Accounting for leadtime in cohort studies: evaluating when to initiate HIV therapies

Stephen R. Cole*,†, Rui Li, Kathryn Anastos, Roger Detels, Mary Young, Joan S. Chmiel and Alvaro Muñoz

*Department of Epidemiology, Johns Hopkins University, Baltimore, MD, U.S.A.
†Lincoln Medical and Mental Health Center, New York, NY, U.S.A.
Department of Epidemiology, UCLA School of Public Health, Los Angeles, CA, U.S.A.
Division of Infectious Disease, Georgetown University Medical Center, Washington, DC, U.S.A.
Department of Preventive Medicine, Northwestern University Medical School, Chicago, IL, U.S.A.

SUMMARY

Commonly reported comparisons of differences in disease progression according to disease staging at therapy initiation may be subject to bias if they do not account for the time it took the deferred group to reach the latter stage (that is, leadtime) and for previous events in those who initiate therapy at late stage (that is, unseen fast progressors). To estimate the impact of deferring initiation of highly active antiretroviral therapies (HAART) on time to clinical AIDS in the context of data from observational cohort studies, we describe a method that capitalizes on data from a pre-HAART period to multiply impute estimated leadtimes and the unseen events among fast progressors. After accounting for leadtime and the unseen events, data from two large cohort studies (N = 739) indicate that deferring HAART initiation until CD4 is below 200 cells/mm³ was detrimental compared to initiating between 201 and 350 (hazard ratio = 1.97; 95 per cent confidence interval [CI] 1.09, 3.54), and that failure to account for leadtime resulted in a 38 per cent higher hazard ratio. In contrast, initiating HAART between 201 and 350 did not increase the hazard of AIDS compared to initiating with CD4 between 351 and 500 cells/mm³ (hazard ratio = 0.70; 95 per cent CI 0.35, 1.42). Methods presented here offer an approach to analysing prevalent cohort studies and provide procedures to maximize the usefulness of observational data. Copyright © 2004 John Wiley & Sons, Ltd.

KEY WORDS: bias; cohort study; multiple imputation; HIV; AIDS

1. INTRODUCTION

In individuals infected with the human immunodeficiency virus (HIV), the T-cells that express CD4 molecules are progressively depleted, so that the CD4 cell count provides a measure...
with which to stage the progression of disease due to HIV infection. With the advent of highly active antiretroviral therapies (HAART) in the mid-1990s [1,2] and recognition of the infrequent but serious adverse sequelae of HAART, the question of whether to initiate HAART at high CD4 cell counts versus deferring therapy until lower CD4 cell counts is of primary clinical interest [3]. Several observational studies have recently reported estimates of differences in disease progression by CD4 cell count at HAART initiation [4,5]. A major limitation acknowledged by these reports is the fact that comparisons of individuals beginning follow-up at heterogeneous disease stages may require adjustment for leadtime [6]. Therefore, the optimal time at which to initiate HAART remains unclear.

Leadtime is defined here as the prior unobserved person-time among individuals who began observation at a later disease stage. For example, when comparing individuals who initiate HAART at lower CD4 cell counts to individuals who start at higher CD4 cell counts, the leadtime from the earlier disease stage (that is, higher CD4 cell count) should be taken into account so that there is a common origin for the time-to-event analysis. In addition, methods accounting for leadtime must account for the unseen sample [7]. The unseen sample comprises fast progressors, namely the individuals with events who were unobserved due to the rapidity of their disease course. In this paper, we provide methods that incorporate both the leadtime and the unseen sample of fast progressors in an observational analysis examining when to initiate HAART.

2. THE MULTICENTER AIDS COHORT AND WOMEN’S INTERAGENCY HIV STUDIES

Data for this analysis were collected as part of two ongoing prospective cohort studies of the natural history of HIV infection that have systematically gathered detailed information on treatments and clinical outcomes prior to and after the widespread introduction of HAART around 1996. The first is the Multicenter AIDS Cohort Study (MACS) [8], which has followed 2761 HIV-positive homosexual men beginning in 1984 at sites located in Baltimore, Chicago, Pittsburgh and Los Angeles. The second cohort is the Women’s Interagency HIV Study (WIHS) [9], which has followed 2070 HIV-positive women beginning in 1994 at sites located in New York, Chicago, Los Angeles, San Francisco and Washington DC. Institutional review boards approved all protocols and informed consent forms, which were completed by study participants in both cohorts. Analyses presented here are based on 312 men and 427 women who were HIV-positive and initiated HAART prior to the onset of clinical AIDS between June 1995 and October 2000 with CD4 counts \( \leq 500 \) cells/mm\(^3\) and on whom the date of HAART initiation could be estimated to within a one-year interval. Follow-up was administratively censored at the earlier of either 31 December 2000 or 3.5 years after HAART initiation. Because the natural history [10] and response to HAART [4,5] are similar in men and women and the AIDS-event rate is low in the HAART era, the data from these two cohorts were combined to increase precision.

Planned study visits were conducted biannually and included an extensive interviewer-administered questionnaire and phlebotomy for the determination of CD4 cell count. Plasma HIV-1 RNA level has been measured concurrently since 1996. T-cell subsets were determined by immunofluorescence using flow cytometry in laboratories participating in the NIAID NIH Quality Assurance Program [11]. CD4 cell count at HAART initiation was considered to be
the CD4 measurement immediately prior to HAART initiation. HAART initiation date is defined as the midpoint of the interval between the last visit HAART-free and the first visit on HAART. The definition of HAART was based on guidelines from the IAS-USA [12] and DHHS/Kaiser panels [13]. Specific antiretroviral medications were self-reported by participants with the help of photo-medication cards at each visit. The number of nucleoside reverse transcriptase inhibitors (NRTIs), non-NRTIs (NNRTIs) and protease inhibitors (PIs) the participant was taking during the previous six months was calculated. HAART was defined according to combinations of the numbers of reported NRTIs, NNRTIs and PIs. Specifically, 92 per cent of the HAART regimens consisted of ≥2 NRTI + (≥1 NNRTI or ≥1 PI), and the remaining HAART regimens consisted of: 1 NRTI + ≥1 NNRTI + ≥1 PI; 1 NRTI + 0 NNRTI + ≥2 PI (including ritonavir and saquinavir); ≥3 NRTI + 0 NNRTI + 0 PI (containing abacavir).

When the exact date of clinical AIDS onset was unknown, it was considered to be the midpoint of the interval from the visit last seen AIDS-free and the first visit seen with AIDS. We used the 1993 CDC clinical criteria to define clinical AIDS [14], so that participants with no clinical conditions and CD4 counts ≤200 cells/mm$^3$ were not considered to have clinical AIDS. Physician or hospital records confirmed most reported AIDS cases in men, while women self-reported clinical AIDS. Deaths that occurred before onset of clinical AIDS ($n=28$) were censored at the time of death.

3. METHODS

Let $T_i$ be the random variable denoting the (possibly unobserved) time from HAART initiation to AIDS and $C_i$ be the random variable denoting the (possibly unobserved) time from HAART initiation to censoring for the $i$th individual. Let $T_i^* = (T_i \wedge C_i)$, $\Delta_i = I[T_i \leq C_i]$ and $D_i = I[CD4_i \leq k]$, where CD4 is measured immediately prior to HAART initiation and $k$ is a constant cutpoint such that $D_i$ is an indicator of ‘deferred’ HAART initiation. The observed data consist of independent and identically distributed (iid) replicates of $\{T^*, \Delta, D\}_i$. Below, we adopt usual conventions of representing realizations of random variables by lower-case letters and suppressing the individual index $i$ as all random variables are assumed iid.

In our example, we make two comparisons corresponding to setting $k=200$ and $k=350$. We restrict the observed data for $k=200$ so that the study population comprises those with CD4 counts below 351 cells/mm$^3$, and for $k=350$ we restrict those with CD4 count between 201 and 500 cells/mm$^3$ (inclusive), such that our CD4 group comparisons are ≤200 versus 201–350 and 201–350 versus 351–500 for ‘deferred’ versus ‘immediate’ HAART, respectively. Note that the subset of men and women who initiated HAART with CD4 between 201 and 350 cells/mm$^3$ appear in both comparisons: in the first as the ‘immediate’ group and in the second as the ‘deferred’ group. Below, we will describe our proposed method using the first comparison ($k=200$) without a loss of generality.

To fix ideas, consider a clinical trial to test the null hypothesis that deferral to CD4 ≤200 cells/mm$^3$ does not compromise the maintenance of AIDS-free status relative to immediately starting HAART while CD4 count is between 201 and 350 cells/mm$^3$. Such a deferment trial might randomize HIV-positive AIDS-free patients with CD4 count between 201 and 350 cells/mm$^3$ into either an immediate arm (initiate HAART at enrolment) or a deferral arm (initiate HAART when CD4 first reaches ≤200 cells/mm$^3$). Let $R=1$ indicate being randomized to deferred treatment with HAART. Those with $R=0$ are randomized to start HAART.
immediately. Let $S$ and $S^*$ be analogous to $T$ and $T^*$ above with time now measured from randomization date to AIDS, so that $S = T$ for the immediate group but $S = L + T$, where $L$ is the leadtime (that is, time to reach 200 CD4 cells), for the deferred group. The null hypothesis can be tested with an intent-to-treat analysis by comparing the time from randomization to AIDS between the deferred and immediate treatment arms, using a Cox proportional hazards model [15] of the form

$$\lambda_S(s) = \lambda_0(s) \exp(zR)$$

(1)

where $\lambda_0(s)$ is an unspecified hazard of AIDS among those who initiated HAART immediately with CD4 counts between 201 and 350 cells/mm$^3$ and $\exp(\hat{z})$ is the estimated hazard ratio for those randomized to deferred HAART. Note that, because randomization (asymptotically) ensures that the two study groups will have similar characteristics, $\exp(\hat{z})$ carries the causal interpretation of interest, namely a contrast between deferred versus immediate HAART regimes.

In the absence of such a trial, an unadjusted analysis of the observational data $\{T^*, \Delta, D\}_i$ may include fitting a Cox proportional hazards model such as

$$h_T(t) = h_0(t) \exp(\hat{\beta}D)$$

(2)

where $h_0(t)$ is an unspecified hazard of AIDS among those who chose to initiate HAART with CD4 counts between 201 and 350 cells/mm$^3$ and $\exp(\hat{\beta})$ is the estimated hazard ratio for those who chose to initiate HAART with CD4 counts $\leq 200$ cells/mm$^3$.

It is well understood that, in general, $\exp(\hat{\beta}) \neq \exp(\hat{z})$. First, this inequality arises because in the observational data the randomization instrument $R$ no longer exists [16]. Therefore, any comparison of persons who chose to initiate HAART with CD4 counts between 201 and 350 cells/mm$^3$ versus persons who chose to defer initiation until CD4 $\leq 200$ cells/mm$^3$ may be confounded by factors that both (i) predispose to early HAART initiation, and (ii) are independent predictors of time to AIDS. Therefore, any analysis must contend with an assumption that, conditional on measured covariates, no confounding remains [17,18]. Second, note that the time scale differs between models 1 and 2; specifically, in model 1 time is measured from CD4 between 201 and 350 for all participants but in model 2 time is measured from HAART initiation with CD4 $\leq 200$ cells/mm$^3$ for those who defer. Unless we are in the unlikely setting of constant hazards, for the observational analysis to be valid (that is, $\beta = z$), the unobserved leadtime among observed HAART-deferred individuals and the unseen sample of events among unobserved fast progressors must be appended to the observed data. Note that the deferment arm of the hypothetical trial described above would have recorded both the events of fast progressors and this leadtime. In such a trial, there would be individuals in the deferred arm who develop AIDS before becoming eligible to start therapy. These are the fast progressors who develop AIDS without reaching a defined level of immunosuppression indicative of initiating HAART. If deferment is a sound strategy, the number of such individuals will likely be small (that is, few individuals in the deferred arm develop AIDS with CD4 $> 200$ cells/mm$^3$) with the balance of individuals in the deferred arm reaching CD4 $\leq 200$ cells/mm$^3$ at $L$ years after randomization, at which time HAART is initiated.

Since the time of enrolment for the deferred group coincides with the initiation of HAART, the observational data $\{T^*, \Delta, D\}_i$ correspond to what Wang et al. [19] termed prevalent
treatment studies. Given that our primary interest is centred on the effect of deferring HAART initiation, ‘...external data on hazards for failure without treatment are necessary to derive any measure of treatment efficacy’ [19]. Therefore, the observed deferred group must be augmented with the unseen fast progressors. Further, this augmentation must come from an external information source corresponding to the hazard of AIDS in the absence of HAART.

The method we propose capitalizes on the existence of a period of HIV natural history where therapies did not appreciably alter the course of disease [2]. Specifically, using data collected from July 1989 to December 1995 (the $6\frac{1}{2}$ years prior to the introduction of HAART) on individuals who were AIDS-free with CD4 cell counts between 201 and 350 cells/mm$^3$, we use maximum likelihood (ML) on a parametric (that is, log-normal) analogue to the non-parametric cumulative incidence estimator of Kalbfeisch and Prentice [20] to estimate the proportion who reach CD4 $\leq 200$ cells/mm$^3$ before $t$ years without AIDS, denoted by $(1−π)×F(t)$, and the proportion who develop AIDS before $t$ years without reaching CD4 $\leq 200$ cells/mm$^3$, denoted by $π×G(t)$. For this mixture of $F$ and $G$, the contributions to the likelihood for (i) those who transition to the CD4 threshold $k$ before AIDS, (ii) those who incur AIDS before transitioning, and (iii) those who are censored without either event are (i) $(1−π)×f$, (ii) $π×g$, and (iii) $(1−π)×(1−F) + π×(1−G)$, respectively. The mixture proportion, $π$, corresponds to the overall proportion of fast progressors (that is, individuals who will develop AIDS before reaching CD4 $\leq 200$ cells/mm$^3$). Since AIDS was recorded on a continuous basis but CD4 cell count was measured at semiannual visits, for those few participants ($n=6$ and 2 for $k=200$ and 350, respectively) who ceased to attend visits without reaching CD4 $\leq 200$ and with a subsequent report of clinical AIDS, we assumed that CD4 crossed below 200 at the midpoint between the last observed CD4 and the onset of AIDS.

We then use the parametric ML estimates of the log-normal distributions $F$ and $G$ as well as the ML estimate of $π$ to multiply impute [21] the leadtime and the unseen sample under the assumption that imputed leadtimes and events are ignorable conditional on CD4 measured at HAART initiation. We chose to impute from a parametric log-normal distribution because it has been shown that this distribution appropriately describes the incubation period of AIDS [22,23]. In addition, using a parametric distribution facilitated the incorporation of the uncertainty due to the sampling variability for the imputation of the leadtime and unseen fast progressors. If $n$ is the number seen (that is, sample size of deferred group), $u$ is the number unseen (that is, number of fast progressors), $N$ is the total (that is, $n+u$), and $π$ is the proportion unseen, then $n=N×(1−π)$, so the number unseen is $u=n×π[1−π]^{-1}$ [24].

The completed (that is, imputed) data comprised three elements, namely $\{T^*,Δ,D\}_i$. For the $m$ individuals in the immediate group, $T^*_i=t_i$, $Δ_i=δ_i$ and $D_i=0$ for $i=1,...,m$. For the $n$ individuals in the deferred group after adding the leadtime, $T^*_i=F^{-1}(U_{i1})+t_i$, $Δ_i=δ_i$ and $D_i=1$ for $i=m+1,...,m+n$. To account for the $n×π[1−π]^{-1}$ unseen fast progressors in the deferred group we augment the data by $T^*_i=G^{-1}(U_{i2})$, $Δ_i=1$ and $D_i=1$ for $i=m+n+1,...,m+n+n×π[1−π]^{-1}$. Our analysis then consisted of fitting a Cox model to the data $\{T^*,Δ,D\}_i$ with

$$
\dot{λ}_T(\hat{t}^*) = \dot{λ}_0(\hat{t}^*) \exp(\gamma D)
$$

where $\exp(\hat{γ})$ is the estimated hazard ratio comparing deferred versus immediate HAART.
Table I. Characteristics of 312 men and 427 women HIV-positive with CD4 counts ≤500 cells/mm³ at HAART initiation.

<table>
<thead>
<tr>
<th>Characteristics at HAART initiation</th>
<th>Total (N = 739)</th>
<th>Men (N = 312)</th>
<th>Women (N = 427)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, median years</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caucasian, (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CD4, median; cells/mm³</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤200</td>
<td>32%</td>
<td>30%</td>
<td>54%</td>
</tr>
<tr>
<td>201–350</td>
<td>38%</td>
<td>39%</td>
<td>20%</td>
</tr>
<tr>
<td>351–500</td>
<td>30%</td>
<td>31%</td>
<td>26%</td>
</tr>
<tr>
<td>Log HIV-1 RNA, median; copies/ml*</td>
<td>4.4</td>
<td>4.5</td>
<td>5.0</td>
</tr>
<tr>
<td>HIV-1 RNA categories, (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤80 (= limit of detection)</td>
<td>3%</td>
<td>2%</td>
<td>0%</td>
</tr>
<tr>
<td>81–10000</td>
<td>34%</td>
<td>23%</td>
<td>15%</td>
</tr>
<tr>
<td>&gt;10000</td>
<td>63%</td>
<td>75%</td>
<td>85%</td>
</tr>
</tbody>
</table>

*Race was missing for 15 women; HIV-1 RNA was missing for 6 men and 26 women.
†AIDS events occurred within 3 ½ years of HAART initiation.

The entire analysis procedure was repeated $J = 50$ times. We estimated $\exp(\gamma)$ by the antilogue of the average of the $J\tilde{\gamma}$'s, namely $\tilde{\gamma} = J^{-1}\sum_{j=1}^{J} \tilde{\gamma}_j$. We estimated the variance of $\gamma$ by the canonical imputation variance formula of Rubin [21], namely

$$\hat{V}(\gamma) = J^{-1}\sum_{j=1}^{J} \hat{V}(\tilde{\gamma}_j) + (1 + J^{-1})(J - 1)^{-1}\sum_{j=1}^{J} (\tilde{\gamma}_j - \tilde{\gamma})^2.$$  

The result of implementing the proposed methods to account for leadtime is a hazard ratio comparing deferred versus immediate treatment, with deferral defined by the chosen cutpoint $k$. This leadtime adjusted hazard ratio will approximate the hazard ratio obtained from a randomized deferral trial (model in equation (1)) if the parametric imputation model is correct, the imputed data is missing at random conditional on CD4 cell count category, and there are no unmeasured common causes of therapy delay and AIDS. To assess the impact of accounting for leadtime, we contrast the leadtime adjusted estimate with the naive hazard ratio obtained from model in equation (2).

4. RESULTS

During the 3 ½ years after HAART initiation, 35 (11 per cent) of the 312 men and 51 (12 per cent) of the 427 women developed clinical AIDS. Ninety per cent of the men were Caucasian, while only 17 per cent of the women were Caucasian (the majority of women were African-American). Median CD4 cell counts at HAART initiation were comparable between men and women (Table I).

For the estimation of $F$ and $G$, we used data collected during the 6 ½ years prior to the introduction of HAART in the MACS. A description of the 981 HIV-positive men who comprise the natural history experience is provided in Table II. In the period between July 1989 and December 1995 we observed 545 AIDS events. The 545 individuals with incident
Table II. Characteristics of 981 HIV-positive MACS participants AIDS-free with CD4 count between 201 and 500 cells/mm³: July 1989 through December 1995.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Total</th>
<th>AIDS-free</th>
<th>Incident AIDS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N = 981</td>
<td>N = 436</td>
<td>N = 545</td>
</tr>
<tr>
<td>Age, median (quartiles) years</td>
<td>38 (34,43)</td>
<td>38 (33,43)</td>
<td>38 (34,42)</td>
</tr>
<tr>
<td>Caucasian, N (%)</td>
<td>860 (88%)</td>
<td>373 (86%)</td>
<td>487 (89%)</td>
</tr>
<tr>
<td>CD4, median (quartiles) cells/mm³</td>
<td>359 (290,426)</td>
<td>385 (316,450)</td>
<td>341 (268,403)</td>
</tr>
<tr>
<td>201–350</td>
<td>452 (46%)</td>
<td>163 (37%)</td>
<td>289 (53%)</td>
</tr>
<tr>
<td>351–500</td>
<td>529 (54%)</td>
<td>273 (63%)</td>
<td>256 (47%)</td>
</tr>
<tr>
<td>Log HIV-1 RNA, median (quartiles) copies/ml*</td>
<td>4.4 (3.9,4.8)</td>
<td>4.1 (3.6,4.5)</td>
<td>4.7 (4.2,5.0)</td>
</tr>
<tr>
<td>HIV-1 RNA &gt;10000 copies/ml, N (%)</td>
<td>289 (70%)</td>
<td>99 (55%)</td>
<td>190 (82%)</td>
</tr>
</tbody>
</table>

* HIV-1 RNA was unavailable for 571 individuals.

AIDS were not appreciably different in age or ethnicity, but had lower CD4 cell counts and higher HIV-1 RNA, when compared with the 436 who remained AIDS-free during this period.

Of the 289 men with initial CD4 count between 201 and 350 cells/mm³ who developed AIDS during follow-up, 28 (10 per cent) transitioned to AIDS with CD4 >200 cells/mm³ (that is, progressed to clinical disease before reaching a severe degree of immunodeficiency) and of the 163 who remained AIDS-free, 82 (50 per cent) were observed to reach CD4 ≤200 cells/mm³ with 81 (50 per cent) remaining AIDS-free and with CD4 >200 cells/mm³ by the end of follow-up. Of the 256 men with initial CD4 count between 351 and 500 cells/mm³ who developed AIDS during follow-up, 14 (5.5 per cent) transitioned to AIDS with CD4 >350 cells/mm³, and of the 273 who remained AIDS-free, 204 (75 per cent) were observed to reach CD4 ≤350 with 69 (25 per cent) remaining AIDS-free and with CD4 >350 cells/mm³ by the end of follow-up. The left-hand panel of Figure 1 shows the non-parametric estimates of the proportion that reached CD4 ≤200 cells/mm³ before t years without AIDS, and the proportion that developed AIDS before t years without reaching CD4 ≤200 cells/mm³. The right-hand panel of Figure 1 shows the non-parametric estimates of the proportion that reached CD4 ≤350 cells/mm³ before t years without AIDS, and the proportion that developed AIDS before t years without reaching CD4 ≤350 cells/mm³. The estimated value \( \hat{\pi} \) was 0.116 (SD = 0.029) and 0.030 (SD = 0.008) for \( k = 200 \) and \( k = 350 \), respectively. In addition, for \( k = 200 \), the fitted log-normal model for \( F \) had a median time of 0.96 years (quartiles 0.35, 2.6) and for \( G \) a median of 4.5 years (quartiles 1.1, 18.5). For \( k = 350 \), the fitted log-normal model for \( F \) had a median time of 0.88 years (quartiles 0.28, 2.7) and for \( G \) a median of 1.1 years (quartiles 0.48, 2.6). Figure 1 overlays the fitted curves from these log-normal models, scaled by \( \pi \).

For \( k = 200 \), we observed 278 individuals to start HAART immediately (that is, at CD4 between 201–350) and 237 who deferred to when CD4 ≤200. The median leadtime for the 237 deferred individuals was calculated for each imputation. The average of these medians was 1.0 (SD = 0.141) years for the transition from CD4 between 201 and 350 to CD4 ≤200 cells/mm³. The average estimated number of fast progressors (that is, those who developed AIDS before reaching CD4 ≤200) was 31.1 (SD = 10.084). For \( k = 350 \), we observed 224 individuals to start HAART immediately (that is, at CD4 between 351–500) and 278 who deferred to when CD4 was between 201 and 350. The average of the median leadtimes for the 278 deferred
Figure 1. Non-parametric cumulative incidence estimates of the time to the CD4 transition without AIDS (thick solid line), and the time to AIDS without CD4 transition (dashed line). Left panel is for CD4 cutpoint of 200 cells/mm$^3$ ($n = 452$) and right panel is for CD4 cutpoint of 350 cells/mm$^3$ ($n = 529$). Thin solid overlaid lines are from fitted log-normal models.

Figure 2. Survival curves for the time to AIDS, comparing those who initiate HAART with CD4 counts between 201 and 350 (thick solid line) to those who initiate with CD4 counts $\leq 200$ cells/mm$^3$ (dashed line) and accounting for leadtime (thin solid line).

The average estimated number of fast progressors (that is, those who developed AIDS before reaching CD4 $\leq 350$) was 0.9 (SD = 0.114) years for transition from CD4 between 351 and 500 to CD4 $\leq 350$ cells/mm$^3$. The average estimated number of fast progressors (that is, those who developed AIDS before reaching CD4 $\leq 350$) was 8.52 (SD = 2.922).

Figure 2 presents the unadjusted and leadtime adjusted Kaplan–Meier curves for time to AIDS by initiating HAART with CD4 between 201 and 350 versus deferring to CD4 $\leq 200$ cells/mm$^3$. The naive hazard ratio for AIDS among those initiating HAART with CD4 $\leq 200$ versus with CD4 between 201 and 350 cells/mm$^3$ was 2.71 (95 per cent CI 1.61, 4.54). After accounting for leadtime, initiating HAART with CD4 $\leq 200$ versus with CD4 between 201 and 350 cells/mm$^3$ was associated with a two-fold increased hazard of AIDS (adjusted hazard ratio = 1.97; 95 per cent CI 1.09, 3.54). The analysis not accounting for leadtime overestimated this adjusted hazard ratio by 38 per cent.
Figure 3 presents the unadjusted and leadtime adjusted Kaplan–Meier curves with corresponding hazard ratios for time to AIDS by initiating HAART with CD4 between 351 and 500 versus deferring to CD4 between 201 and 350 cells/mm$^3$. The naive hazard ratio for AIDS among those initiating HAART with CD4 between 201 and 350 versus CD4 between 351 and 500 cells/mm$^3$ was 0.86 (95 per cent CI 0.46, 1.60). After accounting for leadtime, initiating HAART with CD4 between 201 and 350 versus CD4 between 351 and 500 cells/mm$^3$ appeared to carry a small benefit in AIDS-free survival (adjusted hazard ratio = 0.70; 95 per cent CI 0.32, 1.39). However, we did not have the precision to rule out chance as a reasonable alternative explanation for this finding.

5. DISCUSSION

We demonstrated a method to estimate the effect of initiating HAART at varying biomarker values on the time to AIDS. Accounting for leadtime among persons who chose to defer treatment adds time to individuals’ observed survival and therefore elevates the survival curve for the deferred group. However, correctly accounting for the unseen fast progressors lowers the survival function to which the leadtimes had been added. The balance between elevating the survival curve by adding the leadtime and lowering it by including the unseen fast progressors determines the net change in the adjusted estimator compared to the unadjusted analysis of the observed data. After accounting for leadtime and unseen fast progressors, data from two large cohort studies indicate that deferring HAART initiation until CD4 ≤ 200 cells/mm$^3$ is detrimental with respect to the onset of clinical AIDS, but 38 per cent less than the naive analysis used in previous reports [5]. On the other hand, initiating HAART with CD4 between 201 and 350 appears similar (and perhaps slightly preferable) to initiating with CD4 between 351 and 500 cells/mm$^3$.

Time since seroconversion may be a more appealing metric with which to make decisions about therapy timing. However, our method is based on the use of a primary marker of disease progression (that is, CD4 cell count) because the date of HIV seroconversion is not known in most HIV infected individuals. As with all biomarkers, observed CD4 cell count is an imperfect measure due to inherent and non-negligible variability between subjects and within subjects over time. Extending our method to combinations of biomarkers (for example,
CD4 cell count and plasma HIV-1 RNA) is straightforward. We assumed throughout that the endpoint of interest is clinical AIDS, but viral failure, death or some combined study endpoint could easily be substituted.

The methods we propose require external data to provide estimates of the functions of (i) the time to crossing the CD4 threshold while remaining AIDS-free, (ii) the time to AIDS while remaining above the CD4 threshold, and (iii) the mixture proportion at which these two functions meet. We capitalized on the existence of a period of natural history of HIV infection, July 1989 to December 1995, where therapies did not appreciably alter the course of disease [2]. We used this period because it corresponds to the years immediately prior to the introduction of HAART, making data collection methods more similar to those during the HAART era. A major assumption of our analysis is that leadtimes are missing at random within levels of CD4 cell count measured at HAART initiation. In other words, we assume that the deferred group is exchangeable with the external data used to estimate (i)–(iii) stated above. This assumption will not hold if those who are deferred are a selected group at the time of deferral.

Our assumption that leadtime is missing at random given CD4 count categories has two implications. First, individuals in the immediate group represent (that is, are a random sample of) HIV infected persons with CD4 counts between 201 and 350 cells/mm$^3$ for $k = 200$, and between 351 and 500 cell/mm$^3$ for $k = 350$. Indeed, within the CD4 count categories, the distribution of CD4 counts was similar for the HAART initiators and men during the natural history period. Specifically, for $k = 200$, the 278 HAART initiators with CD4 between 201 and 350 cells/mm$^3$ had a median CD4 of 275, while the 452 men from the natural history period with CD4 in that same range had a median CD4 of 287. For $k = 350$, the 224 HAART initiators with CD4 between 351 and 500 cells/mm$^3$ had a median CD4 of 415, while the 529 men from the natural history period with CD4 in that same range had a median CD4 of 421. A second implication of our assumption is that individuals seen in the deferred group represent HIV infected individuals with CD4 counts between 201 and 350 cells/mm$^3$ who reach a CD4 of 200 while AIDS-free. To explore the sensitivity of our findings, we repeated the analysis limiting the range of the natural history data used to estimate $F$ and $G$ to 230 to 350 and 380 to 500 cells for $k = 200$ and $350$, respectively. This exclusion of the lower 20 per cent of the 150 cell range of CD4 values is consonant with the hypothesis that individuals in the lowest 30 cells of each category are less likely to be deferred. Specifically, for $k = 200$, our sensitivity analysis amounts to assuming that none of the 237 deferred participants (who started HAART below 200 cells) were between 200 and 230 cells at a preceding time when they were similar to the 278 participants who we observed to start between 200 and 350. Our analyses were robust to this restriction (hazard ratios = 1.98 [versus 1.97] and 0.65 [versus 0.70], for $k = 200$ and 350, respectively). Of course, as with all observational analyses, our results will be biased if there exists a strong unmeasured common cause of HAART initiation and clinical AIDS above and beyond CD4 count.

We further assumed that within levels of CD4 cell count, women from the WIHS would have had similar HIV disease progression as their male counterparts in the MACS. The WIHS cohort began follow-up in 1994, thus we do not have completely parallel data for the cohort of women. However, this assumption was checked and supported for the 1 $\frac{1}{7}$ years where we had concurrent data on men and women (data not shown) [25]. Furthermore, this gender equivalence is consonant with the striking similarity of HIV disease progression between men and women [10, 26].
We have shown methods to implement a semi-parametric proportional hazards regression. The use of a parametric approach is not only straightforward but also provides a measure of relative times (for example, the factor by which AIDS-free times are extended if therapies are started immediately). The contributions to the log-likelihood function by the individuals in the immediate group and in the deferred group after adding the leadtime correspond to the standard form for right-censored data. The contribution by the unseen fast progressors to the log-likelihood function is given by \( \sum_{j=1}^{n \times n[1-n]} \log f [\hat{G}^{-1}(U_j)] \), where \( f \) corresponds to the density function of the deferred group. In the context of the log-normal distribution and using analogous imputation methods, the relative times to AIDS were 0.51 (95 per cent CI 0.27, 0.96) and 1.47 (95 per cent CI 0.73, 2.96) for \( k = 200 \) and 350, respectively, after accounting for leadtime. In other words, deferral to CD4 ≤ 200 is consonant with a halving of further AIDS-free time. However, waiting until CD4 reaches 350 cells/mm\(^3\) to start HAART alludes to an extension of AIDS-free time. These parametric results are consistent with inferences from the semi-parametric models.

This proposed method is particularly useful in the absence of randomized evidence. In the presence of randomized evidence, this method may serve as the basis to compare inference from observational and randomized studies. This procedure may be of use in other contexts where longitudinal data on a clinical event of interest and therapy, which is largely determined by a biomarker, are measured along with a relevant natural history period from which imputations may be derived. Grant et al. [27] recently described an alternative method to account for leadtime in cohort studies, which provided similar inferences but is dependent upon the measurement of individual biomarker trajectories.

It will be difficult for a randomized deferment trial to definitively answer the question of when to initiation HAART, in part because the components of HIV therapies evolve rapidly. While randomization (asymptotically) ensures that study participants will have similar characteristics at entry, it is likely that therapies received by the deferred group would be more efficacious and less toxic than those received by the immediate group. Thus, in such a deferment trial it may be impossible to distinguish whether a failure to reject the null hypothesis of equal AIDS-free survival in the two arms is due to the appropriateness of deferral or due to more efficacious therapies received by the deferred group.

There is no substitute for well-conducted randomized trials, however, observational cohort studies collecting comprehensive data on therapies and outcomes at regularly scheduled visits provide valuable information to guide clinical practice. Here we have used a novel compilation of statistical methods to provide evidence that AIDS-free survival is sensitive to the timing of HAART initiation. Methods presented here offer an approach for analysing prevalent cohort studies and provide procedures to maximize the usefulness of data from cohort studies.

ACKNOWLEDGEMENTS

The authors wish to thank the reviewers for thoughtful and constructive suggestions. The Multicenter AIDS Cohort Study (MACS), with centres (Principal Investigators) at The Johns Hopkins Bloomberg School of Public Health (Joseph B. Margolick, Alvaro Muñoz), Howard Brown Health Center and Northwestern University Medical School (John Phair), University of California, Los Angeles (Roger Detels, Beth Jamieson), and University of Pittsburgh (Charles Rinaldo), and the Women's Interagency HIV Study (WIHS) Collaborative Study Group with centres (Principal Investigators) at New York City/Bronx
REFERENCES


