HIV Infection, Antiretroviral Therapy, and CD4+ Cell Count Distributions in African Populations

Brian G. Williams,1 Eline L. Korenromp,3,5 Eleanor Gouws,4 George P. Schmid,2 Bertran Auvert,4,8 and Christopher Dye1

1Stop TB Department and 2HIV/AIDS Department, World Health Organization; 3Global Fund to Fight AIDS, Tuberculosis and Malaria, and 4Policy, Evidence and Partnerships Department, Joint United Nations Programme on HIV/AIDS, Geneva, Switzerland; 5Department of Public Health, Erasmus MC, University Medical Centre Rotterdam, Rotterdam, The Netherlands; 6INSERM U687 and 7University of Versailles–Saint Quentin, Faculté de Médecine Paris-Ile-de-France-Ouest, Saint Maurice, and 8AP-HP, Hôpital Ambroise Pare, Boulogne, France

Background. The variability in CD4+ cell counts within and among human immunodeficiency virus (HIV)–positive and –negative African populations has not been explained but has important implications for understanding the incidence of HIV-related opportunistic infections, especially tuberculosis, in both individuals and populations.

Methods. In HIV-negative African adults, CD4+ cell counts vary within populations (interquartile ranges [IQRs], 169–603 cells/μL) and among populations (means vary from 699 to 1244 cells/μL), with similarly wide variations in HIV-positive adults. We developed dynamic mathematical models to predict the distribution of CD4+ cell counts in HIV-positive adults using the distribution in HIV-negative adults.

Results. Under the assumption that survival is independent of the CD4+ cell count before seroconversion, we fitted the observed distributions in HIV-positive adults. At a CD4+ cell count of 200 cells/μL, the median life expectancy of HIV-positive Zambians (4.0 years) was predicted to be 1.7 times that of HIV-positive South Africans (2.3 years).

Conclusions. The model provides a way to estimate the changing distribution of CD4+ cell counts and, hence, the changing incidence of HIV-related opportunistic infections as the epidemic matures. This could substantially improve the planning of health services, including the need and demand for antiretroviral therapy. Better data are needed to test the model and its assumptions more rigorously and to fully understand the variability in CD4+ cell counts within and among populations.

CD4+ cells orchestrate the immune response to attack by HIV, but HIV invades CD4+ cells and uses them to replicate itself. Soon after infection with HIV, CD4+ cells counts decrease by approximately one-quarter and then decrease slowly thereafter [1]. When the CD4+ cell count reaches ~350 cells/μL, HIV-related opportunistic infections become evident; at 200 cells/μL, the person is classified as having AIDS regardless of the appearance of other opportunistic infections; and at ~50 cells/μL, patients die. CD4+ cell counts have been studied as markers of the progression of HIV infection [2–4], as a measure of the relative risk of developing opportunistic infections [5], to estimate the impact of HIV and the use of antiretroviral drugs on the epidemiological progression of tuberculosis (TB) [1], and to estimate the proportion of malaria that is attributable to HIV in sub-Saharan Africa [6]. The World Health Organization (WHO) recommends antiretroviral therapy (ART) for people in WHO clinical disease stage IV, regardless of CD4+ cell count, those in stage III with CD4+ cell counts >350 cells/μL, and those with CD4+ cell counts <200 cells/μL, regardless of clinical stage [7].

CD4+ cells counts in HIV-negative people are influenced by genetic, immunological, physiological, and behavioral factors [8–14], and they vary widely both within and among populations. In the present article,
we consider the implications of this variation on the distribution of CD4+ cells within and among populations of HIV-positive people as HIV epidemics mature. We have assembled data on CD4+ cell count distributions and examined variability among and within HIV-negative and -positive populations. We have evaluated competing models of CD4+ cell count decrease after infection with HIV and have shown that the distribution of CD4+ cell counts in HIV-positive adults in Africa can be predicted using the distribution in HIV-negative adults, time trends in HIV prevalence, and the survival distribution after infection, assuming only that survival is independent of the preinfection CD4+ cell count. We have used the best-fitting model to investigate the survival of HIV-positive people in Zambia and South Africa in relation to their CD4+ cell counts.

**METHODS**

To account for the effect of the maturity of the HIV epidemic on CD4+ cell counts in people infected with HIV, we used data on time trends in HIV prevalence and on survival after seroconversion to estimate time trends in HIV incidence and mortality (appendix, eqq. [1]–[3]). Knowing how many people were infected during each calendar year and the survival distribution, we calculated the distribution of time since infection at any given time (appendix, eqq. [4] and [5]). From the distribution of time since infection, we estimated the distribution of CD4+ cell counts (scaled as in the appendix, eq. [6]), which depended on assumptions that we made about the decrease in CD4+ cell count with time since infection (appendix, eq. [7]), and then the distribution of the scaled CD4+ cell counts as a function of time since infection (appendix, eq. [8]) for a given preinfection CD4+ cell count. Allowing for variation in the preinfection CD4+ cell count gave the actual distribution of CD4+ counts as a function of the time since infection (appendix, eq. [9]). Finally, we were able to calculate the distribution of CD4+ cell counts at any given time (appendix, eq. [10]).

In applying this model, we assumed that survival is independent of preinfection CD4+ cell counts, but we also considered a competing model in which survival depended entirely on the preinfection CD4+ cell count. In both models, we assumed that CD4+ cell counts decreased linearly with time after the initial rapid decrease between infection and seroconversion.

We used data from population surveys in Africa about CD4+ distributions in HIV-positive and -negative people, data about HIV prevalence from studies of women attending antenatal clinics, and data about the pattern of decrease in CD4+ cell counts from infection to death in HIV-positive people from clinical cohorts.

**Survey data.** Population surveys of CD4+ cell count distributions among HIV-negative and -positive adults in Africa were identified in PubMed, using the keywords “CD4,” “HIV,” “survey,” “Africa,” and “adult.” References in relevant publications were checked to identify additional studies. Only studies of adults >15 years old were included, because CD4+ cell counts are higher in children than in adults and are higher in young children than in older children [8], so lower age cutoffs would have made it more difficult to compare studies.

A study in Orange Farm, Gauteng, South Africa [15], gave individual CD4+ cell counts and a study in Lusaka, Zambia [16], gave the distribution of CD4+ cell counts in intervals of 100 cells/μL [16]. These studies were used to validate the model. Of the remaining studies, most reported only median and interquartile ranges (IQRs) for CD4+ cell counts, although some reported mean values and certain deciles (table 1). To estimate the proportion of people in each CD4+ cell count category, we fitted the cumulative distributions to log-normal functions [25].

Both automated [8, 9, 15–17, 19, 21–23] and manual [20, 26] techniques were used to measure CD4+ cell counts, and this may have affected the measurements [26]. However, the 2 methods used in most of the studies on which we report (FACScan and FACScount) gave consistent results [27, 28], and laboratory variations were generally much smaller than variations due to physiological factors within and among individuals [14]. Both the Zambian [16] and South African [15] studies used the FACScount technique, which is a single-platform direct-measurement technique. Measurement techniques were unlikely to account for the differences in the CD4+ cell count distributions in Zambia and South Africa or for much of the variation within each study population.

**Pattern of CD4+ cell count decrease.** Several studies supported the assumption that survival is independent of the CD4+ cell count before HIV seroconversion. In an Italian cohort, the CD4+ cell count soon after seroconversion was not associated with progression to AIDS [29]. In a multicenter study, HIV-positive people who progressed quickly, slowly, or not at all to AIDS over the course of 2–6 years had similar CD4+ cell counts at seroconversion [30, 31]. In a Swiss cohort, the rate of decrease was directly proportional to the initial CD4+ cell count [32], as was predicted by our model. The results of a meta-analysis suggested that the rate of decrease was higher at CD4+ cell counts <200 cells/μL, but the overall trend with time was close to linear [1].

CD4+ cell counts decrease significantly immediately after infection [2, 33–35]. In men in the United States, CD4+ cell counts decreased by 203 cells/μL (24% of the preinfection level) [2], 221 cells/μL (22%) [35], and >350 cells/μL (34%) within 6 months of seroconversion [34]. In a Swiss cohort, the median CD4+ cell count at a median 5.5 months after the estimated time of seroconversion was 546 cells/μL [32], which was considerably lower than the typical value in HIV-uninfected Swiss adults [36]. In Zambia, the CD4+ cell count in HIV-seropositive individuals at 15 years old were included, because CD4+ cell counts are higher in children than in adults and are higher in young children than in older children [8], so lower age cutoffs would have made it more difficult to compare studies.

A study in Orange Farm, Gauteng, South Africa [15], gave individual CD4+ cell counts and a study in Lusaka, Zambia [16], gave the distribution of CD4+ cell counts in intervals of 100 cells/μL [16]. These studies were used to validate the model. Of the remaining studies, most reported only median and interquartile ranges (IQRs) for CD4+ cell counts, although some reported mean values and certain deciles (table 1). To estimate the proportion of people in each CD4+ cell count category, we fitted the cumulative distributions to log-normal functions [25].

Both automated [8, 9, 15–17, 19, 21–23] and manual [20, 26] techniques were used to measure CD4+ cell counts, and this may have affected the measurements [26]. However, the 2 methods used in most of the studies on which we report (FACScan and FACScount) gave consistent results [27, 28], and laboratory variations were generally much smaller than variations due to physiological factors within and among individuals [14]. Both the Zambian [16] and South African [15] studies used the FACScount technique, which is a single-platform direct-measurement technique. Measurement techniques were unlikely to account for the differences in the CD4+ cell count distributions in Zambia and South Africa or for much of the variation within each study population.

**Pattern of CD4+ cell count decrease.** Several studies supported the assumption that survival is independent of the CD4+ cell count before HIV seroconversion. In an Italian cohort, the CD4+ cell count soon after seroconversion was not associated with progression to AIDS [29]. In a multicenter study, HIV-positive people who progressed quickly, slowly, or not at all to AIDS over the course of 2–6 years had similar CD4+ cell counts at seroconversion [30, 31]. In a Swiss cohort, the rate of decrease was directly proportional to the initial CD4+ cell count [32], as was predicted by our model. The results of a meta-analysis suggested that the rate of decrease was higher at CD4+ cell counts <200 cells/μL, but the overall trend with time was close to linear [1].

CD4+ cell counts decrease significantly immediately after infection [2, 33–35]. In men in the United States, CD4+ cell counts decreased by 203 cells/μL (24% of the preinfection level) [2], 221 cells/μL (22%) [35], and >350 cells/μL (34%) within 6 months of seroconversion [34]. In a Swiss cohort, the median CD4+ cell count at a median 5.5 months after the estimated time of seroconversion was 546 cells/μL [32], which was considerably lower than the typical value in HIV-uninfected Swiss adults [36]. In Zambia, the CD4+ cell count in HIV-seropositive individuals at
Table 1. CD4+ cell count distributions among HIV-negative and -positive adults in Africa.

<table>
<thead>
<tr>
<th>Group, location [reference]</th>
<th>Time</th>
<th>HIV prevalence, %*</th>
<th>Sample type</th>
<th>Median age, years (range)</th>
<th>Sample size</th>
<th>CD4+ cell count, cells/μL</th>
<th>Fitted upper and lower quartiles, cells/μL</th>
<th>Mean</th>
<th>Median</th>
<th>CD4+ cell count, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIV-negative</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bandim, Guinea Bissau [18]</td>
<td>1990–1992</td>
<td>0.5</td>
<td>HM B</td>
<td>&gt;15</td>
<td>51</td>
<td>1151 ± 182</td>
<td>1000 ± 1050 ± 134</td>
<td>778–1381</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Canchungo, Guinea Bissau [19]</td>
<td>1989–1991</td>
<td>0.3</td>
<td>HM B</td>
<td>&gt;15</td>
<td>133</td>
<td>892 ± 107</td>
<td>819 ± 783 ± 73</td>
<td>552–1098</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Orange Farm, South Africa [15]</td>
<td>2002</td>
<td>22</td>
<td>HM M</td>
<td>26 (15–49)</td>
<td>707</td>
<td>1179 ± 36</td>
<td>1116 ± 1115 ± 27</td>
<td>845–1341</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Shirar, Tanzania [21]</td>
<td>1993</td>
<td>13</td>
<td>HM B</td>
<td>37 (&gt;15)</td>
<td>147</td>
<td>980 ± 25</td>
<td>968 ± 958 ± 23</td>
<td>876–1045</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>East Uganda [8]</td>
<td>2002</td>
<td>4.3</td>
<td>HM W</td>
<td>&gt;19</td>
<td>207</td>
<td>819 ± 60</td>
<td>762 ± 757 ± 42</td>
<td>577–987</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Lusaka, Zambia [16]</td>
<td>1999</td>
<td>35</td>
<td>HM M</td>
<td>34 (18–84)</td>
<td>113</td>
<td>869 ± 60</td>
<td>780 ± 792 ± 50</td>
<td>627–992</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>HIV-positive</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Orange Farm, South Africa [15]</td>
<td>2002</td>
<td>22</td>
<td>HM B</td>
<td>26 (15–49)</td>
<td>196</td>
<td>537 ± 82</td>
<td>475 ± 443 ± 41</td>
<td>239–640</td>
<td>18</td>
<td>26</td>
</tr>
</tbody>
</table>

NOTE. B, both men and women; FW, factory workers; HM, household members; M, men; MW, mine workers; VCT people attending voluntary counseling and testing services; W, women.

* HIV prevalence is that found in the study, or, if not reported in the study, an estimate for the surrounding area.
the first measurement (389 cells/µL) was much lower than that in uninfected individuals (774 cells/µL) [16, 33]. The weighted average of the US studies gave a CD4+ cell count decrease of 26% between HIV infection and seroconversion.

There have been few data on CD4+ cell count at death in HIV-infected African adults who do not receive ART. Two hospital-based studies in Abidjan, Cote d’Ivoire, found medians of 58 and 88 cells/µL [37, 38] and a mean of 141 cells/µL [37] at the last measurement before death from AIDS. A population-based study in rural Uganda found a median of 61 cells/µL at the last measurement within 6 months before death [39]. We assumed that death occurs at a CD4+ cell count of 50 cells/µL (C in the appendix, eq. [6]).

RESULTS

Data on CD4+ cell count distributions in HIV-negative and -positive adults from population surveys performed in Africa between 1988 and 2004 are given in table 1 and figure 1. Among 12 HIV-negative populations, median CD4+ cell counts varied from 670 cells/µL (Wonji, Ethiopia) [9] to 1160 cells/µL (Kampala, Uganda) [23], whereas, within populations, the upper limits of the IQRs were up to twice the lower limits. In Chongwe, Lusaka [16] had median CD4+ cell counts of 1290 cells/µL (Addis Ababa, Ethiopia) [17] to 598 cells/µL (Shirati, Tanzania) [21]; 13% of Tanzanians in Shirati [21] but 62% of Zambians in Lusaka [16] had CD4+ cell counts <350 cells/µL (table 1). CD4+ cell count distributions also varied considerably among HIV-positive populations (table 1 and figure 1), with medians ranging from 296 cells/µL (Kamal, Ethiopia) [9, 17] to 598 cells/µL (Shirati, Tanzania) [21]; 13% of Tanzanians in Shirati [21] but 62% of Zambians in Lusaka [16] had CD4+ cell counts <350 cells/µL. The median CD4+ cell count in HIV-positive people in Shirati, Tanzania [21], was similar to that in HIV-negative people in Ethiopia [9, 17] (table 1 and figure 1).

Time trends of HIV prevalence for Lusaka, Zambia [40], and Gauteng, South Africa [41], are shown in figure 2A. In Lusaka, HIV prevalence had reached 27% in 1992 and remained steady until 2002, whereas, in Gauteng, it increased from ~2% in 1992 to 32% in 2002. From time trends in HIV prevalence and under the assumption of Weibull survival with a median of 9.0 years and a shape parameter of 2.25 [1, 41], we estimated the incidence of infection (appendix, eq. [3]), as illustrated for South Africa in figure 2B, and then the probability distribution of times since seroconversion (appendix, eq. [5]), as shown in figure 2C. At the start of the epidemic in South Africa, most people had been recently infected (1990, purple line; 1995, blue line). Because the incidence of HIV in South Africa peaked in 1996, in 2000 (green line), people were most likely to have been infected 4 years earlier, whereas, in 2005 (orange line), they were most likely to have been infected 9 years earlier. The probability distribution of the scaled CD4+ cell counts could then be calculated (appendix, eq. [8]). Figure 2D shows that, at the start of the epidemic in South Africa, most infected people had high scaled CD4+ cell counts (appendix, eq. [6]), relative to the initial value, which was set to 1. In 2000, the frequency distribution of CD4+ cell count peaked at ~70%; in 2005, it peaked at ~40% of its initial value. Once the epidemic stabilizes, all scaled CD4+ cell count values become equally likely, relative to their initial values (2010, red line; 2040, black line).

Finally, we used the observed distribution of CD4+ cell count in HIV-negative people to predict the actual distribution of CD4+ cell counts in HIV-positive people (appendix, eq. [10]), as shown in figure 3A and 3C for South Africa and in figure 3B and 3D for Zambia. The predicted CD4+ cell count distributions among HIV-positive adults (figure 3C and 3D, blue lines) agreed well with the observed distributions. The decrease in CD4+ cell count between infection and seroconversion, as estimated from published studies, is 26%. By varying this parameter, we were able to determine a maximum-likelihood fit to the CD4+ cell distribution of HIV-positive people in South Africa and, thus, to estimate its value. This gave an initial decrease in the CD4+ cell count of 26.5% ± 5.9%, which is in agreement with the value obtained from published studies [2, 33–35].

The CD4+ cell count distributions in HIV-negative people were higher in South Africa than in Zambia, and the epidemic in South Africa is at an earlier stage than that in Zambia (figure 2A). For both reasons, the CD4+ cell count distributions among HIV-negative people were higher in South Africa than in Zambia (figure 3C and 3D). However, survival in our model was independent of the initial CD4+ cell count, so that the average rate of decrease in CD4+ cell count in HIV-positive people was
Figure 2.  

A, Prevalence of HIV over time in Orange Farm, Gauteng, South Africa [41] (red), and Lusaka, Zambia [40] (blue), from national antenatal clinic surveys. Points give data with 95% confidence limits, and lines are fitted logistic curves. B, HIV incidence (red line) and mortality (black line), estimated from HIV prevalence (green line), for South Africa. C, Distribution of time since seroconversion for HIV-positive South African adults in 1990 (purple), 1995 (blue), 2000 (green), 2005 (orange), 2010 (red), and 2040 (black). For the purpose of illustration, the area under each curve is scaled to the HIV incidence in that year. D, Distribution of scaled CD4+ cell counts (appendix, eq. [6]) among HIV-positive South Africans, relative to values before seroconversion (the value after the initial decrease in CD4+ cell count is set to 1, and the value at death is set to 0). Lines as in panel C.

higher in South Africa (85 cells/μL/year) than it was in Zambia (75 cells/μL/year), and, at a given CD4+ cell count, survival was shorter in South Africa than in Zambia (figure 4). At a CD4+ cell count of 200 cells/μL, for example, the median life expectancy predicted by the model was 2.3 years (IQR, 1.5–3.0 years) in South Africa but was 4.0 years (IQR, 2.7–5.4 years) in Zambia (figure 3), and the survival distribution of South Africans with a CD4+ cell count of 500 cells/μL was close to the survival distribution of Zambians with a CD4+ cell count of 350 cells/μL.

We considered an alternative model in which CD4+ cell counts decrease at the same rate for all individuals in a population, so that life expectancy is determined by the preinfection distribution of CD4+ cell count and people whose initial CD4+ cell count is high survive for longer than those for whom it is low. Varying the rate of decline to obtain the best fit, we were unable to fit the South African data (P < .001) (figure 3C, green line). Furthermore, the survival distribution, fitted to a Weibull curve, gave a shape parameter of 3.0 ± 0.3 for South Africa and 3.3 ± 0.5 for Zambia, which was significantly higher in both cases than the accepted value of 2.25 ± 0.05 [1] and implied a cumulative mortality 5 years after infection of ~4%, rather than 17%.

DISCUSSION

There are large variations in CD4+ cell count within and among HIV-negative populations. CD4+ cell counts decrease with age [8] and are, on average, ~100 cells/μL higher in women than in men [8–10]. Infections other than HIV (bacterial, protozoan, and helminthic) may affect CD4+ cell counts [11, 12]. In individuals, CD4+ cell counts may vary by up to 200 cells/μL in readings performed a few weeks apart [13, 14]. Intra- and interlaboratory variation may also contribute to this variation [13, 42]. Ideally, when monitoring CD4+ cell counts for the management of HIV-infected patients, one should perform repeated measurements over time, to reduce the influence of short-term and random fluctuations and to establish the trend in CD4+ cell counts in individual patients.

Despite the substantial variability in CD4+ cell counts within and among HIV-positive and -negative populations, we were able to predict the distribution of the former from that of the latter. The key assumption is that life expectancy is independent of the initial CD4+ cell count, and this was supported by the model predictions for South Africa and Zambia (figure 3). Furthermore, a study in Ethiopia found that, although CD4+ cell counts in HIV-positive people were lower than those elsewhere (figure 1), the life expectancy was similar to that in The Netherlands [43], which suggests that the rate of CD4+ cell count decrease in Ethiopia is proportionately slow and confirms the prediction of our model.

The large interpopulation variability in CD4+ cell counts may also help explain a recent observation in HIV-positive South Africans among whom the clinical disease stage was a stronger

1454 • JID 2006:194 (15 November) • Williams et al.
predictor than CD4+ cell count for the risk of both AIDS and death [44]. CD4+ cells as a percentage of total lymphocyte count, a value that is not affected by laboratory variations related to the measurement of total lymphocyte counts [26, 45, 46], might be considered as an alternative measure of immune status.

Two caveats should be borne in mind when interpreting our findings. First, some studies have suggested that the rate of CD4+ cell loss increases during the last 2 years before the onset of AIDS, and others have suggested that the square root of the CD4+ cell count declines linearly with increasing time, such that the rate of loss decreases with time since seroconversion [1]. Better data on time trends of CD4+ cell counts could help to refine our assumption of a linear decline. Second, we have assumed that the distribution of survival times after seroconversion in these African settings is the same as that in developed countries. The limited available data [39, 41] have suggested that this is so, but different survival curves would affect the analysis.

The model presented here could be used in different ways. First, CD4+ cell count data among HIV-negative populations and trends in HIV prevalence could be used to forecast CD4+ cell counts in HIV-positive populations. To the extent that CD4+ cell counts determine ART needs, we could more accurately predict the likely future demand for antiretroviral and other drugs. Given that much is known about the relationship between the risk of HIV-related opportunistic infections and CD4+ cell counts [5], this model could be used to predict future changes in the incidence of opportunistic infections, including TB, as the prevalence of HIV changes over time and CD4+ cell counts decrease. Second, the model shows how CD4+ cell count distributions among HIV-positive populations are influenced by the stage of the local HIV epidemic (figure 2D), and, hence, provide information about time trends of HIV prevalence and incidence during the preceding 10–20 years. In India [47], there is still some debate concerning the maturity and the likely future course of the epidemic. There is evidence that the prevalence of HIV has been declining in Harare, Zimbabwe, at ~9% per

The present study, which was based primarily on a comparison of data collected from 2 African countries, highlights the need to further explore and understand the factors that determine CD4+ cell count distributions in HIV-negative populations and their decrease over time in HIV-positive populations. However, to test the predictions of this model, better data are needed about (1) the distribution of CD4+ cell counts in HIV-negative populations and the reasons for this variation, (2) the initial and subsequent rate of decrease in CD4+ cell counts after seroconversion, and (3) the variation in survival times to AIDS among people in different populations. If such data were available and if the predictions of the present model were borne out, it may be possible to use CD4+ cell count data were available and if the predictions of the present model to refine the criteria for providing ART to individuals within these populations.

**Acknowledgments**

We thank Nico Nagelkerke, Kevin De Cock, and Peter Ghys, for comments on the manuscript. This project was initiated by the South African Centre for Epidemiological Modelling and Analysis.

**APPENDIX**

We developed a dynamic model of the decline in CD4+ cell count in HIV-positive people. Model inputs were (1) CD4+ cell counts in HIV-negative adults, (2) the decrease in CD4+ cell count immediately after infection, (3) the survival distribution after infection with HIV, and (4) time trends in HIV prevalence.

**Estimating incidence from prevalence.** Time trends in the prevalence of HIV infection, \( P(t) \), were obtained by fitting logistic curves to data about the prevalence of HIV among women attending antenatal clinics. If no one died of AIDS, \( I_0(t) \), the incidence of HIV at calendar time \( t \) would be equal to the time derivative of the prevalence, \( P^*(t) \), so that

\[
I_0(t) = \frac{dP^*(t)}{dt} . \quad (1)
\]

To correct for AIDS-related mortality, we needed the survival distribution, \( W(t) \), which we took to be a Weibull function with median of 9.0 years and a shape parameter of 2.25 [1]. Then, \( D(t) \), the probability (per unit of time) that people die \( t \) years after initial infection, was

\[
D(t) = -\frac{dW(t)}{dt} . \quad (2)
\]

Because deaths reduce prevalence, we added deaths from previous incident infections to our initial estimate of incidence, so that

\[
I(t) = I_0(t) + \int I(t - \tilde{t}) D(\tilde{t}) d\tilde{t} \quad (3)
\]

(in eq. [3] and below, \( t \) and \( \tilde{t} \) indicate calendar time, and \( \tilde{t} \) indicates time since seroconversion.) We solved equation (3) iteratively, because the integrand depended only on the incidence at times preceding time \( t \). Mortality followed directly from incidence and survival.

**Distribution of time since infection as a function of time.** The number of people infected at time \( \hat{t} \) before time \( t \) and surviving to time \( t \) was

\[
N(\hat{t}|t) = I(t - \hat{t}) W(\hat{t}) , \quad (4)
\]

so that the probability density function (PDF) of time since infection, \( \hat{t} \), at time \( t \) was

\[
P(\hat{t}|t) = N(\hat{t}|t)/\int N(\hat{t}|t) d\hat{t} . \quad (5)
\]

**CD4+ cell count distribution as a function of time.** We assumed that the CD4+ cell count decreases to a value of \( C_i \) immediately after seroconversion and then decreases linearly with time [1] to a value of \( C_f \) when the person dies. We assumed initially that survival was independent of the preinfection CD4+ cell count [29–31]. To simplify the analysis, we scaled the CD4+ cell count, \( C \), to get a new variable, \( c \), where

\[
c = \frac{C - C_f}{C_i - C_f} , \quad (6)
\]

so that, as \( C \) decreases from \( C_i \) to \( C_f \), the scaled variable \( c \) decreases from 1 to 0. (We introduced \( c \) purely as a mathematical device to simplify the analysis; precisely the same final result could be obtained without introducing this new variable.)
Then, if a person dies at time \( \hat{t} \) after seroconversion, the scaled CD4\(^+\) cell count \( c \) at time \( \hat{t} \) after seroconversion is

\[
c = 1 - \frac{\hat{t}}{\hat{t}} .
\]  

(7)

The probability density function of \( \hat{t} \) could thus be given by \( D(\hat{t}) \), as defined in equation (2), so that the probability density of \( c \), at time \( \hat{t} \) could be given by \( P(c|\hat{t}) \, dc = D(\hat{t}) \, d\hat{t} \). From equation (7), we then obtained

\[
P(c|\hat{t}) = D(\hat{t}) \frac{d\hat{t}}{dc} = D\left(1 - \frac{\hat{t}}{c}\right) \frac{\hat{t}^2}{\hat{t}} .
\] 

(8)

We also needed to allow for the distribution of the initial CD4\(^+\) cell counts \( C_i \). Let the PDF of the CD4\(^+\) cell count in HIV-negative people be \( P(C) \). Then the PDF of the CD4\(^+\) cell count among those people who seroconverted at time \( \hat{t} \) before the present time could be given as

\[
P(C|\hat{t}) = \int_0^c P(c|\hat{t}) P(C_i|dC_i) .
\] 

(9)

Finally, from equation (5), the PDF of the CD4\(^+\) cell count in all HIV-seropositive people at time \( t \) was

\[
P(C|t) = \int_0^\infty P(C|\hat{t}) P(\hat{t}|t) \, d\hat{t} .
\] 

(10)

We evaluated equation (10) numerically using Visual Basic in Excel (version 2003; Microsoft).

**Distribution of survival at a given CD4\(^+\) cell count.** Consider an HIV-infected person whose initial scaled CD4\(^+\) cell count is \( a \), who seroconverted at time \( \hat{t} \) before the present, and who dies at time \( \hat{t} \) after seroconversion. If their present scaled CD4\(^+\) cell count is \( c \), their present life expectancy is \( \tau = \hat{t} - \hat{t} \), where

\[
\tau = \frac{\hat{c}}{a} .
\] 

(11)

The PDF of the present life expectancy, \( \tau \), for a given value of \( c \) could then be determined using the formula

\[
P(\tau) \, d\tau = \int_0^\infty \int_0^\infty P(a) P(\hat{t}) P(\hat{t}) \, da \, d\hat{t} \ ,
\] 

(12)

where the integral is over the surface defined by equation (11) with \( \tau \) and \( c \) fixed. Equation (12) was evaluated using Monte Carlo integration. We generated sets of points \( a, \hat{t}, \) and \( \hat{t} \) from the distribution appropriate to each; kept all points for which the present CD4\(^+\) cell count lay within a small range of \( c \) (typically ±10 cells/\( \muL \)); and, for each point, calculated the corresponding value, and hence, the PDF of \( \tau \).

**References**