*Research Article*

# Posterior Estimates of Dynamic Constants in HIV Transmission Modeling

**Yingqing Chen,[1] Renee Dale,[2] Hongyu He,[3] and Quoc-Anh T. Le[3]**

[1]*Fred Hutchinson Cancer Research Center, Seattle, WA, USA*
[2]*Department of Biological Sciences, Louisiana State University, Baton Rouge, LA 70803, USA*
[3]*Department of Mathematics, Louisiana State University, Baton Rouge, LA 70803, USA*

Correspondence should be addressed to Hongyu He; hhe@lsu.edu

In this paper, we construct a linear differential system in both continuous time and discrete time to model HIV transmission on the population level. The main question is the determination of parameters based on the posterior information obtained from statistical analysis of the HIV population. We call these parameters dynamic constants in the sense that these constants determine the behavior of the system in various models. There is a long history of using linear or nonlinear dynamic systems to study the HIV population dynamics or other infectious diseases. Nevertheless, the question of determining the dynamic constants in the system has not received much attention. In this paper, we take some initial steps to bridge such a gap. We study the dynamic constants that appear in the linear differential system model in both continuous and discrete time. Our computations are mostly carried out in Matlab.

## 1. Introduction

Patients infected with Human Immunodeficiency Virus (HIV) are very likely to develop Acquired Immunodeficiency Disease Syndrome- (AIDS-) related diseases that are usually fatal if not treated with effective antiretroviral therapies. Since the discovery of HIV in 1983, an efficacious vaccine is yet to be developed to fight the deadly virus. Although highly active antiretroviral therapies (HAART) invented in mid-1990s have saved millions of lives and deterred the disease progression of those infected, HIV infection remains a public health threat. Reducing the risk of HIV transmission is of top priority.

One particular challenge in HIV prevention is its long period of latency period. The average time of an HIV infected patient to become symptomatic with AIDS-related diseases can be more than 10 years [1]. In the sexual transmission of HIV, many of the HIV infected patients may not be aware of their HIV infection status, and the virus continues spreading to their HIV negative partners. Therefore, an in-depth understanding of HIV transmission is the key to successful HIV prevention.

HIV dynamics have long been studied in the field of mathematical epidemiology using linear and nonlinear models [2, 3]. The classic model in epidemiology is the SIR model, which considers the dynamics of the susceptible, infected, and recovered populations [4]. This model is not useful for HIV dynamics, as there is no recovered population. An extension of this is the SEIR model, which includes the population of individuals who are exposed but not yet infected. The period between exposure and infectiousness in HIV lasts about two to four weeks [1]. Since a recovered population does not exist, we can consider this period to have a negligible effect on population dynamics.

Hierarchical models are common in HIV modeling due to the high correlation between risky behavior and HIV incidence [5]. In this paper we will incorporate risk indirectly by considering diagnosed and undiagnosed populations.

Intuitively, diagnosed individuals would modify their behavior relative to their behavior prior to the diagnosis.

In this paper, we shall form two models: a continuous time linear differential model and a discrete time differential model. These models are the most fundamental among their kinds. The focus will then be given to determination of the parameter estimates, the dynamic constants in these models. As we will show in this paper, the estimates of the dynamic constants depend on the type of model as well as the qualitative properties of the models.

There are two important dynamic constants in our model, namely, the transmission rates for diagnosed HIV population and for undiagnosed HIV population. One important finding in our study is that the transmission rates for the diagnosed and undiagnosed infected populations are comparable. This leads to our conclusion that the transmission rates should be attached to different groups of susceptibles based on their risk level.

## 2. General Nonlinear Differential Model

One of the frequently used mathematical models for HIV population dynamics can be described as follows. Let $S(t)$ be the susceptibles. We divide the HIV positive population into two groups: $N_0$ is the populations that are unaware of the infection; $N_1$ is the populations that are aware of the infection. Let $\epsilon_i$ be the mortality rate for the group $N_i$. Let $r$ be the growth rate of the susceptibles. Let $\gamma_0$ be the transmission rate of $N_0$ group and let $\gamma_1$ be the transmission rate of $N_1$ group. Then we have the following nonlinear differential equations:

$$
\begin{aligned}
\frac{dS(t)}{dt} &= S(t)\left(r - \gamma_0 N_0(t) - \gamma_1 N_1(t)\right); \\
\frac{dN_0(t)}{dt} &= (1-\beta)\gamma_1 N_1(t) S(t) + \gamma_0 N_0 S(t) - \delta N_0 \\
&\quad - \epsilon_0 N_0; \\
\frac{dN_1(t)}{dt} &= \beta \gamma_1 N_1(t) S(t) + \delta N_0 - \epsilon_1 N_1.
\end{aligned}
\tag{1}
$$

Here $\gamma_1 N_1(t) S(t)$ counts for those who are infected by group $N_1$ (per unit time), and among them $\beta$ is the proportion of those who are aware of their infection. The constant $\delta$ denotes the rate of the HIV positive population in $N_0$ group who become aware of their infection (per unit time). So there is a flow of $\delta N_0(t)$ from group $N_0$ to $N_1$ once a member from $N_0$ finds out his/her infection through HIV testing.

Many variations of this nonlinear dynamic model have been considered and appeared in the literature to study the HIV population dynamics. For example, in [6], mortality rate of the susceptibles is considered and appears in the differential equation of $S(t)$. In addition, the parameters are allowed to change but are *piecewise* constant.

In our differential equation model, we have a few constants: $\beta$, $\delta$, $\gamma_0$, $\gamma_1$, $\epsilon_0$, and $\epsilon_1$. These constants essentially determine the qualitative and quantitative properties of the mathematical model. We shall call these constants the *dynamic constants* of the model. Notice that some of the

constants, like $\gamma_i$, may have prior estimates, based on the data collected directly from the groups $N_i$ and $S$. Some of the constants, like $\epsilon_i$, will have posterior estimates. The constants $\delta$, $\beta$ may have prior estimates. Our main focus here is to give posterior estimates of these constants.

We shall remark here that the dynamic constants are model-dependent. This might not be obvious. Even though many of them can be estimated statistically without reference to any models, applying these estimates directly to the model may be problematic, as we shall see in the next section. In this paper, we take some initial steps to estimate the model-based dynamic constants.

## 3. Linear Differential Model and Preliminary Discussions

We shall now build a simpler linear model. The main assumption is that the susceptible population is a lot larger than $N_0$ and $N_1$. The change of susceptible population, due to HIV infection, is quite small, comparing with the overall size of susceptible. Therefore, we may ignore the dynamics of susceptible population, by assuming that the susceptible population is a constant. This more or less justifies the use of linear system only involving $N_0$ and $N_1$.

Let us start with the HIV transmission rate estimates by Pinkerton [7]. The estimates of transmission rates are

$$
\begin{aligned}
\gamma_0 &\cong 0.0927, \\
\gamma_1 &\cong 0.0268.
\end{aligned}
\tag{2}
$$

$\gamma_i$ are estimated in terms of infection transmitted per person per year. Since the overall susceptible population is a lot larger that $N_i$, we can assume that HIV transmission events are proportional to the size of $N_1$ and $N_0$. Based on this hypothesis, we may model HIV transmission by linear differential equations:

$$
\begin{aligned}
\frac{dN_0}{dt} &= (1-\beta)\gamma_1 N_1 + \gamma_0 N_0 - \delta N_0 - \epsilon_0 N_0, \\
\frac{dN_1}{dt} &= \beta \gamma_1 N_1 + \delta N_0 - \epsilon_1 N_1.
\end{aligned}
\tag{3}
$$

The dynamic constants $\delta$ and $\beta$ remain unchanged. It is also known that $\epsilon_1 \cong 1.9\%$ [8]. There is no statistics done on $\epsilon_0$. So we can assume $\epsilon_0 \cong 1.9\%$ as well.

Next, we shall apply the known estimates and study our linear differential model. Notice that $\beta$ remain unknown at this moment. According to [8], $\delta$ is somewhere around $1/4$. We may tentatively set $\delta = 1/4$. Utilizing the estimates of dynamic constants directly from [7, 8], let us consider several cases.

TABLE 1: Population of undiagnosed individuals with HIV from 2007 to 2013.

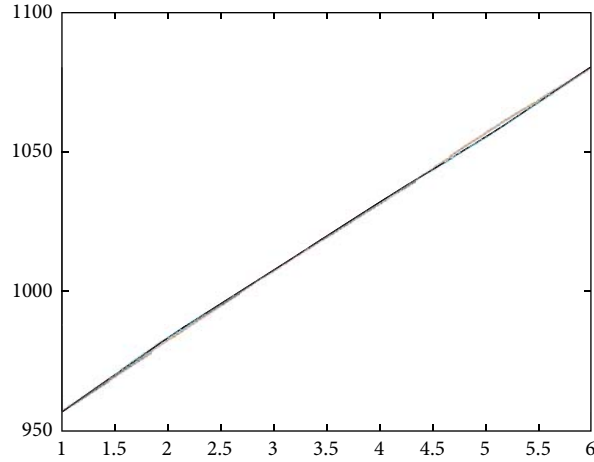| Year | Diagnosed | Undiagnosed | Percentage of total |
|------|-----------|-------------|---------------------|
| 2007 | 929.3 | 183.777 | 16.5 |
| 2008 | 956.9 | 178.1165 | 15.7 |
| 2009 | 982.4 | 170.6282 | 14.8 |
| 2010 | 1007.6 | 165.3507 | 14.1 |
| 2011 | 1031.6 | 162.4248 | 13.6 |
| 2012 | 1057.2 | 160.7760 | 13.2 |
| 2013 | 1080.5 | 161.46 | 13.0 |

Source: [8].



FIGURE 1: Comparison of $\mathbf{TN}_t$ and $\mathbf{N}_{t+1}$: $N_1$ group.

*3.1. $\beta = 4/5$.* We start by assuming that $\beta$ takes the value of the overall portion of those who are aware of their infection. Now we have the following linear equations:

$$\frac{dN_0}{dt} = \frac{0.0268}{5}N_1 + 0.0927N_0 - \frac{1}{4}N_0 - 0.019N_0.$$

$$\frac{dN_1}{dt} = \frac{0.0268 \times 4}{5}N_1 + \frac{1}{4}N_0 - 0.019N_1. \tag{4}$$

We found that the two linear independent solutions have growth rate of

$$\lambda_+ = 0.01,$$
$$\lambda_- = -0.18. \tag{5}$$

However, we know that the growth of $N_0 + N_1$ is about 0.048. Hence our assumption $\beta = \eta = 4/5$ is not valid. Even if we ignore the mortality rate, we have

$$\lambda_+ = 0.021,$$
$$\lambda_- = -0.16. \tag{6}$$

This is still far below the estimated 4.8% growth rate.

*3.2. $\beta = 1$ or $\beta = 0$.* One extreme is that $\beta = 1$, meaning that the population infected by $N_1$ gets tested and becomes aware of their infection (within the first year). We have

$$\frac{dN_0}{dt} = 0.0927N_0 - \frac{1}{4}N_0 - 0.019N_0,$$

$$\frac{dN_1}{dt} = 0.0268N_1 + \frac{1}{4}N_0 - 0.019N_1. \tag{7}$$

Under this assumption $N_0$ will decrease at the rate of $-0.268$, which means that the population $N_0$ will gradually vanish in a few years. This cannot be true.
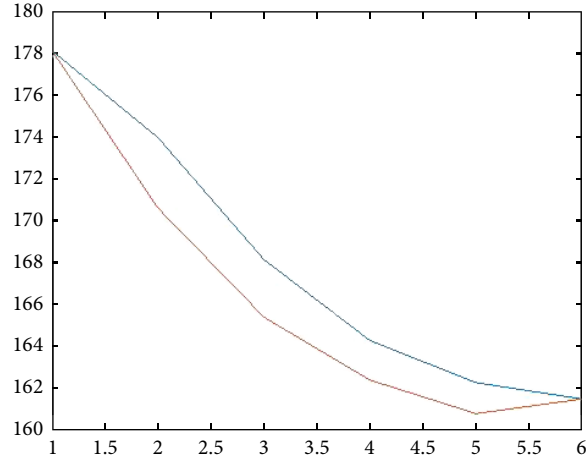
Another extreme is that $\beta = 0$, meaning that the population infected by $N_1$ will be initially unaware of their infection (within the first year). We have

$$\frac{dN_0}{dt} = 0.0268N_1 + 0.0927N_0 - \frac{1}{4}N_0 - 0.019N_0,$$

$$\frac{dN_1}{dt} = \frac{1}{4}N_0 - 0.019N_1. \tag{8}$$

We have

$$\lambda_+ = 0.016,$$
$$\lambda_- = -0.211. \tag{9}$$

The overall HIV population growth will be less than 0.016. This is quite small comparing with the estimate that the growth rate is about 0.048.

Figure 2: Comparison of $\mathbf{TN}_t$ and $\mathbf{N}_{t+1}$: $N_0$ group.

*3.3. $\delta$, $\beta$ Not Fixed.* One might conclude that $\delta$ must be a much smaller number than 1/4, what we have initially assumed. We let $\delta$ and $\beta$ be unfixed. In this case, we have

$$\frac{dN_0}{dt} = (0.0927 - \delta) N_0 + (1 - \beta)(0.0268) N_1$$
$$- 0.019 N_0, \tag{10}$$

$$\frac{dN_1}{dt} = 0.0268\beta N_1 + \delta N_0 - 0.019 N_1.$$

We have the matrix

$$A = \begin{pmatrix} 0.0927 - \delta - 0.019 & (0.0268)(1 - \beta) \\ \delta & 0.0268\beta - 0.019 \end{pmatrix}. \tag{11}$$

We know the growth rates are controlled by the eigenvalues of $A$. In particular, we might assume that $\det(A - \lambda) = 0$ with $\lambda = 0.048$. This will guarantee that the dominant term of the solution will grow at the rate of 0.048 (per year). Hence we obtain

$$(0.0257 - \delta)(0.0268\beta - 0.067) = 0.0268(1 - \beta)\delta. \tag{12}$$

Simplifying it, we have

$$1.56\delta + 0.0268\beta \cong 0.067. \tag{13}$$

Since $0 \le \beta \le 1$, we find that $0.043 \ge \delta \ge 0.0258$. This suggests that there are between 2% to 5% of $N_0$ getting tested. This percentage seems to be too low comparing with the CDC estimate of about 25%.

We shall remark that our discussion is based on the estimates that $\gamma_0 = 0.0927$ and $\gamma_1 = 0.0268$ [7]. As we have seen, directly using these estimates as dynamic constants in differential equation modeling will be inadequate to produce the right kind of outcomes and trend. In this paper, we shall discuss posterior estimate of parameters and hope to find some remedy.

## 4. Posterior Estimate of Parameters

In our earlier discussion, we directly insert the transmission rates from the statistical analysis into the linear differential system. The result is not satisfactory. It is desirable to estimate the transmission rates that will produce the right kind of outcome from the linear differential system model. Let us recall the CDC data from 2007 to 2013 (in thousands) [8].

We first simplify our notation. Let $\mathbf{N} = \begin{pmatrix} N_1 \\ N_0 \end{pmatrix}$. We rewrite our linear system as

$$\frac{d\mathbf{N}}{dt} = \mathbf{M}\mathbf{N}(t), \tag{14}$$

where

$$\mathbf{M} = \begin{pmatrix} \beta\gamma_1 - \epsilon_1 & \delta \\ (1 - \beta)\gamma_1 & \gamma_0 - \delta - \epsilon_0 \end{pmatrix}. \tag{15}$$

The general solution to this system is

$$\mathbf{N}(t) = \mathbf{P}\exp\lambda_1 t + \mathbf{Q}\exp\lambda_2 t. \tag{16}$$

Here $\lambda_1$, $\lambda_2$ are eigenvalues of $M$. They can be both real or complex. There is also a degenerate case $\lambda_1 = \lambda_2$ that we do not treat here. The behavior of the linear differential system is quite different in these two cases. It is not surprising that we need to use two different methods to estimate the matrix $\mathbf{M}$.

*4.1. $\lambda_1$, $\lambda_2$ Real: Simple Curve Fitting.* We try a global optimization curve fitting using Matlab. We have

$$\mathbf{N}(t) = \begin{pmatrix} 14.12 \\ 79.86 \end{pmatrix} \exp - 0.1919t$$
$$+ \begin{pmatrix} 892.3 \\ 115.57 \end{pmatrix} \exp 0.0273t. \tag{17}$$

Let $\mathbf{A} = \left(\begin{smallmatrix} 14.12 & 892.3 \\ 79.86 & 115.57 \end{smallmatrix}\right)$. Then

$$
\begin{aligned}
\mathbf{M} &\cong \mathbf{A} \begin{pmatrix} -0.1919 & 0 \\ 0 & 0.0273 \end{pmatrix} \mathbf{A}^{-1} \\
&= \begin{pmatrix} 0.032 & -0.040 \\ 0.0291 & -0.197 \end{pmatrix}.
\end{aligned}
\tag{18}
$$

Notice that the dominant term $\left(\begin{smallmatrix} 892.3 \\ 115.57 \end{smallmatrix}\right) \exp 0.0273t$ suggested the overall rate of growth of HIV infected population grows at the rate close to 2.73%. This seems to be reasonable. But $\delta$, the rate of flow of population from $N_0$ to $N_1$, is estimated at $-4\%$. This is completely off the mark. One remedy is that we first estimate the dominant term and then estimate the remainder.

### 4.2. $\lambda_1, \lambda_2$ Real: Dominant Term Estimate.

Suppose that $\lambda_2 < \lambda_1$. Then $\mathbf{P} \exp \lambda_1 t$ is the dominant term. We shall have

$$
\|\mathbf{N}(t)\| \cong \|\mathbf{P}\| \exp \lambda_1 t.
\tag{19}
$$

Now

$$
\begin{aligned}
\|\mathbf{N}(t)\| \\
= [947.3, 973.3, 997.1, 1021.1, 1044.3, 1069.4, 1092.5].
\end{aligned}
\tag{20}
$$

Using curve $F(t) = a e^{\lambda t}$ to fit this data, we obtain

$$
\begin{aligned}
a &= 927.7, \\
\lambda &= 0.0236.
\end{aligned}
\tag{21}
$$

### 4.3. $\lambda_1$ Dominant, $\lambda_2$ Real.

Now we can assume $\lambda_1 = 0.0236$ and use curve fitting to find $\lambda_2$, $\mathbf{P}$, and $\mathbf{Q}$. We have

$$
\begin{aligned}
\mathbf{P} &= \begin{pmatrix} 922.1 \\ 115.9 \end{pmatrix}, \\
\mathbf{Q} &= \begin{pmatrix} -20.4 \\ 77.6 \end{pmatrix},
\end{aligned}
\tag{22}
$$

and $\lambda_2 = -0.172$, $\lambda_1 = 0.0236$. It follows that

$$
\begin{aligned}
\mathbf{M} &= \begin{pmatrix} \beta\gamma_1 - \epsilon_1 & \delta \\ (1-\beta)\gamma_1 & \gamma_0 - \delta - \epsilon_0 \end{pmatrix} \\
&= \begin{pmatrix} 0.0173 & 0.0498 \\ 0.0238 & -0.1656 \end{pmatrix}.
\end{aligned}
\tag{23}
$$

We derive that $\gamma_1 - \epsilon_1 \cong 0.041$ and $\delta \cong 0.05$. These parameters seem to be reasonable. However, $\gamma_0 - \epsilon_0 = 0.0498 - 0.1656 = -0.1158$. Hence $\gamma_0$ will be a negative number which is not possible.

### 4.4. $\lambda_1, \lambda_2$ Complex with Fixed Real Part.

Suppose that $\lambda_1$ and $\lambda_2$ are complex. Then $\lambda_1$ and $\lambda_2$ are conjugate to each other. In particular, the real part of $\lambda$, $\Re(\lambda_1) = \Re(\lambda_2)$ should be approximately 0.0236. Write $\lambda_1 = \lambda_0 + i\mu$. We should have

$$
\mathbf{N}(t) = \mathbf{P} \exp \lambda_0 t \cos \mu t + \mathbf{Q} \exp \lambda_0 t \sin \mu t,
\tag{24}
$$

where $\mu$ is sometimes called a phase constant. A simple curve fitting shows that

$$
\begin{aligned}
\mathbf{P} &= \begin{pmatrix} 904.12 \\ 182.7 \end{pmatrix}, \\
\mathbf{Q} &= \begin{pmatrix} -117 \\ 420 \end{pmatrix}
\end{aligned}
\tag{25}
$$

and $\lambda_0 = 0.0236$ and $\mu = -0.017$. Hence

$$
\begin{aligned}
M &\cong [\mathbf{P}, \mathbf{Q}] \begin{pmatrix} 0.0236 & 0.017 \\ -0.017 & 0.0236 \end{pmatrix} [\mathbf{P}, \mathbf{Q}]^{-1} \\
&= \begin{pmatrix} 0.0187 & 0.0325 \\ -0.0089 & -0.0285 \end{pmatrix}.
\end{aligned}
\tag{26}
$$

Let us see what this tells us. We have

$$
\begin{aligned}
\mathbf{M} &= \begin{pmatrix} \beta\gamma_1 - \epsilon_1 & \delta \\ (1-\beta)\gamma_1 & \gamma_0 - \delta - \epsilon_0 \end{pmatrix} \\
&= \begin{pmatrix} 0.0187 & 0.0325 \\ -0.0089 & -0.0285 \end{pmatrix},
\end{aligned}
$$

$$
\gamma_1 - \epsilon_1 = 0.0187 - 0.0089 \cong 0.01,
$$

$$
\gamma_0 - \epsilon_0 = 0.004,
$$

$$
\delta = 0.0325.
\tag{27}
$$

This roughly says that there are about 3.25% of $N_0$ that become aware of their infection every year. The annual transmission rate for $N_1$ is 2.9%. The annual transmission rate for $N_0$ is 2.3%.

### 4.5. $\lambda_1, \lambda_2$ Complex.

We finally use Matlab global optimization to fit the data in the curve

$$
\mathbf{N}(t) = \mathbf{P} \exp \lambda_0 t \cos \mu t + \mathbf{Q} \exp \lambda_0 t \sin \mu t.
\tag{28}
$$

We obtain $\lambda_0 = 0.0088$, $\mu = -0.036$,

$$
\begin{aligned}
\mathbf{P} &= \begin{pmatrix} 902.4 \\ 184 \end{pmatrix}, \\
\mathbf{Q} &= \begin{pmatrix} -476.3 \\ 149.5 \end{pmatrix}.
\end{aligned}
\tag{29}
$$

Hence we obtain the estimate

$$
\begin{aligned}
&\mathbf{M} \\
&\cong \begin{pmatrix} 902.4 & -476.3 \\ 184 & 149.5 \end{pmatrix} \begin{pmatrix} 0.0088 & 0.036 \\ -0.036 & 0.0088 \end{pmatrix} \begin{pmatrix} 902.4 & -476.3 \\ 184 & 149.5 \end{pmatrix}^{-1} \\
&= \begin{pmatrix} -0.0065 & 0.1684 \\ -0.0091 & 0.0241 \end{pmatrix}.
\end{aligned}
\tag{30}
$$

Now we have

$$\beta \gamma_1 - \epsilon_1 \cong -0.0065,$$

$$\delta \cong 0.1684,$$

$$(1 - \beta) \gamma_1 \cong -0.0091,$$

$$\gamma_0 - \delta - \epsilon_0 \cong 0.0241.$$

(31)

It follows that

$$\gamma_1 \cong \epsilon_1 - 0.0156 \cong 0.0034,$$

$$\gamma_0 = 0.193 + \epsilon_1 \cong 0.21.$$

(32)

So $\gamma_1$ is neglectable and $\gamma_0$ is about 21%. This again makes the model invalid.

*4.6. Discussion.* In this section, we choose dynamics constants to fit the temporal data. We have found that these dynamic constants depend on the qualitative properties of the model. Yet, none of the dynamic constants we choose match perfectly with the existing estimates. One reason is that yearly data is not suitable for a continuous time model. Therefore, we shall explore the discrete time model.

## 5. Discrete Dynamic Model

We may regard $\mathbf{N}_t$ ($t = 1, 2, 3, 4, 5, 6, 7$) as a discrete time dynamical system. Let us assume that this discrete dynamics is defined by a transition matrix $\mathbf{T}$:

$$\mathbf{N}_{t+1} = \mathbf{T} \mathbf{N}_t.$$

(33)

In principle, based on our earlier discussion,

$$\mathbf{T} = \mathbf{I} + \begin{pmatrix} \beta \gamma_1 - \epsilon_1 & \delta \\ (1 - \beta) \gamma_1 & \gamma_0 - \delta - \epsilon_0 \end{pmatrix}.$$

(34)

Now we would like to estimate $\mathbf{T}$.

*5.1. Basic Estimates.* The easiest way to find $\mathbf{T}$ is by considering the following matrix equations:

$$[\mathbf{N}_i \mathbf{N}_{i+1}] = \mathbf{T} [\mathbf{N}_{i-1} \mathbf{N}_i].$$

(35)

For example, for $i = 2$, we will have

$$\begin{pmatrix} 956.9 & 982.4 \\ 178.1 & 170.6 \end{pmatrix} = \mathbf{T} \begin{pmatrix} 929.3 & 956.9 \\ 183.7 & 178.1 \end{pmatrix}.$$

(36)

Then we find the following estimate of $\mathbf{T}$:

$$\begin{pmatrix} 0.9775 & 0.2641 \\ -0.0370 & 1.1568 \end{pmatrix}$$

$$\begin{pmatrix} 1.0118 & 0.0800 \\ 0.0302 & 0.7959 \end{pmatrix}$$

$$\begin{pmatrix} 0.9922 & 0.1928 \\ 0.0370 & 0.7563 \end{pmatrix}$$

$$\begin{pmatrix} 1.0481 & -0.1482 \\ 0.0318 & 0.7880 \end{pmatrix}$$

$$\begin{pmatrix} 0.9427 & 0.5214 \\ 0.0639 & 0.5844 \end{pmatrix}.$$

(37)

We can see some consistency among these transition matrices. For example, the $(2, 1)$-th entry has been around 3%. This translates into

$$(1 - \beta) \gamma_1 \cong 3\%.$$

(38)

This is the rate of transmission for group $N_1$. It seems to be consistent with the estimate of [7].

*5.2. (Arithmetic) Average Estimate of* $\mathbf{T}$. Now we may average all $\mathbf{T}$'s and obtain

$$\mathbf{T} \cong \begin{pmatrix} 0.9945 & 0.1820 \\ 0.0252 & 0.8163 \end{pmatrix}.$$

(39)

Hence

$$\mathbf{M} = \begin{pmatrix} \beta \gamma_1 - \epsilon_1 & \delta \\ (1 - \beta) \gamma_1 & \gamma_0 - \delta - \epsilon_0 \end{pmatrix}$$

$$= \begin{pmatrix} -0.0055 & 0.1820 \\ 0.0252 & -0.1837 \end{pmatrix}.$$

(40)

Our estimate yields that $\delta \cong 18\%$; in other words, about 18% of those unaware of their infection will become aware of their infection next year. We also have

$$\gamma_0 - \epsilon_0 = \delta + (\gamma_0 - \delta - \epsilon_0) = -0.0017.$$

(41)

If the mortality rate $\epsilon_0$ is set to be 0.019, then we have $\gamma_0 = 0.017$. Similarly, we have

$$\gamma_1 - \epsilon_1 = \beta \gamma_1 - \epsilon_1 + (1 - \beta) \gamma_1 \cong 0.02.$$

(42)

If the mortality rate $\epsilon_1$ is set to be 0.019, then we have $\gamma_1 = 0.039$. This suggests that the transmission rate of $N_1$ group is twice as large as the transmission rate of $N_0$ group. There may be some truth to it. However, we believe that this estimate is off the mark due to the reason that $[\mathbf{N}_i \mathbf{N}_{i+1}]$ are correlated with each other. Hence each estimate $\mathbf{T}$ will be biased. We shall correct this and give a more robust estimate later.

*5.3. Least Square Estimate of* **T**. Perhaps a good way to estimate **T** is the least square method. We write

$$[\mathbf{N}_2\mathbf{N}_3\cdots\mathbf{N}_7] = \mathbf{T}[\mathbf{N}_1\mathbf{N}_2\cdots\mathbf{N}_6]. \tag{43}$$

Applying the least square method, we find that the least square solution to **T** is

$$\begin{pmatrix} 1.0013 & 0.1406 \\ 0.0245 & 0.8350 \end{pmatrix}. \tag{44}$$

This estimate seems to be better than the arithmetic average, in the sense that, irregularities will have smaller effect on the least square solution. Because we can reorder $\mathbf{N}_i$'s and the least square solution does not change, we also avoid the pitfall that $\mathbf{N}_i$ and $\mathbf{N}_{i+1}$ are correlated. We have our posterior estimates:

$$\begin{aligned} \mathbf{M} &\cong \begin{pmatrix} \beta\gamma_1 - \epsilon_1 & \delta \\ (1-\beta)\gamma_1 & \gamma_0 - \delta - \epsilon_0 \end{pmatrix} \\ &= \begin{pmatrix} 0.0013 & 0.1406 \\ 0.0245 & -0.1650 \end{pmatrix}. \end{aligned} \tag{45}$$

This estimate is similar to the arithmetic average we just computed. The dynamic constant estimates will be very similar. We shall then look for a solution that is more robust. One particular reason that the least square estimate is not satisfactory is that there are additional relations like

$$\mathbf{N}_{i+k} = \mathbf{T}^k\mathbf{N}_i \tag{46}$$

that least square method does not take into consideration. In other words, $\mathbf{T}^2$, $\mathbf{T}^3$ can also be estimated and shall be taken into consideration when we estimate **T**. We shall offer one remedy that avoids this issue.

*5.4. A More Robust Estimate.* One of the problems with our estimate is that $\mathbf{N}_t$ and $\mathbf{N}_{t+1}$ are correlated to each other. As a remedy, we pick $\mathbf{N}_1$ and $\mathbf{N}_6$ as far from each other as possible. We observe that

$$\mathbf{T}[\mathbf{N}_1\mathbf{N}_6] = [\mathbf{N}_2\mathbf{N}_7]. \tag{47}$$

We compute

$$\mathbf{T} = [\mathbf{N}_2\mathbf{N}_7][\mathbf{N}_1\mathbf{N}_6]^{-1} = \begin{pmatrix} 0.9965 & 0.1681 \\ 0.023 & 0.8533 \end{pmatrix}. \tag{48}$$

Then

$$\begin{aligned} \mathbf{M} &= \begin{pmatrix} \beta\gamma_1 - \epsilon_1 & \delta \\ (1-\beta)\gamma_1 & \gamma_0 - \delta - \epsilon_0 \end{pmatrix} \\ &\cong \begin{pmatrix} -0.0035 & 0.1681 \\ 0.023 & -0.1467 \end{pmatrix}. \end{aligned} \tag{49}$$

We have

$$\delta \cong 0.1681,$$
$$\gamma_0 - \epsilon_0 \cong 0.0214, \tag{50}$$
$$\gamma_1 - \epsilon_1 \cong 0.0195.$$

*5.5. Discussion.* This estimate of **M** is quite consistent with the least square estimate. Our estimate seems to suggest that the transmission rates for $N_0$ and $N_1$ may be in the similar range. By [8], assume that $\epsilon_0 = \epsilon_1 = 0.019$. We have

$$\gamma_1 = 0.0385,$$
$$\gamma_0 = 0.04,$$
$$\delta = 0.185; \tag{51}$$
$$\beta = \frac{\beta\gamma_1}{\gamma_1} = \frac{0.019 - 0.0035}{0.0385} = 0.4.$$

Every year about 18.5% of those unaware of their HIV positiveness become aware of their infection due to testing. About 40% of those infected by $N_1$ group become aware of their infection. This seems to be consistent with some of the observations in [7], with one exception; namely, in our estimates, the transmission rates for $N_1$ and $N_0$ are very close. Figures 1 and 2 show the difference between $\mathbf{TN}_t$ and $\mathbf{N}_{t+1}$.

*5.6. Arithmetic Average versus Geometric Average.* Now we may state our problem in greater generality. Given a temporal vector $\mathbf{N}(t)$, suppose that $\mathbf{N}(t+1) = \mathbf{TN}(t)$ with transitional matrix **T**. How should one estimate the matrix **T**?

As we discussed earlier, we can use least squares with the equations

$$\mathbf{N}(t+1) = \mathbf{TN}(t). \tag{52}$$

The least square estimate of **T**, in some sense, is very similar to the arithmetic mean of the transitional matrix **T**. But what makes better sense is a geometric mean. More precisely, we have to take into consideration that

$$\mathbf{N}(t+k) = \mathbf{T}^k\mathbf{N}(t). \tag{53}$$

Suppose that $\mathbf{T}_t$ is the transitional matrix at time $t$. Then a good estimate of **T** should be the "geometric average" of $\mathbf{T}_t$. For scalars, one can define the geometric average of $p_1, p_2, \ldots, p_n$ to be the $n$th root of $\prod p_i$. But matrix multiplications are not commutative and one cannot define the geometric average of matrices. It remains a challenging problem to define computationally a geometric mean of $\mathbf{T}_t$.

*5.7. Roots Estimate.* Tentatively, we may define the geometric mean by taking roots. For example, we may now consider

$$\mathbf{T}^2[\mathbf{N}_1\mathbf{N}_4] = [\mathbf{N}_3\mathbf{N}_6]. \tag{54}$$

Then

$$\mathbf{T} \cong \begin{pmatrix} 1.0038 & 0.1277 \\ 0.0228 & 0.8443 \end{pmatrix},$$
$$\mathbf{M} \cong \begin{pmatrix} 0.0038 & 0.1277 \\ 0.0228 & -0.1557 \end{pmatrix}. \tag{55}$$

We can also consider

$$\mathbf{T}^4[\mathbf{N}_1\mathbf{N}_3] = [\mathbf{N}_5\mathbf{N}_7]. \tag{56}$$

We have

$$\mathbf{T} = \begin{pmatrix} 1.0008 & 0.1430 \\ 0.0301 & 0.8011 \end{pmatrix},$$

$$\mathbf{M} \cong \begin{pmatrix} 0.0008 & 0.143 \\ 0.0301 & -0.2 \end{pmatrix}. \tag{57}$$

Both estimates of $\mathbf{T}$ are consistent with the least square estimate and the estimates in the previous section. Above all, all our estimates point to the same range of transmission rates for both $N_0$ and $N_1$.

## 6. Concluding Remarks

Now we shall compare our dynamic constant estimates in the linear differential model in continuous time and in discrete time.

In the continuous time model, we obtain the transmission rate $\epsilon_1$ of about 4% for the $N_1$ group, those who were aware of their HIV infection. Nevertheless, $\delta$ the rate of flow from $N_0$ to $N_1$ due to HIV testing turned out to be too low and $\epsilon_0$ often came out to be negative, which cannot be the case. The best results are obtained when we assume the two eigenvalues are complex. In this case

$$\gamma_1 \cong 0.029,$$

$$\delta \cong 0.0325, \tag{58}$$

$$\gamma_0 \cong 0.023.$$

Yet $\delta$ seems to be quite low. According to the CDC report [8], $\delta$ is estimated at about 25%.

There are various reasons why our dynamic constants are inconsistent with known estimates. Firstly, the CDC data we use tends to underestimate the $N_0$ and $N_1$ population sizes, particularly in more recent years. The CDC estimates the sizes of the populations infected with HIV by back calculation. The estimates for any given year will increase as new diagnoses are obtained. HIV may go undiagnosed for up to 10 years without causing the death of the patient (Table 1 [5]. Depending on the stage of the disease, the individual will be counted as undiagnosed for a number of years prior to the diagnosis. This causes the estimates of the population sizes to be smaller than the actual size of the population. A new estimate of the HIV prevalence agrees with this conclusion [9]. Although this new estimate is more conservative than the back calculation method, it may still underestimate the $N_0$ population. Both estimates show a downward trend in the data, but this is likely to be erased as more individuals are diagnosed in the later stages of the disease.

Secondly, our computations assume that the susceptible population is much larger than the infected population. However, failing to obtain the right estimates suggests that opposite may be true—the susceptible population could be much smaller. HIV infection is overrepresented in some subpopulations, such as men who have sex with men (MSM) [1]. This subpopulation is only about 5% of the US population, or approximately 15 million individuals. Not all MSM can

be considered to have equal risk of contracting the disease, and since over 1 million individuals are currently living with the disease, the susceptible population may be comparable to the size of the infected population. For this reason, any differential system model of the HIV transmission must include the susceptible population.

The discrete dynamics model seems to be most robust against the bias caused by back calculation and the size of the susceptibles. By ignoring the 2013 year's data, we are able to determine the transmission rates as

$$\gamma_1 \cong 0.0385,$$

$$\gamma_0 \cong 0.04. \tag{59}$$

Also $\delta \cong 0.184$, not too different from the CDC estimate 0.25. We see that the dynamic constants in the discrete time model are less affected by the underestimation caused by back calculation. It is also true that the relative size of susceptibles has less effect on the discrete time model than on the continuous time model.

Finally, our estimates of the transmission rates for diagnosed and undiagnosed HIV population are relatively close. This is very different from the previous estimate, where the transmission rate of the undiagnosed population is about 4 times as high as the diagnosed population [7]. This implies that the transmission rates should be attached to the susceptible population. It makes sense to divide the susceptible population into groups depending on the possible transmission rates for them. We shall pursue this in a different paper.

## Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

## Acknowledgments

## References

[1] CDC HIV Risk Reduction Tool. https://wwwn.cdc.gov/hivrisk/what_is/stages_hiv_infection.html.

[2] R. M. Anderson and R. M. May, "Population biology of infectious diseases: part I," *Nature*, vol. 280, no. 5721, pp. 361–367, 1979.

[3] R. M. May and R. M. Anderson, "Population biology of infectious diseases: part II," *Nature*, vol. 280, no. 5722, pp. 455–461, 1979.

[4] W. O. Kermack and A. G. McKendrick, "A contribution to the mathematical theory of epidemics," *Proceedings of the royal society a: mathematical, physical and engineering sciences*, vol. 115, no. 772, p. 700, 1927.

[5] Centers for Disease Control and Prevention. HIV Surveillance Report, vol. 27. http://www.cdc.gov/hiv/library/reports/hiv-surveillance.html, 2015.

[6] F. Bozkurt and F. Paker, "Mathematical Modelling of HIV epidemic and stability analysis," *Advances in Differential Equations*, vol. 95, 2014.

[7] S. D. Pinkerton, "HIV Transmission rate modeling: A primer, review, and extension," *AIDS and Behavior*, vol. 16, no. 4, pp. 791–796, 2012.

[8] Centers for Disease Control and Prevention, "Monitoring selected national HIV prevention and care objectives by using HIV surveillance data United States and 6 dependent areas," *HIV Surveillance Supplemental Report 2016*, vol. 4, 2014.

[9] R. Song, H. I. Hall, T. A. Green, C. L. Szwarcwald, and N. Pantazis, "Using CD4 Data to Estimate HIV Incidence, Prevalence, and Percent of Undiagnosed Infections in the United States," *Journal of Acquired Immune Deficiency Syndromes*, vol. 74, no. 1, pp. 3–9, 2017.

## Composition Comments

1. We referred to table 1 in the text, please check.