Web appendix 1: Modelling viral load

Viral load in the mathematical model

The viral load value is a key parameter of the model which affects disease progression and probability of HIV transmission. The viral load is defined on two levels, the underlying viral load and the observed viral load:

Underlying viral load

The underlying viral load is a continuous function over time from ART start until time of death (t_d), which is defined as a piecewise linear (on log10 scale) function, depending on events that the patient can experience. Three key events are defined for each patient which determine the viral load trajectory: first-line true treatment failure (t_1), switch to second-line ART (t_S) and second-line true treatment failure (t_2). The failure times are sampled from specific hazard functions following ART start (1st-line) or switch (2nd-line), and the time to switch depends on the monitoring strategy. Three set point values are also defined for the viral load: baseline (V_B), suppressed (V_S) and failing (V_F). Each value is sampled from a normal distribution (on log10 scale) with given mean and standard deviation.

The viral load starts from V_B, and unless first-line failure occurs within the first 6 months, decreases linearly to V_S within the first 3 months. At the time of failure, it increases to the level of V_F within one month. The same procedure is repeated when the patient switches to 2nd-line ART. Formally, the viral load trajectory can be defined as follows:

$$V(t) = \begin{cases} V_B - 4(V_B - V_S)\{t_1 \geq 0.5\}, & 0 \leq t < 0.25 \\ V_S, & 0.25 \leq t < t_1 \\ V_S + 12(V_F - V_S), & t_1 \leq t < t_1 + 0.083 \\ V_F, & t_1 + 0.083 \leq t < t_S \\ V(t_S) - 4(V(t_S) - V_S)\{t_2 - t_S \geq 0.5\}, & t_S \leq t < t_S + 0.25 \\ V_S, & t_S + 0.25 \leq t < t_2 \\ V_S + 12(V_F - V_S), & t_2 \leq t < t_2 + 0.08 \\ V_F, & t_2 + 0.25 \leq t < t_d \end{cases}$$

We assumed no variability in the viral load value within the sections.

Observed viral load

The second level of viral load is the observed viral load. This is a discrete variable which is defined only in strategies with viral load monitoring. A measurement *i* at time t_i will have an observed viral load value v_i, the type of which depends on the type of the viral load test. In this study, we use a qualitative viral load test with a limit of detection (LOD) V_{LOD}, i.e. $v_i \in \{0,1\}$ for all *i*, where 0 represents undetectable and 1 detectable. The value depends on the value of the underlying viral load at the same time, $V(t_i)$ in the following way:

$$v_i \begin{cases} = 1, & V(t_i) \geq V_{LOD} \\ \sim \text{Bin}\left(1, p(V_{LOD}) + (1-p)\dfrac{V(t_i) - V_S}{V_{LOD} - V_S}\right), & V_S \leq V(t_i) < V_{LOD} \end{cases}$$

where $p(V_{LOD})$ is the probability of having a detectable viral load with the particular LOD if the underlying viral load is suppressed. In other words, the viral load test is assumed to be 100% sensitive, but not specific, for viral load values that are above the LOD. The probability of having a detectable viral load when the underlying viral load is below the LOD grows gradually from a specified parameter (in case of the underlying viral load being suppressed) up to 1. This demonstrates both the currently

unexplained short-term low-level viraemia (‘blips’ [1, 2]) and temporary peaks in the viral load due to e.g. non-adherence. We call these viral load measurements ‘detectable viral loads of unknown origin’ (DVU).

The observed viral load values are used to decide when the patient should be switched to second-line ART. According to our definition of virologic failure, two successive detectable measurements are needed. However, it is also possible to have two consecutive DVUs, and therefore viral load monitoring is not a 100% specific strategy to detect true treatment failure.

Statistical analyses of the viral load

Most parameters we used for the model are based on statistical analyses that were used to parameterise two earlier versions of the model. The methods and results of these analyses have been presented elsewhere [3]. Since the focus of this paper is on the detection of treatment failure with viral load tests, we updated our true treatment failure estimates and analysed the viral load data in more detail.

Each viral load value on first-line ART after at least 3 months from ART start was classified into the following categories:

- Suppressed (<125 copies/ml)
- Failing (≥125 copies/ml, at least 2 successive measurements, remaining above 125 copies/ml until patient switches to second-line therapy)
- DVU (≥125 copies/ml, returns below 125 copies/ml before switching)
- Censored (≥125 copies/ml in last measurement, preceded by suppressed VL)

We performed a regression analysis for true treatment failure. We assumed that treatment failure happened between the last suppressed and the first failing

measurement, at a time point which was calculated assuming a linear (log10 scale) increase in viral load between the observed values (Supplementary Table 2, including also parameter estimates from our previous analyses and the literature). After that, we calculated the probabilities of DVU with different LODs (Supplementary Table 1).

Web appendix 2: Modelling HIV transmission

We modelled the number of expected new infections by defining the viral load trajectories for each patient over time (see Web appendix 1 for details) and calculating the probability of infecting the partner during each sex act. We assumed a prevalence of 15% for new partners (i.e. partnerships that start during ART). Among the current partners at the time of ART start, we assumed the prevalence to be 30%. The mean duration of the partnership was chosen to be 1 year for the first partner (from time of ART start) and 2 years for the subsequent partners, and we assumed a constant rate of partner change. Each patient was randomly assigned to one of four sexual behaviour groups with frequencies of 10, 25, 50 or 100 sex acts per year.

We used a relation of individual viral load and transmission probability introduced by Quinn *et al* [4] to estimate the number of new infections each simulated patient could cause. The method calculates the probability of transmission p in a single sex act as

$$p(vl) = p(vl_0)k^{\log_{10}\frac{vl}{vl_0}} \qquad (1)$$

where vl is the viral load on the absolute scale at the time of transmission, vl_0 a reference viral load for which the transmission probability is known, and k a constant. According to Wilson *et al* [5], we chose $k = 2.45$, $vl_0 = 10$ copies/ml and $p(vl_0) = 4.3*10^{-5}$ (male to female) or $2.2*10^{-5}$ (female to male).

Supplementary Table 1. Probability of detectable viral load of unknown origin with different limits of detection.

LOD (copies/ml)	Probability of DVU
125	6.7%
400	4.4%
1 000	3.3%
5 000	2.2%
10 000	1.9%

Data from Gugulethu and Khayelitsha antiretroviral therapy (ART) programmes (Cape Town, South Africa) were analysed. Detectable viral load of unknown origin (DVU) was defined as a viral load above the detection limit which returned to below 125 copies/ml while still on first-line ART.

LOD, limit of detection

Supplementary Table 2. Model parameters.

Outcome	Source	Statistical model	Starting	Value (95% CI)	Dimension	Risk
1) Time to virologic failure						
(a) First-line ART; second-line ART with immediate switch	Cohorts	Parametric Weibull	3 months from ART start	0.57 (0.52-0.63) 2.75 (2.29-3.31)	Shape Scale (100 years)	3.4% fail by 1 year after ART start
(b) Resistance penalty	[6]	*)	n/a	0.05 (0.00-0.20)	Decrease in efficacy	n/a
2) Time to immunologic failure						
(a) After virologic failure	Cohorts	Parametric exponential	Virologic failure	0.08 (0.06-0.10)	Rate ($years^{-1}$)	7.6% fail by 1 year after virologic failure
(b) Before virologic failure	Cohorts	Parametric Weibull	3 months from ART start	0.22 (0.20-0.25) 5.46 (3.14-9.51)	Shape Scale (10^6 years)	3.0% fail by 1 year after ART start
3) Time to clinical failure						
(a) Without virologic or immunologic failure	[7]	Parametric exponential	ART start	0.004	Rate ($years^{-1}$)	0.4% fail by 1 year after ART start
(b) Extra hazard after immunologic failure	[7, 8]	Cox regression	Immunologic failure	3.3	HR, constant over time	n/a
(c) Extra hazard after virologic failure	[7]	Cox regression	Virologic failure	2	HR, constant over time	n/a
4) Time to death and LTFU						
(a) Non-HIV related mortality	ASSA2008 [9]	No specific model (age- and gender-specific rates)	n/a	**)	n/a	21% of men and 13% of women die by age of 50
(b) Observed mortality	Cohorts	No specific model (competing risk analysis)	ART start	***)	n/a	6.5% die by 1 year after ART start
(c) Observed LTFU	Cohorts	No specific model (competing risk analysis)	ART start	***)	n/a	4.0% die by 1 year after ART start
(d) Mortality among LTFU	Analysis 4c, [10]	No specific model (theoretical calculation)	n/a	***)	n/a	n/a
(e) HIV-related mortality	Analyses 4a-4d	Theoretical calculation, double Weibull****)	ART start	0.88 (0.88-0.90) 0.35 (0.32-0.39) 1.00 (1.00-1.00) 64.60 (54.52-76.55) 0.08 (0.08-0.08)	Shape 1 Scale 1 (years) Shape 2 Scale 2 (years) Weight (1^{st} component)	8.8% die by 1 year after ART start
(f) Extra hazard after clinical failure	assumption	Cox regression	Clinical failure	2	HR, constant over time	n/a
(g) Extra hazard after immunologic failure	Cohorts	Cox regression	Immunologic failure	1.76 (1.16-2.68)	HR, constant over time	n/a
(h) Extra hazard after virologic failure	Cohorts	Cox regression	Virologic failure	1.26 (0.86-1.85)	HR, constant over time	n/a

CI, confidence interval; ART, antiretroviral therapy; HR, hazard ratio; ASSA, Actuarial Society of South Africa; LTFU, loss to follow-up; n/a, not applicable

*) Relative decrease in second-line efficacy per year spent on failing first-line ART

**) Age- and sex-specific mortality rates

***) Observed mortality and LTFU rates on successful first-line ART were calculated from the data and used, together with background mortality and expected mortality among patients LTFU, to calculate the corrected HIV-related mortality for the cohort

****) Weighted sum of two Weibull distributions

Supplementary Table 3. Costs of antiretroviral therapy, appointments, measurements and new infections.

	Cost	Source
Antiretroviral therapy (US$/year)		
ZDV+3TC+NVP (300/150/200)*)	134	[11]
TDF+3TC+EFV (300/300/600)*)	159	[11]
(ZDV+3TC)+LPV/r (300/150/200/50)*)	483	[11]
(TDF+3TC)+LPV/r (300/300/200/50)*)	448	[11]
Appointments (US$/visit)		
incl. staff, capital, admin	9.83	[12]
Monitoring (US$/test)		
CD4 test (current practice)	5.43	[13]
Viral load test, LOD <1 000 copies/ml	10-20	assumption
Viral load test, LOD $\geq$1 000 copies/ml	5-20	assumption
Transmission (US$)		
Average lifetime cost of one new infection	6400	assumption**)

All costs have been converted to 2012 US dollars using a 3% annual inflation rate.

ZDV, zidovudine; 3TC, lamivudine; NVP, nevirapine; TDF, tenofovir; EFV, efavirenz; LPV/r, ritonavir-boosted lopinavir; LOD, limit of detection

*) The averages of the most common first-line (ZDV+3TC+NVP, TDF+3TC+EFV) and second-line (ZDV+3TC+LPV/r, TDF+3TC+LPV/r) regimens were used in the model.

**) Assuming each newly infected patient lives 30 years after HIV infection and starts ART 5 years after infection.

Supplementary Table 4. Summary of published studies on cost-effectiveness of viral load monitoring. We included all relevant studies from a previous review [14] and selected relevant studies published thereafter. We chose from each study the CD4 and viral load monitoring strategies that were closest to our study.

Study	Design	CD4 monitoring strategy		VL monitoring strategy		Lifetime lived/lost (months/patient)			Costs (USD/patient)		ICER (USD/ year)	Factors included	
		Δt	Failure criteria	Δt	Failure criteria	Type	CD4	VL	CD4	VL		Adherence	Transm.
Bendavid *et al* [15]	model	6m	-50%	6m	1000	LE	73.7	75.7	3653	4552	5421	no	no
Bishai *et al* [16]	model	6m	<200/µl, no +50	1+6m*	BL/1stm<10	QALE	56.5	57.1	2500	3230	14600	no	no
Braithwaite *et al* [17]	model	6m	WHO	6m	500	QALE	132.0	137.4	12777	17087	9682	n/r	no
Hamers *et al* [18]	model	6m	-50% or <BL	6m	400	LE	64.9	66.3	3397	3407	86	n/r	no
Kahn *et al* [13]	RCT	3m	persistent decline	3m	500	DALT	-41.9	-38.6	6267	7692	5180	no	no
Kimmel *et al* [6]	model	6m	-50%	6m	+1log$_{10}$ or BL	LE	149.0	231.4	10980	14190	1793	no	no
Phillips *et al* [7]	model	6m	-33% in 6m	6m	10000	QALE	110.4	115.4	2792	3950	2757	yes	no
Vijayaraghavan *et al* [19]	model	6m	-50% or <200/µl	3m	400	QALE	116.4	138.4	10624	20349	5314	no	yes

VL, viral load (monitoring); RCT, randomised controlled trial; Δt, frequency of measurements (in months); BL, baseline; LE, life expectancy; QALE, quality-adjusted life expectancy; DALT, disability-adjusted lifetime; ICER, incremental cost-effectiveness ratio of viral load compared to CD4 monitoring; n/r, not reported; Transm., Transmission

*First test 1 month after start, after that 6 month intervals

References

1. Nettles RE, Kieffer TL, Kwon P, Monie D, Han Y, Parsons T, *et al*. Intermittent HIV-1 viremia (Blips) and drug resistance in patients receiving HAART. *JAMA* 2005; **293**:817-829.
2. Gallant JE. Making sense of blips. *J Infect Dis* 2007; **196**:1729-1731.
3. Estill J, Aubrière C, Egger M, Johnson L, Wood R, Garone D *et al*. Viral load monitoring of antiretroviral therapy, cohort viral load and HIV transmission in Southern Africa: A mathematical modelling analysis. *AIDS* 2012; **26**:1403-1413.
4. Quinn TC, Wawer MJ, Sewankambo N, Serwadda D, Li C, Wabwire-Mangen F, *et al*. Viral load and heterosexual transmission of human immunodeficiency virus type 1. *N Engl J Med* 2000; **342**:921-929.
5. Wilson DP, Law MG, Grulich AE, Cooper DA, Kaldor JM. Relation between HIV viral load and infectiousness: a model-based analysis. *Lancet* 2008; **372**:314-320.
6. Kimmel AD, Weinstein MC, Anglaret X, Goldie SJ, Losina E, Yazdanpanah Y, *et al*. Laboratory monitoring to guide switching antiretroviral therapy in resource-limited settings: clinical benefits and cost-effectiveness. *J Acquir Immune Defic Syndr* 2010; **54**:258-268.
7. Phillips AN, Pillay D, Miners AH, Bennett DE, Gilks CF, Lundgren JD. 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* 2008; **371**:1443-1451.
8. Walensky RP, Wood R, Ciaranello AL, Paltiel AD, Lorenzana SB, Anglaret X, *et al*. Scaling up the 2010 World Health Organization HIV Treatment Guidelines in resource-limited settings: a model-based analysis. *PLoS Med* 2010; **7**:e1000382.
9. ASSA2008 Model. Actuarial Society of South Africa; 2011. http://aids.actuarialsociety.org.za/ASSA2008-Model-3480.htm. [Accessed 21 September 2012].
10. Egger M, Spycher BD, Sidle J, Weigel R, Geng EH, Fox MP, *et al*. Correcting mortality for loss to follow up: A nomogram applied to antiretroviral treatment programmes in sub-Saharan Africa. *PLoS Med* 2011; **8**:e1000390.
11. Antiretroviral (ARV) Ceiling Price List. Clinton Health Access Initiative; 2012. http://d2pd3b5abq75bb.cloudfront.net/2012/07/12/15/03/07/163/CHAI_ARV_Ceiling_Price_List_May_2012.pdf. [Accessed 4 October 2012].
12. Medina Lara A, Kigozi J, Amurwon J, Muchabaiwa L, Nyanzi Wakaholi B, Mujica Mota RE, *et al*. Cost effectiveness analysis of clinically driven versus routine laboratory monitoring of antiretroviral therapy in Uganda and Zimbabwe. *PLoS One* 2012; **7**:e33672.
13. Kahn JG, Marseille E, Moore D, Bunnell R, Were W, Degerman R, *et al*. CD4 cell count and viral load monitoring in patients undergoing antiretroviral therapy in Uganda: cost effectiveness study. *BMJ* 2011; **343**:d6884.
14. Walensky RP, Ciaranello AL, Park JE, Freedberg KA. Cost-effectiveness of laboratory monitoring in sub-Saharan Africa: a review of the current literature. *Clin Infect Dis* 2010; **51**:85-92.
15. Bendavid E, Young SD, Katzenstein DA, Bayoumi AM, Sanders GD, Owens DK. Cost-effectiveness of hiv monitoring strategies in resource-limited settings: A southern african analysis. *Arch Intern Med* 2008; **168**:1910-1918.
16. Bishai D, Colchero A, Durack DT. The cost effectiveness of antiretroviral treatment strategies in resource-limited settings. *AIDS* 2007; **21**:1333-1340.
17. Braithwaite RS, Nucifora KA, Yiannoutsos CT, Musick B, Kimaiyo S, Diero L, *et al*. Alternative antiretroviral monitoring strategies for HIV-infected patients in east Africa: opportunities to save more lives? *J Int AIDS Soc* 2011; **14**:38.

18. Hamers RL, Sawyer AW, Tuohy M, Stevens WS, de Wit TF, Hill AM. Cost-effectiveness of laboratory monitoring for management of HIV treatment in sub-Saharan Africa: a model-based analysis. *AIDS* 2012; **26**:1663-1672.
19. Vijayaraghavan A, Efrusy MB, Mazonson PD, Ebrahim O, Sanne IM, Santas CC. Cost-effectiveness of alternative strategies for initiating and monitoring highly active antiretroviral therapy in the developing world. *J Acquir Immune Defic Syndr* 2007; **46**:91-100.