**Supplementary Appendix**

This supplementary Appendix accompanies the article Cost-Effectiveness of HIV Monitoring Strategies in Resource-Limited Settings (*Archives of Internal Medicine*, September 22, 2008), and includes a detailed discussion of the methods.

**Methods**

**Overview & Model Structure**

We developed a Monte Carlo simulation model of HIV disease that followed the natural history of HIV in an individual patient from the time of presentation to care until death. Assuming a modified societal perspective, the model followed the benefits of various treatment and monitoring interventions and costs over time. The main model outputs include life expectancy from the time of presentation to care and total lifetime costs in expressed in 2007 US dollars. The comparative value of the alternative strategies was expressed in terms of incremental cost effectiveness ratios of each strategy compared with the next less effective strategy.

**Model Structure**

The model follows patients in one month intervals from the time of presentation until death. Patients may present at any clinical stage of disease and with any laboratory parameters. That is, on entry, a patient may or may not be ill with an opportunistic disease (OD), and may have any CD4 or viral load. The CD4 distribution of patients at entry was taken from published literature of the Cape Town AIDS Cohort, and their risk of presenting with an opportunistic infection was dependent on their CD4 at
presentation. However, all patients were assumed to be antiretroviral-naïve, and we assumed no patients were known to have contra-indications to first-line therapy such as resistance to non-nucleoside reverse transcriptase inhibitors.

Figure 1 shows a flow diagram of patient care processes during routine outpatient care. At presentation and at each clinic visit, patients who were not on HAART were evaluated whether they met criteria to start HAART. The criteria for choosing to start HAART depended on the monitoring strategy. Patients who were on first-line HAART also arrived to the clinic in regular intervals, and were evaluated for continued efficacy and possible toxicities of first-line HAART. Patients were switched from first-line HAART to second-line HAART for two reasons: confirmed or suspected virologic failure, and drug toxicity severe enough to necessitate a change in regimen.

Finally, patients on second-line HAART were also evaluated in clinic at regular intervals for signs of medication toxicity. Treatment was stopped in patients who had severe medication toxicity to second-line HAART or were otherwise unable to tolerate the medications. However, patients who were on second-line HAART and had evidence of virologic failure were maintained on second-line HAART despite proven or suspected virologic failure due to the independent survival advantages of a non-suppressive regimen compared with HAART cessation.

The model evaluated all patients in one month intervals. While the model tracks all patient parameters – including CD4, viral load, HAART regimen, medication toxicities, and development of opportunistic diseases – most parameters are only available to decision makers only during regular clinic visits. That is, while the model tracks an individual’s CD4 monthly, that patient’s CD4 is only available for treatment
decisions if it is measured and if that patient presents to clinic that month. Figure 1 shows the flow of patient care in the absence of severe opportunistic diseases.

If a patient experiences a severe opportunistic disease, the model assumes they do not present to clinic that month; instead, their risk of death rises based on their CD4, and the costs incurred reflect hospital charges rather than clinic charges. If they survive the acute illness, they return to routine care as shown in Figure 1. Patients were followed until death from HIV or another cause.5-8 Thus, we were able to follow the lifetime costs and benefits of HIV care delivery for a cohort of simulated patients using clinical and utilization data of cohorts cared for in the Cape Town region of South Africa.

**Disease Progression**

We followed the disease progression of patients from the time of presentation based on the following parameters: age, CD4, HIV RNA viral load, HAART regimen, HAART duration, history of opportunistic diseases, virologic failure, and medication toxicity. We monitored all parameters for all patients monthly, but only a subset of them were available to decision-makers and providers at any point.

Upon entry to care, each patient was assigned an initial CD4, viral load, and age from a distribution that was calibrated to generate a population similar to the two study cohorts.1,9-11 Each patient’s immunologic status and risk of clinical events were determined primarily by their CD4 count. The CD4 count was modeled as a continuous variable that varied based on the viral load, HAART, and treatment failure. While the patient’s viral load was not suppressed, their CD4 declined at a rate that depended on their viral load.12 Given the controversial relationship between the viral load and CD4 change, we allowed two non-linear determinants of CD4 decline: random variability
around the regression line, and a slower rate of decline with lower CD4, both guided by recent published data.12-14

Once a patient was started on a successful first-line HAART regimen, their CD4 rose at a rate that depended primarily on their CD4 at the time of treatment initiation. While there is some data to support an age-related effect of CD4 rise, the strongest reproducible predictor of CD4 rise on effective HAART is the CD4 at the time of treatment initiation.15-18 Published data on CD4 rise were extracted using the graph digitizing program DigitizeIt v.1.5 (Braunschweig, Germany), and monthly CD4 increments were determined based on baseline CD4 at treatment initiation.

The principal activity of HAART is suppression of viral replication, and we used viral suppression to undetectable levels as the principal marker that allowed CD4 to rise after treatment initiation. Since the rates of viral resistance in treatment-naïve individuals in sub-Saharan Africa are low, we assumed all patients had the potential to respond to first-line therapy.19 While on successful treatment, viral load was undetectable at a threshold of less than 400 copies/ml.

Treatment failure was reflected in failure to suppress virologic replication and a gradual return of the viral load to detectable levels. Adherence to a treatment regimen and development of virologic resistance, two principal reasons for treatment failure, are closely related to one another. As a result, we used data on overall rates of virologic failure without separating adherence and resistance.11, 20, 21 Patients with virologic failure who were continued on HAART had a lower viral “set point,” and their rate of CD4 decline was slower.4 Virologic failure was either observed in patients who had viral load monitoring, or inferred in patients with CD4 or symptom-based monitoring. Without
viral load monitoring, detection of treatment failure was either delayed (for example, assuming CD4 less than half the highest measured CD4 is an indication of treatment failure takes longer than direct measurement) or made erroneously (for example, a second or third opportunistic disease suggests treatment failure in the symptom-based strategies, but may occur without treatment failure).

The risk of opportunistic diseases was dependent on the current CD4. We calculated the risk of a severe opportunistic disease based on the risk of developing a World Health Organization Stage 4 disease plus the risk of pulmonary TB as opportunistic diseases.\textsuperscript{1,2} We assumed patients received co-trimoxazole prophylaxis, which reflected the clinical data in the South African cohorts.

**Treatment Options**

Consistent with World Health Organization guidelines and with standard practice in many parts of sub-Saharan Africa, we modeled two lines of antiretroviral regimens. First line HAART included two older nucleoside reverse transcriptase inhibitors and one non-nucleoside reverse transcriptase inhibitors, either nevirapine or efavirenz. Second line HAART included a blend of newer and older nucleoside reverse transcriptase inhibitors and a protease inhibitor.\textsuperscript{22} Switching from first to second line therapy occurred due to treatment failure, as detailed above, or due to medication toxicities. Examples of severe drug toxicities include lactic acidosis, life-threatening anemia, and hypersensitivity reactions.

The likelihood of developing a severe medication-related toxicity was highest in the first five months of treatment, when more than half of medication-related toxicities related occur.\textsuperscript{23} Severe medication toxicities were considered as separate clinical events,
and patients had their regimen changed at the time of the event rather at the next clinic visit. Severe toxicities include events such as severe anemia with zidovudine, severe neuropathy with stavudine, or a rash with nevirapine or efavirenz.\textsuperscript{23} We used the survival curve of “Severe or dose-modifying toxicity” from ACTG 384 in order to generate the relative month-to-month risk of severe toxicity. We then calibrated the rates to reflect rates observed in resource-limited settings.\textsuperscript{23-25} In developing countries, individuals switch from first-line regimens after a relatively short duration of time, and we used data from developing countries only to obtain the relative rate of decay. To avoid counting drug substitution within the same regimen (e.g. zidovudine for stavudine after neuropathy), we used rates of switching entire drug classes.\textsuperscript{24, 25} Unlike in the US and Europe where several lines of HAART are available beyond the second line, there are virtually no options for people who are unable to tolerate second-line HAART. Inability to tolerate second-line HAART was the only reason for discontinuation of second-line treatment.

**Monitoring Strategies**

We examined three basic types of monitoring strategies: two symptom-based strategies, four CD4-based strategies, and four CD4-viral load strategies. The symptom-based strategies differed in the threshold of opportunistic diseases used to determine treatment failure (second or third opportunistic disease) and switch HAART. The CD4 and CD4-viral load were each examined when HAART was started at 200 or 350 cells/µl and when laboratory values were monitored every 3 or 6 months, bringing the total number of strategies to 10 (four each of the CD4 and CD4-viral load strategies and a symptom-based strategy).
All patients were monitored clinically. This included regular clinic visits for monitoring development of opportunistic diseases, measurement of laboratory data, and, once on HAART, detecting medication toxicities. In the symptom-based strategies, HAART was started when a patient developed the first severe opportunistic disease, and the first-line regimen was changed if that patient developed medication toxicity or if the patient developed a second (or third) severe opportunistic disease after starting HAART. If the patient developed toxicity to second-line medication as well, HAART was discontinued altogether.

In the CD4 strategies, patients had their CD4 monitored at regular intervals (every 3 or 6 months) starting at presentation. The CD4 count was used for two reasons: for determining HAART initiation, and for timing change from first-line to second-line therapy. We examined two thresholds for HAART initiation (200 cells/µl and 350 cells/µl), and used the observed CD4 (i.e. the patient’s CD4 at the time of a clinic visit) as the trigger for starting HAART. When monitoring CD4 alone, first-line HAART was changed when the measured CD4 dropped to below half of the highest measured CD4 in the patient’s record after starting HAART. We tested two other strategies mentioned in the literature for HAART regimen change using CD4 alone.\textsuperscript{26-28} In one strategy, HAART was changed when the CD4 dropped below the level measured at the time of first-line initiation. The second alternative employed the CD4 slope as a surrogate for virologic failure, and HAART regimen was switched after four consecutive drops in the CD4. While there are no good studies for using CD4 slope as a surrogate for virologic failure, it is used in some regions of Africa where CD4 is available and viral load is not. Additionally, since we allowed an imperfect relationship between viral load and CD4
change, observed CD4s could have a positive or flat slope despite virologic failure, or a negative slope in the absence of virologic failure, further confounding the decision to change regimens.

Patients had their CD4 monitored even after moving to second-line HAART even though no further HAART options were available. This was done for two reasons: first, most places with access to CD4 counts never stop monitoring CD4, since there are additional benefits of knowing a recent CD4 level such as a differential diagnosis of opportunistic infections; and second, we assumed patients were placed on co-trimoxazole prophylaxis using CD4 levels as one of the initiation criteria.

Finally, in the CD4-viral load strategy, patients had both parameters monitored regularly. Viral load was not used to determine timing of HAART initiation, mostly because the guidelines for resource-limited regions recognize the relative scarcity of viral load monitoring. Consequently, it was monitored only after HAART was started, and monitoring was stopped after a patient switched to second line therapy, since there were no further decisions based on the viral load level. CD4 was measured from the time of presentation until death, as in the CD4 strategy. After initiation of first-line HAART, an observed rise in the viral load to greater than 1,000 copies was considered virologic failure and triggered a switch to second-line therapy.

**Benefits and Costs**

All benefits were measures in life expectancy from the time of presentation. We measured both discounted and undiscounted life years and compared the gain in life expectancy among the various strategies.
We considered direct costs of care from a societal perspective in this study. We broke down the costs of care into inpatient costs, outpatient costs, cost of HAART, and costs of laboratory testing. Inpatient and outpatient clinic costs were taken from a detailed study of tertiary and district hospitals and community clinics in the Cape Town area. Inpatient and outpatient care included costs of non-clinical staff such as counselors, clinical staff, non-HAART medicines, and capital costs.

Costs of HAART were estimated from South African studies as well as Médecins Sans Frontières. For second line therapy, we used a weighted average of the costs of second-line medications and their frequency of use. Finally, the cost of testing was reflected in a single per-test cost that included the cost of reagents, labor, parts, data management, maintenance, and the cost of CD4 or viral load enumeration equipment. The up-front fixed cost of purchasing diagnostic equipment may be a substantial impediment to implementation, and we considered several approaches to calculating this cost, including rental, acquisition, and amortization schedules. Over the lifetime of infected individuals, the overall testing costs did not vary substantially among the different calculation methods, and in our base case analysis we include only a single per-test cost.

**Sensitivity Analysis**

Our sensitivity analysis included several one-way analyses, second-order simulations, and changes in the monitoring strategies. We varied uncertain clinical parameters such as the relationship between viral load and CD4 change, rates of virologic failure, the risk of death from opportunistic diseases, and rates of medication toxicity. In addition, we varied all cost and utilization inputs independently, and in second-order
sensitivity analysis we varied all costs simultaneously using probability distributions, looking for changes in our estimate of the incremental cost-effectiveness ratio in settings where costs differ from our cost estimates.

Finally, we varied the algorithms used to determine when to switch from first-line HAART to second-line HAART when CD4 alone is monitored, as mentioned above.

**Model Validation**

We validated the outcomes of the model by comparing model predictions to observed rates of development of severe opportunistic diseases. As well, we calibrated our model against other models of HIV in resource-limited settings and regional cost studies.\(^1,28,32\) Table 1 shows a comparison of model outcomes demonstrating identical life expectancy estimates, similar rates of severe opportunistic diseases, and lifetime costs estimates about 10% lower than previous estimates.
Figure 1: Model flow of routine patient care management

Squares represent states or processes, and diamonds represent decision nodes. For example, newly diagnosed HIV+ patients are seen in clinic, and evaluated whether they meet criteria to start HAART. If they meet criteria, they are started on first-line HAART, and if they do not meet criteria, the model evaluates them again next month. The model does not show the development of acute clinical events such as severe opportunistic diseases or some medication toxicities, which may occur at any time.
Table 1: Model Validation and Calibrations

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Previous</th>
<th>Model</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Life years gained(^{19})</td>
<td>5.8</td>
<td>5.8</td>
<td>Using CD4 of 200 or a single opportunistic disease to start treatment.</td>
</tr>
<tr>
<td>Discounted lifetime costs(^{20})</td>
<td>$5,088</td>
<td>$4,552</td>
<td>Using a strategy where CD4 and viral load were checked every 6 months.</td>
</tr>
<tr>
<td>Rate of severe opportunistic diseases (per 100 person-years)(^{14})</td>
<td></td>
<td></td>
<td>Pulmonary TB and most WHO stage IV opportunistic diseases were considered severe.</td>
</tr>
<tr>
<td>&lt;50</td>
<td>133.1</td>
<td>125.5</td>
<td></td>
</tr>
<tr>
<td>51-200</td>
<td>31.7</td>
<td>31.6</td>
<td></td>
</tr>
<tr>
<td>201-350</td>
<td>13.6</td>
<td>12.9</td>
<td></td>
</tr>
<tr>
<td>&gt;350</td>
<td>3.1</td>
<td>3.4</td>
<td></td>
</tr>
</tbody>
</table>
References


