



Published in final edited form as:

J Acquir Immune Defic Syndr. 2012 November 1; 61(3): 326–333. doi:10.1097/QAI.0b013e31826be75e.

Factors affecting timing of antiretroviral treatment initiation based on monitoring CD4 counts

Farzad NOUBARY, Ph.D.^{1,2} and Michael D. HUGHES, Ph.D.¹

¹Center for Biostatistics in AIDS Research, Harvard School of Public Health, Boston, Massachusetts, USA

²Division of General Medicine, Massachusetts General Hospital, Boston, Massachusetts, USA

Abstract

Objective—To evaluate factors affecting antiretroviral therapy (ART) start time when triggered by a CD4 count < 350 cells/ μ l while monitoring counts over time. Measurement frequency, requirement for confirmatory counts, and precision and accuracy of CD4 enumeration technology were considered.

Methods—Using a model of CD4 count trajectories among seroconverters in the Multicenter AIDS Cohort Study, sequences of counts were simulated for a large hypothetical population monitored for 5 years from seroconversion. Time of first count < 350 cells/ μ l was defined as ART start time. The simulation was adapted to evaluate the effect of the above factors on these times. ART initiation was considered “very late” among patients whose underlying trajectory declined below 200 cells/ μ l during the period simulated if no previous observed count was < 350 cells/ μ l.

Results—For 12-, 6-, 4- and 3-monthly measurements, median start time was 48, 36, 32 and 30 months after seroconversion and proportion of patients starting ART very late was 11.5%, 1.6%, 0.2% and 0.1%. For 6-monthly measurements, requiring confirmation increased the median to 49 months and proportion to 8.9%. Changes in standard deviation of short-term variability in counts of 25% and measurement bias for a novel technology of \pm 10% changed median time by \pm 6 months with modest change in the proportion very late (range 0.5% to 3.2%).

Conclusion—6-monthly measurements appear adequate in achieving low rates of very late ART whereas confirmation affects rates adversely. Studies comparing new versus standard measurement technologies should focus on ruling out modest bias, particularly proximal to important thresholds for treatment management.

Keywords

HIV; CD4+ T-lymphocyte; Patient Management; Method Agreement; Flow Cytometry

Introduction

Monitoring CD4+ T-lymphocyte counts (CD4 counts) is important in antiretroviral treatment (ART) guidelines, with ART initiation recommended when a patient’s CD4 count

Corresponding Author: Farzad Noubary, 50 Staniford St, 9th Floor, Boston, MA 20114 USA. fnoubary@partners.org.

Conflict of Interest

The authors have no conflict of interest related to this work.

Human Subjects Research Approval

This work uses de-identified publicly-available data and was deemed exempt from review by the Harvard School of Public Health IRB.

declines below a defined threshold [1; 2]. The choice of threshold varies depending on expert opinion and resource availability: United States (US) guidelines use a threshold of 350 cells/ μl or, with less strength of opinion, 500 cells/ μl [1]; WHO guidelines use 350 cells/ μl [2] though resource limitations mean that ART programs in some countries use 200 cells/ μl following earlier WHO guidance [3]. We evaluate some issues that might impact timing of ART initiation based on monitoring CD4 counts over time focusing on the 350 threshold but also considering the 200 and 500 thresholds.

We first consider two issues related to differences between US and WHO guidelines [1; 2]. One concerns frequency of CD4 count measurement: US guidelines suggest every 3–4 months, whereas WHO guidelines suggest every 6 months and more frequently as the threshold to initiate ART is approached. The other issue concerns confirmation of an initial count below the threshold before starting ART as recommended in WHO but not US guidelines. We then consider issues related to evaluation of cheaper or simpler alternatives to flow cytometry for CD4 enumeration. Though cross-sectional method comparison studies have been conducted to assess agreement between new technologies and flow cytometry [4; 5; 6; 7], we evaluate how differences in precision and accuracy of technologies might impact timing of ART initiation.

Methods

Approach

We use a combination of statistical modeling of CD4 count trajectories using data from a natural history study and computer simulation to generate trajectories for a large hypothetical patient population followed from HIV seroconversion. Using the simulated trajectories, we identify, for each patient, time of first count below the threshold for starting ART. By varying parameters of the simulation study, we evaluate how measurement frequency, confirmatory values, and precision and accuracy in measurement, might affect the distribution of ART starting times in the simulated population. We briefly describe the study population, statistical model and simulation study; further details are in the Supplemental Digital Content and biostatistical literature [8].

Study Population

We used publicly-available data from HIV seroconverters in the US Multicenter AIDS Cohort Study (MACS) [9]. The MACS design, including standardization and quality control of T-lymphocyte measurements using flow cytometry, have been described [10; 11]. Beginning in 1984, the MACS included enrollment of HIV seronegative homosexual and bisexual men with semiannual testing for HIV infection and CD4 counts. We used data from men who seroconverted, using the date of their first seropositive test as time=0 in modeling CD4 count trajectories following seroconversion. We excluded patients whose time of seroconversion was not well established, i.e. their last negative and first positive tests were more than two scheduled visits apart. Because we are interested in modeling CD4 counts in the absence of ART, follow-up through to 1989, when ART was introduced, was included. The resulting study population included 330 seroconverters followed for up to five years.

Statistical Model for CD4 Count Trajectories

A linear mixed effects model was used for CD4 count trajectories, as previously used in the MACS and other cohorts, [12; 13; 14; 8]. A patient's square-root CD4 count, $\sqrt{CD4(t)}$ (transformed to satisfy model assumptions), at time t years after their first seropositive test, was modeled as $\eta_t + \varepsilon_t$ where η_t represents a patient's underlying "latent" square-root CD4 count and ε_t measures the departure of the observed square-root count from that latent square-root count. The latent count can be viewed as the average of patient's counts when

monitored frequently over a short period (e.g. a few days) close to time t . In the model, each patient's latent $\sqrt{CD4}$ (i.e. η_t) follows a long-term linear trend over time (with rate of change which varies among patients) but with shorter periods of faster or slower changes. The term ε_t is a combination of measurement error (e.g. from laboratory environment/equipment) and very short-term biological variation (e.g. diurnal variation), assumed to be normally distributed with zero mean.

Simulation Study

Using the statistical model, we simulated sequences of semi-annual CD4 measurements for a hypothetical population of 50,000 untreated patients followed for 60 months from seroconversion (the maximum follow-up in the study population), identified each patient's ART start time as the time of their first count below a given threshold, and determined the distribution of ART starting times in the simulated population. This simulation provides a "benchmark" based on monitoring (using flow cytometry) a patient population with characteristics and CD4 trajectories similar to the MACS study population.

The simulation was then adapted to evaluate how ART starting times might change for different measurement frequencies and when confirmatory values are required. The same statistical model was used to simulate monthly latent and observed CD4 counts for each patient, selecting the appropriate counts for the desired frequency. If confirmation of a count below the desired threshold was required, the count one month later was used (extending the simulation to 61 months to allow confirmation for counts at 60 months).

We further adapted the simulation by multiplying the standard deviation of ε_t by a factor δ to evaluate how ART starting times might change if a hypothetical CD4 enumeration technology with different precision from flow cytometry were used. We also considered what might happen if there was measurement bias in a novel technology (i.e., systematic difference between novel technology and flow cytometry measurements) equal to -10% or $+10\%$ of a patient's latent CD4 count.

As measures of monitoring and treatment resource utilization for the simulated population, we obtained mean number of CD4 counts before ART initiation and number of months of ART during the 61 months simulated. We also determined the proportion of patients starting ART "late" or "very late" during this period. "Late" initiation was based on a concept that the ideal ART start time might be when a patient's latent count first declines below the desired threshold, approximated by time of their first simulated monthly latent count below the threshold. A patient not having an observed count (with confirmation if required) below the threshold by that time was considered as starting "late". "Very late" initiation was based on a similar concept among patients having a latent count well below the desired threshold (specifically, latent counts < 350 , < 200 and < 100 cells/ μl for the 500, 350 and 200 thresholds) before meeting the criteria for starting ART during the period simulated.

Results

ART Initiation Times with 6-Monthly Measurements

Figure 1 shows the estimated cumulative proportion of patients in the simulated population who had a CD4 count below the 500, 350, and 200 thresholds by time since first seropositive test based on 6-monthly measurements. Also shown are the corresponding Kaplan-Meier estimates for the 330 seroconverters in the study population. The simulation-based and Kaplan-Meier estimates were similar, showing that the simulation provides a good representation of the distribution of threshold-crossing times. Median time until a first

count below 500 and 350 cells/ μl was 18 and 36 months; for the 200 threshold, it was > 60 months.

Table 1 summarizes operating characteristics for monitoring algorithms with different frequencies of measurement and confirmatory testing requirements. For the benchmark monitoring algorithm (no confirmatory measurement and 6-monthly measurements), during the 61-month period simulated, the mean ART duration was 37.3, 24.0, and 11.1 months and the mean number of pre-treatment measurements was 4.9, 7.1, and 9.2 for the 500, 350 and 200 thresholds.

During the period simulated, 77%, 60%, and 39% of patients had latent counts below 500, 350, and 200 cells/ μl . The proportion having observed counts below these thresholds was higher, 87%, 73%, and 49%, indicating a tendency for earlier ART initiation with 6-monthly monitoring than if latent counts could be monitored. Some patients, however, had a first observed count below a given threshold after their latent count first decreased below that threshold and so would have started ART “late”; for the 350 threshold, during the period simulated, 35% did so. Few patients started ART very “late,” though: 1.6% for the 350 threshold.

Frequency of Measurement

Figure 2(a) shows a trend toward earlier threshold-crossing and hence ART initiation as measurement frequency increases from 12- to 6-, 4- and 3-monthly. This occurs because greater frequency yields more opportunities for fluctuations about the latent count (due to biological variation and measurement error) to give observed counts below a threshold. For the 350 threshold, median ART initiation times were 48, 36, 32, and 30 months, respectively, leading to higher pre-treatment CD4 monitoring costs (means of 4.5, 7.1, 9.6, and 11.8 measurements during the period simulated; Table 1) and higher ART costs (means of 19.0, 24.0, 26.5, and 28.4 months of ART). Conversely, the percentage of patients starting ART “late” decreased (57.6%, 35.3%, 22.4%, and 15.1%) as did the percentage starting “very late” (11.5%, 1.6%, 0.2%, and 0.1%). There were similar patterns for the 500 and 200 thresholds.

As WHO guidelines suggest 6-monthly measurements but then increased frequency when counts approach the threshold to initiate ART [2], for the 350 and 200 thresholds, we evaluated adaptive algorithms involving 6-monthly measurements until an observed count < 500 cells/ μl , and 4- or 3-monthly thereafter (referred to as adaptive 6/4 and adaptive 6/3). These algorithms have operating characteristics intermediate between 6-monthly and 4- or 3-monthly monitoring, respectively. For the 350 threshold, the percentage starting ART late changed from 35.3% for continuous 6-monthly monitoring to 18.9% for adaptive 6/3 monitoring (versus 15.1% for 3-monthly monitoring throughout), and percentage starting ART very late from 1.6% to 0.1% (0.1%) with modest changes in mean number of pre-treatment measurements and time on ART.

Confirmatory Measurements

For 6-monthly monitoring, Figure 2(b) shows that requiring a confirmatory count below a threshold delays ART initiation substantially. For the 350 threshold, median ART initiation time increased from 36 months without confirmation to 49 months with confirmation, with a corresponding increase in mean number of pre-treatment measurements from 7.1 to 9.6 and decrease in mean time on ART from 24.0 to 15.9 months during the period simulated. This delay in starting ART increased the proportion starting ART late (from 35.3 to 76.0%) or very late (1.6 to 8.9%). Table 1 shows similar changes for other thresholds and measurement frequencies.

Measurement Precision and Bias

To explore the effect of different levels of precision, we defined a factor δ which multiplies the standard deviation of differences between observed and latent square-root CD4 counts in our statistical model. The standard deviation reflects a combination of measurement variability and very short-term biological variability, and so is decreased when using a more precise technology than flow cytometry ($\delta < 1$) and increased when using a less precise ($\delta > 1$) technology. Table 2 shows the operating characteristics for 6-monthly monitoring when using measurement technologies having values of δ from 0.5 to 1.5. There was a trend for earlier ART initiation as δ increases. For the 350 threshold, median ART starting time changed from 48 to 36 and 30 months for $\delta = 0.5, 1, \text{ and } 1.5$. This arises because greater variability leads to increased chance of obtaining a measurement below the threshold, and so lower CD4 monitoring costs (means of 8.0, 7.1, and 6.2 pre-treatment measurements during the period simulated) and greater treatment costs (means of 18.7, 24.0, and 29.8 months on ART). Despite this, the proportion of patients starting ART very late showed modest change (0.5%, 1.6%, and 2.2%). These patterns held for all three thresholds.

Table 2 also shows operating characteristics for 6-monthly monitoring using a technology which systematically measures low by 10% (bias = -10%) or high by 10% (bias = +10%) versus flow cytometry. For all three thresholds, technologies with -10% and +10% bias gave median ART starting times 6 months earlier and 6 months later versus no bias, respectively. For the 350 threshold, the median ART starting time was 30, 36, and 42 months for -10% bias, no bias, and +10% bias, giving mean pre-treatment number of CD4 count measurements of 6.4, 7.1, and 7.7 during the period simulated and mean time on ART of 28.2, 24.0, and 20.5 months. Negative bias gave a lower proportion starting late or very late, while positive bias gave a higher proportion. For the 350 threshold, the proportion starting very late was 0.5%, 1.6%, and 3.2% for -10%, no, and +10% bias.

To contrast effects of imprecision and bias on ART starting times, Figure 3 shows the cumulative percentage of patients starting ART for a technology having $\delta = 1$ and -10% bias, versus technologies having $\delta = 1.25$ and 1.5 but no bias. The effect of -10% bias versus increased variability was more pronounced at higher thresholds. For example, for the 200 threshold, the distribution of ART initiation times was similar for a technology with -10% bias and $\delta = 1$ to one having no bias but $\delta = 1.25$, whereas for the 500 threshold, it was similar to one having no bias but $\delta = 1.5$.

Discussion

Our study provides important information concerning factors impacting timing of ART initiation when triggered by a CD4 count below a defined threshold. This has implications for treatment guidelines and evaluation of new technologies for enumerating CD4 counts. We have shown that increased frequency of CD4 measurement, reduced measurement precision, and systematic negative measurement bias lead to earlier initiation on average, while use of confirmatory counts and systematic positive bias lead to later initiation.

US treatment guidelines recommend monitoring CD4 counts every 3–4 months [1], though WHO guidelines recommend every 6 months initially [2]. At the 350 and 500 thresholds, as the immediate risk of HIV-associated clinical events is low [15], our findings suggest that the benefit of earlier ART initiation from monitoring more frequently than every 6 months is likely small and may not justify the increase in monitoring and treatment costs. This is illustrated, for the 350 threshold, by the low proportion of patients in our simulated population starting ART very late with 6-monthly monitoring (1.6%) against the analogous proportions with 4- or 3-monthly monitoring (0.2 and 0.1%, respectively). In contrast, for the 200 threshold where the immediate risk of clinical events is higher, the proportion

starting very late with 6-monthly monitoring (2.6%) may not be acceptable and 4-monthly (0.7%) or 3-monthly (0.1%) monitoring might be preferred. For all three thresholds considered, decreasing the frequency to 12-monthly led to a potentially unacceptable increase in the proportion starting ART very late (e.g., 11.5% for the 350 threshold), consistent with a study which suggested 12-monthly monitoring only at very high counts (e.g. > 900 cells/ μ l when using the 350 threshold) [16]. For the 200 and 350 thresholds, if the proportion starting very late with 6-monthly monitoring is considered too high, our results for adaptive algorithms involving 6-monthly and then 3- or 4-monthly monitoring after an initial count below 500 cells/ μ l appears to perform about as well as 4- or 3-monthly monitoring throughout with similar or lower monitoring and treatment costs. These results concur with cost effectiveness studies which suggest more frequent monitoring after a count close to the threshold, e.g. < 500 or 450 cells/ μ l when using the 350 threshold [17; 18]. Our results provide a formal basis for the concept in WHO guidelines [2] that monitoring be 6-monthly and then more frequent as a patient approaches the threshold.

WHO guidelines recommend obtaining a confirmatory CD4 count if a major decision rests on the value [2]. Requiring such confirmation leads to marked delay in ART initiation and associated increase in proportion of patients starting ART very late (e.g. from 1.6 to 8.9% for the 350 threshold and, notably, from 2.6 to 14.9% for the 200 threshold in our simulation). Requiring confirmation may not, therefore, be desirable particularly at lower thresholds with higher immediate risk of clinical events. The motivation for confirmation arises from high intra-subject variability such that having two measurements below the threshold more reliably indicates a patient's poor immunologic status. Thus, the main value of confirmatory testing is likely for settings with limited treatment availability such that identifying patients most likely to have latent counts below a threshold is important.

An important impetus for this study was to evaluate how differences in performance of a novel CD4 enumeration technology versus flow cytometry might change timing of ART initiation. Typically, however, new and standard technologies are only compared in cross-sectional method comparison studies to assess bias and precision [19]. We found that modest bias, specifically -10% or $+10\%$, led to modest decreases or increases in median time to ART initiation and proportion of patients starting ART very late (0.5, 1.6, and 3.2% for -10% , 0% , and $+10\%$ bias). The elimination of such systematic bias largely concerns calibrating well a novel technology against flow cytometry. It suggests, however, that method comparison studies should be well-powered to rule out biases of the order of $\pm 10\%$, particularly in CD4 count ranges proximal to key thresholds for treatment management rather than on average over a broad range of counts.

The issue of variability in CD4 count measurements is complex because of the numerous sources of variation including not only technological measurement error but also sample processing/shipping, intra-subject biological variation, etc [20; 21], all of which contribute to the "error" in the statistical model in our study. The relative importance of these may differ for a novel technology versus flow cytometry, particularly if the former is for point of care (POC) use whereas the latter is laboratory-based. Our finding was that increases in variability lead to earlier treatment but the effect may be modest. For example, we found that a 25% increase in the standard deviation of the error ($\delta=1.25$ in Table 2 and Figure 3) has less impact on timing of ART initiation than does, for example, choice between 3- versus 6-monthly measurements or a -10% systematic bias in a novel technology. Conversely, a novel technology might give increased precision versus flow cytometry, thereby resulting in a modest delay in ART initiation. A 25% increase (or decrease) in standard deviation across all sources of variation requires, however, a larger than 25% increase (or decrease) in variability in a component source. Because of the difficulties in separating out sources of variability, method comparison studies should provide a

comparison of the total variability across all sources in the settings to be used in practice. For example, such variability might be evaluated by taking pairs of measurements on the same patients a few days apart using each of the novel and standard technologies. For a technology intended for use, this would involve two POC measurements a few days apart whereas for a technology run in a reference laboratory, it would involve two blood draws a few days apart with each transported and processed separately at the laboratory.

Our study has some possible limitations. Our use of a seroincident cohort provides a well-defined time origin for modeling CD4 count trajectories and likely reduces selection biases affecting the distribution of trajectories because subjects were followed from before seroconversion. Study of seroprevalent cohorts would, however, be valuable, ideally involving follow-up from diagnosis of infection so as to include the full period in which a patient would be monitored. Results for a seroprevalent cohort might differ from our findings for two main reasons that may act in opposite directions. Many patients with newly diagnosed infection will have an initial CD4 count below the threshold for starting ART, and so differences in ART starting times arising from the factors explored may be smaller. Conversely, the patient group having high CD4 counts in a seroprevalent cohort will likely be enriched with patients with slower rates of CD4 decline, potentially increasing the impact of these factors. Our study may also be limited by use of data from the 1980s, necessitated by the need for untreated natural history data. CD4 counts measured then may be more variable than counts obtained now using single-platform flow cytometry [22]. However, we found that improvement in precision of a measurement technology was not a major factor in determining timing of ART initiation, likely reflecting the many other sources of variability in counts. The data are also from a US study of homosexual/bisexual men, and so our results may not apply to other populations, for example due to differences in viral or host genetics or environmental or socioeconomic factors [23]. Unfortunately, there are limited long-term data available in other settings [24], though studies have found that healthy African and Asian populations have lower average CD4 counts than their Western counterparts [25; 26; 27; 28]. Conversely, some studies have described slower rates of decline in CD4 count in patients infected with the non-B HIV subtypes predominating in many resource-limited settings [29; 23; 30].

In conclusion, we evaluated factors that are important in monitoring CD4 counts to determine when to initiate ART. In developing guidelines for when to start ART, 6-monthly monitoring appears adequate when using the 350 or 500 thresholds. An adaptive algorithm which increases the frequency of measurement from 6- to 4- or 3-monthly when a count is first observed below 500 cells/ μ l might be preferred if using the 200 threshold and possibly the 350 threshold to reduce the proportion of patients starting ART very late to very low levels. Use of confirmatory counts delays treatment markedly such that an unacceptably high proportion of patients may start ART very late. In developing new CD4 enumeration technologies, method comparison studies should focus on ruling out modest measurement bias, particularly in ranges proximal to important thresholds for treatment management.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Source of Funding

This work was supported in part by grant numbers AI 24643, AI 68634, AI 42006, AI 60354 from the United States National Institutes of Health (NIH). The content is solely the responsibility of the authors and does not represent the official views of the NIH.

References

1. Panel on Antiretroviral Guidelines for Adults and Adolescents. Guidelines for the use of antiretroviral agents in HIV-1-infected adults and adolescents. Department of Health and Human Services; Washington DC, USA: Oct 14. 2011 Available at <http://www.aidsinfo.nih.gov/ContentFiles/AdultandAdolescentGL.pdf>
2. World Health Organization. Antiretroviral therapy for HIV infection in adults and adolescents: recommendations for a public health approach - 2010 revision. Geneva, Switzerland: Nov 30. 2010 Available at <http://whqlibdoc.who.int/publications/2010/9789241599764eng.pdf>
3. World Health Organization. Antiretroviral therapy for HIV infection in adults and adolescents: recommendations for a public health approach - 2006 revision. Geneva, Switzerland: Aug 7. 2006 Available at <http://www.who.int/entity/hiv/pub/guidelines/artadultguidelines.pdf>
4. Mtapuri-Zinyowera S, Chideme M, Mangwanya D, Mugurungi O, Gudukeya S, Hatzold K, et al. Evaluation of the PIMA point-of-care CD4 analyzer in VCT clinics in Zimbabwe. *JAIDS*. 2010; 55:1–7. [PubMed: 20622679]
5. Li X, Breukers C, Ymeti A, Pattanapanyasat K, Sukapirom K, Terstappen LWMM, et al. Clinical evaluation of a simple image cytometer for CD4 enumeration on HIV-infected patients. *Cytometry Part B: Clinical Cytometry*. 2010; 78:31–36.
6. Jani IV, Siteo NE, Chongo PL, Alfai ER, Quevedo JI, Tobaiwa O, et al. Accurate CD4 T-cell enumeration and antiretroviral drug toxicity monitoring in primary healthcare clinics using point-of-care testing. *AIDS*. 2011; 25:807–812. [PubMed: 21378535]
7. Diaw PA, Daneau G, Coly AA, Ndiaye BP, Wade D, Camara M, et al. Multisite evaluation of a point-of-care instrument for CD4+ T-cell enumeration using venous and finger-prick blood: the PIMA CD4. *JAIDS*. 2011; 58:e103–e111. [PubMed: 21909029]
8. Noubary F, Hughes MD. Assessing agreement in the timing of treatment initiation determined by repeated measurements of novel versus gold standard technologies with application to the monitoring of CD4 counts in HIV-infected patients. *Statistics in Medicine*. 2010; 29:1932–1946. [PubMed: 20680986]
9. Multicenter AIDS Cohort Study (MACS). Public Data Set: Release PO4. Springfield, VA: National Technical Information Service; 1995. Available at <http://www.statepi.jhsph.edu/macspdt.html>
10. Kaslow RA, Ostrow DG, Detels R, Phair JP, Polk BF, Rinaldo CR Jr. The Multicenter AIDS Cohort Study: rationale, organization, and selected characteristics of the participants. *American Journal of Epidemiology*. 1987; 126:310–8. [PubMed: 3300281]
11. Giorgi JV, Cheng HL, Margolick JB, Bauer KD, Ferbas J, Waxdal M, et al. Quality control in the flow cytometric measurement of T-lymphocyte subsets: the Multicenter AIDS Cohort Study experience. *Clinical Immunology and Immunopathology*. 1990; 55:173–186. [PubMed: 1969782]
12. Taylor JMG, Cumberland WG, Sy JP. A stochastic model for analysis of longitudinal AIDS data. *Journal of the American Statistical Association*. 1994; 89:727–736.
13. Boscardin WJ, Taylor JMG, Law N. Longitudinal models for AIDS marker data. *Statistical Methods in Medical Research*. 1998; 7:13–27. [PubMed: 9533259]
14. Wolbers M, Babiker A, Sabin C, Young J, Dorrucci M, Chêne G, et al. Pretreatment CD4 cell slope and progression to AIDS or death in HIV-infected patients initiating antiretroviral therapy—the CASCADE Collaboration: a collaboration of 23 cohort studies. *PLoS Medicine*. 2010; 7:e1000239. [PubMed: 20186270]
15. Sterne J, May M, Costagliola D, De Wolf F, Phillips A, Harris R, et al. Timing of initiation of antiretroviral therapy in AIDS-free HIV-1-infected patients: a collaborative analysis of 18 HIV cohort studies. *The Lancet*. 2009; 373:1352–63.
16. Buclin T, Telenti A, Perera R, Csajka C, Furrer H, Aronson JK, et al. Development and validation of decision rules to guide frequency of monitoring CD4 cell count in HIV-1 infection before starting antiretroviral therapy. *PloS one*. 2011; 6:e18578. [PubMed: 21494630]
17. Hallett TB, Gregson S, Dube S, Garnett GP. The impact of monitoring HIV patients prior to treatment in resource-poor settings: insights from mathematical modelling. *PLoS medicine*. 2008; 5:e53. [PubMed: 18336064]

18. Kimmel AD, Goldie SJ, Walensky RP, Losina E, Weinstein MC, Paltiel AD, et al. Optimal frequency of CD4 cell count and HIV RNA monitoring prior to initiation of antiretroviral therapy in HIV-infected patients. *Antiviral Therapy*. 2005; 10:41–52. [PubMed: 15751762]
19. Altman DG, Bland JM. Measurement in medicine: the analysis of method comparison studies. *The Statistician*. 1983; 32:307–317.
20. Stevens W, Gelman R, Glencross DK, Scott LE, Crowe SM, Spira T. Evaluating new CD4 enumeration technologies for resource-constrained countries. *Nature Reviews Microbiology*. 2008; 6:S29–S38.
21. Mandy FF, Nicholson JK, McDougal JS. Guidelines for performing single platform absolute CD4 T-cell determinations, with CD45 gating for persons infected with human immunodeficiency virus. *MMWR*. 2003; 52:1–13. [PubMed: 12583540]
22. Barnett D, Granger V, Whitby L, Storie I, Reilly JT. Absolute CD4+ T-lymphocyte and CD34+ stem cell counts by single-platform flow cytometry: the way forward. *British Journal of Haematology*. 1999; 106:1059–1062. [PubMed: 10520014]
23. Muller V, Wyl V, Yerly S, Bo'ni J, Klimkait T, Burgisser P, et al. African descent is associated with slower CD4 cell count decline in treatment-naïve patients of the Swiss HIV Cohort Study. *AIDS*. 2009; 23:1269–1276. [PubMed: 19461503]
24. Wandel S, Egger M, Rangsiri R, Nelson KE, Costello C, Lewden C, et al. Duration from seroconversion to eligibility for antiretroviral therapy and from ART eligibility to death in adult HIV-infected patients from low and middle-income countries: collaborative analysis of prospective studies. *Sexually Transmitted Infections*. 2008; 84:i31–i36. [PubMed: 18647863]
25. Clerici M, Butto S, Lukwiyab M, Saresellac M, Declichd S, Trabattoni D, et al. Immune activation in Africa is environmentally-driven and is associated with upregulation of CCR5. *AIDS*. 2000; 14:2083–2092. [PubMed: 11061648]
26. Howard RR, Fasano CS, Frey L, Miller CH. Reference intervals of CD3, CD4, CD8, CD4/CD8, and absolute CD4 values in Asian and non-Asian populations. *Cytometry Part B: Clinical Cytometry*. 1996; 26:231–232.
27. Messele T, Abdulkadir M, Fontanet AL, Petros B, Hamann D, Koot M, et al. Reduced naïve and increased activated CD4 and CD8 cells in healthy adult Ethiopians compared with their Dutch counterparts. *Clinical and Experimental Immunology*. 1999; 115:443–450. [PubMed: 10193416]
28. Lewden C, Thi'ebaut R, Boufassa F, Coulibaly A, Malateste K, Seng R, et al. Comparison of Early CD4 T-Cell Count in HIV-1 Seroconverters in Cote d'Îvoire and France: The ANRS PRIMO-CI and SEROCO Cohorts. *JAIDS*. 2010; 53:260–265. [PubMed: 19745754]
29. Keller M, Lu Y, Lalonde RG, Klein MB. Impact of HIV-1 viral subtype on CD4+ T-cell decline and clinical outcomes in antiretroviral naïve patients receiving universal healthcare. *AIDS*. 2009; 23:731–737. [PubMed: 19279446]
30. May M, Wood R, Myer L, Taff'e P, Rauch A, Battegay M, et al. CD4+ T cell count decreases by ethnicity among untreated patients with HIV infection in South Africa and Switzerland. *Journal of Infectious Diseases*. 2009; 200:1729–1735. [PubMed: 19848608]

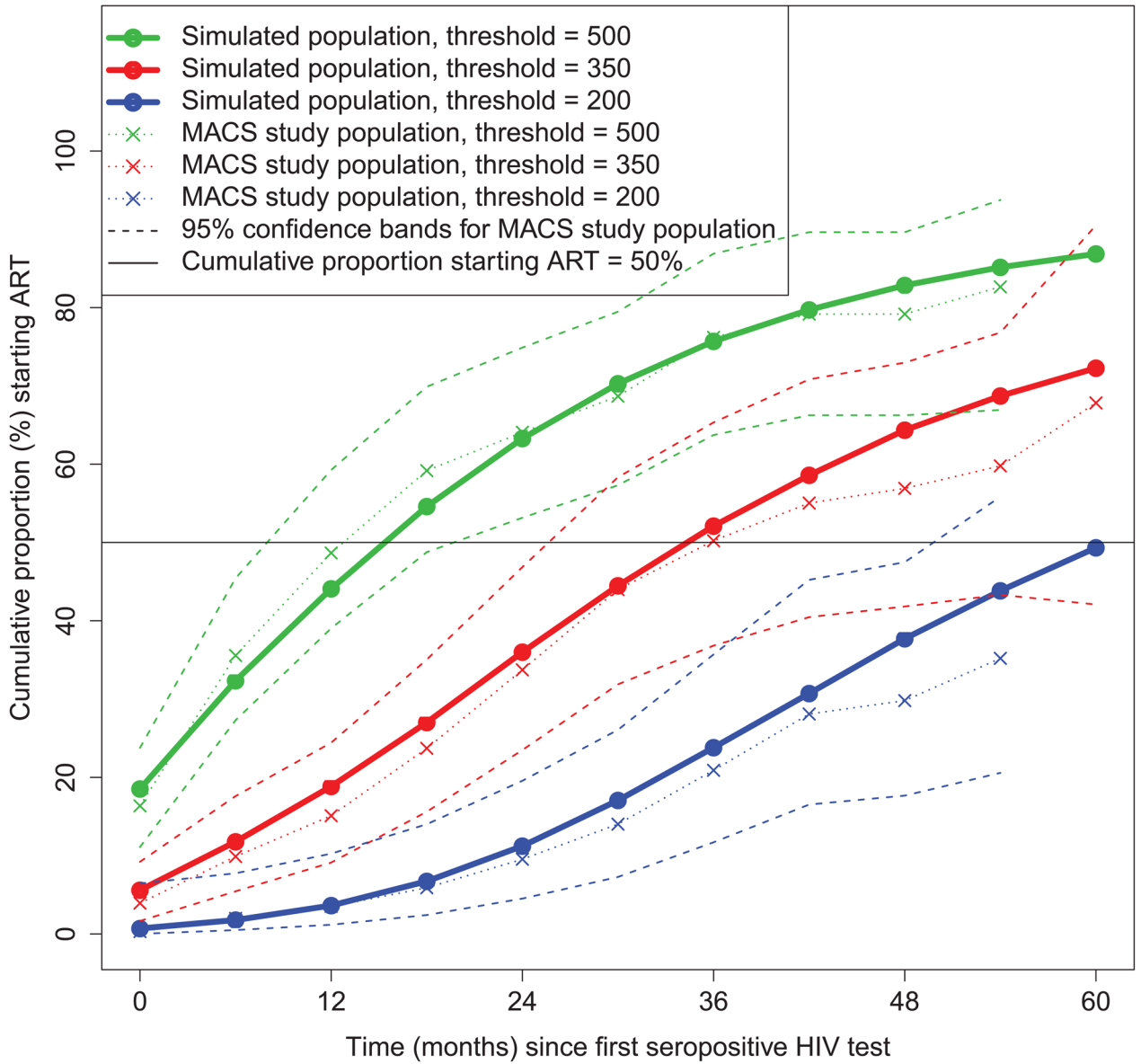


Figure 1. Cumulative proportions of patients starting antiretroviral therapy over time for CD4 count thresholds of 500, 350 and 200 cells/ μ l in simulated data compared with the corresponding estimates (with 95% confidence bands) based on data from the Multicenter AIDS Cohort Study (MACS).

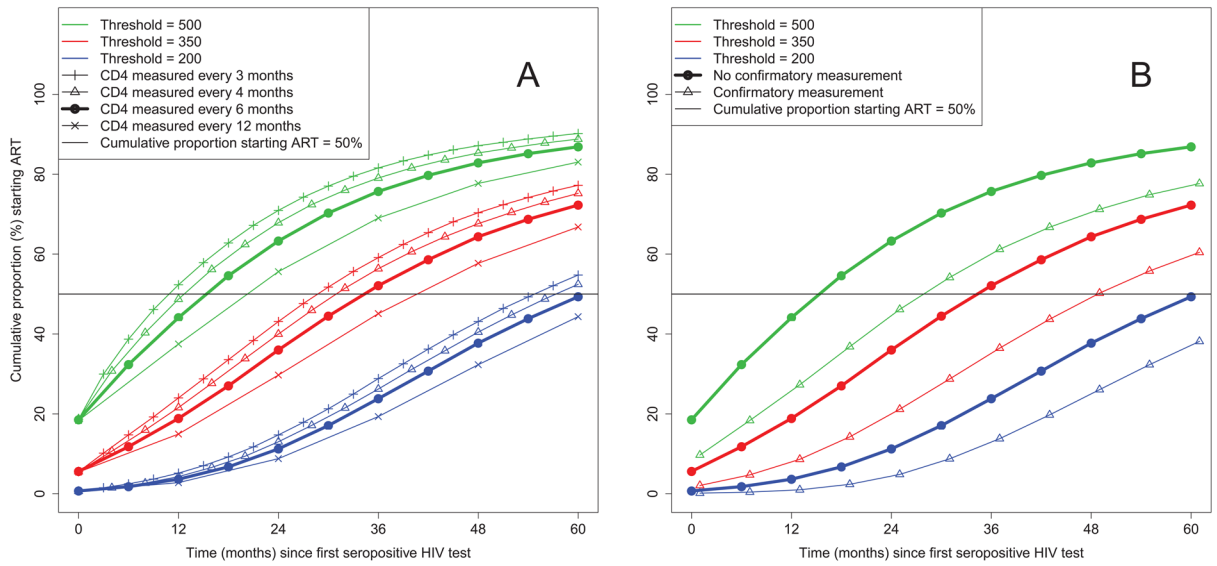


Figure 2. Cumulative proportions of patients starting antiretroviral therapy over time for CD4 count thresholds of 500, 350 and 200 cells/ μ l and (A) different frequencies of CD4 count measurements, and (B) without and with a requirement for a confirmatory measurement.

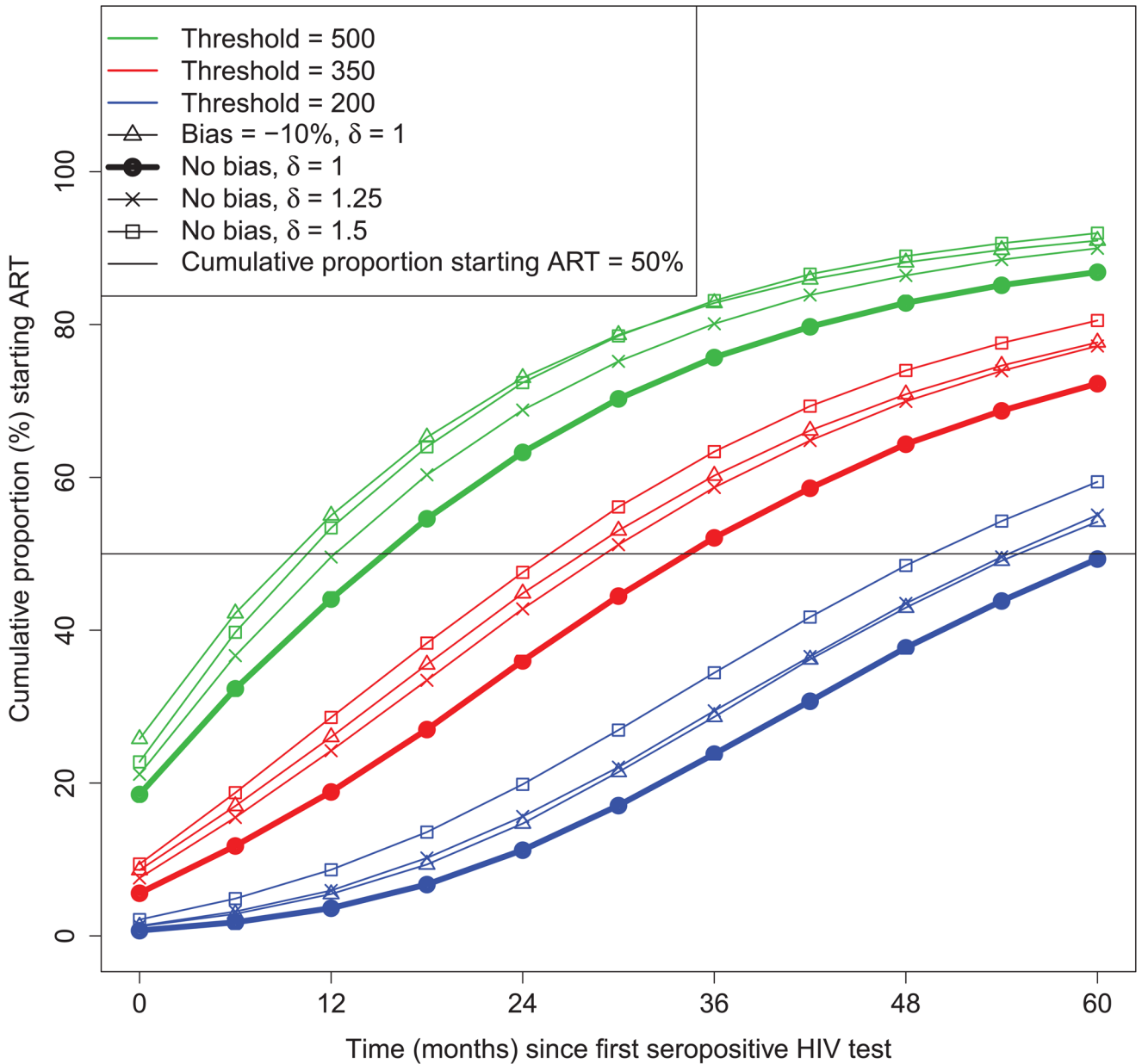


Figure 3.

Cumulative proportions of patients starting antiretroviral therapy over time for CD4 count thresholds of 500, 350 and 200 cells/ μ l based on measurements from the benchmark CD4 count enumeration technology (flow cytometry) as well as from biased and less precise technologies. δ measures the relative change in the standard deviation of short-term variability of CD4 counts about a subject's latent CD4 count when using a novel CD4 count enumeration technology versus the flow cytometry.

Table 1
Operating characteristics for CD4 monitoring algorithms with different frequencies of measurement and confirmatory testing requirements

Confirmatory measurement?	CD4 threshold cells/ μ l	Time between measurements (months)	Median ART start time (months)	Mean over 61 months per patient		% starting ART	
				Pre-treatment CD4 measurements	Months on ART	late ¹	very late ²
No	200	12	>60	5.4	8.1	62.7	18.3
		6	>60	9.2	11.1	38.0	2.6
		4	60	13.0	12.7	24.1	0.7
		3	57	16.5	14.0	16.1	0.1
		adaptive 6/4 ³	>60	11.0	12.3	28.5	1.3
		adaptive 6/3 ⁴	57	12.6	13.9	15.9	0.1
	350	12	48	4.5	19.0	57.6	11.5
		6	36	7.1	24.0	35.3	1.6
		4	32	9.6	26.5	22.4	0.2
		3	30	11.8	28.4	15.1	0.1
		adaptive 6/4	36	7.9	25.3	28.1	0.7
		adaptive 6/3	33	8.3	27.0	18.9	0.1
500	200	12	24	3.4	31.8	52.0	15.7
		6	18	4.9	37.3	34.4	3.6
		4	16	6.2	40.0	24.0	1.2
		3	12	7.4	41.8	17.8	0.8
		12	>61	6.2	4.7	88.6	41.5
		6	>61	10.7	6.6	77.3	14.9
	350	12	61	5.9	12.5	87.0	29.3
		6	49	9.6	15.9	76.0	8.9
		4	49	13.0	17.7	66.6	3.3
		3	46	16.3	19.0	59.2	1.3
		12	37	5.2	23.4	87.5	39.7
		500	37	5.2	23.4	87.5	39.7

Confirmatory measurement?	CD4 threshold cells/ μ l	Time between measurements (months)	Median ART start time (months)	Mean over 61 months per patient		% starting ART	
				Pre-treatment CD4 measurements	Months on ART	late ¹	very late ²
		6	31	7.8	28.0	78.3	18.8
		4	29	10.1	30.3	70.1	10.2
		3	25	12.2	31.9	63.3	7.5

¹The percentage starting ART late is calculated as the number who start ART at any time after their latent CD4 count is first below the threshold divided by the number who have a latent CD4 count below the threshold at any time during the 60 months simulated.

²The very late thresholds for ART initiation thresholds of 500, 350 and 200 cells/ μ l are defined as 350, 200 and 100 cells/ μ l, respectively. The percentage starting ART very late is then calculated as the number who start ART at any time after their latent CD4 count is first below the very late threshold divided by the number who have a latent CD4 count below the very late threshold at any time during the 60 months simulated.

³The adaptive 6/4 monitoring protocol calls for 6-monthly CD4 measurement until a patient crosses a threshold of 500 and a subsequent 4-monthly schedule thereafter until ART initiation.

⁴Same as above but with a subsequent 3-monthly CD4 measurement schedule.

Table 2
Operating characteristics for CD4 monitoring technologies with different measurement precision and bias

CD4 threshold (cells/ μ l)	Bias	Relative change in SD of short-term variability (δ)	Median ART start time (months)	Mean over 61 months per patient		% starting ART late ¹	% starting ART very late ²
				Pre-treatment CD4 measurements	Months on ART		
200	0%	0.5	>60	9.7	8.0	58.6	1.1
	0%	0.75	>60	9.5	9.3	47.3	2.2
	0%	1	>60	9.2	11.1	38.0	2.6
	0%	1.25	60	8.8	13.6	31.0	3.2
	0%	1.5	54	8.5	15.9	25.7	3.3
	-10%	1	60	8.9	13.3	25.7	1.3
350	+10%	1	>60	9.5	9.3	50.7	5.4
	0%	0.5	48	8.0	18.7	54.3	0.5
	0%	0.75	42	7.6	21.2	43.2	1.0
	0%	1	36	7.1	24.0	35.3	1.6
	0%	1.25	30	6.6	27.3	29.3	1.9
	0%	1.5	30	6.2	29.8	24.9	2.2
500	-10%	1	30	6.4	28.2	21.5	0.5
	+10%	1	42	7.7	20.5	50.1	3.2
	0%	0.5	24	5.8	31.9	47.9	1.2
	0%	0.75	24	5.4	34.6	39.8	2.4
	0%	1	18	4.9	37.3	34.4	3.6
	0%	1.25	18	4.5	39.9	30.5	4.3
600	0%	1.5	12	4.2	41.7	27.9	4.8
	-10%	1	12	4.1	42.1	19.6	1.5
	+10%	1	24	5.7	32.8	49.8	7.8

¹The percentage starting ART late is calculated as the number who start ART at any time after their latent CD4 count is first below the threshold divided by the number who have a latent CD4 count below the threshold at any time during the 60 months simulated.

²The very late thresholds for ART initiation thresholds of 500, 350 and 200 cells/ μ l are defined as 350, 200 and 100 cells/ μ l, respectively. The percentage starting ART very late is then calculated as the number who start ART at any time after their latent CD4 count is first below the very late threshold divided by the number who have a latent CD4 count below the very late threshold at any time during the 60 months simulated.