The Cost-Effectiveness of Monitoring Strategies for Antiretroviral Therapy of HIV Infected Patients in Resource-Limited Settings: Software Tool

Janne Estill1*, Luisa Salazar-Vizcaya1, Nello Blaser1, Matthias Egger1,2, Olivia Keiser1

1 Institute of Social and Preventive Medicine (ISPM), University of Bern, Bern, Switzerland, 2 Centre for Infectious Disease Epidemiology and Research, University of Cape Town, Cape Town, South Africa

* janne.estill@ispm.unibe.ch

Abstract

Background
The cost-effectiveness of routine viral load (VL) monitoring of HIV-infected patients on antiretroviral therapy (ART) depends on various factors that differ between settings and across time. Low-cost point-of-care (POC) tests for VL are in development and may make routine VL monitoring affordable in resource-limited settings. We developed a software tool to study the cost-effectiveness of switching to second-line ART with different monitoring strategies, and focused on POC-VL monitoring.

Methods
We used a mathematical model to simulate cohorts of patients from start of ART until death. We modeled 13 strategies (no 2nd-line, clinical, CD4 (with or without targeted VL), POC-VL, and laboratory-based VL monitoring, with different frequencies). We included a scenario with identical failure rates across strategies, and one in which routine VL monitoring reduces the risk of failure. We compared lifetime costs and averted disability-adjusted life-years (DALYs). We calculated incremental cost-effectiveness ratios (ICER). We developed an Excel tool to update the results of the model for varying unit costs and cohort characteristics, and conducted several sensitivity analyses varying the input costs.

Results
Introducing 2nd-line ART had an ICER of US$1651-1766/DALY averted. Compared with clinical monitoring, the ICER of CD4 monitoring was US$1896-US$5488/DALY averted and VL monitoring US$951-US$5813/DALY averted. We found no difference between POC- and laboratory-based VL monitoring, except for the highest measurement frequency (every 6 months), where laboratory-based testing was more effective. Targeted VL monitoring was on the cost-effectiveness frontier only if the difference between 1st- and 2nd-line costs remained large, and if we assumed that routine VL monitoring does not prevent failure.
Conclusion

Compared with the less expensive strategies, the cost-effectiveness of routine VL monitoring essentially depends on the cost of 2nd-line ART. Our Excel tool is useful for determining optimal monitoring strategies for specific settings, with specific sex-and-age-distributions and unit costs.

Introduction

The latest World Health Organization (WHO) guidelines recommend routine viral load monitoring for HIV-infected patients on antiretroviral therapy (ART) in resource-limited settings [1]. Routine viral load (VL) monitoring is the gold standard for detecting treatment failure and deciding when patients should switch to 2nd-line ART. VL monitoring may also support adherence and prevent HIV transmission, thus offering advantages beyond patient survival [2]. Most ART programs in resource-limited settings currently rely on CD4 or clinical monitoring [3], and the debate over the long-term benefit of routine VL monitoring still continues. It centers on the high cost of VL and on the logistical constraints that may make it infeasible to implement this recommendation [4–8].

We have shown that monitoring VL with a qualitative, moderately-sensitive POC VL test benefits the patient, and may be cost-effective compared with CD4 or clinical monitoring [5]. Rapid and affordable point-of-care (POC) tests for VL are already in development [9,10]. POC tests may improve on laboratory-based monitoring by simplifying the process. Patients could get same-day test results, adherence counseling, and/or make treatment decisions, all during the same visit. Though cheaper, POC tests may not be as accurate for diagnosis as sensitive and fully quantitative tests. Nevertheless, in settings where VL monitoring is still unavailable, POC tests may offer an affordable entry into VL monitoring.

The feasibility and cost-effectiveness of VL monitoring with a qualitative POC test depends on many factors, and these can vary substantially among settings. We built on our earlier mathematical simulation model of patients on ART, to create a flexible and user-friendly tool that allows users to evaluate the cost-effectiveness of a wide range of monitoring and switching strategies with varying assumptions.

Methods

Simulation model

We modeled cohorts of HIV-infected patients from ART initiation until death. We used three indicators to define the progression of HIV infection qualitatively: virological; immunological; and, clinical. All three indicators have two possible values: normal and failing. We assumed that all patients started in the normal virological, immunological and clinical stages. This represents the successful introduction of ART: VL decreases rapidly to undetectable values and remains suppressed; CD4 cell count increases; and, the patient will have no more clinical (WHO stage 3 or 4 defining) symptoms. The patient can proceed to virological failure at any time: this represents a rebound in VL to a detectable value of >1000 copies/ml (or, at the early stages of treatment, the failure to suppress). Progression to immunological failure represents a CD4 cell count decline to a level that meets the WHO immunological criteria of treatment failure [1]. Progression to clinical failure represents the occurrence of WHO stage 3 or 4 defining symptoms. Both immunological and clinical progression can occur either as a consequence of
virological failure ("concordant"), or independent of the patient’s virological stage ("discordant"). Clinical progression also depends on the patient’s current immunological status. The parameterization of hazard functions related to disease progression was the same as in previous modeling studies [2,4,5], which were mostly based on two routine ART programs from the Cape Town area, Gugulethu and Khayelitsha. The characteristics of these cohorts were described in a previous publication [2]. The structure of the simulation model is presented in Fig. 1 and the key parameters are listed in Table 1.

The failing virological, immunological and clinical stages will persist, unless the patient switches to a 2nd-line ART regimen. When the patient switches, the failing virological as well as concordant failing immunological and clinical stages will return to normal. Failing discordant immunological and clinical stages will remain failing after switching. During 2nd-line therapy, the patient is again at risk of proceeding to the failing virological, immunological or clinical stage. The parameterization was the same as for 1st-line therapy, except that the risk of proceeding to virological failure was scaled up with a resistance penalty factor, which depended on the time the patient spent on virologically failing 1st-line therapy.

Mortality consists of two components: HIV-free background mortality and HIV-related mortality. We assumed the risk of mortality increased for patients in the failing virological, immunological or clinical stage. Although we assumed that all patients are retained in care from ART initiation until death, we accounted for the expected high mortality among patients lost to follow-up when we estimated the HIV-related mortality rates [4].

We considered a total of 13 monitoring and switching strategies (Table 2). With clinical monitoring, the treatment failure is observed at the next regular appointment after the patient proceeds to the advanced clinical stage, and the patient switches 3 months later. With CD4 monitoring, the failure is observed at the next appointment when the CD4 cell count is measured (which depends on the measurement frequency), and the patient switches 3 months later. Both clinical and CD4 monitoring are assumed to be fully specific and sensitive for detecting clinical or immunological failure. With routine VL monitoring, observing the failure depends on the definition of failure and the test itself. With a laboratory VL test, a failure is observed at the first monitoring appointment after the failing virological stage begins, and we assumed that CD4s are measured simultaneously. With a qualitative POC VL test, a failure may be observed at any visit, and the probability of detecting a failure depends on the detection limit of the test: we assumed it was 5000 copies/ml throughout the study. The test is repeated three months after observing the failure; if the failure is confirmed, the patient switches to 2nd-line ART. With targeted VL monitoring, only CD4 counts are measured routinely. A POC VL test is performed immediately if an immunological failure is detected. If this test is also positive, a second VL test is performed 3 months later.

We modeled two scenarios. In Scenario A, we assumed that the risk of treatment failure did not depend on the monitoring and switching strategy, and thus, for all strategies, we used the rates estimated from the South African data (where VL is monitored regularly). In Scenario B, we assumed that VL monitoring can prevent treatment failure, and that the South African rates underestimate the risk in strategies where VL monitoring is unavailable. We assumed this because routine viral load monitoring can detect poor adherence. Many patients with a detectable viral load can re-suppress viral load on first-line ART, after an adherence intervention [11–14]. Since poor adherence is a major predictor of treatment failure [15,16], routine viral load monitoring may prevent treatment failures caused by poor adherence. For these strategies, we assumed that the hazard would be twice as high across the entire follow-up period.

The model was constructed in three steps. In the first step, we simulated cohorts of 100,000 patients for all 13 monitoring and switching strategies, without baseline characteristics or background mortality. We implemented the model using ‘gems’, an R package for generalized
multistate simulation models [17,18]. 'Gems' models disease progression as a series of events (e.g., diagnosis, treatment and death) that can be displayed in a directed acyclic graph (DAG). The vertices correspond to disease states and the directed edges represent events. Similar models that use the package 'gems' or a similar algorithm have been published elsewhere [2,4,5,19]. For strategies without routine VL monitoring we modeled two cohorts with different failure rates (scenarios A and B). This resulted in a total of 20 cohorts.

Fig 1. Progression of patients in the mathematical model. Panel A shows the progression of the patient's treatment regimen and observed failure status. Within each compartment of panel A, the patient will proceed according to the underlying treatment progression shown in Panel B. The type of failure that can be detected depends on the monitoring strategy. After switching to second-line therapy, the patient will start either in the successful ART compartment (if he/she had no or concordant immunological/clinical failure) or in the clinical and/or immunological failure compartment (if he/she had a discordant failure of the corresponding type). See the main text for definitions of concordant and discordant failures.

doi:10.1371/journal.pone.0119299.g001
Table 1. Key parameters related to disease progression.

| Outcome | Source | Statistical model | Starting | Value (95% CI) | Dimension | Risk |
|---------|--------|-------------------|----------|----------------|-----------|------|
| 1) Time to virological failure |
| (a) First-line ART; second-line ART with immediate switch | Cohorts | Parametric Weibull | 3 months from ART start/switch | 0.57 (0.52–0.63) | Shape | 3.4% fail by 1 year after ART start |
| (b) Resistance penalty | [37] | * | n/a | 0.05 (0.00–0.20) | Decrease in efficacy | n/a |
| 2) Time to immunological failure |
| (a) After virologic failure | Cohorts | Parametric exponential | Virological failure | 0.08 (0.06–0.10) | Rate (years⁻¹) | 7.6% fail by 1 year after virologic failure |
| (b) Independently of virologic failure | Cohorts | Parametric Weibull | 3 months from ART start | 0.22 (0.20–0.25) | Shape | 3.0% fail by 1 year after ART start |
| 3) Time to clinical failure |
| (a) Without virologic or immunologic failure | [6] | Parametric exponential | ART start | 0.004 | Rate (years⁻⁶) | 0.4% fail by 1 year after ART start |
| (b) Extra hazard after immunologic failure | [6,38] | Cox regression | Immunologic failure | 3.3 | HR, constant over time | n/a |
| (c) Extra hazard after virologic failure | [6] | Cox regression | Virologic failure | 2 | HR, constant over time | n/a |
| 4) Time to death (HIV-related mortality) |
| (a) Observed mortality | Cohorts | No specific model (competing risk analysis) | ART start | ** | n/a | 6.5% die by 1 year after ART start |
| (b) Observed LTFU | Cohorts | No specific model (competing risk analysis) | ART start | ** | n/a | 4.0% lost by 1 year after ART start |
| (c) Mortality among LTFU | Analysis 4b, [39] | No specific model (theoretical calculation) | n/a | ** | n/a | n/a |
| (d) HIV-related mortality | Analyses 4a-4c | Theoretical calculation, double Weibull*** | ART start | 0.88 (0.88–0.90) | Shape 1 | 8.6% die by 1 year after ART start |
| | | | | 0.35 (0.32–0.39) | Scale 1 (years) |
| | | | | 1.00 (1.00–1.00) | Shape 2 |
| | | | | 64.60 (54.52–76.55) | Scale 2 (years) |
| | | | | 0.08 (0.08–0.08) | Weight (1st component) |
| (a) Extra hazard after clinical failure | assumption | Cox regression | Clinical failure | 2 | HR, constant over time | n/a |
| (b) Extra hazard after immunologic failure | Cohorts | Cox regression | Immunologic failure | 1.76 (1.16–2.68) | HR, constant over time | n/a |

(Continued)
In the second step, these cohorts were updated to account for differences in background mortality. Thirty-two copies of each of the 20 cohorts were created, for the different sexes (male and female), baseline age groups (15–24, 25–34, 35–44 and 45–54 years), and four different scenarios of background mortality. In the first three scenarios, the background mortality rates represent the overall HIV-free mortality in the general populations of Malawi and Zimbabwe [20], and in Africans in the Western Cape [21]. In the fourth scenario, we assumed that the HIV-free life expectancy from birth would be 75 years for all patients. Each patient was assigned an age, sampled from a uniform distribution within the given range. A time of HIV-unrelated death was sampled for each patient, based on the gender, baseline age and the background mortality rate.

In the third step, we analyzed the outcomes of interest. The main outcomes were disability-adjusted life-years (DALY) lost to HIV, total cost, and cost-effectiveness ratios of the intervention compared to current practice as well as to the next less expensive strategy (incremental cost-effectiveness ratio, ICER). Definitions are given in S1 Text.

**Excel tool**

We developed an Excel spreadsheet tool to adapt the model outputs to specific scenarios. The Excel table contains all outputs of the simulation and presents the results according to the scenario defined by the user. The user can vary the following input variables continuously (Table 3): size of cohort; unit costs for clinic visit, VL and/or CD4 test, one year of 1st- and 2nd-line ART; and, disability weight of symptomatic and asymptomatic HIV. The user can specify the age and gender distribution of the cohort by giving proportions for each of the eight age and gender groups. Finally, the user can specify the failure rate scenario (A or B), background mortality (Malawi, Zimbabwe, Western Cape, or constant life expectancy 75 years) and discounting (0%, 1%, 2%, 3%, 4%, or 5%). The main results are then updated based on the assumptions. The Excel table also shows the ICERs, and graphically presents the costs and averted DALYs of each scenario.

We present the results in this manuscript for a set of input parameters (Table 3). We assumed the costs of 1st-line ART were US$99/year and 2nd-line ART were US$280/year.
tests (both POC and laboratory-based) were assumed to cost US$10, and CD4 tests US$5. We did not consider the cost of clinic appointments. Disability coefficients were 0.135 for asymptomatic HIV and 0.369 for symptomatic HIV. The disability coefficient for symptomatic HIV was the product of the coefficients of asymptomatic HIV and tuberculosis, the most common opportunistic infection [22]. All results are presented per one patient. Costs and DALYs are discounted annually by 3%. Two separate analyses were conducted: one in which virological failure rates were identical (Scenario A); and one in which it was twice as high in strategies without routine VL monitoring than in strategies with routine VL monitoring (Scenario B). We also present 10 sensitivity analyses in the S2 Text and S1 Table, in which the unit costs of tests and ART and the discounting rate were varied.
**Results**

**Excel tool**

The Excel tool is presented in [S1 Excel File](#).

**Scenario A: Failure rate is identical in all monitoring and switching strategies**

In the absence of 2\textsuperscript{nd}-line monitoring, the average lifetime cost of ART was US$1419 per patient (Table 4). On average, each patient lost 7.3 DALYs to HIV. Adding 2\textsuperscript{nd}-line ART with clinical monitoring (2.1) increased the costs to US$1563 per patient and averted 0.09 DALYs. The costs of strategies with CD4 monitoring ranged from US$1690 (3.2) to US$1812 (3.5). CD4 monitoring averted 0.05 to 0.09 DALYs more than clinical monitoring: the benefit was higher for Strategies 3.1 and 3.5 than for Strategies 3.2, 3.3 and 3.4. Targeted VL monitoring (3.5) cost, on average, an additional US$12 per patient, but saved about US$61 per patient for ART over the same strategy without VL tests (3.4) by not switching of patients with suppressed...
VLs. The total costs for strategies with VL monitoring ranged from US$1840 (4.1) to US$2216 (5.3). VL monitoring averted 0 to 0.05 DALYs more than the most effective CD4 monitoring strategy (3.5), and 0.09 to 0.14 more than clinical monitoring (2.1). There were no clear differences between POC and laboratory-based VL monitoring in the number of averted DALYs. The POC VL strategy with most frequent (every 6 months) monitoring (4.2) averted slightly fewer DALYs than POC VL monitoring every 12 months (4.3).

No 2nd-line (1.1), clinical monitoring (2.1), irregular CD4 monitoring every 6 months (3.1), CD4 monitoring every 6 months with targeted VL monitoring (3.5), and POC VL monitoring every 12 months (4.2) were on the cost-effectiveness frontier (Fig. 2A). The cost-effectiveness ratio of clinical monitoring alone (2.1) compared with no 2nd-line (1.1) was US$1651/DALY averted. Compared with clinical monitoring alone (2.1), CD4 monitoring averted a DALY at costs between US$2123 (3.1) and US$5488 (3.5). VL monitoring strategies prevented a DALY at costs between US$3478 (4.1) and US$2494 (4.1) to US$813 (4.3) as compared with Strategy 2.1.

Scenario B: Failure rate is twice as high without VL monitoring as with VL monitoring

When only 1st-line ART was available (1.1), the average lifetime cost of ART was US$1401 per patient (Table 5). On average, each patient lost 7.5 DALYs to HIV. When 2nd-line ART with clinical monitoring was added (2.1), this increased the costs to US$1569 per patient and...
averted 0.10 DALYs. CD4 monitoring strategies cost between US$1794 (3.2) and US$1929 (3.4). CD4 monitoring averted 0.10 to 0.12 DALYs more than clinical monitoring: the benefit increased with monitoring frequency and was highest for targeted VL monitoring. Targeted VL monitoring (3.5) cost, on average, an additional US$13 per patient, but ART cost about US $58 less per patient than in Strategy 3.4, because patients with suppressed VLs did not need to be switched. The results of strategies that used VL monitoring (4.1 to 5.3) are the same as those in Scenario A.

Except for the no 2nd-line strategy (1.1), only POC VL monitoring every 24 (4.1) and 12 months (4.2) were on the cost-effectiveness frontier (Fig. 2B). Clinical monitoring was weakly dominated by POC VL monitoring every 24 months. Strategies based on CD4 monitoring cost about as much as those that used VL monitoring, but averted fewer DALYs. The cost-effectiveness ratio of clinical monitoring alone (2.1) compared with no 2nd-line (1.1) was US$1766/DALY averted. Compared with clinical monitoring alone (2.1), CD4 monitoring averted a DALY at costs between US$1896 (3.5) and US$2540 (3.1). VL monitoring strategies prevented a DALY at costs between US$48 (4.1) and US$1875 (5.3) as compared with Strategy 3.1 or US $951 (4.1) and US$2097 (5.3) as compared with Strategy 2.1.
Table 5. Model outcomes: main analysis assuming that treatment failure rate is twice as high in strategies without routine viral load monitoring (1.1 to 3.5) as in strategies with routine viral load monitoring due to improved adherence.

| Strategy                           | No 2nd-l. | 2.1 | 3.1 | 3.2 | 3.3 | 3.4 | 3.5 | 4.1 | 4.2 | 4.3 | 5.1 | 5.2 | 5.3 |
|-----------------------------------|-----------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Life-years                        |           |     |     |     |     |     |     |     |     |     |     |     |     |
| Healthy life-years left           | 19.5      | 19.5| 19.5| 19.5| 19.5| 19.5| 19.5| 19.5| 19.5| 19.5| 19.5| 19.5| 19.5|
| Life-years on 1st-line ART        | 14.2      | 13.3| 12.5| 12.5| 12.4| 12.4| 12.8| 12.7| 12.7| 12.5| 12.8| 12.7| 12.7|
| Life-years on 2nd-line ART        | 0.0       | 0.9 | 1.9 | 1.9 | 2.0 | 2.0 | 1.7 | 1.8 | 1.9 | 2.0 | 1.8 | 1.9 | 1.9 |
| Life-years without symptoms       | 13.3      | 13.5| 13.6| 13.6| 13.7| 13.7| 13.7| 13.9| 13.9| 13.9| 13.9| 13.9| 13.9|
| Life-years with symptoms          | 0.8       | 0.7 | 0.7 | 0.7 | 0.7 | 0.7 | 0.7 | 0.7 | 0.7 | 0.7 | 0.7 | 0.7 | 0.7 |
| Life-years lost to HIV            | 5.4       | 5.3 | 5.2 | 5.1 | 5.1 | 5.1 | 5.1 | 5.0 | 4.9 | 5.0 | 5.0 | 5.0 | 4.9 |
| Disability-weighted life-years    | 2.1       | 2.1 | 2.1 | 2.1 | 2.1 | 2.1 | 2.1 | 2.1 | 2.1 | 2.1 | 2.1 | 2.1 | 2.1 |
| **DALYs lost to HIV**             | **7.5**   | **7.4**| **7.3**| **7.3**| **7.2**| **7.2**| **7.2**| **7.1**| **7.1**| **7.1**| **7.1**| **7.1**| **7.1**|
| **Costs**                         |           |     |     |     |     |     |     |     |     |     |     |     |     |
| Cost of 1st-line ART              | 1401      | 1320| 1234| 1237| 1228| 1228| 1263| 1261| 1254| 1238| 1263| 1256| 1255|
| Cost of 2nd-line ART              | 0         | 249 | 527 | 521 | 547 | 558 | 466 | 507 | 534 | 563 | 499 | 521 | 530 |
| Cost of diagnostic tests          | 0         | 0   | 70  | 36  | 72  | 143 | 156 | 73  | 147 | 292 | 107 | 216 | 431 |
| **Total costs**                   | 1401      | 1569| 1831| 1794| 1848| 1929| 1884| 1840| 1935| 2094| 1869| 1993| 2216|

**Cost-effectiveness**

- **CER compared to 1.1**
  - I/e: 1766, 2168, 2047, 2102, 2199, 1849, 1156, 1317, 1924, 1258, 1545, 2019
- **ICER**
  - I/e: 1901, 2400, 1735, 1637, 1899, 1375, 1156, 3175, 1258, 1545, 2019

Please see Table 2 for a detailed description of all monitoring strategies. All costs are given in US$ and cost-effectiveness ratios in US$/DALY averted.

POC-VL, point-of-care viral load; lab-VL, laboratory-based viral load; ART, antiretroviral therapy; DALY, disability-adjusted life-year; CER, cost-effectiveness ratio; ICER, incremental cost-effectiveness ratio; I/e, least expensive and least effective strategy; w/d, weakly dominated; s/d, strongly dominated.

doi:10.1371/journal.pone.0119299.t005

Table 6. Cost-effectiveness of 12-monthly point-of-care viral load monitoring (strategy 4.2) compared with clinical monitoring (strategy 2.1), assuming that the risk of virological failure is either identical in both strategies (Scenario A) or twice as high with viral load monitoring than with clinical monitoring (Scenario B).

| Analysis | Varied input value | Cost-effectiveness ratio (US$/DALY averted) |
|----------|--------------------|-------------------------------------------|
|          |                    | Scenario A                  | Scenario B                  |
| Main     |                    | 2723                        | 1180                        |
| VL1      | Cost of VL test: US$7 | 2400                        | 1037                        |
| VL2      | Cost of VL test: US$5 | 2184                        | 942                         |
| VL3      | Cost of VL test: US$15 | 3262                        | 1417                        |
| FL1      | Cost of 1st-line ART: US$55/year | 3045                        | 1275                        |
| FL2      | Cost of 1st-line ART: US$128/year | 2511                        | 1117                        |
| SL1      | Cost of 2nd-line ART: US$210/year | 2131                        | 905                         |
| SL2      | Cost of 2nd-line ART: US$140/year | 1539                        | 720                         |
| SL3      | Cost of 2nd-line ART: US$350/year | 3315                        | 1409                        |
| DI1      | No discounting      | 2015                        | 892                         |

All performed sensitivity analyses are listed in S1 Table and their full results in S2 to S11 Tables of the Supplementary Material. The results of Analysis CD1 are not shown since the cost of CD4 test does not influence either strategy. The input values used in the main analysis are shown in Table 3.

doi:10.1371/journal.pone.0119299.t006
Sensitivity analyses

The results of the sensitivity analyses are presented in detail in Table 6 and S2 Text and S2 to S11 Tables. Of note is the dependence of the benefit of targeted routine VL monitoring on the ratio of the costs between 1st- and 2nd-line regimens. If the prices of 1st- and 2nd-line regimens were similar, targeted VL monitoring was dominated; if 2nd-line ART was assumed to be substantially more expensive, targeted VL monitoring was on the cost-effectiveness frontier.

Discussion

We simulated cohorts of patients who initiated ART under 13 different ART monitoring and switching strategies. VL monitoring was slightly more effective than CD4 monitoring (in particular if we assumed that VL monitoring also reduces the risk of treatment failure). CD4 monitoring was more effective than clinical monitoring, and clinical monitoring was more effective than 1st-line ART only. However, differences in the effectiveness of any two strategies were all below 1 DALY per patient. We observed no clear difference between monitoring strategies that measured at different intervals, or between VL monitoring that used fully quantitative, highly specific and sensitive laboratory tests or those that used a qualitative POC test. The cost-effectiveness of POC VL monitoring clearly improved if we reduced the gap between prices of 1st- and 2nd-line ART (Table 6).

Across our analyses, 12-month routine POC VL monitoring was on the cost-effectiveness frontier. The cost-effectiveness ratio of this strategy compared to clinical monitoring varied between US$700 and US$3300 per DALY averted. Cost-effectiveness improved if we assumed that VL monitoring reduces the risk of failure, and when the price of 2nd-line ART and 1st-line ART were close. The cost-effectiveness ratio of US$700 per DALY averted was reached if 2nd-line costs were reduced to minimum and we assumed that routine VL monitoring prevents failure. If we define a cost-effective intervention as having a cost-effectiveness ratio of less than 3 times the local per-capita gross domestic product, a cost-effectiveness ratio of US$700 per DALY averted can be considered cost-effective in any country [23,24]. We found no major differences between fully quantitative laboratory-based VL monitoring with CD4 tests, and qualitative POC VL monitoring. The potential disadvantage of a POC test that we assumed was the possibility of “false positive” switches to 2nd-line ART, i.e. switching patients who do not have a persistent detectable viral load. For example, if a patient has two successive detectable viral load values, and if the exact values of the VLs and the CD4 cell count are known, the clinician may be able to distinguish a patient failing therapy and a patient with blips, poor adherence, measurement errors, etc. The qualitative test only gives a positive or negative result, upon which a decision must be based. Although we did not find differences between laboratory-based and POC VL monitoring overall, increasing the frequency of POC VL monitoring from 12 to 6 months slightly decreased, rather than increased life-years. This was not the case for laboratory-based VL monitoring. We think this is caused by false positive failures with POC tests: increasing the measurement frequency increases the number of patients who switch unnecessarily, and this may affect their future treatment options.

The cost-effectiveness of targeted VL monitoring varied changed with the input parameters. We found that if the price of 2nd-line ART is considerably higher than for 1st-line ART, and if we assume no additional benefits for routine VL monitoring, it may be cost-effective to conduct routine CD4 monitoring with targeted VL testing. This strategy uses routine CD4 tests to detect patients who may be on a failing 1st-line regimen. Patients are then given VL tests to confirm their status. This strategy reduces the total costs by restricting use of 2nd-line ART to patients who need it most (those with low CD4 cell counts and a high risk of mortality), and by not switching patients with suppressed VLs. However, if we assume that routine VL
monitoring can also reduce the risk of failure, for example, or if the cost difference between the regimens is small, routine VL monitoring is preferable.

Optimal monitoring strategy has been investigated in a number of mathematical modeling studies. Walensky et al. published a review of five modeling studies that assessed the cost-effectiveness of different strategies for monitoring HIV-infected patients on ART [25]. Four of the models investigated VL monitoring [6,26–28]. After the review was published, similar modeling studies have appeared [29,30]. A recent article [31] systematically compared three models. All these studies suggest that VL monitoring may moderately improve the outcomes of ART programs, but cost-effectiveness estimates vary substantially between them. The variance may be caused by the different assumptions in the input values. We believe our study is the first to include a tool that allows users to easily vary input values.

Limitations

Our analysis has several limitations. First, we only included time on ART. Diagnostic tests, and, in particular, CD4 cell counts, are usually recommended for monitoring patients before ART is initiated. If assessment of ART eligibility continues to depend on CD4 cell counts, CD4 monitoring will continue to be important. However, there is a growing tendency towards simpler rules for ART initiation, such as “Option B+”, in which all pregnant and breastfeeding women start lifelong ART [32], or universal “test and treat” [33]. We anticipate that, in many settings, there will be a decreased need to measure CD4 cell count to assess eligibility for ART.

We did not include loss to follow-up (LTFU). High rates of LTFU are a serious problem in most ART programs in resource-limited settings [34]. LTFU is a combination of unregistered deaths, unregistered transfers, and cases in which a patient has stopped ART [35]. We took the effect of LTFU on mortality into account in the parameterization of mortality. Patients who transferred to another ART clinic can be expected to take ART as recommended. We did not take include the effect of patients who stopped ART. POC monitoring may offer advantages over laboratory-based monitoring by making it easier to manage patients and by shortening wait times. This may, in turn, improve retention [36], and is another argument for POC monitoring, but it is not studied in the current analysis.

We did not model onward transmission. VL monitoring reduces the time the patient spends on a failing regimen. VL monitoring can prevent about 30% of transmissions from patients on treatment [2,5], and this may substantially improve the cost-effectiveness of VL monitoring. Therefore, the long-term population-level benefits of VL monitoring may be larger than we estimated.

We assumed that laboratory-based VL monitoring, together with CD4 tests, is 100% sensitive and specific to detect true treatment failure. Although this is a simplifying assumption, we wanted to include two distinct types of viral load test: one of these is as accurate as possible, and one has a lower sensitivity and specificity. The POC tests were assumed to be neither fully sensitive nor specific to detect true treatment failure. The users of the Excel tool can therefore choose either laboratory-based or POC viral load monitoring, depending on the diagnostic capacities of the viral load test they want to investigate. We also assumed CD4 tests are fully sensitive and specific. But since CD4 cell count poorly predicts true treatment failure, this assumption should not have much effect on the results.
The Excel tool allows users to vary input costs and several other input parameters. However, parameters related to disease progression cannot be changed in the Excel tool. Most key parameters for HIV progression are based on two large ART cohorts from the Cape Town area: Gugulethu and Khayelitsha. The results of the model reflect the characteristics of these routine ART programs, typical of southern Africa: the majority of patients are women, and most patients start ART with low CD4 cell counts and advanced clinical symptoms. We believe the results of our model can be generalized widely for ART programs in sub-Saharan Africa. However, there are also important differences between the cohorts we drew on, and others: the Cape Town cohorts had access to frequent laboratory measurements, and there is a continuous tendency to start ART earlier. Our model may not be able to catch all site-level characteristics of the different settings.

**Conclusion**

POC VL testing appears to be a promising alternative for routinely monitoring ART in resource-limited settings, especially if we assume that viral load monitoring can prevent treatment failure by improving adherence, and that the gap between prices of 1st- and 2nd-line ART can be decreased. Under these conditions, VL monitoring every 12 or 24 months, with an affordable qualitative test, does not considerably increase the costs above those of CD4 monitoring. POC VL testing may offer the same benefit as frequently monitoring VL with a quantitative test. Routine VL monitoring may also have benefits beyond more accurate detection of treatment failure; it may for example be able to prevent treatment failure by improved adherence. Targeted VL monitoring, based on routine CD4 monitoring, may be an option if these potential benefits are small, if 2nd-line ART remains substantially more expensive than 1st-line ART, and if it is expected that 2nd-line ART cannot be provided for everyone failing virologically. Special attention should be paid to the price of 2nd-line ART, which we expect will play a more substantial role than the price of the diagnostic tests. To make routine viral load monitoring cost-effective, efforts must be made to drop the costs of 2nd-line antiretroviral drugs. The optimal monitoring and switching strategy for each setting depends substantially on factors such as the unit costs. Our Excel tool allows researchers and policy-makers to vary important cost and population structure parameters and may be a valuable tool for developing local monitoring guidelines.

**Supporting Information**

S1 Excel File. Excel tool for adapting the model outputs to specific scenarios.
(XLSX)

S1 Table. List of sensitivity analyses.
(DOCX)

S2 Table. Model outcomes: sensitivity analysis VL1 assuming that the cost of a viral load test is US$7.
(DOCX)

S3 Table. Model outcomes: sensitivity analysis VL2 assuming that the cost of a viral load test is US$5.
(DOCX)

S4 Table. Model outcomes: sensitivity analysis VL3 assuming that the cost of a viral load test is US$15.
(DOCX)
S5 Table. Model outcomes: sensitivity analysis CD1 assuming that the cost of a CD4 test is US$2.

S6 Table. Model outcomes: sensitivity analysis FL1 assuming that the annual cost of 1st-line therapy is US$55.

S7 Table. Model outcomes: sensitivity analysis FL2 assuming that the annual cost of 1st-line therapy is US$128.

S8 Table. Model outcomes: sensitivity analysis SL1 assuming that the annual cost of 2nd-line ART is US$210.

S9 Table. Model outcomes: sensitivity analysis SL2 assuming that the annual cost of 2nd-line ART is US$140.

S10 Table. Model outcomes: sensitivity analysis SL3 assuming that the annual cost of 2nd-line ART is US$350.

S11 Table. Model outcomes: sensitivity analysis DI1 assuming no discounting.

S1 Text. Technical description of the definitions related to the cost-effectiveness analyses.

S2 Text. Sensitivity analyses.

Acknowledgments

We thank Cindy Zahnd for helpful comments and suggestions and Kali Tal for editing the manuscript.

Author Contributions

Conceived and designed the experiments: JE LS NB ME OK. Performed the experiments: JE LS NB. Wrote the paper: JE LS NB ME OK.

References

1. World Health Organization (2013) Consolidated guidelines on the use of antiretroviral drugs for treating and preventing HIV infection: Recommendations for a public health approach. Geneva, Switzerland: World Health Organization. Available: http://apps.who.int/iris/bitstream/10665/85321/1/9789241505727_eng.pdf?ua=1. Accessed 18 December 2014.

2. Estill J, Aubrière C, Egger M, Johnson L, Wood R, Garone D, et al. (2012) Viral load monitoring of antiretroviral therapy, cohort viral load and HIV transmission in Southern Africa: a mathematical modelling analysis. AIDS 26: 1403–1413. doi:10.1097/QAD.0b013e328356988 PMID: 22421243

3. Joint United Nations Programme on HIV/AIDS (UNAIDS) and Médecines Sans Frontières (MSF) (2012) Speed up scale-up: Strategies, tools and policies to get the best HIV treatment to more people, sooner. UNAIDS; MSF. Available: http://www.msfaccess.org/sites/default/files/MSF_assets/HIV_AIDS/Docs/AIDS_Report_SpeedUpScaleUp_ENG_2012.pdf. Accessed 18 December 2014.
4. Estill J, Egger M, Johnson LF, Gsponer T, Wandeler G, Davies MA, et al. (2013) Monitoring of antiretroviral therapy and mortality in HIV programmes in Malawi, South Africa and Zambia: mathematical modelling study. PLoS One 8: e57611. doi: 10.1371/journal.pone.0057611 PMID: 23469035

5. Estill J, Egger M, Blaser N, Salazar-Vizcaya L, Garone D, Wood R, et al. (2013) Cost-effectiveness of point-of-care viral load monitoring of antiretroviral therapy in resource-limited settings: mathematical modelling study. AIDS 27: 1483–1492. doi: 10.1097/QAD.0b013e328360a4e5 PMID: 23462219

6. Phillips AN, Pillay D, Miners AH, Bennett DE, Gilks CF, Lundgren JD (2008) Outcomes from monitoring of patients on antiretroviral therapy in resource-limited settings with viral load, CD4 cell count, or clinical observation alone: a computer simulation model. Lancet 371: 1443–1451. doi: 10.1016/S0140-6736(08)60624-8 PMID: 18440426

7. Mermin J, Ekwaru JP, Were W, Degerman R, Bunnell R, Kaharuza F, et al. (2011) Utility of routine viral load, CD4 cell count, and clinical monitoring among adults with HIV receiving antiretroviral therapy in Uganda: randomised trial. BMJ 343: d6792. doi: 10.1136/bmj.d6792 PMID: 22074711

8. Kahn JG, Marseille E, Moore D, Bunnell R, Were W, Degerman R, et al. (2011) CD4 cell count and viral load monitoring in patients undergoing antiretroviral therapy in Uganda: cost effectiveness study. BMJ 343: d6884. doi: 10.1136/bmj.d6884 PMID: 22074713

9. UNITAID (2013) HIV/AIDS Diagnostics Technology Landscape. Semi-annual update. (2013). Geneva, Switzerland: UNITAID. Available: http://www.unitaid.eu/images/UNITAID_2013_Semi-annual_Update_HIV_Diagnostics_Technology_Landscape.pdf. Accessed 18 December 2014.

10. Rousseau C (2012) HIV Point of Care Viral Load Tests Development: Request for Applications. Available: http://mat1.gtimg.com/gongyi/2012/enzhenduandahui/16ChristineRousseau.pdf. Accessed 18 December 2014.

11. Bonner K, Mezochow A, Roberts T, Ford N, Cohn J (2013) Viral load monitoring as a tool to reinforce adherence: a systematic review. J Acquir Immune Defic Syndr 64: 74–78. doi: 10.1097/QAI.0b013e3182905ac PMID: 23774877

12. Gupta RK, Goodall RL, Ranopa M, Kityo C, Munderi P, Lyagoba F, et al. (2014) High rate of HIV resuppression after viral failure on first-line antiretroviral therapy in the absence of switch to second-line therapy. Clin Infect Dis 58: 1023–1026. doi: 10.1093/cid/cit933 PMID: 24352348

13. Hoffmann CJ, Charalambous S, Sim J, Ledwaba J, Schwikgard G, Chaisson RE, et al. (2009) Viremia, resuppression, and time to resistance in human immunodeficiency virus (HIV) subtype C during first-line antiretroviral therapy in South Africa. Clin Infect Dis 49: 1928–1935. doi: 10.1086/648444 PMID: 19911963

14. Orrell C, Harling G, Lawn SD, Kaplan R, McNally M, Bekker LG, et al. (2007) Conservation of first-line antiretroviral treatment regimen where therapeutic options are limited. Antivir Ther 12: 83–88. PMID: 17503751

15. Zoufaly A, Jochem J, Hammerl R, Nasiini N, Raymond Y, Burchard GD, et al. (2014) Virological failure after 1 year of first-line ART is not associated with HIV minority drug resistance in rural Cameroon. J Antimicrob Chemother. In press.

16. Maggioni F, Airoldi M, Kleinloog HD, Callegaro A, Ravasio V, Arici C, et al. (2007) Effect of adherence to HAART on virologic outcome and on the selection of resistance-conferring mutations in NNRTI- or PI-treated patients. HIV Clin Trials 8: 282–292. PMID: 17956829

17. Salazar-Vizcaya L, Blaser N, Gsponer T (2014) gems: Generalized multistate simulation model. Available: http://cran.r-project.org/web/packages/gems/index.html. Accessed 18 December 2014.

18. Blaser N, Salazar-Vizcaya L, Estill J, Zahnd C, Kalesan B, Egger M, et al. (2014) gems: an R package for Simulating from Disease Progression Models. J Stat Softw. In press. PMID: 25400517

19. Estill J, Twywa H, Egger M, Wandeler G, Feldacker C, Johnson LF, et al. (2014) Tracing of patients lost to follow-up and HIV transmission: Mathematical modelling study based on two large ART programs in Malawi. J Acquir Immune Defic Syndr. 65: e179–e186. doi: 10.1097/QAI.0b013e328360a4e5 PMID: 24326599

20. Brinkhof MWG, Boule A, Weigel R, Messou E, Mathers C, Orrell C, et al. (2009) Mortality of HIV-infected patients starting antiretroviral therapy in sub-Saharan Africa: comparison with HIV-unrelated mortality. PLoS Med 6: e1000066. doi: 10.1371/journal.pmed.1000066 PMID: 19399157

21. Actuarial Society of South Africa, Aids Committee (2011) ASSA2008 Model. Actuarial Society of South Africa. Available: http://www.actuarialsociety.org.za/Societyactivities/CommitteeActivities/AidsCommittee/Models.aspx. Accessed 18 December 2014.

22. World Health Organization. Global Burden of Disease 2004 update: Disability weights for diseases and conditions. Geneva, Switzerland: World Health Organization. Available: http://www.who.int/healthinfo/gbd2004_DisabilityWeights.pdf. Accessed 18 December 2014.
23. World Health Organization, Commission on Macroeconomics and Health (2001) Macroeconomics and health: investing in health for economic development. Report of the Commission on Macroeconomics and Health. Geneva, Switzerland: World Health Organization. Available: http://whqlibdoc.who.int/publications/2001/924154550x.pdf. Accessed 18 December 2014.

24. International Monetary Fund (2014) World Economic and Financial Surveys: World Economic Outlook Database. International Monetary Fund. Available: http://www.imf.org/external/pubs/ft/weo/2014/01/weodata/index.aspx. Accessed 18 December 2014.

25. Walensky RP, Ciaranello AL, Park J-E, Freedberg KA (2010) Cost-effectiveness of laboratory monitoring in sub-Saharan Africa: a review of the current literature. Clin Infect Dis 51: 85–92. doi: 10.1086/653119 PMID: 20482371

26. Bendavid E, Young SD, Katzenstein DA, Bayoumi AM, Sanders GD, Owens DK (2008) Cost-effectiveness of HIV monitoring strategies in resource-limited settings: a southern African analysis. Arch Intern Med 168: 1910–1918. doi: 10.1001/archinternmed.2008.1 PMID: 18809819

27. Bishai D, Colchero A, Durack DT (2007) The cost effectiveness of antiretroviral treatment strategies in resource-limited settings. AIDS 21: 1333–1340. PMID: 17545710

28. Vijayaraghavan A, Efrusy MB, Mazonson PD, Ebrahim O, Sanne IM, Santas CC (2007) Cost-effectiveness of alternative strategies for initiating and monitoring highly active antiretroviral therapy in developing world. J Acquir Immune Defic Syndr 46: 91–100. PMID: 17621241

29. Braithwaite RS, Nucifora KA, Yannioutsos CT, Musick B, Kimaiyo S, Diero L, et al. (2011) Alternative antiretroviral monitoring strategies for HIV-infected patients in east Africa: opportunities to save more lives? J Int AIDS Soc 14: 38. doi: 10.1186/1758-2652-14-38 PMID: 21801434

30. Hamers RL, Sawyer AW, Tuohy M, Stevens WS, Rinke de Wit TF, Hill AM (2012) Cost-effectiveness of laboratory monitoring for management of HIV treatment in sub-Saharan Africa: a model-based analysis. AIDS 26: 1663–1672. PMID: 22695297

31. Keebler D, Revill P, Braithwaite RS, Phillips AN, Blaser N, Borquez A, et al. (2014) Cost-effectiveness of different strategies to monitor adults on antiretroviral treatment: a combined analysis of three mathematical models. Lancet Glob Health 2: e35–e43. doi: 10.1016/S2214-109X(13)70048-2 PMID: 25104633

32. Centers for Disease Control and Prevention (CDC) (2013) Impact of an innovative approach to prevent mother-to-child transmission of HIV—Malawi, July 2011-September 2012. MMWR Morb Mortal Wkly Rep 62: 148–151. PMID: 23446514

33. Granich RM, Gilks CF, Dye C, Williams BG (2009) Universal voluntary HIV testing with immediate antiretroviral therapy as a strategy for elimination of HIV transmission: a mathematical model. Lancet 373: 48–57. doi: 10.1016/S0140-6736(08)61697-9 PMID: 19038438

34. Wandelger G, Keiser O, Pfeiffer K, Pestilli S, Fritz C, Labhardt ND, et al. (2012) Outcomes of antiretroviral treatment programs in rural Southern Africa. J Acquir Immune Defic Syndr 59: e9–e16. doi: 10.1097/QAI.0b013e31823edb6a PMID: 22067665

35. Tweya H, Feldacker C, Estill J, Ahan N, Ng'ambi W, Ben-Smith A, et al. (2013) Are they really lost? “true” status and reasons for treatment discontinuation among HIV infected patients on antiretroviral therapy considered lost to follow up in Urban Malawi. PLoS One 8: e75761. doi: 10.1371/journal.pone.0075761 PMID: 24086627

36. Jani IV, Sitoe NE, Alfai ER, Chongo PL, Quevedo JI, Rocha BM, et al. (2011) Effect of point-of-care CD4 cell count tests on retention of patients and rates of antiretroviral therapy initiation in primary health clinics: an observational cohort study. Lancet 378: 1572–1579. doi: 10.1016/S0140-6736(11)61052-0 PMID: 21951656

37. Kimmel AD, Weinstein MC, Anglaret X, Goldie SJ, Losina E, Yazdanpanah Y, et al. (2010) Laboratory monitoring to guide switching antiretroviral therapy in resource-limited settings: clinical benefits and cost-effectiveness. J Acquir Immune Defic Syndr 54: 258–268. doi: 10.1097/QAI.0b013e3181d0db97 PMID: 20404739

38. Walensky RP, Wood R, Ciaranello AL, Paltiel AD, Lorenzana SB, Anglaret X, et al. (2010) Scoping up the 2010 World Health Organization HIV Treatment Guidelines in resource-limited settings: a model-based analysis. PLoS Med 7: e1000382. doi: 10.1371/journal.pmed.1000382 PMID: 21209794

39. Egger M, Spycher BD, Sidle J, Weigel R, Geng EH, Fox MP, et al. (2011) Correcting mortality for loss to follow-up: a nomogram applied to antiretroviral treatment programmes in sub-Saharan Africa. PLoS Med 8: e1000390. doi: 10.1371/journal.pmed.1000390 PMID: 21267057