The Dynamics of CD4⁺ T-Lymphocyte Decline in HIV-Infected Individuals: A Markov Modeling Approach

Ira M. Longini, Jr., W. Scott Clark, Lytt I. Gardner, and John F. Brudnaj

Division of Biostatistics, School of Public Health, Emory University, Atlanta, Georgia and Department of Epidemiology, Division of Preventive Medicine, Walter Reed Army Institute of Research, Washington, D.C.

Summary: We modeled the decline of CD4⁺ T-lymphocytes (T4 cells) in HIV-infected individuals with a continuous-time Markov process. The model partitions the HIV infection period into six progressive T4-cell count intervals (states), followed by a seventh state: a definitive HIV-infection end point, i.e., AIDS diagnosis or Walter Reed stage 6 (opportunistic infections). The Markov model was used to estimate the state-specific progression rates from data as functions of important progression cofactors. We applied the model to data on 1,796 HIV-positive individuals in the U.S. Army. The estimated mean waiting time from seroconversion to when the T4-cell count persistently drops below 500/mm³, but is greater than 349/mm³, is 4.1 years, and the waiting time to a T4-cell count of less than 200/mm³ is estimated at 8.0 years. The estimated rate of T4-cell decline was higher for HIV-infected individuals with initially high numbers of T4 cells, but the estimated rate of decline remains relatively uniform when the T4-cell count dropped persistently below 500/mm³. The opportunistic infection incubation period, i.e., the time from seroconversion to opportunistic infection diagnosis, is estimated at 9.6 years. Age is found to be an important cofactor. The estimated mean opportunistic infection incubation periods are 11.1, 10.0, and 8.9 years for the youngest (≤25 years old), the middle (26–30 years old), and the oldest (>30 years old) age groups, respectively. Moreover, we found that the rate of progression is the same for all three age groups when T4-cell counts are ≥500/mm³, but it is faster for the two older age groups when T4-cell are <500/mm³. Key Words: T-lymphocytes—Stages of infection—Natural history—Markov model—Opportunistic infection—Incubation period.

One of the primary targets of the human immunodeficiency virus (HIV) in the host is CD4⁺ T-lymphocytes (T4 cells), and T4-cell decline is a leading indicator of HIV infection progression. In HIV-infected individuals, T4-cell levels are used as indicators signaling when antiviral treatments and prophylaxis against opportunistic infections (OIs) and other HIV-related illnesses should begin. Thus, it seems logical to base HIV progression rates on T4-cell decline. One of the first HIV staging systems to be devised, the Walter Reed (WR) system (1), is partially based on T4-cell cell levels: WR stages 1–2 indicate that an individual's T4-cell count is ≥400 T4 lymphocytes/mm³, and WR stages 3–6 indicate that the T4-cell count is <400/mm³. Brundage et al. (2) recently used T4-cell count intervals to characterize the progression of HIV disease in 988 HIV-infected individuals in the U.S. Army. Longini (3) recognized that their system could be formulated as a discrete-time Markov chain, allowing the extensive analytic machinery of Markov processes (4)

Address correspondence and reprint requests to Dr. I. M. Longini at Division of Biostatistics, School of Public Health, Emory University, Atlanta, GA 30322, U.S.A.
to be applied to estimating and predicting HIV-infection progression rates. Longini et al. (5,6) had previously applied a continuous-time Markov process based on clinical indicators to model the progression of HIV-infected individuals through the stages of infection. Estimated progression rates have been coupled with the backcalculation method (7) to predict the stage-specific numbers of HIV-infected patients from AIDS incidence curves for selected populations (8–10).

In this research, we employ a continuous-time Markov model to define a system of HIV disease progression based on T4-cell count intervals that we refer to as states of HIV infection. The model includes cofactors that affect the state-specific progression rates. We apply the model to data on 1,796 HIV-infected individuals in the U.S. Army.

METHODS

The Markov Model Based on T4-Cell Decline

We define the eight states of HIV infection based on T4-cell count in Fig. 1. Persons in state 0 are not infected, while those in states 1–7 are infected and antibody positive. State 7 represents an end point in terms of HIV infection, e.g., AIDS diagnosis and an HIV OI diagnosis, which is the same as Walter Reed stage 6—see Redfield et al. (1). Individuals in state 8 are deceased. We assume that infectives flow reversibly through the states and thus the states can be thought of as being stages of infection. Individuals may progress at different rates, depending on the cofactors. We categorize cofactor (e.g., age, gender, and therapy use) levels as \( i = 1, \ldots, I \). Then we define the state-specific, monthly progression (i.e., hazard) rate as \( \lambda_{ik} \) for an individual with cofactor level \( i (i = 1, \ldots, I) \) who is in state \( k \) \((k = 0,1, \ldots, 7)\). The mathematical formulation of the Markov model based on these progression rates is a straightforward extension of our previous models (3,5,6).

The mean waiting time in state \( k \) for an individual with cofactor level \( i \) is \( \mu_{ik} = 1/\lambda_{ik} \), and the mean waiting time to go from state \( j \) to \( k \) is

\[
\mu_{ijk} = \sum_{r=j}^{k-1} \mu_{ir}, \quad 1 \leq j < k \leq 8,
\]

e.g., the OI incubation period is \( \mu_{i17} = \mu_{i1} + \cdots + \mu_{i7} \). We are also interested in the rate of T4-cell decline (per month) within state \( k \). This rate is simply \( r_k = w_k \lambda_{ik} \) (T4 cells/month), where \( w_k \) is the width of the T4-cell interval \( k \), e.g., \( w_2 = 200 \) T4 cells.

Maximum likelihood estimates (MLEs) of the parameters in the arrays \( \lambda_i = (\lambda_{i1}, \lambda_{i2}, \ldots, \lambda_{i7}) \), \( i = 1, \ldots, I \), are estimated from data via previously described methods (5,6). The asymptotic multivariate normality properties of MLEs are used to conduct hypothesis tests concerning the state-specific waiting times.

Perspective Criteria for T4-Cell Decline

T4-cell levels in individuals exhibit considerable variation due to measurement error and fluctuating physiologic conditions (11,12). A single aberrantly low T4-cell count could result in an individual being assigned to a more advanced (later) state of infection than he or she is actually in. We control for some of these fluctuations in the T4-cell count by making the T4-cell categories relatively broad (ranges of 150–200 T4-lymphocytes/mm\(^3\)—see Fig. 1). In addition, we devised persistence criteria to control for aberrant T4-cell counts. These criteria, given below, vary according to whether the individual in question was observed to seroconvert or whether he or she was observed to be seropositive.

---

**FIG. 1.** The flow through the eight states of infection. State 0 represents the pool of HIV-negative individuals who are exposed to HIV infection.
at his or her first exam, i.e., a prevalent seropositive.

For the prevalent seropositives, the higher of their first two T4-cell count measurements determines their initial state. To move to a later state (i.e., a lower T4-cell count level), two consecutive measurements at later states are needed to confirm a true reduction in total T4-cell count. If progression to a later state had occurred, the state to which the individuals had progressed is determined by the minimum of the two later states. For example, an individual seen in state 2 at time 0 may have a T4-cell count of 320/mm³ (state 5) 12 months later, and a T4-cell count measure of 410/mm³ (state 4) 6 months after that. This individual would then be considered to be in state 4 12 months after time 0.

For the seroconverters, if the time between the last seronegative and the first seropositive exam was 1 year or less, then we use the midpoint as the seroconversion time. Otherwise, we use the first seropositive exam time as the starting point for the seroconverter's progression. Then the persistence criteria described above are employed. Seroconverters are started in state 1, provided that their initial T4-cell count total is above 899/mm³; otherwise, they are treated the same as the prevalent seropositives at their first HIV-positive measurement.

RESULTS

We used data on HIV-1-infected individuals in the U.S. Army. Since the U.S. Army uses the WR staging system (1) for categorizing the progression of HIV disease, we use OIs in lieu of AIDS, i.e., state 7 of the Markov process. Virtually all personnel of the U.S. Army have been screened for HIV-1 infection at least once, and those found to be HIV positive are carefully followed (2,12,13). Between June 1985 and April 1990, a total of 1,796 HIV-infected individuals had at least two seropositive exams. Among these individuals, 1,533 were seropositive at their first exam, and 263 seroconverted. There was an average of 4.2 (± 2.0) exams per HIV-positive person, and the average time between exams was 6.9 (± 4.5) months. Because the time of death may not be available for many individuals, we did not include state 8 in our analysis. In addition, we did not have information on HIV therapy.

We first consider progression without cofactors, i.e., $\lambda_{ik} = \lambda_k$. Using the ML procedure, we fitted the model of the infection histories to the 1,796 individuals in the study. Table 1 gives the number of observed transitions in the data, e.g., there were 775 total observed transitions from state 2 to states 2–7, 48 of which were from state 2 to 4. There were 5,451 contributions (observed transitions) to the likelihood function from these 1,796 people. The MLEs of the parameters $\theta$ (= $\lambda_1, \ldots, \lambda_6$) and the MLEs of the state-specific mean waiting times $\mu$ (= $\mu_1, \ldots, \mu_6$) are given in Table 2. Thus, the estimated mean OI incubation period is 9.6 ± 0.2 years. The estimated mean waiting time from seroconversion to when the T4-cell count persistently drops below 500/mm³, but is greater than 349/mm³, is 4.1 ± 0.1 years, and the mean waiting time to a T4-cell count of less than 200/mm³ is 8.0 ± 0.2 years.

The state-specific rates of T4-cell decline, $r_{ik}$ in T4 lymphocytes/mm³/month are given in Table 3. The rate of decrease is highest in state 1 (22.9 ± 1.5 T4 cells/month) and steadily declines to state 4 (6.4 ± 0.3 T4 cells/month). The rate of decrease then remains relatively constant for the remaining two states. The waiting time in state 6 is the time from when an individual’s T4-cell count drops below 200/mm³ to the onset of an OI. The median T4-cell count at the time of OI diagnosis was 70/mm³. Thus, we use a T4-cell range of 130 T4 cells for state 6.

<table>
<thead>
<tr>
<th>From state</th>
<th>T4-cell count (mm³)</th>
<th>To state</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>&gt; 899</td>
<td></td>
<td>245</td>
<td>84</td>
<td>59</td>
<td>19</td>
<td>6</td>
<td>1</td>
<td>1</td>
<td>415</td>
</tr>
<tr>
<td>2</td>
<td>700–899</td>
<td></td>
<td>541</td>
<td>175</td>
<td>48</td>
<td>9</td>
<td>0</td>
<td>2</td>
<td>775</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>500–699</td>
<td></td>
<td>1,067</td>
<td>274</td>
<td>68</td>
<td>12</td>
<td>5</td>
<td>1,426</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>350–499</td>
<td></td>
<td>1,210</td>
<td>266</td>
<td>42</td>
<td>6</td>
<td>1,524</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>200–349</td>
<td></td>
<td>723</td>
<td>153</td>
<td>30</td>
<td>906</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0–199</td>
<td></td>
<td>353</td>
<td>52</td>
<td>405</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>5,451</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

We consider an individual’s age at his or her first exam as a cofactor for progression. We stratify on the following three age groups: \( i = 1, \) age \( \leq 25 \) years; \( i = 2, 25 < \text{age} \leq 30 \) years; and \( i = 3, \) age \( > 30 \) years. Table 4 gives the estimated mean waiting times. There is a marked decrease in the OI incubation period with increasing age. The OI incubation period is longest for the youngest age group (estimated to be 11.1 years), and it is significantly different (longer) \((p < 0.001)\) from the OI incubation period of the oldest age group (estimated to be 8.9 years). The OI incubation period for the middle age group (estimated to be 10.0 years) is significantly different (longer) from that of the oldest age group \((p < 0.05)\) but not from that of the youngest age group. The age-specific cumulative distribution functions (CDFs) for the OI incubation period based on the Markov model are shown in Fig. 2. From the figure, the median OI incubation periods are 10.4, 9.4, and 8.4 years for the youngest, middle, and oldest age groups, respectively. Table 4 shows no significant differences among the age groups for the first three states of infection. However, the waiting times for the last three states decrease significantly with increasing age. The mean waiting time in state 4 for the youngest age group \((\mu_{14} = 2.4 \text{ years})\) is significantly different \((p < 0.01)\) from that for the oldest age group \((\mu_{34} = 1.7 \text{ years})\). For state 5, the mean waiting time for the youngest age group \((\mu_{15} = 2.6 \text{ years})\) is significantly different \((p < 0.05)\) from that for the oldest age group \((\mu_{36} = 1.9 \text{ years})\). For state 6, the mean waiting time for the oldest age group \((\mu_{36} = 1.3 \text{ years})\) is significantly different \((p < 0.05)\) from that for both the youngest and middle age groups \((\mu_{16} = \mu_{26} = 2.0 \text{ years})\).

The estimated mean waiting times for each age group from seroconversion to the point at which the T4-cell count persistently drops below 500/mm³ but is greater than 349/mm³ are 4.1 \( \pm 0.2, 4.1 \pm 0.2, \) and 4.0 \( \pm 0.2 \) years for the youngest, middle, and oldest age groups, respectively. In contrast, the waiting times for a T4-cell count of less than 200/mm³ are 9.1 \( \pm 0.5, 8.0 \pm 0.3, \) and 7.6 \( \pm 0.3 \) years for the youngest, the middle, and oldest age groups, respectively.

**DISCUSSION**

The Markov modeling approach yields a unifying description of T4-cell decline that is consistent with previous analyses. Longini (3) previously used a discrete-time Markov model to analyze the one-step, T4-cell decline transition matrix estimated by Brundage et al. (2) from a subset of the data analyzed here. The continuous-time model employed

---

**TABLE 3. Rate of T4-cell decline—no cofactors**

<table>
<thead>
<tr>
<th>State ( k )</th>
<th>T4-cell count ((/\text{mm}^3))</th>
<th>Rate of T4-cell decline ((\mu_1^{(4)} \text{ SE})) ( (\text{T4 cells/months}) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( &gt;899 )</td>
<td>22.9 (1.5)</td>
</tr>
<tr>
<td>2</td>
<td>700-899</td>
<td>13.3 (0.7)</td>
</tr>
<tr>
<td>3</td>
<td>500-699</td>
<td>10.0 (0.4)</td>
</tr>
<tr>
<td>4</td>
<td>350-499</td>
<td>6.4 (0.3)</td>
</tr>
<tr>
<td>5</td>
<td>200-349</td>
<td>6.1 (0.3)</td>
</tr>
<tr>
<td>6</td>
<td>( 0-199 )</td>
<td>6.9 (0.5)</td>
</tr>
</tbody>
</table>

\( ^a \) We assume that state 1 has a range of 300 T4 cells.

\( ^b \) We assume that state 6 has a range of 130 T4 cells.

---

**TABLE 4. Estimated mean waiting times \( \mu_i \) in each state of infection, with age as a cofactor**

<table>
<thead>
<tr>
<th>State ( k )</th>
<th>T4-cell count ((/\text{mm}^3))</th>
<th>Mean waiting times ( \mu_i ) ((\text{SE})) years, age group ( i )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( &lt;25 )</td>
<td>( 26-30 )</td>
</tr>
<tr>
<td>1</td>
<td>( &gt;899 )</td>
<td>1.1 (0.1)</td>
</tr>
<tr>
<td>2</td>
<td>700-899</td>
<td>1.4 (0.1)</td>
</tr>
<tr>
<td>3</td>
<td>500-699</td>
<td>1.6 (0.1)</td>
</tr>
<tr>
<td>Cumulative ( \geq 500 )</td>
<td>4.1 (0.2)</td>
<td>4.1 (0.2)</td>
</tr>
<tr>
<td>4</td>
<td>350-499</td>
<td>2.4 (0.2)</td>
</tr>
<tr>
<td>5</td>
<td>200-349</td>
<td>2.6 (0.3)</td>
</tr>
<tr>
<td>Cumulative ( \geq 200 )</td>
<td>9.1 (0.5)</td>
<td>8.0 (0.3)</td>
</tr>
<tr>
<td>6</td>
<td>( 0-199 )</td>
<td>2.0 (0.3)</td>
</tr>
<tr>
<td>Cumulative*</td>
<td>11.1 (0.5)</td>
<td>10.0 (0.4)</td>
</tr>
</tbody>
</table>

Number of individuals

\( 555 \)

\( 513 \)

\( 728 \)

\( ^a \) OI incubation period.
here utilizes more information than the discrete-time model, and thus the former model is more efficient in terms of extracting information from the data. Nevertheless, the results from the discrete-time model agree closely with those found here. The estimated rate of T4-cell decline decreases as the state of infection increases from states 1-4 (see Table 3). Then the estimated rate of T4-cell decline remains relatively constant for states 5 and 6. Brundage et al. (12) estimated the rate of T4-cell decline by computing the median slopes of decline for overlapping T4-cell intervals from data on HIV-infected individuals in the U.S. Army and Navy. They also found a decrease in the rate of T4-cell decline as the state of infection increased for T4-cell counts above 500/mm$^3$ (i.e., states 1-3). Their estimates of the magnitudes of the rate of T4-cell decline are close to ours although they have estimated slightly lower rates of decline in the 500-699/mm$^3$ range than we have estimated. As pointed out by Brundage et al. (12), these findings imply that it may be prudent for investigators to target therapeutic development to slow the rate of T4-cell decline (e.g., zidovudine) in the early states of HIV infection when the rate of T4-cell decline is most rapid and the immune system is relatively intact. Prevention of early rapid T4-cell decline could help prevent HIV-infected individuals from progressing rapidly to the later symptomatic states of HIV disease.

We find that age has a strong effect on the rate of HIV infection progression, for which the estimated mean OI incubation periods are 11.1, 10.0, and 8.9 years for the youngest (≤25 years old), the middle (26-30 years old), and the oldest (>30 years old) age groups, respectively. Moreover, we find that the rate of progression is the same for all three age groups when T4-cell counts are ≥500/mm$^3$ but faster for the two older age groups when T4-cell counts <500/mm$^3$. Other investigators have found that the waiting time until the appearance of pre-AIDS symptoms and the length of the AIDS incubation period both decrease with increasing age in individuals infected by infectious blood products (14-19), but they have not identified at what segment of the AIDS incubation period the age-specific change in progression rate occurs.

This phenomenon of increasing progression rates with increasing age in the later states of infection could be due to biological factors associated with aging. One such biological mechanism could involve the interaction of HIV and other latent infectious agents (accumulated with increasing age) with the cellular immune system (20-22). The nature of such an interaction should be the subject of further study. In addition, the phenomenon of increasing
progression rates associated with increasing age in the later states of infection could be partially due to the confounding effect of calendar time. Thus, effective treatment may have been more readily available to younger individuals who tended to be in the later states of infection later in calendar time, while older individuals tended to be in the later states earlier in calendar time before effective treatment was available. In order to control for the effect of calendar time, we stratified on the date January 1988 (results not given here). Although progression rates were slower after January 1988 than before, we found that the older individuals still progressed more rapidly than younger individuals in the later states of infection within both calendar time strata. This basic trend holds for other calendar time cut points after January 1988.

Our results are dependent on the Markov assumption that a person's progression rate to the next state is independent of his or her progression rate through previous states. This is a necessary simplifying assumption that must be made in order to obtain tractable forms for the transition probabilities and thus the likelihood function. Such tractability is needed when the data are as heavily censored as they are here (5,6). However, if the investigator can identify indicators of slower or faster progression, then the model can be stratified accordingly. We have done this for age. We feel that the Markov model provides a good description of HIV disease progression on the group or population level, but the model would be inadequate for describing the progression of a single individual. We note from Fig. 2 that the estimated CDF for the OI incubation period based on the Markov model and the Army data compares well with the CDF for the AIDS incubation period estimated by Brookmeyer and Goedert (19) from a cohort of individuals (>20 years old) with hemophilia.

We have applied an ad hoc method to control for fluctuations in T4-cell counts. An alternative approach would be first to smooth each individual's T4-cell counts via a moving average or some other smoothing device. However, such an approach would be difficult here, since there is an average of only 4.2 exams per individual. Another suggested approach for modeling and smoothing T4-cell decline is to treat some function of T4-cell count as a growth curve model that allows for individual tracking (23). The growth curve model provides a good alternative to the Markov model. Berman (24) modeled T4-cell decline directly (i.e., not in stages) as a stochastic process, which represents another attractive alternative modeling strategy.

The Markov model for T4-cell decline presented here could be used to assess the efficacy of therapies for early-state HIV infection clinical trials. A good measure of such efficacy could be the comparison of the rate of T4-cell decline in the therapy vs. control groups (12). The control group could be directly enrolled in the study, or it could be represented by the expected rate of T4-cell decline in the absence of therapy estimated from other studies. In either case, the state-specific progression rates would be compared within T4-cell count strata among the therapy and control groups. The rate of T4-cell count decline could also be used as a surrogate end point for clinical end points such as AIDS diagnosis in clinical trials involving later-state HIV infection.

A dominant theme in the search for therapies and vaccines against HIV infection has been the demand for a shortening of the testing and approval process that brings drugs and vaccines from development to the consumer. A recent report (25) concluded that the T4-cell count stands out as the leading candidate for a surrogate end point in clinical trials for HIV and AIDS. Machado et al. (26) show that under certain conditions the use of such a surrogate end point can result in substantial time savings in clinical trials. Other biological markers, such as β2-microglobulin (27) and neopterin (28), could be used as surrogate end points.

Acknowledgment: We thank M. Elizabeth Halloran and C. Robert Horshburg, Jr. for their helpful comments, and Mark Millazo and Robin Pomerantz from SRA Technologies, Inc. for preparing the data files that were used in our analysis. We also thank the physicians, nurses, and supporting personnel from the following eight Army Medical Centers (AMCs) (assisted by Sociometrics, Inc.) for collecting, reporting, and transmitting the data: Eisenhower AMC, Ft. Gordon, GA; Brooke AMC, Ft. Sam Houston, TX; Fitzsimmons AMC, Aurora, CO; Letterman AMC, San Francisco, CA; Madigan AMC, Tacoma, WA; Tripler AMC, Honolulu, HI; Walter Reed AMC, Washington, D.C.; and Beaumont AMC, El Paso, TX. This research was partially supported by the Division of HIV/AIDS, Centers for Infectious Diseases, Centers for Disease Control, Public Health Service, U.S. Department of Health and Human Services.

REFERENCES


