Supplementary Web Appendix

This appendix forms part of the original manuscript.

Diana M. Negoescu, Zhenhuan Zhang, Heiner C. Bucher, Eran Bendavid and the Swiss HIV Cohort Study

Differentiated HIV RNA Viral Load Monitoring in Resource Limited Settings: An Economic Analysis

1. Data description and preparation:

The data we use consists of 95414 clinical visits from 5251 HIV infected patients in Switzerland who started a preferred or alternative ART regimen according to the DHHS 2013 guidelines after April 1st.

Inclusion criteria: All naïve patients starting a preferred or alternative antiretroviral therapy (ART) regimen based on the 2013 DHHS guidelines who had at least two viral load and CD4 cell count measurements, and a measurement of CD4 cell count at baseline prior to ART initiation. Regimens considered consisted of a ritonavir boosted protease inhibitor (PI) or a non-nucleoside reverse transcriptase inhibitor (NNRTI) and at least 2 nucleoside reverse transcriptase inhibitors (NRTIs).

Exclusion criteria: Initiating of ART during pregnancy.

Variables for analysis: CD4 cells at baseline and at each subsequent measurement, viral load at baseline (<100’000 versus ≥100’000 copies/ml) and at subsequent measurements,
age, gender, type of regimen (PI versus NNRTI based), type of NRTI backbone (abacavir / lamivudine versus other backbone), self-reported adherence.

Virologic failure was defined as 2 consecutive measurements of a viral load >200 copies/ml or one viral load measurement >1000 copies/ml. Those who do not reach virologic suppression within six months of ART initiation or regimen change are also considered to be in virologic failure. Non-adherence was defined as having missed more than 2 ART doses in the past 6 months prior to a biannual SHCS cohort visit, as self-reported by the patient in an adherence questionnaire.

1.2. Simulation Model:

Our stochastic simulation model is determined by three quantities updated at each monthly time step: patient non-adherence (binary -adherent or non-adherent), failure status (binary – in failure or not in failure), and CD4 cell progression (real-valued).

At each monthly time step in our simulation of a patient, we sample the patient’s non-adherence status. Having this, we then sample the patient’s failure status, and finally, given the failure status we sample the patient’s new CD4 cell count.

We now provide a detailed description for each of the three key components of our model:

1.2.1. Non-Adherence:

We modeled a patient’s current non-adherence status using mixed effect logistic regression fitted to SHCS data, where the binary outcome of current non-adherence was modeled using a logit function of
• Previous non-adherence status
• Age
• Gender
• Time since last measurement

Specifically, we fit a mixed effect logit regression model to our data, such that

\[
\text{Prob}(\text{Patient } i \text{ has been non-adherent during the past month}) = \\
= \frac{1}{1 + \exp[-(\beta_0 + \beta_1 \text{Age} + \beta_2 \text{Sex} + \beta_3 \text{Education} + \beta_4 \\
(\text{Previous_nonadherence}) + \beta_5 (\text{Time between measurements}) + \beta_6 (\text{Previous_non-adherence}) \times (\text{Time between measurements})] + p_i ]},
\]

where

• Age is measured in years,
• Sex is a binary variable (0 = male, 1 = female),
• Previous non-adherence is binary (0 = previously adherent, 1 = previously non-adherent),
• Education is a binary variable (0 = primary education or less, 1 = secondary education or more)
• Time between measurements is measured in days, and enters the regression models as a linearly splined function, with two knots at 70 and 110 days. These values were selected to be clinically meaningful as to have an equal number of observations falling within each spline segment.
• \(\beta_0, ..., \beta_6\) are fixed effects coefficients
• \(p_i\) is the random intercept of patient \(i\), where \(p \sim N (0, \Sigma)\)
When simulating adherence at monthly time steps, we use this logit model with the *Time between measurements* variable fixed at 30 days. We assume that the interaction between patients and physicians during the monitoring visits induces a basic level of adherence, i.e. we do not consider ‘lost cause’ patients who never attempt to take the treatment - we acknowledge that for those patients monitoring viral load is likely unnecessary.

1.2.2. Likelihood of failure:

We are primarily interested in accurately modeling the risk of failure for a patient on a first-line regimen. This is mainly because only two regimens are usually available in resource-limited settings, and therefore once a patient has switched to secondline therapy, detecting a virologic failure brings no further benefits, as the patient does not have the option of switching again. In our simulation, once a patient is switched to secondline therapy, we use monthly probability of failure obtained from the rates reported in [1], which do not depend on individual patient characteristics.

Our model also assumes that, if a virologic failure is generated for a patient previously not in failure, there is a chance (set to 0.58 if patient is adherent, and 0.49 if patient is non-adherent – values obtained from analyzing SHCS data) that patients become resuppressed on the same regimen within a month. If patients do not become resuppressed within a month, they remain in virologic failure unless a switch in regimen occurs.
For a patient previously not in virologic failure and who hasn’t switched regimens, we fit a logistic regression model to predict current failure status to the data from the SHCS coming from patients on first-line regimen, who had reached virologic suppression and did not present virologic failure at the previous measurement. We model the chance that the patient is currently in virologic failure as a function of:

- Current non-adherence status
- Time on current regimen
- Age
- Gender
- Time since last measurement

Specifically

\[
\text{Prob(}\text{Patient is currently in virologic failure}) = \\
\frac{1}{1 + \text{Exp}[-(\beta_0 + \beta_1 \text{Age} + \beta_2 \text{Sex} + \beta_3 \text{(Current non-adherence)} + \beta_4 \\
\text{(Time between measurements)} + \beta_5 \text{(Time on current regimen)})]}
\]

where Time on current regimen is measured in days and enters as a linear spline function, with knots at 200 and 540 days, chosen such that about a third of data points fall with each segment, and the other covariates are defined in the same way as in the adherence model.

1.2.3. CD4 cell progression:

Estimating and simulating CD4 cell progression for each patient through time is very important, since CD4 cell values dictate the risk of opportunistic infections, death and quality of life. The current CD4 cell count of a patient depends on previous CD4 cell
count (and other characteristics), with higher variation in current CD4 values for those with higher previous CD4 cell values. To capture the heteroskedasticity in the data, we use quantile regression models to fit the change in CD4 cell count as a function of

- Current failure status
- Previous CD4 cell count
- Time since ART initiation
- CD4 cell at ART initiation
- Age
- Gender

Specifically, given a set of patient characteristics, our quantile regression models output the $1^{st}$, $5^{th}$, $10^{th}$, $20^{th}$, …, $80^{th}$, $90^{th}$, $95^{th}$, $99^{th}$ percentile of the change in CD4 cell count. We then fit piece-wise linear functions through each percentile and use the resulting functions as distributions from which to sample the value of change in CD4 cells.

We fit two CD4 cell progression models: one for those patients currently in failure, and one for those currently not in failure.

To fit the model for patients currently in failure, we selected the SHCS data coming from patients who had previously reached virologic suppression, and had a measurement along with the previous measurement recording virologic failure without a change in regimen. Since SHCS data is likely biased towards patients who do not experience CD4 cell decline, we considered outliers those measurements where the change in CD4 cell count was more than 25% of the previous CD4 cell value (more than
a 25% increase in CD4 cells despite failure status), or those whose CD4 cell count increased by more than 250 CD4 cells/ mm$^3$ despite failure status.

The $j^{th}$ percentile of the change in CD4 cell count from the previous measurement for a patient in failure was then a function of the form

$$\Delta CD4_j = \beta_0^j + \beta_1^j (Prev\ CD4) + \beta_2^j Age + \beta_3^j Sex + \beta_4^j (Time\ between\ measurements)$$

where $Prev\ CD4$ enters as a linear spline function with knots at 200 and 500 cells/ mm$^3$, age and the other covariates are defined in the same way as in the previous models.

To fit the CD4 cell progression model for patients currently not in failure, we selected the SHCS data corresponding to non-failure measurements. We removed outliers where the increase in CD4 cell counts from two consecutive measurements was more than 1000 cells/ mm$^3$.

The $j^{th}$ percentile of the change in CD4 cell count from the previous measurement for a patient not in failure was then a function of the form

$$\Delta CD4_j =$$

$$\beta_0^j + \beta_1^j (Prev\ CD4) + \beta_2^j Age + \beta_3^j Sex + \beta_4^j (Time\ between\ measurements) +$$

$$\beta_5^j (Time\ since\ ART\ initiation) + \beta_6^j (CD4\ at\ initiation) + \beta_7^j (CD4\ at\ initiation) * (Time\ since\ ART\ initiation)$$

where $Time\ since\ ART\ initiation$ enters as a linear spline function with knots at 90 and 480 days, $CD4\ cells\ at\ initiation$ is the last recorded CD4 value before ART was started, age and the other covariates are defined in the same way as in the previous models.

1.2.4. Modeling mortality
The probability that a patient dies within a given month depends on the patient’s current CD4 cell counts, virologic failure states, gender and age. We use WHO life tables to estimate age and gender-specific mortality. We use [2] to estimate CD4 cell-specific mortality rates if the patient is not currently in virologic failure and [3] if virologic failure is present. Annual mortality rates are shown in Table A1.

Given a patient of a specific failure status, CD4 cell count, gender and age, we compute her yearly mortality rate according to

\[ m_{CD4, \text{age}} = \frac{m_{CD4} \cdot m_{\text{age}}}{\sum_{\text{age}_i} w_i \cdot m_{\text{age}_i}} \]

where \( m_{CD4} \) is the CD4 cell and failure-specific mortality specified in Table 1, \( m_{\text{age}} \) is the age-specific mortality rate as obtained from WHO life tables, and \( w_i \) is the proportion of population in age group \( i \), estimated from WHO life tables.

1.2.5 **Modeling Opportunistic Infections (OIs)**

Patients in any CD4 cell states can develop opportunistic infections, in which case they will present to their clinic and will be monitored. We estimated CD4 cell-specific OI-rates by starting from those reported in [4] and adjusting them in calibration to match OI-rates reported in [5]. Our rates are reported in Table A2.

1.2.6 **Mapping a health state to QALYs**

The health utility (Quality-adjusted life years, QALY) of a patient in a given month depends on the CD4 cell count [6, 7] of the patient during that month according to Table A3, where one QALY corresponding to one year in perfect health.
1.2.7. Mapping a health state to costs

Each month in the life of a patient in our simulation brings about a health cost that depends on the patient’s CD4 cell count (whether it is less than 250 cells/mm$^3$ or not), whether the patient is on first-line or second-line regimen and whether the patient was administered a VL test that month [8-11]. Cost parameters are summarized in Table A4.

1.2.8. Estimating secondary infections

Timely identification of cases of virologic failure is important not only for the patient, but also for the community at large, since patients experiencing virologic failure have a much higher rate of HIV transmission than those on ART with undetectable VL.

Even though we do not explicitly model HIV transmission in our framework, we can estimate the number of secondary infections resulting from a given policy since our simulation model keeps track of how many months each patient spends in virologic failure. If a patient spends $l$ months in failure and $k$ months not in failure during the time horizon of our simulation, we estimate the number of secondary infections generated by this patients using the following equation presented in [12].

$$S = m(1 - \pi) \left(1 - (1 - \alpha_l)^{n_{sex}l}(1 - \alpha_k)^{n_{sex}k}\right)$$

where $m$ is the average number of partners, $\pi$ is the HIV prevalence (i.e., the probability that a partner is HIV-positive), $\alpha_l$ is the probability of HIV transmission per sex act if the patient is in virologic failure, $\alpha_k$ is the probability of HIV transmission per sex act if the patient is not in virologic failure, and $n_{sex}$ is the average number of sex acts per month. The values we use for these parameters are summarized in Table A5 (Sources: [13-15]).
To adjust the costs and QALYs from our simulation so that we account for secondary infections resulting from time spent by patients in failure, we estimate the costs and QALYs during the simulation time horizon of a person from the population if that person had not become infected. To do so, we first average the time spent alive by patients in our simulation ($T_{\text{alive}}$). We then estimate the total discounted QALYs of an uninfected patient during this time by assuming a health utility value of 0.9, i.e. a secondary infection patient would have achieved

$$QALY_{\text{NoHIV}} = 0.9(1 + \alpha + \alpha^2 + \ldots + \alpha^{T_{\text{alive}}}) = 0.9 \frac{1 - \alpha^{T_{\text{alive}}+1}}{1 - \alpha}$$

QALYs had the person not become infected, where $\alpha$ is a yearly discount factor (3%). Similarly, we estimate that the health cost of a person infected by someone in our simulation would have been

$$\text{Costs}_{\text{NoHIV}} = \frac{1}{10} c_{\text{average}}(1 + \alpha + \alpha^2 + \ldots + \alpha^{T_{\text{alive}}}) = \frac{1}{10} c_{\text{average}} \frac{1 - \alpha^{T_{\text{alive}}+1}}{1 - \alpha}$$

where $c_{\text{average}}$ is the average yearly cost incurred by a patient in our simulation, and where we assume that the yearly cost of a non-infected person is 1/10 of that of an infected person.

We then estimate the loss in QALYs for each secondary infection as

$$QALY_{\text{loss}} = QALY_{\text{NoHIV}} - Q_{\text{simulation}}$$

where $Q_{\text{simulation}}$ are the average per patient QALYs from our simulation model; and we estimate the gain in costs for each secondary infection as $\text{Costs}_{\text{gained}} = C_{\text{simulation}} - \text{Costs}_{\text{NoHIV}}$, where $C_{\text{simulation}}$ is the average per patient cost from our simulation model. Then, the QALYs of patient $i$ are adjusted according to $QALY(i)_{\text{adjusted}} = QALY(i)_{\text{unadjusted}} - S(i) \times QALY_{\text{loss}}$, and the costs of patient $i$ are adjusted according to $\text{Cost}(i)_{\text{adjusted}} = \text{Cost}(i)_{\text{unadjusted}} + S(i) \times \text{Costs}_{\text{gained}}$. 

10
1.3. Simulating fixed interval monitoring policies

We use our simulation model to evaluate the performance of five fixed interval monitoring policies: every 1, 3, 6, 12 and 24 months, regardless of medical history or patient characteristics. For all policies, if a patient on first-line regimen is found to be in virologic failure, he/she is asked to come for a follow-up VL monitoring appointment in one month. If the patient is found to be in virologic failure at the follow-up appointment, he/she is prescribed a second-line regimen. We use this follow-up appointment in order to avoid patients being prescribed second-line regimen in the cases in which their virologic failure would spontaneously resolve by next month. We assume that all patients would be present for their appointment. We also assume that patients who experience an opportunistic infection (OI) will self-present and have a VL test even if they are not due for a test at the time of their OI. Once a patient is switched to second-line regimen, we assume he or she will be monitored once a year.

1.4. Adaptive policy optimization

Since different patients can have different risks of failures, our goal was to search for monitoring policies that are specific to the current health state of the patient, the patient’s medical history and personal characteristics. We call such policies “adaptive”. For instance, a patient who reports having missed more than 2 doses of ART in the past month should be called for the next VL test in a shorter time than someone who reported taking all ART doses.
A good adaptive policy provides a balance between “over-monitoring”, meaning monitoring VL earlier than when virologic failure occurs, and “under-monitoring”, meaning monitoring VL later than when VL occurs. Monitoring very frequently brings about more costs from monitoring and a potential treatment change, but improves the QALYs of the patient and limits the number of months that the patient spends in VL failure (period during which HIV transmission is more likely). Since achieving higher QALYs also requires more cost expenditures, at a smaller cost-effectiveness threshold (CET), measured in dollars per/ QALY, monitoring intervals should generally be larger than at higher CET values.

We note that the monitoring decision is only relevant as long as the patient is on first-line regimen, since we assume that in limited resource settings there are only two regimens available, and therefore once a patient is on second-line regimen there are no more decisions to be made.

We take a stochastic optimization approach and develop a policy where, given a patient with specific health attributes (such as time since ART initiation, current adherence status, age, sex, education level), we determine the time in months until the next VL test by maximizing the average per month expected net monetary benefit (NMB), where $NMB = QALYS \times CET - COSTS$. Specifically, we find $T$ that maximizes

$$\max_T \frac{NMB(T)}{T}$$

where $NMB(T)$ is the expected net monetary benefit if the patient is only monitored in $T$ time (months). If failure occurs prior to $T$, the patient will spend the rest of the time in VL failure. We divide by T in order to normalize the NMB values to monthly averages – otherwise, differences in NMB values could be caused simply because of different time
horizons considered. In maximizing NMB, we attempt to balance the costs of “over-monitoring” (monitoring sooner than when VL failure occurs) against the costs of “under-monitoring” (monitoring after failure occurs) while also having in mind the resource setting.

To compute the respective NMBs, we estimate the expected net monitoring benefit in each of the future 36 months using a simplified Markov chain informed by some of the elements from our simulation model. Specifically, the Markov chain has 4 states:

1. Adherent and not in failure (AN)
2. Non-adherent and not in failure (NN)
3. In virological failure (F) and
4. Death (D).

The major difference between our simulation model and this Markov chain is the absence of information on CD4 cell counts. We use this simplification for computational tractability, to keep the state space small when optimizing the monitoring policies. A schematic representation of the model is shown in Figure A1. Individuals may also leave each compartment (except for Death) according to the mortality or transition rates given the patient’s current state.

We define transition probabilities

\[ P_k(i, j) \triangleq \text{probability of transitioning from state } i \text{ to state } j \text{ from } \]

\[ \text{time } k - 1 \text{ to } k \text{ months in the future} \]

These transition probabilities are informed by our logit regressions of adherence and failure, which are time inhomogeneous and also depend on patient’s age, sex, and
educational level and time since ART initiation. Mortality, costs and QALYs in the Markov chain are mapped similarly as in the simulation (Section 1.2), where we assume a CD4 cell count of 500 CD4 cells/ mm$^3$ for non-failure states and a CD4 cell count of 250 cells/ mm$^3$ for the failure state.

To estimate the expected average NMB, we initialize the Markov chain with a 1 in the corresponding current state of the patient and run the chain for 36 months into the future (e.g. $k \in \{1, 2, ..., 36\}$). We thus obtain the probability that the patient is in each of the chain’s 4 states at every month in the future 3 years. Using this information, we can compute the expected number of months the patient will spend in each of the 4 states, in particular in the alive states and the failure state, and thus estimate the cost and QALYS associated (hence the NMB) from the present until the time of monitoring. We then choose the time of monitoring that maximize the normalized NMB.

We now mathematically describe the methodology we mentioned above in a few steps. Let $X_t$ be a random variable denoting the patient health state (one of the 4 possible states) at time (month) $t$; similarly, let $x_t$ be a realization of $X_t$.

Our objective is to find a monitoring interval $T$ which maximizes $NMB(T)/T$. We express the total expected net monetary benefit from the present up to time $T$ as follows:

$$NMB(T) = \sum_{k=1}^{T} E[Q(X_k)] \cdot CET - \sum_{k=1}^{T} E[C(X_k)] - E[S(X)] \cdot (QALY_{loss/infection} \cdot CET + Cost_{gained/infection})$$

(*)

The first term on the right hand side is the total expected QALY gained in the next $T$ months, where $Q(X_k)$ is the health utility associated with spending a month in health state $X_k$. Similarly, $C(X_k)$ maps state $X_k$ to a cost. The last term represents the NMB loss due to
secondary infections (Section 1.2.8), where \( E[S(X)] \) is the expected number of secondary infections generated by this patient until time \( T \) if the sequence of states over the next \( T \) months is \( X = (x, X_1, \ldots, X_T) \). We note that in order to estimate \( E[S(X)] \) we do not require the entire sequence, only how many of the months \( 1 \ldots T \) are spent in the failure state and the death state, as per the expression presented in Section 1.2.8 of this Appendix.

Let \( \pi_k(i) \) be the probability that a patient is in state \( i \) at time \( k \), i.e. \( \pi_k(i) = P[X_k = i] \). The probability distribution over the states of the model at month \( k \) is then a row vector \( \pi_k = (\pi_k(AN), \pi_k(NN), \pi_k(F), \pi_k(D)) \).

Now, suppose we are optimizing the monitoring policy for a patient with certain characteristics (initial age, sex, edu, etc.) who is in state \( x \) at the present time.

- Given initial state \( x \), from the logit regressions and mortality mapping in our original simulation model, we derive the transition probability matrix at time \( k \)

\[
P_k = \{P_k(i,j), i,j \in (AN,NN,F,D)\}, \forall \ k \in \{1,\ldots,T\} ,
\]

where

\[
P_k(i,j) = P[X_k = j \mid X_{k-1} = i]
\]

- We then derive the sequence of state distributions \( \{\pi_k, k \in \{1, \ldots, T\}\} \) over the next \( T \) months by the multiplication \( \pi_{k-1} \times P_k = \pi_k \).

- We calculate the expected health utility and costs incurred in any month \( k \), where \( k \in \{1, \ldots, T\} \)

\[
E[Q(X_k)] = \sum_{i \in \{AN,NN,F,D\}} \pi_k(i) \cdot Q(i)
\]

\[
E[C(X_k)] = \sum_{i \in \{AN,NN,F,D\}} \pi_k(i) \cdot C(i)
\]
Summing the above terms from \( k = 1 \) to \( T \), we obtain the total expected cost and health utility associated with the next \( T \) months.

The last term that we need to compute is the expected number of secondary infections \( (E[S(X)]) \). If among the next \( T \) months, the patient spends \( f \) months in failure and \( n \) months alive and not in failure, as described in Appendix Section 1.2.8, the expected number of secondary infections is

\[
S(f, n) = m(1 - \pi) \left(1 - (1 - \alpha_f)^{n_{sex}f}(1 - \alpha_n)^{n_{sex}n}\right).
\]

This non-Markovian dependency on \( f \) and \( n \) renders the calculation of \( E[S(X)] \) more involved than that of the costs and QALYs.

We proceed to describe the calculation of this expectation, by listing the four possible scenarios involving virologic failure and death in the next \( T \) months and the probabilities with which they appear.

a. The patient remains alive and not in failure by monitoring time \( T \), in which case the expected number of secondary infections is

\[
S(0, T) \cdot \left(\pi_T(AN) + \pi_T(NN)\right)
\]

b. The patient dies at some time \( k \) before \( T \) without having reached virologic failure. Then the number of secondary infections is

\[
\sum_{k=2}^{T} S(0, k - 1) \left(\pi_{k-1}(AN)P_k(AN, D) + \pi_{k-1}(NN)P_k(NN, D)\right)
\]

c. The patient remains alive but falls into virologic failure at a time \( k \) prior to the visit at time \( T \), in which case we have

\[
\sum_{k=2}^{T+1} S(T + 1 - k, k - 1) \left(\pi_{k-1}(AN)P_k(AN, F) + \pi_{k-1}(NN)P_k(NN, F)\right) \prod_{j=k+1}^{T} P_j(F, F)
\]

d. The patient may experience virologic failure at some time \( k \) and die at a time
$j > k$ before time $T$. Similarly as before, we have

$$
\sum_{k=2}^{T} \sum_{j=k+1}^{T+1} S(j - k - 1, k - 1) \left( \pi_{k-1}(AN) P_k(AN, F) + \pi_{k-1}(NN) P_k(NN, F) \right) \prod_{l=k+1}^{j-1} P_l(F, F) P_j(F, D)
$$

Summing up the expressions in cases (a) through (d) we obtain $E[S(X)]$, the expected number of secondary infections occurring from the current patient if monitoring occurs at time $T$. We now have all components in the expression of $\text{NMB}(T)$ (*), and can now find the $T^*$ that maximizes $\text{NMB}(T)/T$.

A patient on first-line regimen whose VL is currently undetectable will therefore be called for the next VL test in $T^*$ months, as determined by the stochastic optimization algorithm just described. If at the time of the next test her VL is undetectable, we compute the next optimized $T'$ interval and ask the patient to come for her next test then. If, on the other hand, his or her VL is elevated, we ask the patient to follow-up in a month. If his or her VL is still elevated at that time, we switch him/her to second-line regimen and from then on perform VL tests once every 12 months. If at the time of follow-up the patient’s VL is undetectable (failure has spontaneously resolved), we again use our algorithm to compute the time of the next VL test.

2. **Model validation**

Our goal in developing this model is to inform monitoring decisions in resource-limited settings where monitoring tests constitute a significant cost burden. Given that our data set is collected in resource-rich settings (Switzerland), a natural concern is the applicability of our model to resource-limited settings of interest. To evaluate this, we used our model to simulate a randomized clinical trial [5] performed in Uganda between 2003 and 2004, where 3321 ART-naïve patients with CD4 counts below 200 cells/ mm$^3$
were randomized to two monitoring arms: laboratory and clinical monitoring (LCM), where patients’ CD4 counts were measured every 3 months, and clinically-driven monitoring (CDM), where CD4 counts were measured only if the patient experienced a WHO stage 4 opportunistic infection (OI) or grade 4 toxicities. Participants switched to second-line ART after new or recurrent WHO stage 4 events in both groups, or CD4 count less than 100 cells per µL (LCM only). We note that, while CD4 cell counts in the calibration stage were initialized to match the baseline CD4 cell counts reported in this clinical trial (< 200 cells/mm$^3$), the baseline CD4 cell counts in the simulation model reflected the higher values from Uganda’s 2011 nationally representative Demographic and Health Survey (average CD4 cells count of 331 cells/mm$^3$) [16].

Table A6 compares the main outcomes reported by the DART trial (in yellow) to our simulation model. As shown, our model closely matches key quantities, usually within the 95% confidence interval. We thus believe that our model can be reliably used in simulating monitoring policies in Uganda.

3. **Additional Sensitivity Analyses**

We performed additional one-way sensitivity analyses to evaluate the robustness of our results with respect to some of our assumptions and changes in parameter values. Figures A2 - A9 display the resulting cost-effectiveness plots. In each analysis, we kept the adaptive monitoring policies the same as in the base case analysis, and re-ran the simulation model with the modified parameter or assumption. This was done in order to estimate the effect of a miss-specified model used in optimizing the adaptive policies.
While the relative benefits of adaptive versus fixed monitoring policies diminished in some cases, we generally observed that the relative ordering of the policies remained the same as in the base-case analysis (most adaptive policies were found to lie on or very close to the cost-effectiveness frontier), suggesting that our results are not very sensitive to perturbations in the tested assumptions.

3.1. Cost-effectiveness results without secondary infections adjustments

When adjustments in costs and QALYs for secondary transmission during simulation are removed (Figure A2), we observe that the relative performance of the policies remained the same as in the base-case analysis, but the QALY and cost values became closer to one another. This is expected, as removing the adjustments also removes some of the benefits of the policies that monitor relatively more frequently.

3.2. Varying discount rate and time horizon considered

Running our simulation model with a discount factor of 5% while keeping the adaptive policies optimized for a 3% discount rate (Figure A3) diminished the benefits from using the adaptive policies relative to the fixed interval policies. However, adaptive policies still remained on the cost-effectiveness frontier, performing similarly to fixed interval policies at low CETs and outperforming (cost-wise) fixed policies at high CETs.

Increasing the simulation horizon to 20 years improved the relative performance of adaptive policies (Figure A4). This can be explained by the fact that the benefits, including in terms of secondary transmission, from early identification of VL failure are long-term, and increasing the simulation horizon allows us to tabulate more of these benefits.
3.3. Varying parameter values for adherence and education, and time to follow-up upon detection

Our results were also robust to perturbations in adherence coefficient values (Figures A5 and A6) and education coefficient values (Figures A7 and A8), as well as in increasing the time to follow-up upon detection of VL failure from one to three months (thus allowing patients in whom re-suppression does not spontaneously occur to remain in VL failure for an additional two months). Table A7 contains the relevant coefficient information.
References

1. Pujades-Rodriguez M, Balkan S, Arnould L, Brinkhof MW, Calmy A, f AWGoM. Treatment failure and mortality factors in patients receiving second-line hiv therapy in resource-limited countries. JAMA. 2010;304(3):303-12.

2. Cohen MS, Chen YQ, McCauley M, Gamble T, Hosseinipour MC, Kumarasamy N, et al. Prevention of HIV-1 Infection with Early Antiretroviral Therapy. New England Journal of Medicine. 2011;365(6):493-505.

3. Mermin J, Lule J, Ekwaru JP, Malamba S, Downing R, Ransom R, et al. Effect of co-trimoxazole prophylaxis on morbidity, mortality, CD4-cell count, and viral load in HIV infection in rural Uganda. The Lancet. 2004;364(9443):1428-34.

4. Holmes CB, Wood R, Badri M, Zilber S, Wang B, Maartens G, et al. CD4 Decline and Incidence of Opportunistic Infections in Cape Town, South Africa: Implications for Prophylaxis and Treatment. JAIDS Journal of Acquired Immune Deficiency Syndromes. 2006;42(4):464-9.

5. Dart Trial Team. Routine versus clinically driven laboratory monitoring of HIV antiretroviral therapy in Africa (DART): a randomised non-inferiority trial. Lancet. 2010;375(9709):123-31.

6. Robberstad B, Olsen JA. The health related quality of life of people living with HIV/AIDS in sub-Saharan Africa - a literature review and focus group study. Cost Effectiveness and Resource Allocation : C/E. 2010;8:5-.
7. Pitter C, Kahn JG, Marseille E, Lule JR, McFarland DA, Ekwaru JP, et al. Cost-Effectiveness of Cotrimoxazole Prophylaxis Among Persons With HIV in Uganda. JAIDS Journal of Acquired Immune Deficiency Syndromes. 2007;44(3):336-43.

8. Besamusca J TK, Kahyarara G, Tingum EN. Wages in Uganda - WageIndicator survey 2012 2012 2012. Available from: http://www.wageindicator.org/documents/publicationslist/publications-2012/AIAS_WI_countryreports_f2f_report_Uganda_final_20121211.pdf.

9. Siedner MJ, Lankowski A, Tsai AC, Muzoora C, Martin JN, Hunt PW, et al. GPS-measured distance to clinic, but not self-reported transportation factors, are associated with missed HIV clinic visits in rural Uganda. AIDS (London, England). 2013;27(9):1503-8.

10. Hyle EP, Jani IV, Lehe J, Su AE, Wood R, Quevedo J, et al. The Clinical and Economic Impact of Point-of-Care CD4 Testing in Mozambique and Other Resource-Limited Settings: A Cost-Effectiveness Analysis. PLoS Med. 2014;11(9):e1001725. PubMed PMID: PMC4165752.

11. Initiative CHA. 2015 CHAI ARV Reference Price List: Clinton Health Access Initiative 2015. Available from: http://www.clintonhealthaccess.org/2015-chai-arv-reference-price-list-2/.

12. Pinkerton SD, Holtgrave DR, Leviton LC, Wagstaff DA, Abramson PR. Model-based evaluation of HIV prevention interventions. Evaluation Review. 1998;22(2):155-74.
13. Wawer MJ, Gray RH, Sewankambo NK, Serwadda D, Li X, Laeyendecker O, et al. Rates of HIV-1 Transmission per Coital Act, by Stage of HIV-1 Infection, in Rakai, Uganda. Journal of Infectious Diseases. 2005;191(9):1403-9.

14. Hughes JP, Baeten JM, Lingappa JR, Magaret AS, Wald A, de Bruyn G, et al. Determinants of Per-Coital-Act HIV-1 Infectivity Among African HIV-1–Serodiscordant Couples. The Journal of Infectious Diseases. 2012;205(3):358-65.

15. Gray RH, Wawer MJ, Brookmeyer R, Sewankambo NK, Serwadda D, Wabwire-Mangen F, et al. Probability of HIV-1 transmission per coital act in monogamous, heterosexual, HIV-1-discordant couples in Rakai, Uganda. The Lancet. 2001;357(9263):1149-53.

16. UDHS I. Uganda demographic and health survey. Uganda Bureau of Statistics, Kampala Uganda 2011, Uganda Bureau of Statistics (UBOS) and ICF International Inc. 2012.
Figure A1: Example Diagram of Markov Chain Model
Figure A2: Cost-effectiveness results without secondary transmission adjustments

![Cost Effectiveness Plot](image)

- Monitor every 24 months
- Monitor every 12 months
- Monitor every 6 months
- Monitor every 3 months
- Monitor every month

Costs

QALYs

Policies
- Fixed
- Adaptive
Figure A3: Cost-effectiveness results with 5% discount rate

Cost Effectiveness Plot
Figure A4: Cost-effectiveness results when simulation time is extended to 20 years

Cost Effectiveness Plot
Figure A5: Cost-effectiveness results when varying adherence coefficient in simulation model to upper 95% CI value

Cost Effectiveness Plot

Costs

- Monitor every 24 months
- Monitor every 6 months
- Monitor every 12 months
- Monitor every 3 months
- Monitor every month

QALYs

- Adaptive 1
- Adaptive 3
- Adaptive 10
- Adaptive 30
- Adaptive 50
- Adaptive 100

Policies

- Fixed
- Adaptive
Figure A6: Cost-effectiveness results when varying adherence coefficient in simulation model to lower 95% CI value

Cost Effectiveness Plot
Figure A7: Cost-effectiveness results when varying education coefficient in simulation model to upper 95% CI value

Cost Effectiveness Plot
Figure A8: Cost-effectiveness results when varying education coefficient in simulation model to lower 95% CI value

Cost Effectiveness Plot
Figure A9: Cost-effectiveness results when follow-up time upon detection is increased from 1 month to 3 months
Table A1: CD4 cell count and VL failure - specific mortality[2, 3]

| CD4 count (cells/mm³) | Annual mortality rate if no virologic failure | Annual mortality rate if virologic failure |
|-----------------------|---------------------------------------------|-------------------------------------------|
| 0-50                  | 0.0673                                      | 0.3880                                    |
| 50-100                | 0.0368                                      | 0.1740                                    |
| 100-150               | 0.0275                                      | 0.0550                                    |
| 150-250               | 0.0202                                      | 0.0404                                    |
| 250-450               | 0.0191                                      | 0.0382                                    |
| 450-500               | 0.0191                                      | 0.0298                                    |
| >500                  | 0.0191                                      | 0.0191                                    |
Table A2: Opportunistic infection rates [4, 5]

| CD4 count (cells/ mm$^3$) | Annual rate of OI |
|---------------------------|-------------------|
| <50                       | 0.9916            |
| 50 – 200                  | 0.0768            |
| 200-350                   | 0.0228            |
| >350                      | 0.0037            |
Table A3: Health Utilities[6, 7]

| CD4 count (cells/mm³) | Health Utility (yearly) |
|-----------------------|-------------------------|
| <50                   | 0.495                   |
| 50-100                | 0.526                   |
| 100-200               | 0.588                   |
| 200-350               | 0.680                   |
| 350-500               | 0.772                   |
| >500                  | 0.865                   |
Table A4: Costs used.[8-11]

| Parameter                                      | Value (US$) |
|-----------------------------------------------|-------------|
| Cost of viral load test                       | 30          |
| Yearly health cost if CD4 cells < 250 cells/ mm³ | 250         |
| Yearly health cost if CD4 cells > 250 cells/ mm³ | 160         |
| Yearly cost of first-line regimen              | 120         |
| Yearly cost of second-line regimen             | 345         |
Table A5: Secondary infection parameters[13-15].

| Parameter | Value |
|-----------|-------|
| $\alpha_l$ | 0.002 |
| $\alpha_k$ | 0.0001 |
| $n_{sex}$ | 9 |
| $m$ | 1.2 |
| $\pi$ | 0.073 |
Table A6: Modeling the DART clinical trial [5]

| Outcome                                      | DART Trial  | Simulation |
|----------------------------------------------|-------------|------------|
| 5-year survival on CDM                       | 0.87(0.85-0.88) | 0.87      |
| 5-year survival on LCM                       | 0.90(0.88-0.91) | 0.89      |
| Opportunistic Infection (OI) free survival on CDM | 0.72(0.7-0.74)  | 0.73      |
| OI-free survival on LCM                      | 0.78(0.76-0.8)  | 0.78      |
| Percentage of CD4s above 200 cells/mm³ at year 5 on CDM | 0.82         | 0.80      |
| Percentage of CD4s above 200 cells/mm³ at year 5 on LCM | 0.86         | 0.83      |
| Median CD4 for last seen alive on CDM        | 339         | 336      |
| Median CD4 for last seen alive on LCM        | 372         | 335      |

CDM- clinical driven monitoring; LCM- laboratory and clinical monitoring
Table A7: Sensitivity analysis coefficient values

| Parameters            | Base Case value | Lower 95% | Upper 95% |
|-----------------------|-----------------|-----------|-----------|
| Previous non-adherence| 1.223           | 0.852     | 1.595     |
| Educational Level     | -0.634          | -1.040    | -0.228    |