



Published in final edited form as:

Epidemiology. 2010 September ; 21(5): 698–705. doi:10.1097/EDE.0b013e3181e97737.

Estimating the Optimal CD4 Count for HIV-infected Persons to Start Antiretroviral Therapy

Bryan E. Shepherd, Cathy A. Jenkins, Peter F. Rebeiro, Samuel E. Stinnette, Sally S. Bebawy, Catherine C. McGowan, Todd Hulan, and Timothy R. Sterling

Vanderbilt University Medical Center, Nashville, Tennessee, USA

Abstract

Background—Optimal timing of antiretroviral therapy in HIV-infected persons is unclear, although two recent large observational studies have improved our understanding of the best CD4 threshold for initiation. These studies compared the effect of starting HAART on mortality and mortality/AIDS between strata defined using broad ranges of CD4 counts. We sought to expand this understanding using a novel statistical approach proposed by Robins and colleagues.

Methods—Using observational data from 1034 antiretroviral-naïve HIV-infected patients from Nashville, Tennessee, we directly estimated the optimal CD4 count for initiation of HAART to maximize patient health 6, 12, 24, and 36 months after the first instance of CD4 falling below 750. We measured health using two outcome metrics, one based on CD4 counts at the end of follow-up and the other based on a published quality-of-life scale; both metrics incorporated death, AIDS-defining events, serious non-AIDS events, and CD4 at the end of follow-up if asymptomatic.

Results—The CD4-based metric estimated that to maximize health 6, 12, 24, and 36 months after study entry, HAART should be initiated within 3 months of CD4 first dropping below 495 (95% confidence interval [CI] = 468 – 522), 554 (459 – 750), 489 (427 – 750), and 509 (460 – 750), respectively. The quality-of-life-based metric produced CD4 initiation threshold estimates of 337 (95% CI = 201–442), 354 (288 – 386), 358 (294 – 750), and 475 (287 – 750) for the same time points.

Conclusions—Our results support early initiation of antiretroviral therapy, although the criterion for starting therapy depends on the choice of health outcome.

Optimal timing of highly active antiretroviral therapy (HAART) initiation in HIV-infected persons is unclear.¹ Current guidelines recommend starting when CD4+ lymphocyte count (CD4) falls below 350 cells/mm³, although these guidelines are based on non-randomized studies potentially subject to lead-time and selection biases.^{2–4}

Two recent studies from the Antiretroviral Therapy Cohort Collaboration⁵ and the North American AIDS Cohort Collaboration on Research and Design⁶ addressed the optimal time to initiate HAART. Both studies used large, multi-site cohorts of observational data, employing statistical methods to avoid lead-time bias and minimize confounding. The former study⁵ employed historical data from the pre-HAART era to account for deaths and AIDS-defining events before HAART initiation. These researchers concluded that HAART should be initiated some time before CD4 drops below 350. The latter study⁶ employed a method proposed by Robins and colleagues^{7,8} to analyze the observational data as if they

Address for Correspondence: Bryan E. Shepherd, Department of Biostatistics, Vanderbilt University Medical Center, 1161 21st Avenue South S2323 MCN, Nashville, TN 37232-2158, phone: 615-343-3496, fax: 615-343-4924, bryan.shepherd@vanderbilt.edu.

SDC Supplemental digital content is available through direct URL citations in the HTML and PDF versions of this article (www.epidem.com).

were from a randomized trial, and comparing treatment strategies for starting HAART when $CD4 > 500$ versus ≤ 500 and starting HAART when $CD4 = 351 - 500$ versus ≤ 350 . This study concluded that HAART should be initiated at $CD4 > 500$.

These studies improve our understanding of the optimal timing of HAART initiation, but they have limitations. First, both studies were non-randomized and used retrospective data. The Antiretroviral Therapy Cohort Collaboration study⁵ assumed that rates of death due to AIDS among untreated patients in the pre-HAART era were similar to those in the HAART era within particular CD4 strata—a strong assumption given temporal changes in AIDS care and outcomes.^{9,10} The other study⁶ compared broad CD4 ranges and was based on a rule defined as starting HAART within 6 months of a particular CD4 level; those who died within the six months before HAART initiation were assigned to the delayed treatment group, which could be a potential source of bias.¹¹ Neither study directly considered serious non-AIDS defining events¹² or incorporated the health of asymptomatic persons without events at the end of follow-up.

In the present study we estimated the optimal CD4 threshold for HAART initiation for a cohort of HIV-infected persons in Nashville, Tennessee, USA. We employed a method proposed by Robins et al,⁸ which is similar to the approach applied by the North American AIDS Cohort Collaboration on Research and Design⁶ in that it analyzes the observational data as if they were from a randomized trial. However, rather than comparing the impact of starting HAART in one stratum (defined using a broad range of CD4 counts) versus another CD4 stratum, our approach directly estimates the optimal CD4 at which to start therapy in order to maximize health at a specified time in the future. For example, in the earlier analysis⁶ the researchers treated their data as if they were from a two-armed clinical trial comparing the rules “start HAART when $CD4 > 500$ ” versus “start HAART when $CD4 \leq 500$.” The method we employed analyzes the data as if they came from a multi-armed clinical trial, where a study subject is assigned to one of 550 possible treatment rules corresponding to “starting HAART within 3 months of the first CD4 measurement below 201, 202, ..., 750.” Rather than computing a single hazard ratio, this approach computes expected health for each treatment rule and then estimates which treatment rule yields the best expected health. We studied multiple measures of health based on mortality, AIDS-defining events, non-AIDS-defining events, CD4, and quality of life.

Methods

Study Population and Definitions

We conducted a retrospective observational cohort study among persons treated at the Comprehensive Care Center, an outpatient clinic in Nashville, TN. The study population included all patients who established care and had at least 2 provider visits from 1998 to 2007. Eligible individuals were those with no prior antiretroviral use at the Center who had at least one CD4 measurement in the range 200–749 before experiencing an AIDS-defining event or initiating HAART. Study entry was the date of the first CD4 level between 200–749. Follow-up ended at death, 31 December 2007, or the last clinic visit for persons lost to follow-up. During the study period, healthcare coverage was available to virtually all HIV-infected Tennesseans.¹³

AIDS-defining events were based on U.S. Centers for Disease Control and Prevention classification criteria, excluding $CD4 < 200$ cells/mm³.¹⁴ Non-AIDS events were based on recommendations of an endpoints committee comprised of infectious disease clinicians. Examples include acute myocardial infarction and cirrhosis of liver; a complete list of non-AIDS events considered for this study is found in the eAppendix (<http://links.lww.com>). Subjects were considered lost to follow-up if they had no clinic encounter for more than a

year before date of death or 31 December 2007, whichever came first. The precision of all measurements involving time is in units of days. HAART was defined as any regimen containing 3 or more active antiretroviral therapy agents.³ The study was approved by the Vanderbilt University Medical Center Institutional Review Board.

Statistical Methods

Given a set of candidate rules, “start HAART within 3 months of first CD4 measured below x ,” where $x = 201, 202, \dots, 750$, we sought to estimate which value of x would result in the best expected patient health k months after study entry. To address this question, we followed the methodology of Robins et al.⁸ This methodology requires specifying a measurement of patient health at k months (the outcome variable), and the treatment rules, x , that are compatible with each patient’s CD4 and HAART initiation history (the explanatory variable). We considered $k = 6, 12, 24$, and 36 months and employed inverse probability weights to account for potential bias due to non-random assignment of treatment rules and patient drop-out.

Outcome Variables

Our analyses required specifying a metric of each patient’s health at time k months. This outcome, y , is termed a “utility” or “health metric” and in our analysis was a function of death, AIDS-defining events, non-AIDS-defining events, and CD4 if asymptomatic. The utility was an arbitrary but reasoned quantity; in our analyses we used the following utilities:

1. CD4 count-based: $y =$ most recent CD4 at or before month k if patient alive and asymptomatic at month k , $y = 100$ if subject had an AIDS or non-AIDS event by month k but was not deceased, $y = 100 \times (t - 12 - k) / 18$ if subject was deceased before month k where $t =$ months to death (see Figure 1A).
2. Quality of life: Based on a validated quality-of-life scale incorporating death, type of AIDS or non-AIDS event, and CD4 at month k (see Figure 1B and eAppendix [<http://links.lww.com>] for formula).¹⁵ In this utility death was assigned $y = 0$, AIDS or non-AIDS events were assigned y between 0.56 and 0.65 depending on the specific type of event, and asymptomatic patients were assigned y between 0.78 and 0.95 depending on their most recent CD4 at or before month k .

For both utilities, a low value of y corresponded to a poor outcome. Utility 1 was elicited a priori from consultations with the Vanderbilt-Meharry Center for AIDS Research Epidemiology/Outcomes group. This utility was based on CD4 count and assigned a patient with an AIDS or non-AIDS event by month k the same utility as an asymptomatic patient with CD4=100; individuals who died were assigned negative utility scores, with those who died earliest given larger negative values. Utility 2 also incorporated death, AIDS and non-AIDS events, and CD4, but was based on a published quality-of-life scale¹⁵ and used type of first AIDS event. In both utilities, if a patient had either type of event and subsequently died before time k , then the death was recorded. If both AIDS and non-AIDS events occurred, the earlier took precedence.

Regimen Rules

Following Robins et al,⁸ consider a multi-armed trial where each patient is randomly assigned a value of x between 201 and 750, and then asked to follow the rule “start HAART within 3 months of first CD4 measurement below x .” In such a trial, suppose a patient was assigned the rule with $x = 400$. If this patient’s first CD4 measurement below (but not equal to) 400 was 350, and if he started HAART within 3 months of this measurement, then this patient was adherent to his assigned rule. Notice that although this patient was randomized to the rule “start HAART within 3 months of the first CD4 measured below 400,” his CD4/

treatment history was also consistent with the rules “start HAART within 3 months of the first CD4 measured below 399, 398, ..., 351.” In contrast, if this patient had not started HAART within 3 months of his CD4 measurement of 350 or had started HAART before his CD4 was measured below 400, he would have been non-adherent to his assigned rule.

Using this model, we examined each patient’s CD4 and HAART initiation history and determined compatible rules for each patient. It should be noted that for the purpose of computing rules, follow-up stopped at the earliest of date of HAART initiation, first AIDS event, last visit, death, or k months.

Table 1 contains some hypothetical examples matching treatment histories to regimen rules. Consider Patient A: his first CD4 measurement was 400, his next was 350 at month 3, and he then started HAART in month 4. This patient’s data were compatible with the rules “start HAART within 3 months of first CD4 measured below $x = 351, \dots, 400$.” Had this patient been assigned the rule with $x = 400$, he would have been compliant because the first CD4 measured below (but not equal to) 400 was 350, and he started HAART one month after this observation. However, this patient’s data were not compatible with the rule “start HAART within 3 months of first CD4 measured below $x = 401$,” because his first CD4 measurement below 401 (CD4 = 400 at month 0) was taken more than 3 months before he started HAART. Similarly, this patient’s data were not compatible with the rule “start HAART within 3 months of first CD4 measured below $x = 350$ ” because he started HAART without ever having a CD4 measured below 350.

Patients whose data were not compatible with any regimen rule were artificially censored at the date their data became incompatible. For example, Patient B did not start HAART within 3 months of CD4=250, but then started HAART in month 4, one month after CD4=300. Therefore, his data were not consistent with any x , and he was artificially censored when he started HAART. Patient C’s data were also inconsistent with all rules as he started HAART more than 3 months after his last CD4 measurement. Patient D failed to start HAART within 3 months of his first CD4 below 201, so his data were not consistent with any rule and he was artificially censored 3 months after his first CD4 below 201. In contrast, Patient E was consistent with $x=201, \dots, 350$.

It should be noted that regimen rules were based on measured rather than actual CD4. For example, Patient F started HAART within 3 months of his first CD4 measured below 750 although it is likely that his CD4 dropped below 750 cells more than 3 months before initiating HAART, but was not observed. Patient G’s history is similar to Patient F’s with the additional CD4 measurement of 250 taken at month 1. Although this measurement may have prompted the initiation of HAART, Patient G also started HAART within 3 months of his first CD4 measured below 750. Similarly, patient H was assigned all rules $x = 201, \dots, 750$. Patient I was also assigned all treatment rules as the study ended less than 3 months before his first CD4 measurement and it is therefore unclear whether he was deferring treatment until a lower CD4 or preparing to start. Finally, Patients J, K, and L never started HAART, but were consistent with the rules “start HAART within 3 months of the first CD4 measured below $x = 201, \dots, 250$ ”—Patient J because he never had a CD4 measurement below 250 and Patients K and L because their next measurements were less than 3 months before $k = 6$. The complete algorithm used for determining compatible rules is given in the eAppendix (<http://links.lww.com>).

Weighted Regression Models

To find the rule for starting HAART that maximized health at k months, we regressed y on x and found the value of x that achieved the maximum predicted value of y . Each individual contributed as many values of (x,y) as the number of rules compatible with their data. For

example, a person who had data compatible with starting HAART within 3 months of their first CD4 below $x = 201, \dots, 250$ contributed 50 data points (x, y) to the analysis: $(201, y), (202, y), \dots, (250, y)$; their outcome y was the same for all x . We fit a curve to the (x, y) pairs of all eligible patients using weighted least squares regression, with x expanded using restricted cubic splines to allow the relationship between x and y to be non-linear. Our splines used 6 knots, at default positions of the Design package¹⁶ of R statistical software version 2.8.1 (<http://www.r-project.org>).

Inverse Probability Weights

Persons whose data were compatible with a given rule may have had characteristics different from those whose data were compatible with other rules. To address this potential source of bias, we employed inverse probability weighting methods in the manner described by Cain et al.¹⁷ Briefly, for months $0, 1, \dots, k$, we estimated the probability of initiating HAART using logistic regression and the covariates age, sex, race, injection drug use as HIV risk factor, most recent CD4, CD4%, and HIV-1 RNA, time since most recent laboratory measurements, and time in care. Months were included in the model using restricted cubic splines. For each compatible treatment rule x per person, we then computed the predicted probability of remaining compatible with x for each month of follow-up. Based on CD4 history, if a patient could not be artificially censored from a particular rule x at a given month, then the probability of remaining compatible with x for that patient in that month was 1, otherwise it was one minus the probability of initiating HAART. Weights were computed as the product of the inverse of these probabilities over the k months of follow-up.

Some patients' health at time k months was unknown, due either to loss to follow-up or end-of-study censoring. Separate stabilized inverse-probability weights to address loss to follow-up and end-of-study censoring were computed. Our final weights were the product of the multiple inverse-probability weights. In order to reduce variability, the product of the weights was truncated at the 2.5th and 97.5th percentiles.¹⁸

Confidence intervals (CIs) for the optimal rule were constructed from the 2.5th and 97.5th percentiles of 500 bootstrap replications. All model-fitting procedures were repeated for each bootstrap replication. Details of all models are in the eAppendix (<http://links.wv.com>), and analysis scripts are posted at <http://biostat.mc.vanderbilt.edu/WhenToStartHaartCode>.

Results

Of 2011 patients with at least two provider visits, 1034 met inclusion criteria. 430 were excluded because their first pre-HAART CD4 < 750 was under 200, 42 had no pre-HAART CD4 < 750, 232 had a prior AIDS-defining event, 240 had been on prior non-HAART antiretroviral therapy, and 33 had both a prior AIDS-defining event and prior antiretroviral therapy.

Of the 1034 included patients, 73% were male, 42% African-American, and 8% had injection drug use as probable infection route. At study entry, the median age was 35 years (interquartile range (IQR) = 28 – 42), and the median CD4 was 403 (301 – 528). The median follow-up was 35 months (14 – 65), and the median number of visits per year was 6.4 (4.5 – 9.5). Sixty percent started HAART during follow-up; among those initiating HAART, the median time to initiation was 4.1 months (1.7 – 17.3). The median CD4 prior to HAART initiation was 342 (264 – 462). Male sex, high CD4, and low HIV-1 RNA were associated with lower odds of starting HAART.

During follow-up, 93 patients died (9%), 82 experienced at least one AIDS event (8%) (25 of these patients later died), and 20 had a non-AIDS event (2%) (7 of these patients later

died). Table 2 contains the number of patients who had an event within 6, 12, 24, and 36 months of study entry. Table 2 also includes the number of persons without an outcome at 6, 12, 24, and 36 months due to loss to follow-up, end-of-study censoring, or artificial censoring because their data were incompatible with all treatment rules. Male sex, younger age, and lower CD4% were generally associated with more loss to follow-up.

Optimal CD4 to initiate HAART

Figure 2 demonstrates estimation of the optimal CD4 to initiate HAART in order to maximize health 12 months after study entry, based on utility 1 (Figures 2A–D) and utility 2 (Figures 2A–B, E–F). We estimated that health 12 months after study entry was maximized by following the rules “start HAART within 3 months of first CD4 measurement below 554” (95% CI = 459–750) and 354 (288 – 386) for utilities 1 and 2, respectively. Notice that the confidence intervals correspond to rules in Figures 2D and 2F where the best-fitting curves were similar to their maximum levels. Note also that most patients were asymptomatic after 12 months. Therefore, most utility 2 scores were between 0.9 and 0.95 (Figure 2E), and a relatively small number of deaths had a large influence on estimates.

Similar analyses were performed for both utilities at $k = 6, 12, 24,$ and 36 months. Estimates and 95% confidence intervals for the optimal CD4 level to start HAART for both utilities and at all time points are given in Figure 3. Results were dependent on the choice of utility and the follow-up period.

To maximize health as defined by utility 1 at k months, the optimal rule was to “start HAART within 3 months of first CD4 measurement below” 495 (95% CI=468–522), 554 (459 – 750), 489 (427 – 750), and 509 (460 – 750) for $k = 6, 12, 24,$ and 36 months, respectively. The confidence intervals widened for increasing k because fewer people were followed for the longer periods of time. In contrast, to maximize utility 2 (quality-of-life) at k months, the optimal rule for starting HAART was to “start within 3 months of first CD4 measurement below” 337 (201 – 442), 354 (288 – 386), 358 (294 – 750), and 475 (287 – 750) for $k = 6, 12, 24,$ and 36 months, respectively.

Secondary analyses considering alternative utilities and regimen rules are reported in the eAppendix (<http://links.lww.com>). Briefly, analyses that did not include non-AIDS events in the utility were similar to those presented above (eg, optimal rule estimated as “start HAART within 3 months of CD4 measured below 563” instead of 554 for CD4-based utility at $k = 12$ months). Analyses that assigned worse health metrics to individuals who had changed regimens favored starting HAART at somewhat lower CD4 counts (eg, 414 for CD4-based utility at $k = 12$ months). Analyses that included only candidate rules in the range $x = 201$ –500 were generally a little lower than primary estimates (eg, 449 for CD4-based utility at $k = 12$ months). Analyses comparing the candidate rules $x = 201$ –500 but restricted to those with at least one pre-HAART $CD4 \geq 500$ generally favored starting HAART at slightly lower CD4 levels (eg, 404 for CD4 -based utility at $k = 12$ months). We also performed secondary analyses to estimate the optimal rule for starting a modern, efavirenz-based regimen, artificially censoring subjects who started other regimens. Our estimated CD4 thresholds for starting efavirenz-based HAART were slightly lower (eg, 509 for CD4-based utility at $k = 12$ months). The results of other secondary analyses using utilities based on survival, ADE-free survival, and AIDS/non-AIDS-events-free survival were quite variable, presumably due to small numbers of events.

Discussion

We have applied a novel approach to directly estimate the optimal CD4 level for initiating HAART. The approach mimicked a series of randomized trials and defined patient health by

more than just death and ADE. Our estimates were sensitive to the choice of patient health metric. Our health metric, which substantially differentiated between asymptomatic patients with widely different CD4 at the end of follow-up, favored starting HAART early, at CD4 levels around 500. In contrast, the quality-of-life health metric, which distinguished very little between asymptomatic patients with low and high CD4 at the end of follow-up, tended to favor initiating HAART later, at lower CD4 levels.

There are several advantages to this analytic approach. Our analyses account for lead-time bias without incorporating historical controls. Prior studies have classified patients into categories based on CD4, and then estimated the optimal CD4 at which to start HAART, using hazard ratios comparing the different categories. In contrast, we have directly estimated the optimal CD4 and computed confidence intervals for this estimate. This approach therefore does not require arbitrary categorization.¹⁹ Our treatment rules were based on starting HAART within 3 months of CD4 measured below a particular level, in contrast to the 6 months used recently by the North American AIDS Cohort Collaboration on Research and Design.⁶ Three months was chosen because it is the typical length of time between visits at our clinic. It is worth noting that with this analytic approach, pre-HAART deaths within 3 months after study entry do not bias results¹¹ (see eAppendix, <http://links.lww.com>). Finally, we also directly incorporated non-AIDS events into our analysis, which may be associated with HAART.^{20,21}

An apparent disadvantage with our analysis was that it required defining a health metric at a specified time of follow-up, and the choice of this metric greatly affected our conclusions. Estimates based on utility 1 were closer to results from the North American AIDS Cohort Collaboration on Research and Design study:⁶ start HAART at high CD4 counts, possibly above 500. In contrast, estimates based on utility 2 using the same outcomes (death, AIDS events, non-AIDS events, and CD4 if asymptomatic) favored starting HAART at lower CD4 levels, which is more consistent with results from the Antiretroviral Therapy Cohort Collaboration study⁵ and current guidelines. We believe both utilities are reasonable measures for defining health, and thus we present both sets of results. Utility 2 was based on a previously published quality-of-life score, where there was substantial separation in the metric between those who died and those who were living but little separation between those who were alive with various CD4 levels. Hence, the relatively few deaths in our study had a large impact on analyses with this utility (see Figure 2). In contrast, utility 1 was elicited from our clinicians/study investigators a priori and put more emphasis on differences between CD4 outcomes in asymptomatic patients. Perhaps if we had data from more patients, particularly those who subsequently died, estimates for the optimal CD4 to start HAART using the two different utilities would converge. But this may not be the case; indeed, how one defines health plausibly has great impact on when one chooses to initiate HAART. We cannot give results under all possible utilities. The arbitrary nature of utility choice may therefore lead one to favor analyses that consider only the hard endpoint of death—which implicitly assigns equal health scores to all living patients. However, keeping patients alive is not the only goal of modern HIV-therapy; it is important also to consider the impact of the timing of treatment initiation on other outcomes.

Our study included all patients who had at least one CD4 within the range 200–750. A randomized controlled trial may include only those who had a CD4 measurement above a specific threshold (eg, 750) and then subsequently dropped below that threshold. In a secondary analysis we limited our study to persons who had pre-HAART CD4 \geq 500, and we defined study entry as the date of their first CD4 $<$ 500. Most patients do not present to care with CD4 \geq 500; therefore the number of patients was limited for this analysis, and results may not be as generalizable. We believe the question of when to start HAART should not be

limited to the small subset of individuals who enter care with $CD4 \geq 500$ (discussed in eAppendix, <http://links.lww.com>).

Our study has other limitations. First, the number of patients was relatively small (particularly the number of deaths), and follow-up time was limited. Analyzing data from more patients followed for a longer time period would improve the precision of our estimates. In addition, the impact of HAART on non-AIDS events may be greater after many years of use. Second, we included patient data only from the southeastern United States, and thus conclusions may not be applicable to other populations. Third, although we controlled for many clinically-important factors, as with all observational studies there may have been residual confounding or model misspecification, thus biasing results. Fourth, our goal was to estimate the rule that maximized health; estimates of maxima are generally quite variable and tend to favor the lower and upper limits of the allowable rules (201 and 750) (see the eAppendix, <http://links.lww.com>). In addition, our candidate rules for starting HAART were defined using only CD4, whereas other factors (eg, injection drug use) should likely be included in such a decision. Because our rules for initiating HAART are based on measured CD4, not actual CD4 counts, the frequency of CD4 measurements might have affected results. Finally, our analyses considered only the health of infected patients and ignored the impact of the timing of HAART on HIV transmission.

In conclusion, we have applied a novel method to estimate the optimal CD4 for initiating HAART. Similar analyses should be performed in larger observational datasets. Ongoing and future randomized trials should consider the effect of the timing of HAART initiation on various health-metrics in addition to AIDS and death.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Sources of Financial Support:

Vanderbilt-Meharry Center for AIDS Research (NIH program 930 AI54999) and National Institutes of Health (grant R01 DA023879-03 to B.E.S., grant K23 AT002508-01 to T.H., grant K24 A1065298 to T.R.S.)

References

1. Hammer SM, Eron JJ Jr, Reiss P, et al. Antiretroviral treatment of adult HIV infection: 2008 recommendations of the International AIDS Society-USA panel. *JAMA*. 2008; 300(5):555–570. [PubMed: 18677028]
2. United States Department of Health and Human Services. AIDSInfo. US Dept. HHS Website; November 3. 2008 Guidelines for the Use of Antiretroviral Agents in HIV-1-Infected Adults and Adolescents. Available at: <http://www.aidsinfo.nih.gov/guidelines/>
3. Egger M, May M, Chene G, et al. Prognosis of HIV-1-infected patients starting highly active antiretroviral therapy: a collaborative analysis of prospective studies. *Lancet*. 2002; 360(9327):119–129. [PubMed: 12126821]
4. Cole SR, Li R, Anastos K, et al. Accounting for leadtime in cohort studies: evaluating when to initiate HIV therapies. *Stat Med*. 2004; 23(21):3351–3363. [PubMed: 15493031]
5. Sterne JAC, May M, Costagliola D, et al. Timing of initiation of antiretroviral therapy in AIDS-free HIV-1-infected patients: a collaborative analysis of 18 HIV cohort studies. *Lancet*. 2009; 373:1352–1363. [PubMed: 19361855]
6. Kitahata MM, Gange SJ, Abraham A, et al. Effect of early versus deferred antiretroviral therapy for HIV on survival. *N Engl J Med*. 2009; 360:1815–1826. [PubMed: 19339714]

7. Hernán MA, Lanoy E, Costagliola D, Robins JM. Comparison of dynamic treatment regimes via inverse probability weighting. *Basic Clin Pharmacol Toxicol*. 2006; 98(3):237–242. [PubMed: 16611197]
8. Robins J, Orellana L, Rotnitzky A. Estimation and extrapolation of optimal treatment and testing strategies. *Stat Med*. 2008; 27(23):4678–4721. [PubMed: 18646286]
9. Palella FJ Jr, Delaney KM, Moorman AC, et al. HIV Outpatient Study Investigators. Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. *N Engl J Med*. 1998; 338(13):853–860. [PubMed: 9516219]
10. Mocroft A, Ledergerber B, Katlama C, et al. EuroSIDA study group. Decline in the AIDS and death rates in the EuroSIDA study: an observational study. *Lancet*. 2003; 362(9377):22–29. [PubMed: 12853195]
11. Hernan MA, Robins JM. Early versus deferred antiretroviral therapy for HIV [letter]. *N Eng J Med*. 2009; 361(8):822–823.
12. The Strategies for Management of Antiretroviral Therapy (SMART) Study Group. CD4+ count-guided interruption of antiretroviral treatment. *N Engl J Med*. 2006; 355(22):2283–2296. [PubMed: 17135583]
13. Bailey JE, Van Brunt DL, Raffanti SP, Long WJ, Jenkins PH. Improvements in access to care for HIV and AIDS in a statewide Medicaid managed care system. *Am J Manag Care*. 2003; 9(9):595–602. [PubMed: 14527105]
14. Centers for Disease Control and Prevention (CDC). 1993 revised classification system for HIV infection and expanded surveillance case definition for AIDS among adolescents and adults. *MMWR Recomm Rep*. 1992; 41(RR-17):1–19.
15. Freedberg KA, Scharfstein JA, Seage GR III, et al. The cost-effectiveness of preventing AIDS-related opportunistic infections. *JAMA*. 1998; 279(2):130–136. [PubMed: 9440663]
16. Harrell FE, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med*. 1996; 15(4): 361–387. [PubMed: 8668867]
17. Cain LE, Robins JM, Lanoy E, et al. When to start treatment? a systematic approach to the comparison of dynamic regimes using observational data. *International Journal of Biostatistics*. 2010; 6(2):18.
18. Cole SR, Hernan MA. Constructing inverse probability weights for marginal structural models. *Am J Epidemiol*. 2008; 168(6):656–664. [PubMed: 18682488]
19. Altman DG, Royston P. The cost of dichotomising continuous variables. *BMJ*. 2006; 332(7549): 1080. [PubMed: 16675816]
20. Palella FJ Jr, Baker RK, Moorman AC, et al. HIV Outpatient Study Investigators. Mortality in the highly active antiretroviral therapy era: changing causes of death and disease in the HIV outpatient study. *J Acquir Immune Defic Syndr*. 2006; 43(1):27–34. [PubMed: 16878047]
21. Lau B, Gange SJ, Moore RD. Risk of non-AIDS-related mortality may exceed risk of AIDS-related mortality among individuals enrolling into care with CD4+ counts greater than 200 cells/mm³. *J Acquir Immune Defic Syndr*. 2007; 44(2):179–187. [PubMed: 17075385]

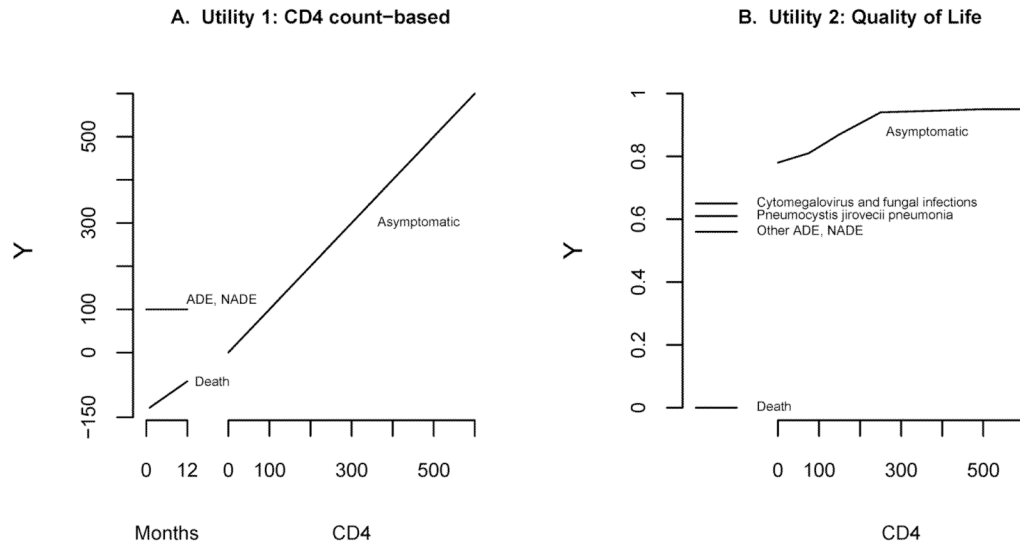
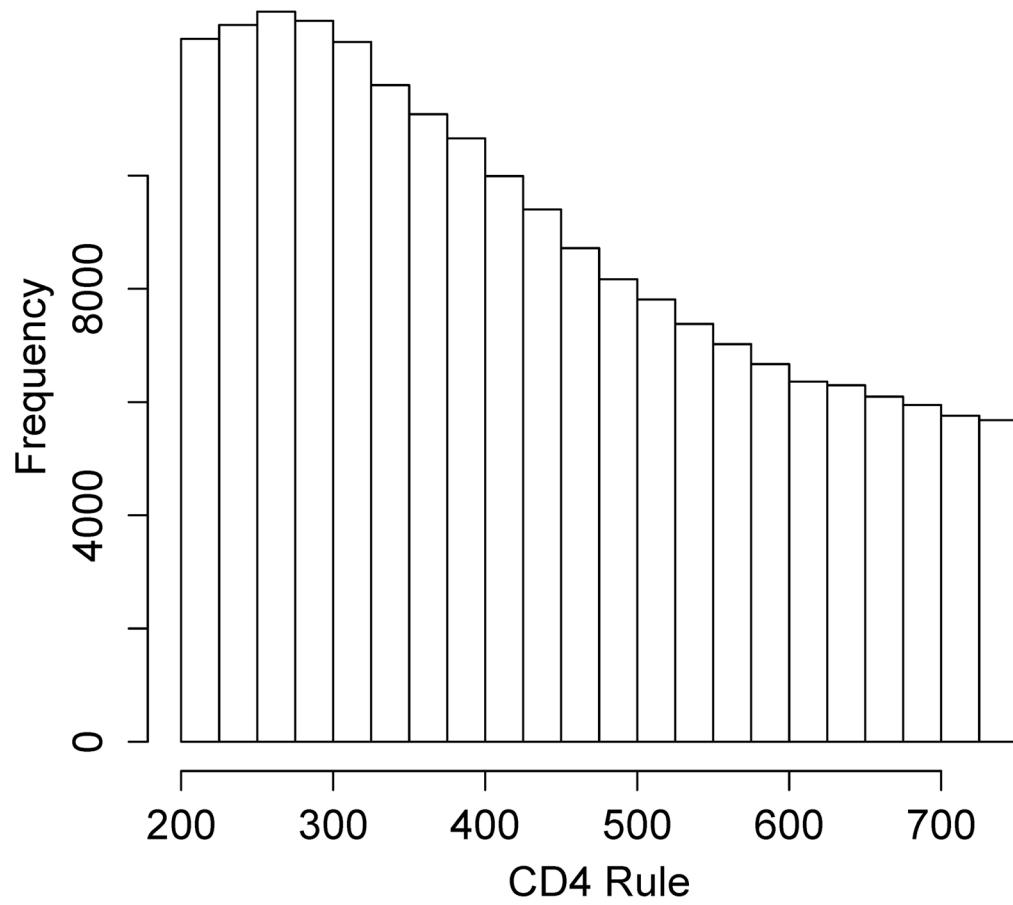
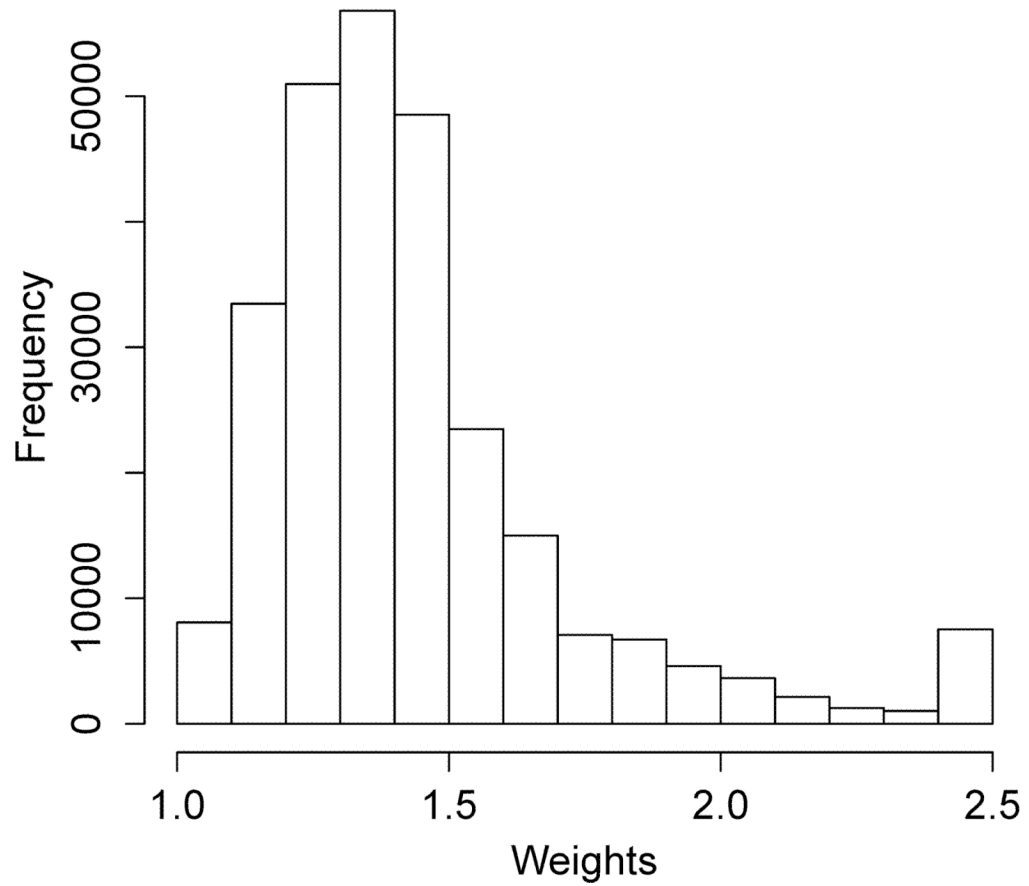


Figure 1. Utility scores (measures of patient health) according to death, AIDS events (ADE), serious non-AIDS events (NADE), and CD4 count if asymptomatic one year after study entry. A. Utility 1 was defined through discussions with the Vanderbilt-Meharry Center for AIDS Research Epidemiology/Outcomes group. B. Utility 2 was based on a published quality-of-life metric.¹⁵

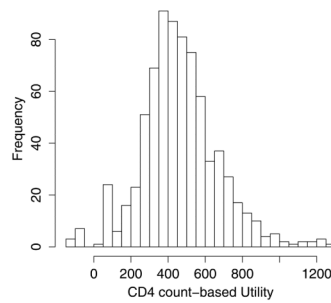
A. Histogram of HAART Initiation Rules



B. Histogram of Inverse Probability Weights



C. Histogram of Utility 1



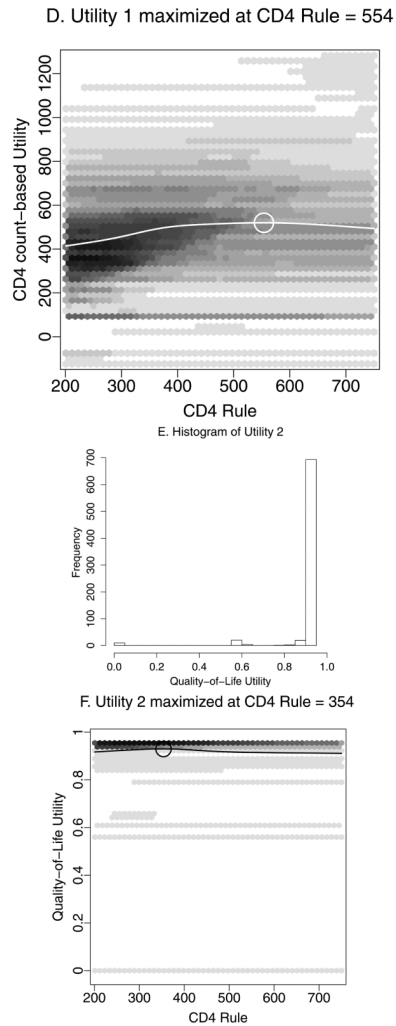


Figure 2.

Estimating the optimal CD4 level to initiate HAART in order to maximize health 12 months after study entry. **A.** The distribution of regimen rules compatible with patients' CD4/HAART histories. The rules are "start HAART within 3 months of first CD4 measurement below x ," where x is between 201 and 750. **B.** The distribution of the truncated inverse probability weights used to account for confounding, lost to follow-up, and artificial censoring. **C.** The distribution of utility 1 (the CD4 count-based measure of patient health) at $k = 12$ months. **D.** A plot of the regimen rules of Figure 2A (x-axis) by the CD4 count-based utility of Figure 2C (y-axis), where the shading represents the weight given a particular region of the graph using the inverse probability weights of Figure 2B. The white line is the best-fitting weighted-least-squares regression curve, and it achieves its maximum at the circled location. The rule which corresponds to this maximum, "start HAART within 3 months of first CD4 measurement below 554," is the estimated optimal rule based on utility 1. **E.** The distribution of utility 2 (the quality-of-life measure of patient health) at $k=12$ months. **F.** A plot of the regimen rules of Figure 2A (x-axis) by the quality-of-life utility of Figure 2E (y-axis), where the shading represents the weight given a particular region of the graph using the inverse probability weights of Figure 2B. The black line is the best fitting weighted least squares regression curve, and it achieves its maximum at the circled location. The rule which corresponds to this maximum, "start HAART within 3 months of first CD4 measurement below 354," is the estimated optimal rule based on utility 2.

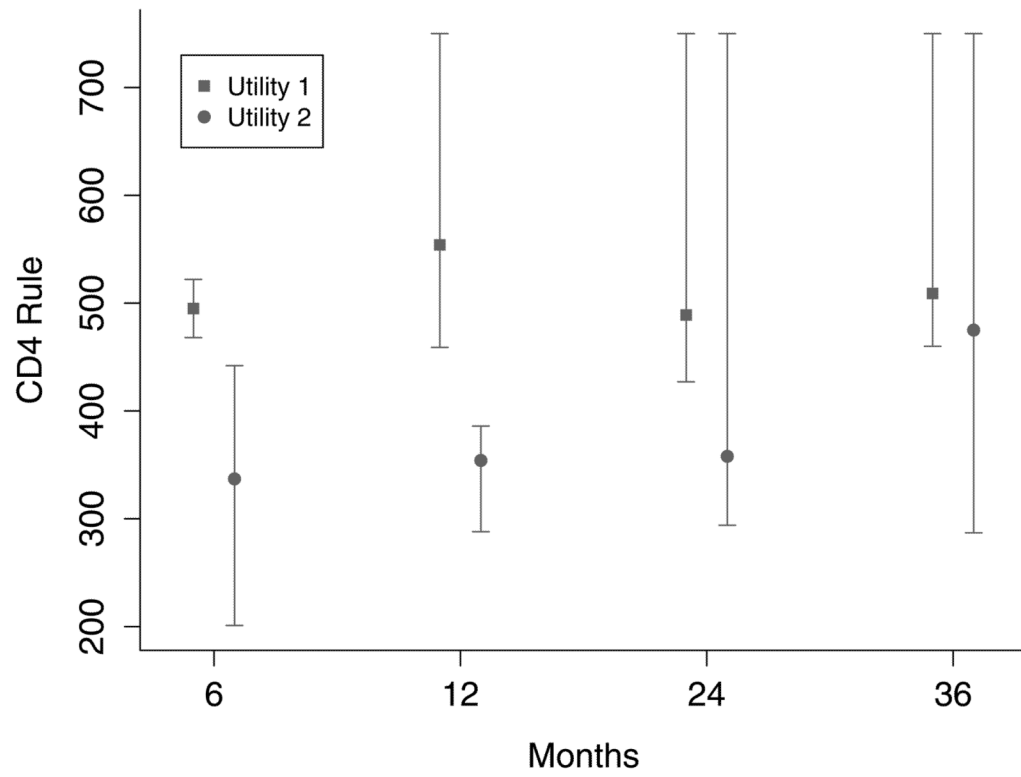


Figure 3. Estimates (squares and circles) and 95% confidence intervals (vertical lines) of the optimal CD4 count at which to initiate HAART in order to maximize both utilities 6, 12, 24, and 36 months after study entry.

Table 1

Assigning regimen rules compatible with patients' data; some examples when $k = 6$ months.

Patient	Month						Compatible Regimens	
	0	1	2	3	4	5		6
A	400			350	HAART			$x = 351, \dots, 400$
B	250			300	HAART			not compatible with any x , censored at month 4
C	400				HAART			not compatible with any x , censored at month 4
D	350	180	180	210		200	HAART	not compatible with any x , censored at month 5
E	350		180	210	HAART			$x = 201, \dots, 350$
F	400		HAART					$x = 401, \dots, 750$
G	400	250	HAART					$x = 251, \dots, 750$
H	400	180	HAART					$x = 201, \dots, 750$
I	400		Study end ^a					$x = 201, \dots, 750$
J	450		250					$x = 201, \dots, 250$
K	450		250			225		$x = 201, \dots, 250$
L	450		250			300		$x = 201, \dots, 250$

^a All other patients in this table were followed for $k = 6$ months.

Table 2

Number of patients who died, had an AIDS-defining event (ADE), a non-AIDS defining event (NADE), or were censored according to follow-up period.

Months of Follow-up (t)	No. Deaths	No. AIDS-defining events	No. Non-AIDS-defining events	No. Asymptomatic	No. Censored	
					Lost to follow up	End of study
6	3	11	8	852	79	40
12	10	14	10	715	115	96
24	18	24	11	519	176	170
36	32	33	14	390	217	211
						Artificially censored
						41
						74
						116
						137