A single-staged (SS) model and a staged progression (SP) model for HIV/AIDS with the same variable contact rate over time were formulated. In both models, analytical expressions for the HIV prevalence were obtained. A comparison of the two models was undertaken. It is shown that prevalence projections from the SS model are lower than projections from the SP model up to and beyond the peak prevalence, although the SS model prevalence may be higher than that of the SP model much later in the epidemic. A switch from faster SP model prevalence changes to faster SS prevalence changes occurs beyond the SP model peak prevalence. Hence using the SS model underestimates HIV prevalence in the early stages of the epidemic but may overestimate prevalence in the declining HIV prevalence phase. Our comparison suggests that the SP model provides better prevalence projections than the SS model. Moreover, the extra parameters that would make the SP model appear difficult to implement may not be sought from national survey data but from existing HIV/AIDS literature.

Copyright © 2007 F. Baryarama and J. Y. T. Mugisha. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. Introduction

The basis for most national HIV projections is usually a simple mathematical model often based on a single-stage (SS) model to fit the observed prevalence patterns [1]. Yet, HIV-dynamics are quite complex. The infectiousness of HIV infected individuals is known to vary with stage of infection, being highly infectious in the first few weeks after becoming infected, then having low infectivity for many years, and finally becoming gradually more infectious as the immune system of those infected with HIV break downs [2–4].
This trend of disease progression may imply that HIV/AIDS exhibits three basic infectious stages. The infection rates for the first and third stages have been documented to be several times higher than those for the second stage [5]. Models that account for such stages are called staged progression (SP) models [6, 7]. The average period from the time of acquiring HIV infection to developing full blown disease (AIDS) has been reported in the range of 8–12 years [8].

SS and SP models have in general used constant parameter values [6, 4]. In particular, the behavioral parameters that account for the contact rates and probabilities of transmission per sexual act have been assumed constant for each stage of infection over long-time periods. Variability of probability of transmission by stage of infection is however the defining property of a staged progression model, that is affected by behavioural changes especially through condom use [9]. Our analyses have explored incorporating behavioural changes through variable parameters for each stage of HIV infection. Behavioural change was accounted for by using a lower bounded function that allows for a gradual reduction in behaviour change [10] and accurate description of the past trends in these parameters.

Although both the SS and the SP models have been extensively studied, a comparison of the two models has not been undertaken. Deviations between projected HIV prevalence using SS models and observed HIV prevalence have been noted. In Uganda, this raised concern about the correctness of observed data around the peak of the HIV epidemic in the early 1990s [11, 12]. However, such deviations between observed data and predicted prevalence from SS model may be due to the simplicity of the models used to make the projections. Simple models are however preferred given that they have fewer model parameters whose values can be established with greater precision. This is not the case with complex models such as SP models that may require estimation of many parameters, some of which may not be available from routine national survey data. Moreover, some countries may not have any reliable data to derive any parameters even for simple models.

In this paper, we formulate an SS model with variable contact rate to cater for behavioural response at any time \( t \), and an SP model that considers three stages of infectiousness with a variable contact rate that is the same as that used in the SS model. We compare the SS and the SP equations for the rates of change of HIV prevalence, and derive implications for using the SS model to approximate the staged progression of HIV dynamics. In particular, the accuracy of the SS model predictions is discussed in the contexts of an increasing HIV prevalence, around the peak HIV prevalence and a declining HIV prevalence. Implications for high- and low-prevalence settings are also discussed.

2. The SS model

Consider an AIDS model in which susceptibles \( S(t) \) are recruited at a rate \( q \) of the total sexually active adult population \( N(t) \). Let \( I(t) \) be the number of infectives in the population. The AIDS cases are assumed not sexually active so that \( N(t) = S(t) + I(t) \). Susceptibles are removed at a constant natural death rate \( \mu \), and through infection to join the infective class at a variable rate of infection \( \omega(t)I(t)/N(t) \) per susceptible. The form
Table 2.1. Table of parameter values used for Figures 2.1 and 2.2.

<table>
<thead>
<tr>
<th>Description of parameters</th>
<th>Values used in SS model</th>
<th>Values used in SP model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unprotected contacts per year: (high prevalence)</td>
<td>$c = \begin{cases} 144, 0 \leq t \leq 8 \ 36/(1 - 0.75 \times 0.97^{(t-8)}), t &gt; 8 \end{cases}$</td>
<td>$c_1 = c_2 = c_3 = c$</td>
</tr>
<tr>
<td>Unprotected contacts per year: (low prevalence)</td>
<td>$c = \begin{cases} 108, 0 \leq t \leq 8 \ 27/(1 - 0.75 \times 0.97^{(t-8)}), t &gt; 8 \end{cases}$</td>
<td>$c_1 = c_2 = c_3 = c$</td>
</tr>
<tr>
<td>Probability of transmission</td>
<td>$\beta = 0.0019$</td>
<td>$\beta = 0.00074$, $\beta_1 = 50\beta_2$, $\beta_3 = 5\beta_2$, $\nu_1 = 8.7$, $\nu_2 = 0.167$, $\nu_3 = 0.5$</td>
</tr>
<tr>
<td>Rates of progression</td>
<td>$\nu = 0.125$</td>
<td></td>
</tr>
<tr>
<td>Rates of recruitment into sexually active adults assumed at 15th birth-day</td>
<td>$q = 0.032$</td>
<td>$q = 0.032$</td>
</tr>
<tr>
<td>Initial HIV prevalence</td>
<td>$\phi_0 = 0.01$</td>
<td>$\phi_0 = 0.01$</td>
</tr>
</tbody>
</table>

The time $t$ refers to the time elapsed since the beginning of the epidemic. We define $q$ as the recruitment of individuals into the susceptible population. HIV prevalence refers to the fraction of infectives in the population excluding those with AIDS symptoms, that is, prevalence $= I(t)/N(t)$. These assumptions lead to the following system of equations:

$$\frac{dS}{dt} = qN(t) - \omega(t) \frac{S(t)}{N(t)} I(t) - \mu S(t), \quad (2.1)$$

$$\frac{dI}{dt} = \omega(t) \frac{S(t)}{N(t)} I(t) - (\nu + \mu) I(t), \quad (2.2)$$

$$\frac{dN}{dt} = rN(t) - \nu I(t), \quad (2.3)$$

where $dN/dt = dS/dt + dI/dt$ and $r = q - \mu$ are the net growth rate of the adults in the absence of the epidemic.

Let $\phi_{ss}(t) = I(t)/N(t)$ be the HIV prevalence in the SS model. Then

$$\frac{d\phi_{ss}}{dt} = \frac{N(dI/dt) - I(dN/dt)}{N^2} = \frac{1}{N} \frac{dI}{dt} - \frac{1}{N} \frac{1}{N} \frac{dN}{dt} \quad (2.4)$$
Figure 2.1. Time trends in number of unprotected sexual contacts for high- and low-prevalence settings used for model projections.

Figure 2.2. Single stage (SS) and staged progression (SP) HIV prevalence projections for a high prevalence setting, $C_0 = 44$.

Using (2.2) and (2.3), we get

$$\frac{d\phi_{ss}}{dt} = \frac{1}{N} \left[ \omega(t) \frac{SI}{N} - (\nu + \mu)I \right] - \frac{I}{N} \frac{1}{N} \left[ rN(t) - \nu I \right]. \quad (2.5)$$
Substituting with $S/N = 1 - I/N$ and replacing $I/N$ with $\phi_{ss}$ give

$$\frac{d\phi_{ss}}{dt} = \omega(t)\phi_{ss}(1 - \phi_{ss}) - (v + \mu)\phi_{ss} - r\phi_{ss} + \nu\phi_{ss}^2. \quad (2.6)$$

Rearranging and factorising, we get

$$\frac{d\phi_{ss}}{dt} = \phi_{ss}[\omega(t) - (v + q) - (\omega(t) - \nu)\phi_{ss}], \quad (2.7)$$

which can be written as

$$\frac{d\phi_{ss}}{dt} = \phi_{ss}(a(t) - b(t)\phi_{ss}), \quad (2.8)$$

where $a(t) = \omega(t) - (v + q)$ and $b(t) = \omega(t) - \nu$.

3. The SP model

3.1. Model assumptions. The SP model presented uses three stages of HIV infection prior to development of AIDS. The first stage which commences soon after infection, lasting on average six weeks, is highly infectious. The second stage lasts for an average of 6 years. It is characterised by a long period of low infectivity. The third stage lasts for an average of 2 years, is also highly infectious and has been explained to be a result of the impaired immune system of the infected person. The ratio of infectiousness of these three stages has been reported as $100 : 1 : 10$ \cite{2–4}.

In the SP model, we assume that the frequency of sexual contacts between susceptibles and infectives varies with the stage of HIV infection. We note that this assumption introduces behavioural changes in response to time since infection. We also assume that the frequency of sexual contacts varies with time since the beginning of the epidemic. Further, we assume that the rate of progression from each stage $i$ ($i = 1, 2, 3$) of HIV infection to the next stage is constant over time.

3.2. Variables and parameters. Let $S(t)$ be the number of susceptibles and let $I_1(t), I_2(t),$ and $I_3(t)$ be the number of infectives in the corresponding stages of HIV infection. We describe the different changes in each of the epidemiological classes as follows.

(a) Susceptibles are recruited from young people on reaching the age of 15 ($a = 15$). Of all the children born free of HIV/AIDS ($t - a$) units of time ago at a per capita birth rate of $\lambda$, a proportion $p = e^{-\mu_1 a}$, are assumed to survive to age $a = 15$ and hence join the adult sexually active age group, where $\mu_1$ is the mortality rate for persons aged below 15 years. Susceptibles are removed through acquiring HIV infection or by natural death $\mu$. If infectives on average have $c_i$ ($i = 1, 2, 3$) sexual contacts, then a proportion of these contacts are with the susceptible population. Assuming proportional mixing, the proportion of contacts that are with susceptibles is equal to the prevalence of the susceptible class in the population, $S/N$. Hence, $c_i(S/N)$ is the average number of sexual contacts between infectives in stage $i$ and the susceptible population. Assuming proportional mixing, the probability of transmission per sexual act with infectives in stage $i$ is $\beta_i$ then $\sum_{i=1}^3 \beta_i c_i I_i/N$ is the number of susceptibles infected per susceptible per unit time. Multiplying this number by $S(t)$ gives
the total number of susceptibles infected per unit time that move on to the first stage of HIV infection. (b) The infectives joining the first stage are removed through progression to stage two at a rate \( \nu_1 \) and through a natural death rate \( \mu \). Infectives in stage two progress to stage three at a rate \( \nu_2 \) (and die of natural death at a rate \( \mu \)). Infectives in stage three progress to the AIDS stage at a rate \( \nu_3 \).

The SP model is hence described by the following system of differential equations:

\[
\begin{align*}
\frac{dS}{dt} &= p\lambda N - \left[ \frac{\beta_1 c_1 I_1}{N} + \frac{\beta_2 c_2 I_2}{N} + \frac{\beta_3 c_3 I_3}{N} \right] S - \mu S, \quad (3.1) \\
\frac{dI_1}{dt} &= \left[ \frac{\beta_1 c_1 I_1}{N} + \frac{\beta_2 c_2 I_2}{N} + \frac{\beta_3 c_3 I_3}{N} \right] S - (\nu_1 + \mu) I_1, \quad (3.2) \\
\frac{dI_2}{dt} &= \nu_1 I_1 - (\nu_2 + \mu) I_2, \quad (3.3) \\
\frac{dI_3}{dt} &= \nu_2 I_2 - (\nu_3 + \mu) I_3, \quad (3.4)
\end{align*}
\]

where \( N = S + \sum_{i=1}^{3} I_i \) and \( I = I_1 + I_2 + I_3 \).

### 3.3. Model analysis.

The prevalence equation for the SP model is obtained as follows. Let \( \phi_{sp} = \sum_{i=1}^{3} I_i/N, \phi_i = I_i/N \), then

\[
\frac{d\phi_{sp}}{dt} = \sum_{i=1}^{3} \left( \frac{N(dI_i/dt) - I_i(dN/dt)}{N^2} \right) = \sum_{i=1}^{3} \left( \frac{1}{N} \frac{dI_i}{dt} - \phi_i \frac{1}{N} \frac{dN}{dt} \right). \quad (3.5)
\]

From (3.2) to (3.4)

\[
\begin{align*}
\sum_{i=1}^{3} \frac{dI_i}{dt} &= S \sum_{i=1}^{3} \beta_i c_i \frac{I_i}{N} - \nu_3 I_3 - \mu \sum_{i=1}^{3} I_i, \quad (3.6) \\
\frac{dN}{dt} &= rN - \nu_3 I_3, \quad r = p\lambda - \mu,
\end{align*}
\]

we have

\[
\frac{d\phi_{sp}}{dt} = \frac{1}{N} \left( S \sum_{i=1}^{3} \beta_i c_i \frac{I_i}{N} - \nu_3 I_3 - \mu \sum_{i=1}^{3} I_i - \phi(rN - \nu_3 I_3) \right). \quad (3.7)
\]

Since \( S/N = 1 - \phi_{sp} \), then

\[
\frac{d\phi_{sp}}{dt} = (1 - \phi_{sp}) \sum_{i=1}^{3} \beta_i c_i \phi_i - \nu_3 \phi_3 - \mu \phi_{sp} - \phi_{sp} (r - \nu_3 \phi_3). \quad (3.8)
\]
If we let $\sum_{i=1}^{3} \beta_i c_i \phi_i = \overline{\beta \bar{c}} \phi_{sp}$, where $\overline{\beta \bar{c}}$ are the weighted transmission and contact rates of the three infectious stages, then after rearranging, we have

$$\frac{d\phi_{sp}}{dt} = (1 - \phi_{sp}) \overline{\beta \bar{c}} \phi_{sp} - q \phi_{sp} - \nu_3 \phi_{sp} (1 - \phi_{sp}).$$  (3.9)

4. Analytical comparison Of SS and SP models

In the special case where $\phi_{sp} = k \phi_{sp}$, that is, the ratio of the prevalence of infectives in the third stage of HIV infection to the total prevalence of infectives in stages 1 to 3 is a constant, then the SP model reduces to

$$\frac{d\phi_{sp}}{dt} = \phi_{sp} (\overline{\beta \bar{c}} - q - \nu_3 k - (\overline{\beta \bar{c}} - \nu_3 k) \phi_{sp}).$$  (4.1)

In this special case, the prevalence equation of the SP model reduces to the prevalence equation (2.8) of the SS model. However, the prevalence of infectives in stage three is not likely to have a linear relationship with the total prevalence over an extended epidemic period as required by $\phi_{sp} = k \phi_{sp}$. This may mean that the SS model is a poor approximation of the SP model in all conceivable cases.

In the general case, subtracting the SP prevalence equation (3.9) from the SS prevalence equation (2.8) gives the difference denoted by $d$ as

$$d = \phi_{ss} [a(t) - b(t) \phi_{ss}] - [(1 - \phi_{sp}) \overline{\beta \bar{c}} \phi_{sp} - (r + \mu) \phi_{sp} - \nu_3 \phi_{sp} (1 - \phi_{sp})].$$  (4.2)

Substituting for $a(t), b(t), r$ and setting $\overline{\beta \bar{c}} = \omega(t)$ gives

$$d = \phi_{ss} \omega - \phi_{ss} q - q \phi_{ss} - \omega \phi_{ss}^2 + \nu \phi_{ss}^2 - \omega \phi_{sp} + \omega \phi_{sp}^2 + q \phi_{sp} + \nu_3 \phi_{sp} - \nu_3 \phi_{sp} \phi_{sp}.$$  (4.3)

Simplifying, we get

$$d = (\phi_{ss} - \phi_{sp}) [\omega (1 - (\phi_{ss} + \phi_{sp})) - q] + \nu_3 \phi_{sp} (1 - \phi_{sp}) - \nu \phi_{ss} (1 - \phi_{ss}).$$  (4.4)

As $\phi_{ss} \to 0, \phi_{sp} \to 0$, the product terms in $\phi_{ss}, \phi_{sp}, \phi_{sp}$ vanish. This gives

$$d \approx (\phi_{ss} - \phi_{sp}) (\omega - q) + \nu_3 \phi_{sp} - \nu \phi_{ss}.$$  (4.5)

In particular, at the beginning of the epidemic, both the SS and SP models use the same initial prevalence and noting that the prevalence of infectives in the third stage of HIV infection is initially negligible ($\phi_{sp} \approx 0$), then from (4.5), $d$ is negative. This implies that $\phi_{sp}$ initially increase at a faster rate than $\phi_{ss}$. Towards the end of the epidemic, we assume that most of the few remaining infectives are in stage three. Hence $\phi_{ss} \approx \phi_{sp} \approx \phi_{sp}$. Denote this prevalence by $\phi_f$. Then again from (4.5), $d \approx (\nu_3 - \nu) \phi_f$. But $\nu_3 \gg \nu$. Hence $d$ is a positive quantity implying that the projected SS prevalence decreases at a faster rate towards the end of the epidemic than the SP prevalence.

Since $d < 0$ at the beginning of the epidemic and $d > 0$ at the end of the epidemic, from the intermediate value theorem, it follows that there exist a $\phi_{ss}$ and a $\phi_{sp}$ such that $d = 0$. 
Let this correspond to an SP prevalence of $\phi_e$ and an SS prevalence of $\kappa\phi_e$, where $\kappa$ is a constant. Then,

$$
\frac{d}{dt} = \phi_e(\kappa - 1)(\omega(1 - \phi_e(\kappa + 1)) - q) + \nu_3\phi_3(1 - \phi_e) - \nu_3\phi_e(1 - \kappa\phi_e).
$$

(4.6)

At equilibrium, (4.5) gives $\phi_3 = \nu\phi_e/\nu_3$, hence

$$
\frac{d}{dt} = \phi_e(\kappa - 1)(\omega(1 - \phi_e(\kappa + 1)) - q) + \nu\phi_e(1 - \phi_e) - \nu\kappa\phi_e(1 - \kappa\phi_e).
$$

(4.7)

Setting $d = 0$ and solving for $\phi_e$ give

$$
\phi_e = \frac{\nu + q - \omega}{(\kappa + 1)(\nu + \omega)}
$$

(4.8)

provided $\omega < \nu + q$.

This means that the switch from faster changes in SP prevalence (to faster changes in SS prevalence) can only happen when the parameter $\omega$ has reduced to values less than $\nu + q$. Note that $\nu + q$ is the net removal due to development of AIDS symptoms ($\nu$) plus the removal effect resulting from recruitment of HIV negative persons reaching 15 years of age. Hence we deduce that the switch is more likely to happen after the peak prevalence since from (2.8) the maximum SS prevalence of $1 + q/(\nu - \omega)$ occurs when $\omega > \nu$. This occurs before $\omega$ was reduced to less than $\nu + q$ that is required to reach $\phi_e$. Hence $\phi_e$ is obtained after reaching the peak SP prevalence. Moreover model simulations indicate that the peak prevalences for the SP and SS models occur about the same time as shown in Figures 2.2, 4.1 except when an inflated probability of transmission is used in the SS model.

It can be noted that $\phi_e$ decreases with the increasing $k$ (since $\kappa$ is a particular case of $k$). In the situation where $0 < k < 1$ implying that the SS prevalence is less than the SP prevalence, which we have shown that it is the most likely scenario. This may imply that when the SS and SP prevalences are close to each other (as $k \to 1$), the equilibrium point corresponding to $d = 0$ is attained earlier. In addition, a sensitivity analysis indicates that $\phi_e$ is higher for lower values of $\omega$. This may suggest that a faster reduction in the parameter $\omega$ may produce a much greater discrepancy between the prevalences from the SS and SP models projections.

5. Simulations

Table 2.1 shows parameter values that were used in model simulations. The frequency of sexual contact $c$ was assumed constant for the first eight years and later assumed to decrease according to the lower-bounded function. For high-prevalence settings, these contacts decrease from 144 in the eighth year to approach a lower bound of 36 contacts per year as $t \to \infty$. This fits available data that suggests that $c$ was initially 144 in the early 1980s [13] but decreased to 107 by the mid 1990s [14] in some communities in Uganda. For the low-prevalence setting, unprotected sexual contacts decrease from 108 in the eighth year to approach a lower bound of 27 contacts per year as $t \to \infty$. The probabilities of infection per sexual contact for the SP model were chosen to give an SP peak prevalence of 16% (older than 15 years) as observed in antinatal clinic data in Uganda [15]. The equilibrium
SS probability of transmission of 0.019 was derived from the SP probabilities at the equilibrium prevalence level which is equal to \( \frac{\beta_1/\nu_1 + \beta_2/\nu_2 + \beta_3/\nu_3}{1/\nu_1 + 1/\nu_2 + 1/\nu_3} \). The elevated SS probability of transmission was assigned to achieve similar initial SS and SP prevalence projections.

Figure 2.2 shows the resulting prevalence projections for the SS and SP models for high-prevalence settings (initial number of 144 unprotected sexual acts). Figure 4.1 corresponds to a low-prevalence setting with an initial number of 108 unprotected sex acts per year.

The graphs show that for the same behaviour parameter changes in \( c \) and comparable probabilities of transmission, the SP model gives much more marked peak prevalence compared to the SS model both in low- and high-prevalence settings. The results show that the magnitude of the peak prevalence is grossly underestimated by the SS model. Attempts to obtain accurate prevalence estimates for the initial years of the epidemic using the SS model by using an elevated probability of transmission result in an inaccurately higher peak prevalence well beyond the SP prevalence and an overestimated epidemic during its declining phase. However, both the SS and SP models give similar positioning of the year at which the peak prevalence is reached in both low- and high-prevalence settings.

6. Discussion

The analytical results show that from the same SS and SP initial HIV prevalences, the SP prevalence increase at a faster rate (and hence reaches a higher peak) than the SS prevalence during the initial phase of the HIV epidemic. Hence the SP model gives a steeper
rise in HIV prevalence than the SS model. Assuming that the SP model gives a better HIV prevalence projection (since the SP model represents the HIV dynamics more accurately than the SS model), then HIV prevalence projections based on SS models underestimate the true HIV prevalence during the initial epidemic growth phase and consequently give a lower peak prevalence than the SP model.

The analytical results further show that towards the end of the epidemic, the SS prevalence has a steeper decline than the SP prevalence. It follows that sometimes during the epidemic there is a switch from faster SP prevalence changes to faster SS prevalence changes. This occurs at a prevalence that depends on the removal rate $\nu$ and the recruitment rate $q$. It follows from both analytical results and simulations that the prevalence projections by the SS and the SP differ greatly. Moreover, fitting an SS model to the existing prevalence data in the rising phase of the epidemic may grossly overestimate the peak prevalence and the entire declining phase of the epidemic.

Our findings are however subject to several limitations. First, the probability of transmission applied to the SS model assumed equilibrium conditions. This assumption may not be appropriate to the earlier phases of the epidemic due to more infections in the first stage of HIV infection leading to a higher initial probability of transmission for the SS model. However this cannot be easily accounted for by an elevated probability of transmission since the equilibrium conditions are later attained, hence requiring changing to a lower probability of transmission. Second, the time trend in the number of unprotected sexual contacts may not be as smooth as the function for the contact rate used in the model suggests. However, this function is more realistic given its lower boundedness. The function is also particularly realistic for Uganda since the initial interventions in the late 1980s were highly intense resulting in immediate reductions in number of sexual partners [16]. Lastly, we assumed several constant parameters over the fifty-year simulation period including rates of HIV progression ($\nu_1$, $\nu_2$, and $\nu_3$), mortality and birth rates and we also ignored the effects of mother-to-child transmission on population size.

These findings have significant implications on the choice of the HIV model to use for national HIV prevalence projections. Our comparison of the SS and SP models suggests that the SP model provides better projections than the SS model. Moreover extra parameters required by the SP model are not sought from national survey data but from existing literature.

Acknowledgment

The authors thank the referees for the valuable comments, and East African Universities Mathematics programme (EAUMP) for funding the article processing charges.

References


F. Baryarama: Department of Mathematics, Makerere University, P.O. Box 7062, Kampala, Uganda

*Email address*: feb3@ug.cdc.gov

J. Y. T. Mugisha: Department of Mathematics, Makerere University, P.O. Box 7062, Kampala, Uganda

*Email address*: jytmugisha@math.mak.ac.ug