HIV and the trade-off hypothesis: considering the effect of HAART

Author: David Hodgson

Supervisors:
Dr. Katherine Atkins
Dr. Stephane Hue
Dr. Jasmina Panovska-Griffiths

April 24, 2015
1 Introduction

1.1 AIDS and the discovery of HIV

Throughout the summer of 1981, clinicians in New York and San Francisco became addled by an emerging medical phenomenon: clusters of young, previously healthy homosexual men exhibiting a sudden decline in health owing to the contraction of a combination of very rare diseases [27, 5]. Of these observed diseases, two were significantly prevalent; one, an aggressive form of cancer, known as Kaposi’s sarcoma, and the other a type of pneumonia caused by the opportunistic yeast-like fungus, *Pneumocystis jirovecii*. It was not until a few months later that it became apparent that these patients suffered from an unspecified type of deficiency in cell-mediated immunity, as a healthy person would expunge the pathogens which caused such infections with ease [34, 29].

Over the next year different, distinct groups within the American population began to contract similar aggressive diseases as the originally observed homosexual cohorts; what unified all these persons however, was an aberrant immune response, leading to the United States Center for Disease Control and Infection (CDC) coining the term “acquired immunodeficiency syndrome” (AIDS) to describe the infections. As groups which exhibit similar behaviour initially expressed the symptoms of AIDS, speculation that specific behavioural and environment mechanisms give rise to AIDS initially became the popular consensus, whereas from a medical prospective, it seemed more plausible that an undetermined, infectious pathogen was the cause of the infection [24]. Evidence of this came to fruition several years later in 1984 when reports of a hematologically isolated retrovirus from patients with AIDS were published in various medical journals [10, 37, 32], thus providing compelling evidence that AIDS is the result of a virus, called HIV-1. This also initiated the development of a large-scale screening method for infection.

In subsequent years, cases of AIDS, thus HIV-1 infections, were being reported in persons who were not considered to be in the previously defined high risk groups, and there was a rapid transmission of the disease to places including sub-Saharan African and eastern Asian [4]. Since, the pandemic of HIV infection has become a defining medical and public health issue of our generation and is ranked as one the most significance infectious disease scourges in history [22]. Commensurate to the magnitude of the HIV pandemic has been the scientific onslaught to delineate the virology, pathogenesis and epidemiology of the disease.

1.2 Pathogenesis of HIV

As a lentivirus, HIV has three distinct phases of infection, the initial acute phrase occurs 2–4 weeks after initial infection and infected persons generally develop severe flu-like symptoms [36]. Following this phase the virus enters a asymptomatic period of infection which can last between two and twenty years if untreated [30]. Though asymptomatic the virus is still very active at this time and eventually the virus enters the final stage of infection, AIDS. At this point HIV has infected a large amount of immune cells and as such the body is no longer capable of generating a suitable immune response, leading to opportunistic infections occurring such as Kaposi’s sarcoma. Once a person has progressed to AIDS, the life expectancy of a patient is no longer than three years [16].

Immune cells are produced as part of the body’s response initiated upon the detection of a foreign entity within the body. HIV infects a type of immune cell, T cells and using specific receptors found on its surface — CD4, CCR5 and CXCR4 — HIV infiltrates the cell and then manipulates the replicative machinery within to produce more mature virions [15, 31]. In other words, the virus takes advantage of the human body’s immune response mechanism, utilising the increased availability of immune cells, leading to an accelerated progression of the viral infection. This explains the paradoxical phenomenon observed within the body of a patient suffering from AIDS whereby the individual has an obvious immune deficiency, but simultaneously has an aberrantly activated immune response.
Modern management of HIV.

2.1 The Pandemic Today

In 2013 it was estimated that 32.1 (29.1 – 35.2) million people were living with HIV-1, with 2 (1.7 – 2.4) million new infections occurring in that year alone [2]. The epidemiological picture of HIV varies across the globe. In sub-Saharan Africa for example, a region which accounts for 69% of all HIV infections worldwide, 49 out of 1,000 people are infected and heterosexual transmission accounts for around 70% of new infections [2]. In the UK and Canada, where HIV prevalence is relatively low (3.2 and 2.1 out of 1,000 respectively), men who have sex with men (MSM) are the largest risk group, accounting for 46.6% and 54% of all diagnoses in 2013 respectively [20, 3]. In Ukraine however, where HIV is most prevalent in Europe (10 out of 1,000), people who inject with drugs (IDUs) account for nearly 40.1% of all new HIV infections, a mode of transmission which accounts for less than 1% of HIV infections in the UK [1]. The heterogeneity between the incidence in risk groups, and modes of transmission across the world makes the control of the HIV pandemic difficult, requiring a combination of global health and country specific policies to try to curtail transmission.

2.2 HAART.

One of the essential global health policies which has shown the greatest success in HIV prevention is the use of Highly Active Antiretroviral Treatments (HAART). HAART is the name given to the uptake of multiple antiviral drugs which impede the progression of HIV. Though incapable of eradicating the virus from within a host, [19, 40, 39] they have shown a great success of two ways. Firstly, patient uptake of HAART has significantly beneficial results with regards to the prognosis of an individual infected with HIV; effectively stagnating the progression of the disease, allowing the time until AIDS progression becoming longer than the lifetime of most patients [14]. HAART also reduces the infectiousness of an infected person [16], so from an epidemiological prospective it has proven beneficial in reducing incidence. It is for these reasons that global health initiatives have promoted a “test and treat” programme in some regions whereby persons who are diagnosed with HIV are placed on HAART shortly after diagnosis in order to reduce the infectiousness of the patient as soon as possible and also improve their prognosis. Before this the general medical consensus was to wait until an infected person’s CD4 cell count fell below a threshold (~350 cells/µl) to delay the emergence of drug resistance as much as possible. Though the results of implementing test and treat have so far shown an increases prognosis in patients and a reduction in transmission in comparison to other treatment strategies, there is much debate over the cost-effectiveness of such treatments and the effect on drug resistance [18].

2.3 Viral Evolution.

The virulence of a pathogen the severity of the disease it causes, in the case of HIV, a highly virulent strain is said to shorten the period of asymptomatic infection, thus leading to AIDS and subsequently death more quickly. That is, for the purposes of this study, virulence is time it takes to kill the host. As virulence varies between genetically different strains of pathogens, these strains are subject to natural selection. Understanding how and why the virulence of a pathogen evolves is arguably one of the most important questions in evolutionary biology [7]. It is well established that the way in which the virus is transmitted has a large effect on the level of virulence that will evolve [13, 21, 12, 33]. Effectively, it is not possible for a parasite to increase the duration of an infection without paying a cost. For example, if a strain is to evolve to become more transmissive, the evolutionary cost of this is an increase in virulence, in accordance to Darwinian theory of natural selection. Therefore it is expected that a highly virulent pathogen must be highly transmissible, and a pathogen which has a low virulence may not be transmissible but can survive within a host for a longer time. However it is important to consider that pathogen evolutionary trajectory will be subject to different physical constraints. This concept of virulence and transmissibility playing off against one another is known as the “Trade-off hypothesis” and is one of the most popular and controversial theories that try to delineate the evolution of pathogenic virulence for lethal pathogens.
Generally an intermediate level of virulence and transmissibility is compromised, the result is referred to as the optimal virulence.

In the example of HIV, it has been suggested that there may exists a negative relationship between the duration of asymptomatic infection and viral load, and a positive relationship between transmissibility and viral load. This leads to the observation that the transmission potential is optimal for an intermediate viral loads, showing a trade-off between duration of infection and transmissibility as viral load (or virulence) increases, assuming a constant rate of transmission throughout the period of infection [25].

If virulence is an evolvable trait, one asks, does the virulence of a pathogen change when a new environment and if so how? For example, it has been shown that the use of vaccines diminishes selection against virulent pathogens. The subsequent evolution leads to higher levels of intrinsic virulence and hence to more severe disease in unvaccinated individuals [28]. A direct example of this is Marek’s Disease virus (MDV) in industrialised poultry agriculture. It has been shown that MDV has become more virulent over the last few years owing to the effect of intense vaccinating, prolonging of the infectious period of more virulence strains of MDV [9].

This study asks whether the effect of administrating HAART can lead to the evolution of more virulent strains of HIV. Using an estimate for the virulence of the virus (Section 3.1), a function measuring the time of undiagnosed asymptomatic infection will be estimated and then the transmission potential will be calculated as done in previous papers [25]. By comparing the optimal reproductive fitness with and without HAART, I will evaluate the impact of HAART on HIV virulence selection.

3 Methods.

3.1 Viral Load

The amount of virus within an infected person is known as their viral load, which can be measured as of the number of HIV virions per millilitre of peripheral blood. It fluctuates between patients and within a host throughout the different stages of infection, peaking during the acute phase, before an antibody response can take effect. During the asymptomatic period of infection however, the viral load stays relatively constant, and is known as a person’s “set-point” viral load, or SPVL. The SPVL has been shown to be an extremely useful way of quantifying the severity of a HIV infection. Patients with higher SPVL will be more infectious [38, 23], but will generally progress to death faster [17, 35]. Therefore the SPVL is a good indication of the virulence of the infection. For the purpose of this project it can be assumed that a patient with a higher SPVL has a more virulent strain of the virus.

Previous Work

3.2 Functions in Fraser et al paper

I will set out the definitions and functions previously developed in Fraser et al paper, [25].

3.2.1 Duration of asymptomatic infection as a function of SPVL

\( D(V) \) is a continuous function which gives the mean duration of asymptomatic infection as a function of SPVL, \( V \), given no external medical treatment has been administered. The period of asymptomatic was defined by Fraser et al. to be the time between 6 months after first seropositive sample and the first AIDS-defining symptoms. This definition is maintained throughout this study.
Due to the biological expectation that the asymptomatic duration will decrease as the SPVL increases, the mean period of asymptomatic infection was modelled by a decreasing Hill function:

$$D(V) = D_{\text{max}} \frac{D_{50}^k}{D_k + D_{50}^k}$$  

(1)

where $D_{\text{max}}$ is the maximum duration of asymptomatic infection in years, $D_{50}$ is the SPVL at which the duration is half its maximum and $D_k$ is the steepness of the decrease (Hill coefficient.)

In order to gain the full profile of duration of asymptomatic period for a given SPVL as oppose to just a mean, the time of the asymptomatic infection was assumed to follow a Gamma distribution.

**The Gamma Distribution** If a random variable $X$ follows a Gamma distribution, with shape parameter $\rho$ and scale parameter $\theta$, $X \sim \text{Gamma}(\rho, \theta)$, then $E(X) = \rho \theta$, and the cumulative density function (CDF) is given by

$$\frac{\gamma(\rho, \frac{T}{\theta})}{\Gamma(\rho)}$$  

(2)

where $\gamma$ is the lower, incomplete gamma function and $\Gamma$ is the standard gamma function. One can rewrite this CDF in terms of the expectation (or mean) by noticing that $\frac{T}{\theta} = \frac{\rho T}{E(X)}$.

Thus the probability a person is still asymptomatic at time $T$ is a gamma distribution with mean $D(V)$ and shape parameter $\rho$. Thus, the probability that an individual is still alive $T$ years after infection with virus strain with SPVL, $V$ is given by:

$$P(V, T) = 1 - \exp(-\beta(V)T)$$  

(3)

After this model was fitted to SPVL data for the Amsterdam and Zambian cohorts (Appendix 1), it was shown that the parameter values that maximise the likelihood are: $D_{\text{max}} = 25.4$ years, $D_{50} = 3.058$ copies per millilitre of blood, $D_k = 0.41$ and $\rho = 3.56$.

### 3.2.2 Transmission Rate as a function of SPVL

$\beta(V)$ is a continuous function which gives the mean infection hazard (rate of infection per infected contact per year) of infecting another person as a function of the SPVL, $V$. For a single susceptible individual, $\beta(V)$ is the mean transmission rate.

Due to the biological expectation that the transmission rate will increase as the SPVL increases, Fraser et al defined the transmission rate as increasing Hill function:

$$\beta(V) = \beta_{\text{max}} \frac{\beta_{50}^k}{V \beta_k + \beta_{50}^k}$$  

(4)

where $\beta_{\text{max}}$ is the maximum infection per year, $\beta_{50}$ is the SPVL at which the transmission potential is half its maximum and $\beta_k$ is the steepness of the increase (Hill coefficient.)

To gain a full profile of the transmission potential of a person over a time $T$ for a given SPVL, the exponential distribution is used. Assuming a constant hazard of transmission to a single sexual partner throughout the entire asymptomatic stage, Fraser et al defined the probability that the sexual partner was to become infected after $T$ years from the first infection with a SPVL, $V$ to be:

$$P(V, T) = 1 - \exp(-\beta(V)T)$$  

(5)
After this model was fitted to SPVL data for the Amsterdam and Zambiam cohorts (Appendix 1), it was shown that the parameter values that maximise the likelihood are: \( \beta_{max} = 0.317 \) per year, \( \beta_{50} = 13,938 \) copies per millilitre of blood and \( \beta_k = 1.02 \).

### 3.2.3 Transmission Potential as a function of SPVL.

\( TP(V) \) was defined to be a continuous function which gives the mean transmission potential as a function of the SPVL. The mean transmission potential can be seen as the mean number of persons a patient with a SPVL \( V \) will infect over the entire asymptomatic period.

The transmission potential, \( TP(V) \) is defined to be the product of the transmission rate and the duration of asymptomatic infection;

\[
TP(V) = D(V) \beta(V)
\]

#### New Work

### 3.3 Data extraction of confidence intervals from Fraser et Al

I extracted data for the duration of asymptomatic infection, transmission rate and transmission potential from the plots published in [25] using image processing software (ImageJ, [http://imagej.nih.gov/ij/](http://imagej.nih.gov/ij/)). A discrete set of SPVL was used, this set will be referred to as \( V^* \).

\( D(V^*), \beta(V^*) \) and \( TP(V^*) \) and the derived 95% confidence intervals at each value of \( V^* \) are plotted. (Figure 3.1, Figure 3.2 and Figure 3.3)

![Plot of mean and confidence intervals for duration function \( D(V) \).](image)

**Figure 3.1:** Graphic showing the mean duration of asymptomatic period, \( D(V) \), with confidence intervals for various SLVP, \( V^* \), extracted from Fraser et al.
### 3.4 Estimating distribution for $\mathcal{D}(V)$, $\beta(V)$ and $TP$

In Section 3.2 the three mean value functions were defined, $\mathcal{D}(V)$, $\beta(V)$ and $TP(V)$ as a function of SPVL, $V$. Then in Section 3.3 I extracted the confidence intervals for a range of SPVL, $V^*$. Using the both the mean value and the upper and lower confidence intervals, I fit a Gamma distribution for the duration of asymptomatic infection, the transmission rate and the transmission potential for each SPVL, $V^*$.

#### 3.4.1 Model Fitting.

I used the following method to fit the Gamma distribution; using the period of asymptomatic infection as an example. For a SPVL, $V^*$, first the average value is determined; $\mathcal{D}(V^*)$, one then proposed that the distribution is given by a gamma distribution of the form $\text{Gamma}(\alpha, \frac{\mathcal{D}(V^*)}{\alpha})$, due to the property that $E(\text{Gamma}(\alpha, \frac{\mathcal{D}(V^*)}{\alpha})) = \mathcal{D}(V^*)$. In order to determine the value of $\alpha$, the lower and upper bounds of the confidence intervals at SPVL $V^*$ are used. If $X \sim \text{Gamma}(\alpha, \frac{\mathcal{D}(V^*)}{\alpha})$, and $x_{v_1}$, $x_{v_2}$ are the lower and upper confidence intervals respectively,
then minimising the objective function;

\[ R(\alpha) = \sum_{i=1,2} (F_i(\alpha) - x_{v_i})^2 \]  

(7)

where \( F_i(\alpha) \) is the CDF of the Gamma distribution at the point \( x_{v_i} \), one finds a value for \( \alpha \). The value for \( R^2 \), the different between model and data is also calculated.

### 3.5 Including HAART treatment.

In order to implement the effect of HAART within our previously described model, I extend the previously defined duration function, \( D(V) \), which determines the duration of infectiousness for a given patient with SPVL, \( V \). I define two durations under drug treatment for a given SPVL; one, the time until the asymptomatic period has ended and the patient has thus progress to AIDS under no treatment (\( \mathcal{D} \)) and two, the day from infection until a person receives HAART (\( T_d \)). The minimum of these two durations will be taken and thus the duration of asymptomatic infection for patients who are treated, (\( \mathcal{D}t \)), is determined.

Consistent with “test and treat” policies it can be assumed that the time until HAART treatment is the same as the time until detection. Using data from Public Health England, I quantified the annual rate of HIV detection, \( \tau \):

\[ \tau = \frac{\text{The number of people diagnosed with HIV in a given year}}{\text{The number of people undiagnosed with HIV within the population}} \]  

(8)

thus, it is possible to estimate the time until diagnosis by assuming that \( T_d \sim \text{Exp}(\tau) \). A mathematical definition for \( \mathcal{D}t(V,\tau) \) can then be defined to be:

\[ \mathcal{D}t(V,\tau) = \min\{\mathcal{D}(V),\mathcal{E}(T_d)\} \]  

(9)

Analogously, the mean value of the transmission potential including the possibility of heart HAART, in a continuous function of \( V \) and \( \tau \) and is defined by:

\[ \mathcal{TP}(V,\tau) = \mathcal{D}u(V,\tau)\beta(V) \]  

(10)

I calculated \( \tau \) for three groups within the UK population; men who have sex with men (MSM), black-African women (BAW) and heterosexual men (HM). The rate of infection detection for each subpopulation is denoted \( \tau_{\text{MSM}}, \tau_{\text{BAW}} \) and \( \tau_{\text{HM}} \) respectively. The prevalence of the infection is analysed in each subpopulation, by using data given in [20] whereby the number of persons diagnosed and living with HIV is documented, and through complex CD4 back-counting methods, the number of people living with HIV who are undiagnosed are also estimated [11, 6]. The results are summarised in Figure 3.4.

**MSM.** In the UK, the subpopulation with the highest prevalence of HIV infection is men who has sex with men (MSM.) In 2013 it is has been predicted that around 43,500 (40,210 – 48,160) MSM are living with HIV in the UK, 16 (10 – 25)% of whom are unaware of the infection...

**BAW.** In the UK, the subpopulation with the second highest prevalence of HIV infection is black African women. In 2013 it is has been predicted that around 25,060 (22,360-28,870) black African women are living with HIV, 31 (23 – 40)% of whom are unaware of the infection.

**HM.** In 2013 it is has been predicted that around 23,980 (21,610-27,410) heterosexual men are living with HIV in the UK, 27 (18 – 36)% of whom are unaware of the infection.

The number of people diagnosed annually in each category will be analysed from 2004 until 2013. A summary of the figures are shown in Figure 3.5.
Prevalence of HIV in subpopulations in the UK (2013).

Figure 3.4: Bar chart showing the number of diagnosed and undiagnosed HIV MSM and HM living in the UK in 2013. The red bars represent 95% confidence intervals.

**MSM.** New HIV diagnoses in MSM accounted for 54% of all reported diagnoses in 2013, totalling 3,250 new diagnoses. There has been a steady increase in incidence of HIV in MSM over the last 10 years.

**BAW.** There were approximately 950 new diagnoses for BAW in 2013, around 16% of all new diagnosis.

**HM.** There were approximately 968 new diagnoses in 2013. Accounting again for around 16% of all new diagnosis. Both BAW and HM have had a steady fall in the yearly diagnosis of HIV in the last decade.

Following this, the rate of detection $\tau_j$ subpopulation, $j$, is calculated by taking an average of the number of persons diagnosed with HIV annually for the last ten years, then dividing this number by the estimated number of undiagnosed persons in 2013.

I assume that the delay between infection and detection for risk group, $j$, follows an exponential distribution with rate $\tau_j$. This assumption results in a mean time to detection of $\frac{1}{0.3482} = 2.9$ years for MSM, consistent with previous estimates ??.
Figure 3.5: Line graph showing the number of new diagnosis of HIV in MSM, BAW and HM from 2004-2013.

4 Results

4.1 Estimating distributions

4.1.1 Duration of asymptomatic infection

Utilising the model fitting procedure explained in Section 3.4.1, Figure 4.1 shows table of the fitted $\alpha$ parameter values dependent on $V^*$ and the $R^2$ value for the fit. Plots for the resulting cumulative density function for SPVL ($V^*$), compared to the observed values of the confidence intervals to show goodness of fit. Using these fitted distributions one can define a random variable, $D_T(V^*)$ of the time of asymptomatic infection, dependent on SPVL $V^*$ assuming there is no HAART treatment.
For example, when $v = \log_{10}(V) = 3$, the distribution of duration of asymptomatic infection is given by $D_T(10^3) \sim \text{Gamma}(14.2, \frac{D_T(10^3)}{14.2}) = \text{Gamma}(14.2, 1.0986)$.

### 4.1.2 Transmission Rate

Following the same procedure as above, one fits Gamma distributions to the transmission rate for SPVL, $V^*$. Figure 4.2 shows a table of the estimated $\alpha$ values for each SPVL and the goodness of fit given by $R^2$. 

![CDF of the fitted models for Infectiousness](image-url)
Using these fitted distributions, for a set SPVL, $V^*$, the distribution of the transmission potential is found from the random variable $B_\beta(V^*) \sim \text{Gamma}(\alpha, \frac{\beta(V^*)}{\alpha})$.

### 4.1.3 Finding the Transmission Potential using fitting distributions.

To test the accuracy of these distributions, I sampled $10^4$ values of $D_T(V^*)$ and $B_\beta(V^*)$ from the derived Gamma distributions in Section 4.1.1 and 4.1.2 and multiplied for each SPVL, $V^*$, to provide a simulated range of values for the transmission potential at said SVLP (Figure 4.3).

![Mean and confidence intervals for transmission potential](image)

Figure 4.3: The mean and 95\% confidence intervals for the data derived from the data extraction from Fraser paper (blue), and from fitted distributions (red).

### 4.2 Estimating the infection detection rate.

Recalling the definition of the infection detection rate for a given subpopulation, $\tau_j$, given in Section 3.5, and the mean time until diagnosis, $T_d$, one uses the above prevalence and incidence data to find the value of $\tau_j$ for the three subpopulations, $j$.

**MSM** In 2013, for MSM, $\tau_{MSM} = 0.399$ (95\% confidence intervals (0.242, 0.722)), corresponding to a average time until detection to be 2.504 (95\% confidence intervals (1.385, 4.133). Thus, $T_d \sim \text{Exp}(0.399)$.

**BAW** In 2013, for BAW, one gains a value of $\tau_{BAW} = 0.212$ (95\% confidence intervals (0.315, 0.143)), corresponding to a average time until detection to be 4.717 (95\% confidence intervals (3.171, 6.980).) Thus, $T_d \sim \text{Exp}(0.212)$. 

11
Viral Load, Log$_{10}$(V).

Duration, years.

$\mathcal{D}(V, \tau_j)$ for different subpopulations, j

$\tau_{HM} = 0.172$
$\tau_{BAW} = 0.220$
$\tau_{MSM} = 0.399$

No HAART, $\mathcal{D}(V)$

Figure 4.5: Plots of $\mathcal{D}(V, \tau_j)$ for derived values of $\tau_j$.

HM In 2013, for HM, one gains a value of $\tau_{HM} = 0.172$ (95% confidence intervals (0.242, 0.142)), corresponding to an average time until detection to be 5.812 (95% confidence intervals (4.140, 6.949).) Thus, $T_d \sim \text{Exp}(0.172)$.

<table>
<thead>
<tr>
<th>Mean value for $\tau$</th>
<th>Mean</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSM</td>
<td>0.399359</td>
<td>(0.722166, 0.241961)</td>
</tr>
<tr>
<td>BAW</td>
<td>0.21199</td>
<td>(0.31537, 0.143276)</td>
</tr>
<tr>
<td>HM</td>
<td>0.172051</td>
<td>(0.2412554, 0.143909)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mean time until diagnosis.</th>
<th>Mean</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSM</td>
<td>2.50401</td>
<td>(1.38472, 4.13289)</td>
</tr>
<tr>
<td>BAW</td>
<td>4.71721</td>
<td>(3.17088, 6.97954)</td>
</tr>
<tr>
<td>HM</td>
<td>5.81224</td>
<td>(4.13985, 6.94886)</td>
</tr>
</tbody>
</table>

(a) Values for $\tau_j$, yearly detection rate.  (b) Values for the mean time until diagnosis.

Figure 4.4: Tables summarising the mean rate of detection rate and mean time diagnosis with confidence intervals for the three studied subpopulations

4.3 Transmission potential under HAART

4.3.1 Finding the period of undiagnosed asymptomatic infection, for given SPVL, $V^*$

In Section 3.5 the function which models the mean duration of undiagnosed asymptomatic infection as a function of SPVL was defined $\mathcal{D}(V, \tau)$.

Figure 4.5, gives the profiles of this function as a continuous function of $V$ for the three derived values of $\tau_j$ given in Section 4.2. Notice how $\mathcal{D} \to \mathcal{D}$ as $\tau \to 0$.

For the set SPVL, $V^*$, confidence intervals are determined by sampling the random variables $T_d$ (defined in Section 3.5) and $D_T(V^*)$ for a diagnosis rate $\tau$ and taking the minimum of the two. After repeating sampling, one determines confidence intervals for each subpopulation compared to the duration when no HAART is administrated (Figure 4.6, 4.7 and 4.8).
Figure 4.6: Confidence intervals for $\mathcal{D}_u(V^*, \tau_{MSM})$ when $\tau_{MSM} = 0.399$, (red) and the confidence intervals for $\mathcal{D}(V^*)$ (blue) at SPVL, $V^*$.

Figure 4.7: Confidence intervals for $\mathcal{D}_u(V^*, \tau_{BAW})$ when $\tau_{BAW} = 0.220$, (red) and the confidence intervals for $\mathcal{D}(V^*)$ (blue) at SPVL, $V^*$.

Figure 4.8: Confidence intervals for $\mathcal{D}_u(V^*, \tau_{HM})$ when $\tau_{HM} = 0.172$, (red) and the confidence intervals for $\mathcal{D}(V^*)$ (blue) at SPVL, $V^*$.
### 4.3.2 Transmission potential with possibility of HAART.

A function for the mean transmission potential, \( TP(V, \tau) \), including the possibility of HAART was defined in Section 3.5, using the derived values of \( \tau_j \) for populations \( j \). Figure 4.9 plots the profile of this function for each subpopulation.

Given the fitted Gamma distributions for the transmission rate function \( \beta \) at a specific SPVL, \( V^* \) given in Section 4.1.2 and the sampling of the period of untreated function, \( D(V^*) \) given in Section 4.3.1, one can sample each function at said SPVL and multiply the results to find a distribution for the transmission potential, including HAART, at a specific SPVL, \( V^* \). Figures 4.10, 4.11 and 4.12 shows the results for each subpopulations, compared to the case where HAART is no implemented.
3.0 3.5 4.0 4.5 5.0 5.5 6.0
0.0
0.5
1.0
1.5
2.0
2.5
3.0
Viral load, (Log10)
Duration in Years
Plot of mean and confidence intervals for duration function \( \mathcal{TP} \).

Figure 4.11: Confidence intervals for \( \mathcal{TP}(V, \tau_{BAW}) \) when \( \tau_{BAW} = 0.220 \), (red) and the confidence intervals for \( \mathcal{TP}(V^*) \) (blue) at SPVL, \( V^* \).

3.0 3.5 4.0 4.5 5.0 5.5 6.0
0.0
0.5
1.0
1.5
2.0
2.5
3.0
Viral load, (Log10)
Duration in Years
Plot of mean and confidence intervals for duration function \( \mathcal{TP} \).

Figure 4.12: Confidence intervals for \( \mathcal{TP}(V, \tau_{HM}) \) when \( \tau_{HM} = 0.172 \), (red) and the confidence intervals for \( \mathcal{TP}(V^*) \) (blue) at SPVL, \( V^* \).

4.4 General effect of changing \( \tau_j \) on the optimal virulence.

Given a value of \( \tau_j \), the rate of infection detection, can one use the above model to say something about the transmission potential at certain SPVL. Due to evolutionary consideration it is of interest to study the optimal virulence for each value of \( \tau_j \), that is the SPVL that maximises the transmission potential give a rate \( \tau_j \). Figure 4.13 shows the maximum transmission potential for a set of values of \( \tau_j \) and legended is the SPVL at which this maximum occurs.

5 Discussion

In this project we applied previous mathematical methods to explore effect of HAART on the transmission potential of the HIV virus assuming a test and treat policy. Specifically we focused on quantifying the transmission potential of HIV-1 virus as a function of set-point viral load. This poses an important question as if the hypothesis that a viral load that maximises transmission potential occurs and that the mean of the distribution of viral loads coincide because of adaptive evolution, then this study would imply that the use of HAART in a test and treat regime would allow more virulent strains of HIV to be more transmissive, thus increasing the prevalence of virulent
Our results (Figure 4.3) suggest that the viruses which have intermediate SPVL have the highest transmission potential. The maximum of the transmission potential $TP$ is given at the SPVL, $V = \log_{10}4.52$ and this is close to the observed means within two cohorts in Appendix I (given by 4.36 and 4.74 log$_{10}$ copies per millilitre for the Amsterdam and Zambian cohorts respectively). We note that is an important finding that could potentially be the outcome of natural selection acting on HIV-1 virus to maximise its transmission. The distribution of SPVL given in Appendix I for the two cohorts correlates to the mean points for the transmission potential in Figure 4.3 and the maximum transmission potential and means of the two cohorts are particular close which seems to support this finding. The other essential biological observation that must hold for this evolutionary hypothesis to be true is that there is a correlation in SPVL between the transmitter and recipient. Recent research suggests that a proportion of SPVL is inheritable and linked to transmissibility and this suggests that the virus has evolved and may continue to evolve to maintain transmission fitness. However, what proportion of this inheritability comes from the host and and how much comes from the virus is unclear, and estimates tend not to agree [26]. Further work is required to explore this.

Our results in Figure 4.13 suggest that a higher rate of infection detection, $\tau$, leads to lower transmission potentials. But the peak of the transmission potentials occurs at higher SPVL, thus a higher value of $\tau$ leads to a higher optimal virulence. Following the same argument as before, if the transmission potential peaks at higher SPVL, then these higher viral loads will be more transmissible and because there exists a correlation between the transmitter viral loads and recipient viral load, this implies that strains of HIV-1 displaying higher SPVL will become more prevalent over time. However, as HAART is effective at stopping the transmission of the virus the incidence of more virulent viral strains would inevitably fall and as all infected individuals are placed on HAART there is very low mortality rate. The problem occurs with the undiagnosed population, if HIV becomes more virulent then persons newly obtaining the infection will die quicker, which could be devastating in groups such as the black African women and heterosexual men where the rate of infection detection is low.

There are a number of limitations within my model. Mainly, the transmission rate is assumed to be constant through the period of infection. This may be incorrect as some studies suggest that transmission rate changes over the course of infection. It is well established that that viral loads are highest during the acute phase of the infection, leading to the possibility that most transmission occurs in the acute phase of infection [30]. If this is
true, and transmission during the asymptomatic phase is extremely low relative to the acute phase, then these results hold less relevance as there would be only a small selection pressure acting on the virus after the acute phase has ended. An extension of this model would be to apply a similar model on patients during the acute stage of infection, however, due to the massive fluctuation in viral loads within a host over the acute phase of infection and difficulty obtaining data for patients in this phase, the model presented in this study would not be appropriate.

Of course it is important to consider that there may be many other factors at play here. Firstly it is probable that transmissibility is not solely dependant on viral load and there may be many in-host factors to consider such as viral resistance, drug resistance etc. It is also difficult to estimate, τ, owing to the fact that the number of undiagnosed infections is not a perfect calculation and often has large confidence intervals associated with it. It is also likely that the rate of infection detection will change with viral load, owing to the fact that people with higher viral loads may have more pathogenic symptoms and thus are more likely to seek medical help. Another factor to consider is the adherence to HAART within the population. This model assumes a 100% adherence to the treatment which in reality is not demonstrated [20]. A lack of adherence can stop the effectiveness of HAART and thus people who are diagnosed may still be infectious.

Currently in the UK, guidelines for receiving HAART do not comply with the test and treat policy assumed within this model. Instead administration occurs when CD4 counts falls below the threshold of 350 cells per µl. An interesting extension to this framework would be to develop a model whereby HAART is administrated when a person falls below this threshold and analyse the effect on viral evolution. If this analysis shows a large difference on the evolutionary pressure placed on the virus, then it would be of public health interest to consider the long term effects on viral evolution when changing to a test and treat policy.

Appendix I

Distribution of Set-Point Viral Loads within populations.

Taking data from two groups of HIV positive cohorts, the logarithm of set-point viral loads are measured and their distribution is analysed.

These distributions are well described by a skewed normal distribution for the logarithm of the viral loads, i.e. the probability density function of \( v = \log_{10}(V) \), denoted \( F(v) \), is

\[
F(v) = \frac{2}{\sigma} z \left( \frac{v - \bar{\mu}}{\sigma} \right) Z \left( \alpha \frac{v - \bar{\mu}}{\sigma} \right)
\]  

(11)

where \( z(.) \) and \( Z(.) \) are the probability density function and the cumulative density function of the standard normal distribution. This distribution has position parameter \( \bar{\mu} \), variability parameter \( \sigma \) and skewness parameter \( \alpha \). The distribution with probability density function \( F(.) \) has mean and variance;

\[
\mu = \bar{\mu} + \frac{2\sigma \alpha}{\sqrt{2\pi(1+\alpha^2)}}
\]  

(12)

\[
\sigma = \bar{\sigma} \sqrt{1 - \frac{2\alpha^2}{\pi(1+\alpha^2)}}
\]  

(13)

Maximum likelihood estimates are \( \mu = 4.36, \sigma = 0.68 \) and \( \alpha = -3.74 \) for the Amsterdam seroconverters and \( \mu = 4.74, \sigma = 0.78 \) and \( \alpha = -3.55 \) for the Zambian cohort. One defines the probability density function for the Zambian Cohort to be \( F_Z \) and for the Amsterdam Seroconverters to be \( F_A \).
Figure 5.1a and 5.1b shows the plots for the these two distributions with corresponding histograms for 1000 randomly generated variables using this distribution for each. Figure 5.2 compares the profile of these two distributions.

Figure 5.2: Graphic comparing the two probability density functions which fit the distribution of the set point viral loads for the two cohorts studied.

References


