Dropout and re-enrollment: implications for epidemiological projections of treatment programs

Daniel J. Klein*, Anna Bershteyn* and Philip A. Eckhoff

Objective: EMOD-HIV v0.8 has been used to estimate the potential impact of expanding treatment guidelines to allow earlier initiation of antiretroviral therapy (ART) in sub-Saharan Africa with current or improved treatment coverage. In generating these results, a model must additionally make assumptions about the rates of dropout and re-initiation into ART programs before and after the program change, about which little is known. The objective of this work is to rigorously analyze modeling assumptions and the sensitivity of model results with respect to relevant mechanisms and parameters.

Methods: We varied key model assumptions pertaining to ART dropout and re-enrollment to analyze their effect on the cost, impact, and cost-effectiveness of expanding treatment guidelines, and of expanding coverage via improved testing and linkage to care. Additionally, we performed a sensitivity analysis of 17 relevant model parameters.

Setting: South Africa.

Results: Allowing re-initiation of ART irrespective of prior treatment doubled the cost and impact of expanding treatment guidelines, as compared with a scenario in which re-initiation could only be triggered by a health event (AIDS symptoms, diagnosis of a partner, or an antenatal care visit). Increasing the probability of ‘voluntary’ re-initiation (not triggered by a health event) was the most cost-effective way to improve the treatment program, especially in the short term because it provided immediate benefits to those who would otherwise have delayed re-initiation until the onset of AIDS symptoms. However, the maximum impact of this change was limited compared with expanding coverage through improvements in testing and linkage to care. Beyond improvements in coverage and re-initiation, further gains could be made by improving retention in care. Only with optimal retention in care was expansion of guidelines cost-saving after 20 years due to reductions in transmission. Re-initiation did not reduce transmission sufficiently to make a guideline change cost-effective due to transmission that occurred while patients were away from care. Sensitivity analysis suggested that enormous health benefits could be attained by improving treatment regimens to have higher efficacy at preventing transmission, increasing the proportion of the population with access to improved healthcare, and reducing ‘leaks’ in the ‘cascade of care.’ Increasing the proportion of individuals who receive CD4+ cell results was particularly cost-effective at baseline levels of coverage, and increasing retention on ART was particularly cost-effective with expanded coverage.

Conclusion: This analysis provides a sense of the magnitude of uncertainty in program cost and impact that policy-makers could anticipate in the face of uncertain future programmatic changes. Our findings suggest that increasing re-initiation is the most cost-effective means of initial program improvement, especially in the short term, but that improvements in retention are necessary in order to reap the full transmission-blocking benefits of a test-and-treat program in the long term.

© 2014 Wolters Kluwer Health | Lippincott Williams & Wilkins

AIDS 2014, 28 (Suppl 1):S47–S59

Keywords: antiretroviral therapy treatment guidelines, computer model, EMOD, epidemiology, HIV, sensitivity analysis
Introduction

Recent scientific findings have demonstrated the efficacy [1] and effectiveness [2] of antiretroviral therapy (ART) for preventing transmission in heterosexual partnerships. Additionally, observational evidence suggests that increases in ART coverage may have reduced HIV incidence in communities where prevalence is high in the general population [3]. Randomized controlled trials are now underway to test the community-level impact of ART on HIV incidence [4]. However, the evidence is already strong enough to demand debate about potentially expanding ART eligibility to maximize its prevention benefits.

Now more than ever before, such debates are being informed by results from mathematical modeling [5]. Mathematical models serve to combine putative mechanisms for disease transmission and/or burden with available epidemiological data, and carry these forward in time to estimate the cost and health impact of different policy and implementation scenarios.

It is informative to consider a wide range of assumptions about dropout and re-enrollment because these have not been well characterized on national scales. Owing to its flexible implementation, EMOD-HIV can be configured for a variety of retention/re-initiation assumptions. Dropout and re-initiation of ART have not been well-measured on national scales. In this study, a wide range of assumptions about dropout and re-initiation were tested a flexible agent-based epidemiological model called EMOD-HIV. We hypothesized that these assumptions could change the strategy for maximizing health benefits for a given investment of resources.

In this analysis, we will focus on key aspects of improving testing and treatment programs under current antiretroviral treatment guidelines, or under expanded guidelines that recommend treatment regardless of CD4+ cell count, referred to here as universal test-and-treat (UTT). As we shall see, our model predicts that improvements in re-enrollment of those lost to follow-up and retention of those in care are the most cost-effective strategies to incrementally improve ART programs. Guideline changes should only be considered after the benefits of re-enrollment and retention have been maximized. However, our model did not take into consideration any secondary costs incurred by improvements in retention and re-enrollment, such as patient outreach.

Similarly to retention and re-enrollment, there are other model parameters that are uncertain, yet could impact the projected cost and benefits of treatment programs. Therefore, we additionally performed a sensitivity analysis of 17 model parameters relevant to ART program implementation. Consistent with our study of dropout and re-initiation, the most sensitive parameters were the probability of linking to ART under current levels of testing/linkage, and the proportion reachable by expanded healthcare access with expanded levels testing/linkage.

Differences in these key assumptions could partially explain the variability in predictions made by different models of ART scale-up [5]. In this analysis, we will show the impact of these uncertainties on the output of a single model: EMOD-HIV. The same analysis may yield different magnitudes of impact in different models, and may not be feasible to perform in all types of models. Nonetheless, the analysis will provide a sense of the magnitude of uncertainty in program cost and impact that policy-makers could anticipate in the face of uncertain future programmatic changes.

Methods

Epidemic simulations were performed using EMOD-HIV v0.8, an individual-based stochastic model of heterosexual HIV transmission [6]. To estimate program cost, epidemiological outputs from EMOD-HIV were postprocessed with unit costs established by a working group of the HIV Modelling Consortium for ART initiation, time on ART, pre-ART care, testing, healthcare, and death, along with proportional markups for supply chain management and programmatic support [7]. All costs are reported in 2012 US dollars.

Interventions were assumed to begin the middle of 2013. For pairs of intervention scenarios, we report total undiscounted change in cost up until the middle of 2018 or 2033, total undiscounted disability-adjusted life-years (DALYs) averted up until mid-2018 or 2033, and the incremental cost-effectiveness ratio (ICER), defined as the incremental cost per incremental DALYs averted between the scenarios with a 3% annual discounting rate on both costs and DALYs. This format was chosen because policy-makers must often consider total, undiscounted cost when considering the affordability of programmatic choices, whereas discounting is generally used by health economists to compare the cost-effectiveness of different strategies. Variations on this analysis, as well as additional epidemiological outputs stratified by most demographic or clinical categories, are available from the authors on request.

Modeling methods

EMOD-HIV v0.8 was developed by the EMOD-HIV v0.8 was developed by the Institute for Disease Modeling at Intellectual Ventures Laboratory. An earlier version of the model, v0.7, has been documented in detail [6]. Additional model features added to create v0.8 are described here and in the Supplementary Methods, http://links.lww.com/HJH/A291.
The most significant new feature set in v0.8 is the HIV treatment 'cascade of care,' defined as the series of health actions involved in reaching and continuing HIV care [8]. The structure of the cascade is depicted graphically in Fig. 1.

In EMOD-HIV, entry into the cascade of care begins with a positive HIV diagnostic test. Two percent of individuals receive a false negative diagnostic test [9,10] and, therefore, do not link to care. Testing occurs as a result of voluntary HIV testing and counseling, antenatal care, partner testing, or healthcare of symptomatic individuals.

Rates of voluntary testing and counseling vary by calendar year and by age, with some individuals beginning regular testing shortly after sexual debut, and others beginning later in life. The rates were set to match the self-reported proportion of individuals ever tested and tested in the last year according to national-level survey data [11,12]. When available, we used data for individuals who tested and received the test results.

Antenatal testing was assumed to occur at 14 weeks gestation and to vary by calendar year. The rates were estimated by multiplying the time-variable coverage of antenatal services by the time-variable rate of HIV testing and counseling in antenatal clinics [13–16], and adjusting as necessary to match overall self-reported testing rates [11,12], rates of ART initiation [17], and CD4⁺ cell count at ART enrolment [18,19].

Couples testing was assumed to always occur during pre-ART monitoring of a treatment-ineligible individual. Specifically, individuals who, with their doctors, choose not to initiate ART when found to have high CD4 counts are recommended to bring a regular partner for testing during a subsequent follow-up visit. If the regular partner is found to be HIV-uninfected, then ART may be recommended for the infected partner in order to prevent transmission. Pre-ART monitoring visits are assumed to occur every 6 months. If the individual has an active sex partner at the time of the pre-ART visit, the probability of bringing the partner for HIV testing is assumed to be 10%. If the individual has multiple partners, the longest standing partner is brought for testing. The partner receives HIV testing and counseling at this time, and thus may enter the treatment cascade via the couples testing modality. Because of low rates of linkage to pre-ART care

---

**Fig. 1.** Graphical depiction of the HIV treatment cascade implemented in EMOD-HIV v0.8. ART, antiretroviral therapy.

*Dropout and re-enrollment implications* Klein et al. S49

---

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.
and the low (10%) baseline rate of partner recruitment, partner testing is not a significant source of ART initiations at baseline.

Finally, individuals receive an HIV test when symptomatic, which is assumed to occur at a CD4\(^+\) cell count below 200 cells/\(\mu l\), chosen randomly between 200 and 100 cells/\(\mu l\) for some individuals and at a CD4 count below 100 cells/\(\mu L\) for other individuals. The proportion who test due to AIDS-related symptoms above versus below 100 cells/\(\mu L\) was adjusted to match CD4\(^+\) cell counts at ART enrolment [18,19].

Only 80% of the population could access all four modes of testing (voluntary, antenatal, couples, and symptomatic). This was the same 80% of the population that, per a standardized assumption across all models, was deemed to be accessible by improvements in testing and linkage rates in the expanded access scenarios. The remaining 20% received only antenatal and AIDS-symptomatic testing, and were not affected by healthcare improvements in the expanded access scenarios.

After receiving a positive HIV test, a proportion of individuals return for a CD4\(^+\) cell result and determination of ART eligibility. The probability of doing so is 59% at baseline [8], increasing to 100% after mid-2013 in scenarios with expanded linkage. Changes in guidelines were assumed to have no effect on the proportion retained in this stage of the cascade, even in scenarios in which eligibility was independent of CD4\(^+\) cell count. EMOD accounts for national guideline changes in 2004, 2010, and 2011, as well as a possible guideline change in the middle of 2013 depending on the scenario.

Eligible individuals may link to ART, whereas ineligible individuals may link to pre-ART. The probabilities of linking to pre-ART and ART are independently configurable and time-variable. They were adjusted along with other parameters to match the total number enrolled in ART over time [17] and the CD4\(^+\) cell counts at ART initiation [18,19]. The probability of returning for each subsequent 6-monthly pre-ART monitoring visit was assumed to be the same as the pre-ART linkage probability. Retention time on ART is exponentially distributed with a mean of 10 years, giving a dropout probability of 9.5% in the first year. In scenarios with expanded testing and linkage, the linkage rates became 100% and the dropout rate became 0% for subset of the population (80%) reachable by the healthcare expansion, although we explored alternatives to this assumption.

After dropout, a proportion of individuals may re-initiate ART by means of a new voluntary test. This proportion was set to 50% for the model comparison, and is one of the key parameters varied in this analysis. This proportion is applied independently for each dropout, regardless of whether an individual has previously dropped out and reinitiated ART. Reinitiation via symptomatic or antenatal testing is always permitted.

Additional changes in v0.8 include an altered fertility rate according to HIV status, changes in prevention of mother-to-child transmission (PMTCT) options over time and their respective transmission probabilities, and historical changes in guidelines over time for adults and children. These changes, along with the time-variable parameter values for the cascade and a list of key model inputs and outputs, are described in the Supplementary Methods, http://links.lww.com/HJH/A291. Unit and overhead costs used for postprocessing were provided by a working group of the HIV Modelling Consortium [7].

After incorporation of these features into v0.8, the model was calibrated to HIV prevalence by gender, age, and year; proportion ever tested and recently tested by gender, age, year; number on ART by gender, age, and year; CD4\(^+\) cell count at ART enrollment by gender, and total population by year. Table S1, http://links.lww.com/HJH/A291 lists the data sources used. A subset of these data, indicated with a * in Table S1, http://links.lww.com/HJH/A291, was provided by the HIV Modelling Consortium, and calibration to this subset of data was a requirement for model participation in the comparison study [7].

Quality of fit to these data was combined into a single likelihood score. Calibration was performed with iterative rounds of Incremental Mixture Importance Sampling [20], a Bayesian technique that returns samples from the posterior distribution given a likelihood function and a prior distribution for the relevant model parameters. After calibration of the preintervention baseline, we used the parameter configuration yielding the maximum a posteriori probability to simulate the scenarios.

**Variation of structural and parametric assumptions**

Our first set of analyses focused on structural model assumptions about the effect of improved healthcare access on retention and re-initiation rates on ART. These assumptions were not standardized across models, making it particularly interesting to systematically vary them within a single model.

For all scenarios, we changed the percentage of individuals who voluntarily re-initiate ART (i.e., without a health event such as antenatal visits, AIDS symptoms, or diagnosis of a partner) from 50 (the value used in the model comparison) to 0 or 100 beginning in the middle of 2013. Pre-2013 re-initiation probabilities remained at 50% to maintain the baseline model calibration.

For scenarios with expanded testing and linkage rates, we examined the effect of additionally expanding retention
to 100% for those reachable by the healthcare expansion (as we assumed for the model comparison) to no change in dropout rates (as was assumed by multiple other models in the comparison).

For a more thorough investigation of the impact of our assumptions about the HIV treatment cascade, we then systematically varied 17 other parameters related to ART initiation, linkage, retention, and effectiveness at viral suppression. These parameters were modified at the beginning of the simulated ART program in 2004, unlike the retention/re-initiation assumptions that were modified in 2013 as part of the intervention scenario definition. Nonetheless, sensitive parameters are relevant when integrating results from several models, and are useful in identifying data gaps.

Large changes in these parameter values would alter the model calibration. For instance, changing testing or linkage rates would alter the number of individuals receiving ART over time, to which the model was calibrated. To circumvent this problem, we did not analyze extreme values of the parameters as with linkage or retention. Instead, we made small variations around the calibrated parameter value and estimated the incremental change in the impact of the intervention scenarios (incremental costs, DALYs averted, and ICER). Thus, the ‘one-dimensional’ effect of changing each parameter in isolation was captured in the form of the local slope of costs and DALYs round the calibrated value. ICER was estimated as the ratio of the change in costs over the change in DALYs. We chose not to offset the effect of a changing parameter by modifying other parameters in order to maintain calibration, because this approach is dependent on the choice of offsetting parameters, and affords less intuitive insight into the structure of the model.

We used Latin hypercube sampling [21,22] to choose 350 parameter values between a minimum and a maximum value for each parameter. Reference simulations were run at the calibrated ‘nominal value’. Detailed definitions and ranges of perturbation for each parameter are listed in Table 1.

Because EMOD-HIV is stochastic, it includes random variability in the trajectory of an epidemic, especially early on when the number of HIV cases is relatively small. For example, due to the stochastic nature of partnership formation and disease transmission, some simulation runs have higher HIV incidence than others, leading to greater HIV prevalence and numbers on ART.

Changes related to the treatment cascade only modify the trajectory of a simulation after ART becomes available in 2004. Even when examining post-2004 simulation outcomes, simulations that shared the same pre-2004 trajectory tend to exhibit less variability than those that had stochastically different pre-2004 trajectories. This reduction in stochastic ‘noise’ enabled analysis of very small perturbations in parameter values, which otherwise would have been overwhelmed by the early randomness in the epidemic.

We, therefore, repeated each sensitivity analysis four times, using four independent pre-2004 trajectories, each with 350 variations on the parameter, for a total of 1400 simulation runs per parameter. The slope of incremental cost and DALYs averted as a function of each parameter was estimated using ordinary least-squares linear regression. The mean of these four slopes was reported as the parameter sensitivity, and the coefficient of variation was used to ensure that the measured sensitivity was consistent regardless of the pre-2004 epidemic trajectory.

**Results**

**Stochastic variability in incremental costs and DALYs averted**

Program cost and impact after 5 or 20 years can be seen in Fig. 2 and Table 2 with expanded guidelines and/or healthcare access, and under different assumptions about retention and re-initiation of ART. For each variation of assumptions, Fig. 2 shows the mean outcome, 50 individual simulation results, and an ellipse of 1.5 standard deviations from the mean. The uncertainty captured here is due to the stochastic nature of the model, and does not include parameter uncertainty.

The uncertainty ellipses tend to be elongated along the diagonal. This diagonal elongation stems from the natural correlation between DALYs averted and costs. When more individuals receive ART due to a stochastic effect in a simulation, both DALYs averted and additional costs are increased.

**Impact of model assumptions about re-initiation of antiretroviral therapy**

Incremental cost and DALYs averted are reported relative to a baseline scenario that assumed no expansion of guidelines or healthcare access beyond present-day tends, with no change in retention and a continuation of 50% voluntary re-initiation of ART after dropout. Decreasing the probability of voluntary re-initiation from 50 to 0% had a negative impact on cost and DALY's, as expected: after 20 years, 1064 thousand fewer DALY's were averted with a $995 million decline in program cost. Likewise, increasing re-initiation to 100% averted 1132 thousand additional DALY's with $879 million additional cost.

Increasing re-initiation was the most cost-effective way to improve the program, especially in the short term, because it provided immediate benefits to those who
Table 1. Parameters varied in sensitivity analysis.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Significance</th>
<th>Unit</th>
<th>Base</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean ART duration</td>
<td>The mean of an exponential distribution that determines how long individuals stay on ART before dropping out.</td>
<td>day</td>
<td>3650</td>
<td>3102</td>
<td>4197</td>
</tr>
<tr>
<td>Resume voluntary testing after ART dropout</td>
<td>The fraction of individuals who uptake voluntary testing after dropping out of ART.</td>
<td>%</td>
<td>0.50</td>
<td>0.43</td>
<td>0.58</td>
</tr>
<tr>
<td>Receive CD4+ cell result baseline</td>
<td>The fraction of individuals at baseline who receive a CD4+ cell staging result after testing positive for HIV.</td>
<td>%</td>
<td>0.59</td>
<td>0.50</td>
<td>0.68</td>
</tr>
<tr>
<td>Receive CD4+ cell result expanded</td>
<td>The fraction of healthcare-accessible individuals who receive a CD4+ cell staging result after testing positive for HIV in a prioritized (PE) or universal (UE) scenario.</td>
<td>%</td>
<td>1.00</td>
<td>0.70</td>
<td>1.00</td>
</tr>
<tr>
<td>Reachable by healthcare system</td>
<td>The fraction of the population who receive increased testing and linkage rates in PE and UE simulation scenarios.</td>
<td>%</td>
<td>0.80</td>
<td>0.68</td>
<td>0.92</td>
</tr>
<tr>
<td>Return for next monitoring test</td>
<td>The probability that an individual returns for each subsequent pre-ART diagnostic test, which are 6 months apart.</td>
<td>%</td>
<td>0.46</td>
<td>0.39</td>
<td>0.53</td>
</tr>
<tr>
<td>Relationship duration considered stable</td>
<td>The number of days a relationship must endure before the participants are eligible for ART under expanded guidelines.</td>
<td>day</td>
<td>365</td>
<td>310</td>
<td>420</td>
</tr>
<tr>
<td>Partner testing at baseline</td>
<td>The fraction of pre-ART patients having a stable partner that brings their partner for testing,</td>
<td>%</td>
<td>0.10</td>
<td>0.00</td>
<td>0.20</td>
</tr>
<tr>
<td>Antenatal testing</td>
<td>A multiplier of each value in a time-varying probability. The probability values, (0, 0.07, 0.58, 0.85), are linearly interpolated in time, with values at years (2000, 2001, 2002, 2006).</td>
<td></td>
<td>1.00</td>
<td>0.85</td>
<td>1.15</td>
</tr>
<tr>
<td>Pre-ART linking ramp max</td>
<td>Pre-ART linking varies sigmoidally in time, and this parameter governs the maximum value.</td>
<td>%</td>
<td>0.46</td>
<td>0.39</td>
<td>0.53</td>
</tr>
<tr>
<td>ART linking Ramp Max</td>
<td>ART linking varies sigmoidally in time, and this parameter governs the maximum value.</td>
<td>%</td>
<td>0.90</td>
<td>0.77</td>
<td>1.00</td>
</tr>
<tr>
<td>Diagnostic testing at debut ramp max</td>
<td>At debut, a sigmoidally-varying-in-time fraction of the population begins voluntary testing. This parameter governs the maximum value.</td>
<td>%</td>
<td>0.25</td>
<td>0.21</td>
<td>0.29</td>
</tr>
<tr>
<td>Diagnostic testing postdebut</td>
<td>Each year, a fraction of the postdebut population begins voluntary testing. This parameter scales probabilities (0, 0.13, 0.04, 0.07, 0.27), which are linearly interpolated corresponding to years (1998, 2000, 2003, 2006, 2009). Note that a fraction of individuals who drop out of ART will never begin/resume voluntary testing.</td>
<td>%</td>
<td>1.00</td>
<td>0.85</td>
<td>1.15</td>
</tr>
<tr>
<td>Symptomatic presentation probability</td>
<td>The probability of presenting for testing when symptomatic with AIDS.</td>
<td>%</td>
<td>1.00</td>
<td>0.70</td>
<td>1.00</td>
</tr>
<tr>
<td>Mean duration between diagnostic tests (baseline)</td>
<td>When testing voluntarily, the duration between tests is exponentially distributed. This parameter is the mean of this distribution.</td>
<td>day</td>
<td>548</td>
<td>465</td>
<td>630</td>
</tr>
</tbody>
</table>
Table 1 (continued)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Significance</th>
<th>Unit</th>
<th>Base</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean duration between diagnostic tests (expanded)</td>
<td>Like above, but used in priority (PE) and universally expanded (UE) scenarios for healthcare-accessible individuals.</td>
<td>day</td>
<td>365</td>
<td>310</td>
<td>420</td>
</tr>
<tr>
<td>ART viral suppression multiplier</td>
<td>When ART has suppressed viral load, this parameter multiplies the per-act transmission probability.</td>
<td>%</td>
<td>0.08</td>
<td>0.04</td>
<td>0.12</td>
</tr>
</tbody>
</table>

Fig. 2. Cost-effectiveness of antiretroviral treatment programs, with cost and health benefits accumulated over five years (a) and twenty years (b), computed with 3% annual discounting. Individual simulation results are plotted as points, with the means indicated by a + symbol, and uncertainty ellipses are drawn 1.5 standard deviations from the mean of the repeated stochastic model runs. Scenarios with expanded access to testing and linkage lie in the upper right of each figure. Line colors indicate the proportion of individuals able to re-enroll in antiretroviral therapy (ART) after dropout (blue: 0%, purple: 50%, green: 100%, red: dropout reduced to zero after 2013). Thick lines indicate present-day ART guidelines for South Africa, whereas thin lines indicate universal test-and-treat (UTT) guidelines. For scenarios at the lower left that project status quo levels of access, dashed lines indicate ‘patched’ UTT, in which CD4+ cell-agnostic guidelines partially improve linkage to care by removing the need to determine CD4+ cell counts prior to ART initiation.
Table 2. Model estimates of the impact, cost, and cost-effectiveness of improving different aspects of South Africa’s antiretroviral therapy program.

2a. Cumulative 5-year impact, cost, and cost-effectiveness of changing different treatment guidelines and re-initiation rates at ‘status quo’ levels of testing and linkage to care.

<table>
<thead>
<tr>
<th>Variation</th>
<th>Re-initiation</th>
<th>Retention</th>
<th>CD4⁺ cell count &lt;350</th>
<th>L-UTT</th>
<th>P-UTT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>DALYs averted (000s)</td>
<td>Costs added ($US M)</td>
<td>$/DALYs (3% disc)</td>
</tr>
<tr>
<td>0%</td>
<td>Maintain</td>
<td>0%</td>
<td>-18.6</td>
<td>-51.9</td>
<td>2820.3</td>
</tr>
<tr>
<td>50%</td>
<td>Maintain</td>
<td>50%</td>
<td>0.0</td>
<td>0.0</td>
<td>N/A</td>
</tr>
<tr>
<td>100%</td>
<td>Maintain</td>
<td>100%</td>
<td>15.2</td>
<td>40.2</td>
<td>2700.3</td>
</tr>
</tbody>
</table>

2b. Cumulative 20-year impact, cost, and cost-effectiveness of changing different treatment guidelines and re-initiation rates at ‘status quo’ levels of testing and linkage to care.

<table>
<thead>
<tr>
<th>Variation</th>
<th>Re-initiation</th>
<th>Retention</th>
<th>CD4⁺ cell count &lt;350</th>
<th>L-UTT</th>
<th>P-UTT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>DALYs averted (000s)</td>
<td>Costs added (M)</td>
<td>$/DALY (3% disc)</td>
</tr>
<tr>
<td>0%</td>
<td>Maintain</td>
<td>0%</td>
<td>-1064</td>
<td>-995</td>
<td>982</td>
</tr>
<tr>
<td>50%</td>
<td>Maintain</td>
<td>50%</td>
<td>0.0</td>
<td>0.0</td>
<td>N/A</td>
</tr>
<tr>
<td>100%</td>
<td>Maintain</td>
<td>100%</td>
<td>1132</td>
<td>879</td>
<td>809</td>
</tr>
</tbody>
</table>

2c. Cumulative 5-year impact, cost, and cost-effectiveness of changing different treatment guidelines and re-initiation rates with improvements in testing and linkage for 80% of the population.

<table>
<thead>
<tr>
<th>Variation</th>
<th>Re-initiation</th>
<th>Retention</th>
<th>CD4⁺ cell count &lt;350</th>
<th>P-UTT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>DALYs averted (000s)</td>
<td>Costs added ($US M)</td>
</tr>
<tr>
<td>0%</td>
<td>Maintain</td>
<td>0%</td>
<td>1240.7</td>
<td>4603.4</td>
</tr>
<tr>
<td>50%</td>
<td>Maintain</td>
<td>50%</td>
<td>1331.2</td>
<td>4920.5</td>
</tr>
<tr>
<td>100%</td>
<td>Maintain</td>
<td>100%</td>
<td>1435.8</td>
<td>5257.9</td>
</tr>
<tr>
<td>50%</td>
<td>Increase</td>
<td>50%</td>
<td>1604.2</td>
<td>6404.4</td>
</tr>
</tbody>
</table>

2d. Cumulative 20-year impact, cost, and cost-effectiveness of changing different treatment guidelines and re-initiation rates with improvements in testing and linkage for 80% of the population.

<table>
<thead>
<tr>
<th>Variation</th>
<th>Re-initiation</th>
<th>Retention</th>
<th>CD4⁺ cell count &lt;350</th>
<th>P-UTT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>DALYs averted (000s)</td>
<td>Costs added (M)</td>
</tr>
<tr>
<td>0%</td>
<td>Maintain</td>
<td>0%</td>
<td>9753</td>
<td>14672</td>
</tr>
<tr>
<td>50%</td>
<td>Maintain</td>
<td>50%</td>
<td>13315</td>
<td>19604</td>
</tr>
<tr>
<td>100%</td>
<td>Maintain</td>
<td>100%</td>
<td>179067</td>
<td>23508</td>
</tr>
<tr>
<td>50%</td>
<td>Increase</td>
<td>50%</td>
<td>20366</td>
<td>29582</td>
</tr>
</tbody>
</table>

L-UTT, ‘leaky’ universal test-and-treat. No change in loss to follow-up at CD4⁺ cell testing, although guidelines are CD4⁺ cell-agnostic. P-UTT, ‘patched’ universal test-and-treat. No loss to follow-up between diagnosis and determination of eligibility.
would otherwise have delayed re-initiation until the onset of AIDS symptoms. Enabling re-initiation provides individuals with additional opportunities to go through the ‘leaky cascade’ at higher CD4⁺ cell counts. With 3% annual discounting, increasing re-initiation from 50 to 100% had an ICER of 2700$/DALY after 5 years and 8088/DALY after 20 years. The cost-efficiency of voluntary re-initiation, rather than antenatal, couples, or symptom-driven re-initiation, is in part because EMOD-HIV accounts for higher mortality among patients who initiate at low CD4⁺ cell counts, due to opportunistic infections and other complications.

Cost–effectiveness strategies for improving treatment programs

The most cost-effective initial improvement to ART programs is increasing re-initiation rates for those who have previously dropped out of ART. This is represented by the thick orange ellipse near the origins of Fig. 2a (5-year horizon) and Fig. 2b (20-year horizon). Though cost-effective, this improvement is limited in its total impact compared with expansion of guidelines or increases in testing/linkage to care. A change from 0 to 100% re-initiation resulted in only 34 thousand DALYs averted after 5 years and 2.2 million DALYs averted after 20 years.

After fully taking advantage of re-initiation, the impact of a program can be expanded either by changing guidelines to UTT (moving from thick lines to thin lines in Fig. 2) or by improving testing and linkage (moving from the left-hand set of thin lines to the right-hand set of thin lines in Fig. 2). Full scale-up of testing and linkage for 80% of the population is associated with greater cost and impact than a change in treatment guidelines. Cost–effectiveness depends on the time horizon of interest: improvements in testing/linkage are more cost-effective than guideline changes over a 5-year horizon, but become equally or more cost-effective after 20 years.

Changing guidelines to UTT with no further increases in testing or linkage (and continuation of increased re-initiation rates) would cost $6.4 billion while averted 1.6 million DALYs after 5 years, and costing $29 billion while averting 19 million DALYs after 20 years (thin lines above ‘Status Quo Access’ in Fig. 2a and b, respectively). Alternatively, if guidelines remained the same while testing/linkage were maximized (universal annual testing and perfect linkage), for 80% of the population such expansion would cost $1.4 billion while averting 114,000 DALYs after 5 years, and cost $5.4 billion while averting 3.6 million DALYs after 20 years (thick lines above ‘Expanded Access in Fig. 2a and b, respectively).

After the potential benefits of re-initiation have been exhausted, the cost–effectiveness of next changing guidelines as compared with next improving testing/linkage depends on whether UTT alone has an effect on linkage rates. EMOD-HIV assumed that 41% of patients are lost from the treatment cascade due to failure to return for a CD4⁺ cell count result. However, this does not necessarily mean that CD4⁺ cell-agnostic treatment guidelines would ‘patch’ this ‘leak’ in the cascade. We examined two extremes: ‘leaky UTT’ (L-UTT), in which 41% of patients continue to ‘leak’ from the cascade at the CD4⁺ cell/staging step, versus ‘patched UTT’ (P-UTT), in which all individuals are retained through this first step in the cascade, although they may still ‘leak’ in subsequent steps, such as failing to link to ART, at the same rate as in L-UTT.

After 5 years, improvements in testing/linkage were more cost-effective ($3742/DALY) than switching to either L-UTT ($12,601/DALY) or P-UTT ($5558/DALY). However, after 20 years, improvements in testing/linkage (1501$/DALY) were less cost-effective than P-UTT ($1241/DALY), though more cost-effective than L-UTT ($1692/DALY). Thus, the choice of investing in expanded guidelines versus improved testing/linkage depends on whether UTT provides inherent improvements in linkage by obviating the need for CD4⁺ cell testing.

Retention is required for universal test-and-treat to become cost-saving

Finally, we analyzed the impact of changing guidelines to UTT in a scenario in which testing and linkage was already maximally expanded for 80% of the population. In the short term, switching to UTT incurred great cost with little health benefit. This is because, with annual testing and perfect linkage, the 80% healthcare-accessible population is expected to initiate ART an average of 1 year of reaching CD4⁺ cell count less than 350 cells/µL. Thus, little immediate health benefit is provided by initiating at higher CD4⁺ cell counts under these optimistic assumptions. Indeed, the benefits of UTT under these assumptions are derived almost entirely from averting transmission through suppression of infectiousness.

By 20 years, the transmission-interrupting potential of UTT with expanded testing/linkage becomes apparent, though still not cost-effective. The cost of UTT was magnified by increasing ART re-initiation, because of increases in patient-time on ART. However, benefits accumulated through the transmission-blocking effect of ART in UTT were limited by the time spent between subsequent re-initiations. Time off ART was exponentially distributed with a mean of 1 year, consistent with the agreed-upon frequency of expanded testing.

To explore the full potential of UTT, we examined the effect of halting all ART dropout for the 80% healthcare-reachable population, alongside the improvements in testing and linkage that were already included in the scenario definition. Without dropout, the effects of
switching from current guidelines to UTT became cost-effective, and indeed, cost-saving: switching to UTT averted 22 million DALYs over 20 years, compared with 20 million DALYs under current guidelines, at comparable costs of $29 billion for UTT and $30 billion for current guidelines. Most of these benefits were accumulated in later years: thus, with 3% annual discounting, the switch to UTT was not cost-saving by 20 years, though nearly cost-neutral and well within stochastic uncertainty (Fig. 2); without discounting, the change was cost-saving (Supplementary Figure 1, http://links.lww.com/HJH/A291).

**Parameter sensitivity analysis**

The changes in incremental cost and DALYs averted per 1-day increase in duration or 1% increase in probability are provided in Table 3 for every parameter for which a statistically significant trend in sensitivity could be measured from 1400 simulation runs. The ICERs reported are the ratio of mean incremental cost to mean DALYs averted as compared with the unchanged parameter value, accumulated up until 2033 with 3% annual discounting. The magnitude of these changes provides an idea for the sensitivity of model results to the parameter values.

For example, under current guidelines and baseline levels of testing and linkage, a 1% increase in the proportion of the population reachable by healthcare would avert 148,000 DALYs at an additional cost of $193 million. After expansion of testing/linkage, the same 1% increase in the reachable proportion would avert 358,000 DALYs at an additional cost of $456 million. Reachability by healthcare enabled individuals to perform voluntary HIV testing, which occurred more frequently after expansion of testing. Reachability also enabled these individuals to link to care with increased probabilities, but only in the scenarios with improved linkage. This explains why sensitivity to the proportion reachable by healthcare was more than doubled by expansion of testing/linkage rates. Indeed, in the scenarios with expanded testing/linkage, this was the most sensitive of the 12 parameters that could be measured as percentages.

Under baseline levels of testing and linkage, the most sensitive parameter was the probability of linking to ART (after establishment of treatment eligibility). A 1% increase in this parameter would avert 532,000 DALYs at an additional cost of $386 million. Under expanded testing/linkage, all models were standardized to assume 100% linkage to ART for the healthcare-reachable population. Reducing this by 1% resulted in only 179,000 fewer DALYs averted with a $154 million decrease in cost. Sensitivity was lower after expanded testing/linkage because frequent testing enabled individuals to re-test and link with relatively high probability.

Table 3 also provides the ratio of the change in DALYs to the change in cost, after both numerator and denominator have received 3% annual discounting. Although this can be called an ICER, it is important to note that this cost-effectiveness is not expected to hold for large changes in parameter values. First, the region of linearity for changes in cost and DALYs may not be large for some parameters. Second, this analysis reflects only a one-dimensional parameter change without offsetting changes in other parameters to ensure that model calibration is maintained. Nonetheless, for small increments of change, such analysis helps to identify current model sensitivities and suggest avenues for potentially cost-effective program improvements.

Increasing the ART viral suppression multiplier, which determines the effectiveness of ART at reducing transmission rates, increased the impact and reduced the cost of the ART program after 20 years, as expected. The unit cost of ART was not altered while changing this parameter, so that this increase enhances the prevention benefits of ART ‘for free,’ eventually reducing program cost by averting infections.

Other than ART viral suppression, the most cost-effective change under current levels of testing/linkage was an increase in the proportion who return for a CD4+ cell result after testing positive. This was consistent with our study of the impact and cost-effectiveness of P-UTT. With expanded testing/linkage, the sensitivity and cost-effectiveness of this parameter were diminished because the parameter value only applied to the 20% of the population who did not receive 100% linkage throughout the cascade.

After expanding testing/linkage, the most cost-effective parameter change became the mean ART duration. The cost-effectiveness of this change was especially impressive for UTT guidelines with expanded testing and linkage, consistent with finding that enhanced retention on ART could make the switch to UTT cost-saving within 20 years. However, the mean duration on ART was not a sensitive parameter: a 1-day increase in the population’s mean duration on ART averted only 700 DALYs by 2033. A 1% change in the proportion reachable by healthcare had the same magnitude of impact as a 1.4-year increase in the mean duration on ART.

Overall, sensitivity analysis suggests that enormous gains could be achieved by improving treatment regimens to have higher efficacy at preventing transmission, increasing the proportion of the population with access to improved healthcare, and reducing ‘leaks’ in the ‘cascade of care.’

**Discussion**

In a detailed model of HIV epidemiology in South Africa, assumptions about retention and re-initiation rates on
ART had enormous implications for the magnitude of cost and impact associated with expanding guidelines or improving testing/linkage. A full increase in voluntary re-initiation rates from 0 to 100% approximately doubled the cost and impact of switching from current guidelines to UTT, and increased the cost of expanding testing and linkage by 80%.

Further, re-initiation was found to be the most cost-effective initial improvement from current programs, especially in the short term before the transmission-blocking benefits of other interventions could be realized. By removing the requirement that ART re-initiation be triggered by a health event such as AIDS symptoms, diagnosis of a partner, or pregnancy, voluntary re-initiation increased the probability of successful treatment by raising CD4⁺ cell counts at ART initiation and providing multiple opportunities for linkage through a leaky cascade of care.

An important caveat to this cost-efficiency result is that the epidemiological and cost models did not account for potential programmatic costs associated with increasing re-initiation rates. For example, improvements in patient tracking and outreach would likely require record-keeping, telecommunications, transportation, and personnel resources not widely available to care providers. However, a similar argument could be made about

<table>
<thead>
<tr>
<th>Variation</th>
<th>Unit</th>
<th>CD4⁺ cell count &lt;350</th>
<th>All HIV+</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>DALYs averted (000s)</td>
<td>Additional cost (millions)</td>
</tr>
<tr>
<td>Mean ART duration</td>
<td>day</td>
<td>2.9</td>
<td>$2.8</td>
</tr>
<tr>
<td>Receive CD4⁺ cell result baseline</td>
<td>%</td>
<td>184.16</td>
<td>$71.2</td>
</tr>
<tr>
<td>Reachable by healthcare system</td>
<td>%</td>
<td>147.51</td>
<td>$193.5</td>
</tr>
<tr>
<td>ART linking after scale-up</td>
<td>%</td>
<td>532.42</td>
<td>$386.0</td>
</tr>
<tr>
<td>Diagnostic testing postdebut</td>
<td>%</td>
<td>49.44</td>
<td>$47.0</td>
</tr>
<tr>
<td>Symptomatic presentation probability</td>
<td>%</td>
<td>337.68</td>
<td>$285.8</td>
</tr>
<tr>
<td>Mean duration between diagnostic tests (baseline)</td>
<td>Day</td>
<td>−4.1</td>
<td>−222.3</td>
</tr>
<tr>
<td>ART viral suppression multiplier</td>
<td>%</td>
<td>−168.7</td>
<td>$142.7</td>
</tr>
</tbody>
</table>

Note that sensitivities and incremental cost-effectiveness ratio (ICER) values stem from changing parameter values in year 2004, when antiretroviral therapy (ART) first became available.
increases in testing and linkage rates used to simulate hypothetical expanded access scenarios in the main collaborative model analysis.

A second important caveat is that EMOD-HIV did not account for an increased risk of drug resistance among individuals who re-initiate ART multiple times [23, 24]. The effect of drug resistance would be to reduce the health benefit of ART with frequent re-initiation, potentially making expanded re-initiation less beneficial and less cost-effective.

Subject to these caveats, our findings suggest that it would be most efficient to switch to expanded guidelines or increased testing/linkage only after the potential benefits of re-initiation have been exhausted. However, the maximum impact of re-initiation is limited compared to UTT guidelines or expanded testing/linkage rates. After taking full advantage of re-initiation, further impact could be made by expanding treatment guidelines, improving testing/linkage rates, or both. In a 5-year timeframe, improvements in testing/linkage are more cost-effective. However, in a 20-year timeframe, the choice depends on the effects of expanded guidelines on the treatment cascade.

Recent studies have found that more than one-third of patients who receive a positive diagnostic test result fail to return for a CD4$^+$ cell result (10, 23–25). Rapid point-of-care CD4$^+$ cell testing could reduce this rate of dropout by more than half [25,26]. It is not known the extent to which CD4$^+$ cell-agnostic treatment guidelines would reduce dropout at this step of the treatment cascade in South Africa. However, the importance of this assumption was supported by our sensitivity analysis, which found the proportion receiving a CD4$^+$ cell result to be the most cost-effective parameter change under baseline testing/linkage levels.

Therefore, we compared scenarios in which UTT had no effect on the cascade (‘leaky UTT’ or L-UTT) to ones in which CD4$^+$ cell-agnostic guidelines completely ‘patched’ this leak (P-UTT). P-UTT increased the magnitude of impact and the cost-efficiency of expanding guidelines. Over a 20-year time horizon with 3% annual discounting, expansion of guidelines was more cost-effective than improving testing/linkage with P-UTT, but not with L-UTT. Thus, the best choice depends on the extent to which UTT ‘patches’ early ‘leaks’ in the treatment cascade.

After expanding testing and linkage, additionally expanding guidelines to UTT is costly and yields little short-term benefit. This is because expanded testing already ensures that individuals enroll soon after becoming eligible, so that earlier initiation provides little in the way of immediate health benefit. The long-term health benefits of UTT come from reductions in incidence due to the prevention effect of ART. Indeed, the proportion of the population with access to expanded linkage/testing and the prevention efficacy of ART were highly sensitive model parameters in these scenarios.

Even after 20 years, expansion to UTT was similarly or less cost-effective than maintaining current guidelines, regardless of the level of voluntary re-initiation of ART. In contrast, improving retention allows UTT to become cost-saving relative to current guidelines within 20 years, while averting an additional 1.3 million DALYs. Thus, in a fully expanded test-and-treat program with excellent linkage, improved retention in care would be necessary in order to reap the benefits of treatment as prevention.

All model projections are associated with uncertainty, either due to uncertain parameter values, stochastic components within the model, or limitations of the structures and mechanisms that compose the model itself. An example of the latter in the current version of our model ($\psi_{0.8}$) is the absence of a relationship between ART re-initiation and the development of drug resistance.

With a flexible and extensible model framework, we hope to better understand the implications of parametric, stochastic, and structural sources of model uncertainty. Such analyses could help guide data collection that will improve the predictive capabilities of models, while encouraging policy-makers to approach models with intense engagement and appropriate caution.

**Conclusion**

Model frameworks that can implement multiple assumptions within the same basic structures will play a key role in the communication between modelers and policy-makers by helping to translate and compare the effect of structural model assumptions. We have shown how model assumptions about re-initiation and retention on ART can dramatically change the predicted cost and impact of expanded guidelines or improvements in testing and linkage to care.

Our findings suggest that increasing re-initiation is the most cost-effective means of initial program expansion, especially in the short term, but that improvements in retention will be necessary in order to reap the full transmission-blocking benefits of a test-and-treat program in the long term.

**Acknowledgements**

The authors thank Bill and Melinda Gates for their active support of this work and their sponsorship through the
Global Good Fund. Members of the HIV Modelling Consortium provided feedback during the development of EMOD-HIV, helped to design the jointly modeled scenarios, provided data for calibration, and developed the cost model used in this analysis.

Conflicts of interest
There are no conflicts of interest.

References
