Eastside have likely reduced $\lambda_u$ significantly from its 1999 value. A drop in syringe sharing behaviour by a factor of approximately three between 1999 and 2007 has been reported [71]. We explore the impact of this reduction in the subsection below.

The parameter $\beta_{inc}$ is obtained by setting HIV prevalence $p_2$ from equation (22) to 0.07 and solving for $\beta_{inc}$. The result is $\beta_{inc} \approx 0.059$ year$^{-1}$. The Downtown Eastside values for all parameters in the model are listed in Table 1.

Evaluating the impact of harm reduction and HAART. Harm reduction programs have the effect of reducing $\lambda_u$, the infectivity of undiagnosed HIV-positive individuals. We examine the impact of expanded HAART coverage on equilibrium HIV prevalence and incidence post-1999 in the Downtown Eastside. All parameters other than $\lambda_u$ are set equal to their Downtown Eastside values in Table 1. The risk propagation coefficient $\beta_{inc}$ is fixed at 0.059 year$^{-1}$ and does not depend on HAART coverage $\tau_e$ in this subsection.

The value of $\lambda_u$ in the Downtown Eastside is likely to have decreased since 1999 as a result of harm reduction. The graphs in Figure 11 show the impact on HIV incidence and prevalence of reducing $\lambda_u$ from 0.33 year$^{-1}$ to 0.13 year$^{-1}$ in increments of 0.05 year$^{-1}$. When $\lambda_u = 0.28$ year$^{-1}$, elimination of the epidemic is theoretically possible by expansion of HAART to 100% coverage of the population in the chronic phase of HIV infection. The inequality (19) gives the threshold value of $\lambda_u$, above which the epidemic cannot be eliminated solely by increasing HAART coverage for individuals in the HIV chronic phase. In the absence of HAART optimism, this threshold value for $\lambda_u$ in the Downtown Eastside is approximately 0.32 year$^{-1}$. Further reduction of $\lambda_u$ beyond this threshold improves the effectiveness of HAART expansion in containing the epidemic. For example, if $\lambda_u$ were 0.13 year$^{-1}$ in the Downtown Eastside, then both equilibrium prevalence and incidence would decrease significantly when HAART coverage exceeds 50% and equilibrium prevalence would be reduced to zero when HAART coverage exceeds approximately 65%. These results highlight the importance of combining harm reduction and HAART expansion to reduce HIV incidence and prevalence.

HAART optimism in the downtown eastside. HAART optimism in the Downtown Eastside is modelled by assuming that the risk propagation coefficient $\beta_{inc}$ depends on HAART coverage.
according to equation (27). For different values of the intensity of HAART optimism $\omega$, the resulting dependence of HIV incidence and prevalence on HAART coverage are plotted in Figure 12. Parameter values for these plots are taken from the Downtown Eastside values in Table 1. For each value of $\omega$ in these plots, the value of $\beta_0$ is chosen so that $\beta_{\text{inc}}$ is equal to 0.059 year$^{-1}$ when $r^* = 0.20$.

The graphs in Figure 12 show that in 1999 HAART optimism would have caused both HIV incidence and prevalence in the Downtown Eastside to increase significantly. The reason for this is that the relatively high untreated HIV chronic phase infectivity $\lambda_w$ would have caused any increase in risk behaviour to drive up HIV incidence and prevalence significantly. However, for incidence, there is no empirical evidence for such an increase after HAART was introduced, on the contrary, incidence was found to decrease [9,70]. This implies that if HAART optimism existed at all, it could have played only a minor role in the HIV epidemic in Vancouver’s Downtown Eastside.

Figure 8. Threshold level of equilibrium HAART coverage for HIV prevalence. This surface plot shows the threshold level $r^*$ of equilibrium HAART coverage $r^*$ above which equilibrium HIV prevalence $p^*$ decreases with increasing HAART coverage, plotted as a function of $\beta_0$, the value of the risk propagation coefficient in the absence of HAART, and the intensity of HAART optimism $\omega$. The rapid decrease in the $r^*$ surface for small values of $\beta_0$ and $\omega$ occurs because for these values of $\beta_0$ and $\omega$ the HIV epidemic can be extinguished for large enough values of the equilibrium HAART coverage $r^*$.

doi:10.1371/journal.pone.0062321.g008
Figure 9. Equilibrium HIV incidence as a function of equilibrium HAART coverage, with HAART optimism. Equilibrium HIV incidence $i_0$, given by equation (25), is shown for three different values of the intensity of HAART optimism $\omega$ and the value of the infectivity in the absence of HAART optimism $\beta_0$.

doi:10.1371/journal.pone.0062321.g009

Discussion

Main Findings

Using a mathematical model for the spread of HIV driven by risk behaviour as a social contagion, we mapped possible outcomes for the HIV epidemic in the context of expanding HAART coverage. We focused specifically on: (1) the potential for Treatment as Prevention to control the HIV epidemic in the context of changing risk behaviours; (2) the combined effects of HAART expansion and interventions targeting risk behaviours; and (3) the impact of HAART optimism.

Our results show that Treatment as Prevention has the potential to be a powerful strategy for controlling the spread of HIV. We find that comprehensive programs securing high HAART coverage could substantially reduce HIV incidence and potentially eliminate the epidemic. While returns on prioritizing high HAART coverage are likely to be great, they do not increase linearly and are influenced by the prevailing risk conditions. Consequently, when high HAART coverage is supported by interventions targeting risk behaviour, these interventions work synergistically to drive the epidemic to lower levels of equilibrium incidence and prevalence. Benefits of interventions accelerate as we approach elimination of the HIV epidemic.

It has been argued that gains due to increased HAART coverage could be undermined by an increase in risk behaviour induced by HAART optimism [4,17,18]. Is this inevitable and if not, under what circumstances does it occur?

The dynamic representation of risk behaviour in our model enables us to take a unique approach to examining the potential epidemiological impact of HAART optimism in the population as a whole. In our model, the spread of risk behaviour in the community has two components: a baseline risk behaviour unrelated to HAART coverage and an additional risk behaviour arising from the intensity of HAART optimism. For low levels of HAART coverage, we find that the preventive benefits of HAART can indeed be overwhelmed by the negative impact of HAART optimism, leading to an increase in equilibrium incidence and prevalence. However, typically there exists a threshold above which the preventive benefits of HAART gain the upper hand and further increases in HAART coverage reduce equilibrium incidence. Likewise, there exists a similar HAART coverage threshold for equilibrium prevalence. These threshold HAART coverage levels have important public health implications.

It is more informative to draw conclusions from HIV incidence, because the effect of HAART on prevalence is confounded by its impact on the life expectancy of treated patients. Equilibrium HIV incidence declines for increasing HAART coverage above the threshold. For high values of the baseline risk behaviour $\beta_0$, the threshold is constant zero. In this case, expansion of HAART will consistently yield public health benefits, regardless of the presence of HAART optimism. For lower values of the baseline risk behaviour, the threshold increases initially with the intensity of HAART optimism, but ultimately levels at a value no higher than approximately 65%. Although for this case low values of HAART coverage can cause an increase in incidence, it should be possible to overcome the effects of HAART optimism at attainable levels of HAART coverage.

The Vancouver Downtown Eastside is an example of a setting, which today can be considered an environment with relatively low levels of risk behaviour propagation according to our definition. Intensive harm reduction beginning in the early 1990s resulted in a major shift in injection behaviours in this neighbourhood. Most IDU are now aware of the risks of HIV infection and avoid sharing syringes [72]. Applying our model to this setting, we showed that HIV incidence and prevalence would have had to increase substantially with behavioural disinhibition if HAART optimism existed in the community. This has not been observed empirically. In fact, HIV incidence has decreased over time [9] and a recent study found no evidence of HAART optimism in the Downtown Eastside [73]. Whether or not social norms operate to limit the resurgence of risk behaviours in environments with low levels of risk behaviour propagation, such as the Downtown Eastside, HAART optimism may exist in this context and addressing risk behaviours is likely to be important.

Our results also suggest that the potential HIV prevention benefits of HAART in high-risk environments could be substantial, despite any behavioural shifts due to HAART optimism and this could be a factor to consider in planning interventions. Sufficiently aggressive expansion of HAART access in these settings would reduce incidence. Furthermore, additional harm reduction, behavioural or substance use interventions are expected to boost and stabilize these effects. This could have relevance for some IDU and MSM groups. HIV infection rates among IDU in parts of Eastern Europe and Central Asia continue to be very high today due to high rates of syringe sharing and unprotected sex [74]. Containing the HIV epidemic in MSM groups where risk...
behaviours continue to be frequent, remains a challenge and resurgent epidemics are seen [92,75]. Our model suggests that achieving and maintaining high HAART coverage in these types of settings could have major benefits in reducing incidence, especially in combination prevention strategies. HAART optimism is not likely to substantially diminish these benefits.

![Figure 10](image)

**Figure 10.** Threshold level of equilibrium HAART coverage for HIV incidence. This surface plot shows the threshold level $\tau$ of equilibrium HAART coverage $\tau$, above which equilibrium HIV incidence $i_2$ decreases with increasing HAART coverage, plotted as a function of $\beta_0$, the value of the risk propagation coefficient in the absence of HAART, and the intensity of HAART optimism $\omega$. The rapid decrease in the $\tau$ surface for small values of $\beta_0$ and $\omega$ occurs because for these values of HAART and $\omega$ the HIV epidemic can be extinguished for large enough values of the equilibrium HAART coverage $\tau$. The region for which the $\tau$ surface is identically 0 corresponds to $i_2$ being a strictly decreasing function of $\tau$, for $0 \leq \tau \leq 1$.

doi:10.1371/journal.pone.0062321.g010

![Figure 11](image)

**Figure 11.** HIV prevalence and incidence in Vancouver’s Downtown Eastside as a function of equilibrium HAART coverage for different values of untreated HIV chronic phase infectivity. The impact of equilibrium HAART coverage $\tau$ on HIV prevalence and incidence is shown for different values of the untreated chronic phase infectivity $\lambda_0$. These graphs illustrate the potential impact of reducing $\lambda_0$ through harm reduction programs. For these plots, it is assumed that there is no HAART optimism.

doi:10.1371/journal.pone.0062321.g011
Strengths and Limitations

One of the strengths of our method is that we are able to draw broadly valid qualitative conclusions about the epidemiological effects of changing risk behaviours. This has been possible because the model is simple enough to be mathematically tractable, which enabled us to prove the existence and describe the nature of all equilibria. Another strength of our approach is that we did not restrict much of our analysis to parameter values for the HIV epidemic in any specific setting, but examined model behaviour over the entire parameter space.

Through linking risk behaviour dynamically to HIV transmission, we were able to incorporate aspects of social influence in the production of HIV risk into the model. Representing risk behaviour as a social contagion has been used for modelling health issues such as obesity, smoking, drug use, and alcohol consumption [76]. The dynamic risk behaviour component of the model allowed us to study the effects of HAART optimism on the HIV epidemic in a novel way. Our analysis of risk behaviour encompasses the wider population, including both HIV negative and HIV infected subpopulations.

Although a strength, the generality of our analysis is also a limitation because it forces us to focus on the equilibrium states of the model. Non-equilibrium analysis can provide useful information on the short-term response of the epidemic to interventions. However, this type of analysis must consider a specific setting, because knowledge of the current state of the epidemic is necessary to generate reliable projections.

Although risk behaviour interacts with other components of the model dynamically, our treatment of it is still quite simplistic. Propagation of risk behaviour is assumed to occur only through peer interaction. Moreover, we have assumed that different types of risk behaviour and their ensuing transmission channels can be aggregated into a single generic risk behaviour and transmission channel. Furthermore, once risk behaviour is initiated, it can never stop in the model, making our predictions more pessimistic than might be the case with more realistic assumptions about risk behaviours.

HAART optimism is incorporated into the model by assuming that higher levels of HAART coverage lead to an increase in the risk behaviour propagation rate. In the absence of data, we assume a simple linear relationship between these two parameters. Even with a more general dependence of the risk behaviour propagation rate on HAART coverage, a linear relationship would still be a good approximation as long as there are no large changes in HAART coverage. Mathematically, this means that the linear relationship corresponds to the first two terms in a Taylor series expansion of the risk behaviour propagation rate as a function of HAART coverage.

We neglected drug resistance to retain model simplicity. This assumption is supported by findings in British Columbia where rates of acquired and primary resistance have been declining since 1996, despite marked expansion of HAART [77]. Due to this assumption, our conclusions should be interpreted with caution in the context of settings where drug resistance is an important problem. Other simplifications that we make include omitting details on risk groups, variability in risk behaviours, social network features, testing strategies and combining the latent and AIDS stages of infection into a single chronic infection phase.

Conclusions

Our analysis suggests that Treatment as Prevention has significant potential for containing the HIV epidemic. However, substantial gains in reducing HIV incidence and prevalence are only achieved at or near critical coverage levels for HAART or other interventions. Therefore, determining critical HAART coverage levels may help develop more effective Treatment as Prevention programs. Prioritizing sufficiently high HAART coverage and incorporating interventions to reduce risk behaviour will amplify the preventive impact, possibly even eliminating the HIV epidemic. Local epidemiological conditions play a strong role in how much HAART coverage is required for controlling the HIV epidemic. The requirement for sufficiently high HAART coverage may extend to undiagnosed infections, necessitating comprehensive testing to be a part of a total prevention package. While HAART optimism promotes the growth of the HIV epidemic, this can be controlled by increasing HAART coverage and behavioural interventions. Therefore, our findings do not confirm previously predicted overwhelming negative impacts of HAART optimism on the benefits of Treatment as Prevention. In conclusion, an important public health lesson that emerged from this modelling exercise is that with Treatment as Prevention, half measures do not necessarily mean half results, while full commitment to comprehensive
programs has substantial potential for controlling the HIV epidemic.

Supporting Information

Mathematical Supplement S1. (PDF)

References


