An Algorithm to Optimize Viral Load Testing in HIV-Positive Patients With Suspected First-Line Antiretroviral Therapy Failure in Cambodia

Lutgarde Lynen, MD, Sokkab An, MD,† Olivier Koole, MD,* Sopheak Thai, MD,† Seilavath Ros, MD,† Paul De Munter, MD,‡ Delphine Sculier, MD,* Line Arnould, MD,* Katrien Fransen, MSc,* Joris Menten, MSc,* Marleen Boelaert, MD, PhD,* Jef Van den Ende, MD, PhD,* and Robert Colebunders, MD, PhD*§*

Objective: To develop an algorithm for optimal use of viral load testing in patients with suspected first-line antiretroviral treatment (ART) failure.

Methods: Data from a cohort of patients on first-line ART in Cambodia were analyzed in a cross-sectional way to detect markers for treatment failure. Markers with an adjusted likelihood ratio <0.67 or >1.5 were retained to calculate a predictor score. The accuracy of a 2-step algorithm based on this score followed by targeted viral load testing was compared with World Health Organization criteria for suspected treatment failure.

Results: One thousand eight hundred three viral load measurements of 764 patients were available for analysis. Prior ART exposure, CD4 count below baseline, 25% and 50% drop from peak CD4 count, hemoglobin drop of ≥ 1 g/dL, CD4 count <100 cells per microliter after 12 months of treatment, new onset of papular pruritic eruption, and visual analog scale $<95\%$ were included in the predictor score. A score ≥ 2 had the best combination of sensitivity and specificity and required confirmatory viral load testing for only 9% of patients. World Health Organization criteria had a similar sensitivity but a lower specificity and required viral load testing for 24.9% of patients.

Conclusion: An algorithm combining a predictor score with targeted viral load testing in patients with an intermediate probability of failure optimizes the use of scarce resources.

Key Words: Cambodia, first-line antiretroviral treatment, low- and middle-income countries, second-line ART, targeted viral load, treatment failure

(*J Acquir Immune Defic Syndr* 2009;52:40–48)

INTRODUCTION

Since 2003, access to antiretroviral therapy (ART) for persons living with HIV in low- and middle-income countries (LMICs) has improved dramatically.[1] In some countries, the coverage of ART for adults and children with advanced HIV is as high as 80%.[2] In Cambodia, a country with one of the highest HIV prevalence in Southeast Asia,[3] 26,664 (67% of estimated need) persons living with HIV were on ART by the end of 2007.[1]

In high-income countries, treatment failure is detected mainly by monitoring viral load.[4,5] In resource-limited settings, many obstacles make this kind of monitoring difficult.[6,7] Indeed, viral load testing is costly and technologically complex.[8–10] Therefore, the World Health Organization (WHO) has proposed guidelines for switching to second-line ART based on clinical and immunological criteria and assessment of adherence.[11,12] Preliminary studies have shown that immunological and clinical criteria alone are not sensitive enough and have a low positive predictive value (PPV).[13–15] We and others have tried to develop an alternative set of criteria with higher diagnostic accuracy.[16,17] We proposed an algorithm based on early clinical indicators, CD4 count, drug history, and adherence data,[16] but this empirical algorithm did not perform well in South Africa.[18]

We did further work on data obtained from an ART cohort study in Cambodia and developed a scoring system that identifies patients with a low, intermediate, and high probability of treatment failure and, as such, allows restricting viral load testing to those patients with an intermediate risk of failure. The goal of this scoring system is to minimize the number of inappropriate switches to second-line treatment while limiting monitoring costs. In this article, we present the development of the scoring system, as well as an evaluation of its accuracy.

Received for publication September 17, 2008; accepted January 5, 2009.

From the *Departments of Clinical Sciences, Public Health, and Microbiology, Institute of Tropical Medicine, Antwerp, Belgium; †Infectious Diseases Department, Sihanouk Hospital Center of HOPE, Phnom Penh, Cambodia; ‡Department of General Internal Medicine, University Hospitals Leuven, Belgium; and §Faculty of Medicine, University of Antwerp, Belgium.

This research was funded by the European Commission project Europe AID, CONTRACT SANTE/2002/045-809: Clinical research/activity II B 5 and by the Belgian DGDC (Directory General of Development Cooperation) through a framework agreement with ITM (Institutional Strengthening project HOPE Cambodia 95900 and 95901).

Parts of the data were presented at the 4th IAS Conference on HIV Pathogenesis and Treatment, 22–25 July, 2007, Sydney, Australia.

Conflict of interest: none.

Correspondence to: Lutgarde Lynen, MD, Institute of Tropical Medicine, Nationalestraat 155, 2000 Antwerpen, Belgium (e-mail: llynen@itg.be).

METHODS

The study was conducted in Phnom Penh, Cambodia, at the Sihanouk Hospital Center of HOPE (SHCH), a private not-for-profit hospital that, by the end of 2007, was providing care for 2024 patients with HIV, including 1503 on ART. Patients were seen every month and treated according to WHO guidelines.[11,12] CD4 counts were performed every 6 months by FACSCount (Becton Dickinson, Franklin Lakes, NJ). Viral load monitoring was not part of routine follow-up.

Inclusion

All adult patients (≥18 years old) on first-line ART who presented at SHCH between December 2005 up to and including May 2007 were eligible. During this period, a viral load was performed every 6 months together with the CD4 count for every patient included. For the analysis, we included all patients who had been on ART for at least 12 months, but we excluded those patients who had only a viral load at 6 months.

Data Collection

Since 2003, data for every patient visit at SHCH were entered into an electronic database (Microsoft Access 2003). These data included age; sex; WHO stage; previous exposure to ART; previous and current opportunistic infections; any WHO stage 2, 3, or 4 condition (as defined in the 2003 WHO staging system)[19]; prophylactic medications; body weight before illness; current body weight and body mass index [weight in kilogram ÷ (height in meter)^2]; basic laboratory data (total lymphocyte count, CD4 count, hemoglobin, renal function, and liver function tests); ART regimen; side effects, reasons for substituting or halting drugs; and adherence indicators. Adherence was measured by trained counselors using a questionnaire adapted from the Simple Medication Adherence Questionnaire (SMAQ),[20] which consisted of 7 binary questions concerning adherence during the previous 3 months. The adapted SMAQ score ranged from 0 to 7, with 0 corresponding to 100% adherence. Second, counselors asked patients to indicate on a 10-cm-long visual analog scale (VAS) how many drugs he/she took during the previous 30 days.[21]

Laboratory Analysis

Hemoglobin and total lymphocyte counts were measured every 3 months using a Sysmex KX-21 (Sysmex Corporation, Kobe, Japan). Every 6 months, a CD4 count was performed at the National Institute of Public Health in Phnom Penh, Cambodia, by FACSCount (Becton Dickinson).

Samples for viral load were taken at 6, 12, 18, 24, 30, 36, 42, and 48 months, stored at −70°C, and sent on dry ice to the Institute of Tropical Medicine, Antwerp, Belgium, where the viral load was measured using the Cobas Ampliprep/Cobas Amplicor HIV1 Monitor test (v1.5) (Roche Diagnostics Corporation, Branchburg, NJ). Viral load was expressed as copies per milliliter.

Data Analysis

Virological failure was defined as a viral load above 1000 copies per milliliter at least once, instead of the classical cutoff of 50 or 400 copies per milliliter. The cutoff 1000 copies per milliliter was chosen to avoid inclusion of viral load "blips" as treatment failure. Variables that were assessed as possible predictors for failure were (1) changes, over the previous 6 months, in hemoglobin, body weight, total lymphocyte count, and (re)appearance of clinical symptoms (WHO stage 2, 3, and 4 conditions); (2) percent decrease in CD4 count from peak value, absolute CD4 count, adherence (VAS and SMAQ), and duration of ART treatment; and (3) sex, age, previous ART exposure, and number of drug substitutions ever (irrespective of interruption of ART).

To develop a scoring system, we used the Spiegelhalter and Knill-Jones method adapted by Berkley et al.[22–24] Continuous variables were dichotomized as guided by Receiver Operating Characteristic curves, with the optimal cutoff at the point with the highest sum of sensitivity and specificity. The cutoffs were rounded to values that are easy to use in clinical practice. In case no cutoff was superior, we used the median. For some variables, we retained 2 cutoff points to capture the full amount of information. For all categories of possible predictors of treatment failure, the crude likelihood ratio (LHR) was calculated using a continuity correction for standard formulas.[25] Variables associated with treatment failure with a crude LHR ≥ 2 or ≤0.5 were selected for inclusion in the scoring system. The LHRs were then adjusted for correlations between the predictors in a multivariate logistic regression model, in which we retained variables with an adjusted LHR of >1.5 or <0.67.

We developed a scoring system by first computing a predictor score defined as the natural logarithm of the adjusted LHR for each predictor category (with a value of 0 assigned to missing data) and by rounding this result to the nearest integer. Summing the predictor scores of all the risk factors presented by a patient gave his/her total predictor score. A more extensive explanation of the Spiegelhalter and Knill-Jones method has been published elsewhere.[22,23] We used visits as units of observation. We assessed the diagnostic accuracy of this new scoring system by calculating the proportion of virological failure and the observed sensitivity, specificity, PPV, and negative predictive value of total predictor scores in our sample. Confidence intervals for the LHRs, sensitivity, specificity, PPV, and NPV were calculated using robust standard errors, taking into account the intra-patient clustering.

We distinguished 3 probability categories of failure based on PPV: low (<10%), intermediate (10%–90%), and high (>90%). For comparison, we also calculated the diagnostic accuracy of the clinical and immunological WHO criteria for treatment failure in our sample. Two definitions were used: "Stringent" WHO criteria (for patients on ART for at least 6 months): new or recurrent WHO stage 4 conditions, a CD4 count below baseline, a CD4 decrease of 50% from the peak CD4 count during treatment, or a CD4 count below 100 cells per microliter at ≥12 months. Lenient WHO criteria: same as the stringent criteria plus new or recurrent WHO stage 3 conditions.

With an adapted nominal group technique,[26] we developed, with SHCH clinicians, a 2-step algorithm using the newly developed scoring system followed by targeted viral load testing. All statistical analyses were performed using

Stata software, version 9.2. (Stata Corporation, College Station, TX).

Ethical Considerations

Only patients who gave written informed consent were enrolled in the study. The study protocol was approved by the Institutional Review Board of the Institute of Tropical Medicine in Antwerp and by the National Ethics Committee in Cambodia.

RESULTS

Characteristics of the Study Population

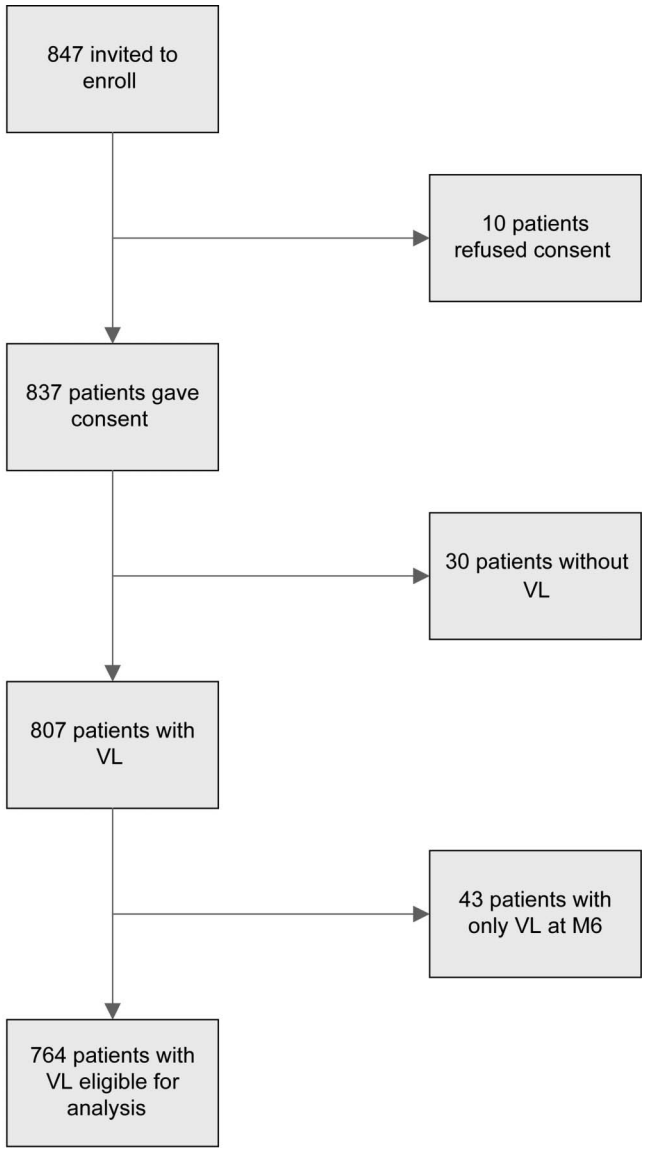
Between November 2005 and May 2007, 847 patients were invited to participate in the study and 837 (99%) gave informed consent (Fig. 1). Thirty patients without any viral load measurement and 43 patients who had only a viral load at month 6 were excluded from the analysis. Characteristics of the study population (n = 764) are summarized in Table 1. Nineteen patients died and 13 were lost to follow up over the 18-month study period. The median time on ART at the time of the viral load assessment was 18.3 months (interquartile range: 12.3–24.4 months). A total of 764 patients and a total of 1803 patient visits were included in the development of the scoring system. Sixty patients had at least 1 detectable viral load >1000 copies per milliliter (7.9%) at 1 time point in the follow-up. Eighty-seven samples from 1803 viral load measurements (4.8%) had a viral load >1000 copies per milliliter.

Predictors of Virological Failure

The frequency of virological failure by risk factor and the crude LHRs are shown in Table 2. Table 3 shows the adjusted LHRs and the resulting score. Predictors of a viral load >1000 copies per milliliter were prior ART exposure, CD4 count below baseline, a 25% or 50% drop from the peak CD4 count, a hemoglobin drop of ≥1 g/dL, an absolute CD4 count of <100 cells per microliter, a new onset of papular pruritic eruption, and VAS <95%. SMAQ was not predictive of treatment failure.

The total predictor score per patient ranged from 0 to 5. The percentage of virological failure in the different score groups is shown in Figure 2. Using as cutoff for virological failure a viral load >10,000 copies per milliliter (as proposed by WHO),[12] the same predictors and predictor scores were obtained (data not shown).

The sensitivity, specificity, and PPV of the different score cutoffs to predict viral loads >1000 copies per milliliter in comparison with the stringent and lenient WHO criteria are shown in Table 4. A score ≥2 seems to provide the optimal combination of sensitivity and specificity if we assume that the medical consequences of a false-positive or false-negative classification have a similar weight. This cutoff has a sensitivity (41.4%) similar to the WHO lenient criteria but has a better specificity (92.6% versus 75.9%). Based on the PPV, we distinguish 3 risk categories for virological failure: score 0–1, low probability (n = 1640, 91%); score 2–4, intermediate probability (n = 160, 8.7%); and score ≥5, high probability (n = 3, 0.2%).



VL = Viral load

M = Month

FIGURE 1. Flow chart of patients enrolled in the study.

In a consensus meeting with SHCH physicians, we developed a 2-step algorithm based on a predictor score (Fig. 3A) to target the viral load testing and applied this algorithm to the original dataset (Fig. 3B). Two patient categories were considered not to require viral load measurements: those with a low probability (score 0 or 1) and those with a high probability (score ≥5) of virological failure. Together, they accounted for 91% of all observations. For the remaining 9% with an intermediate probability of failure, it was suggested to perform viral load testing to confirm who is on a truly failing regimen. The sensitivity of this algorithm was similar to the lenient WHO criteria (41.4%) but required less viral load measurements (9% versus 24.9% following WHO lenient criteria, Fig. 3C).

TABLE 1. Overall Patient Demographics and Baseline (Pre-ART) Characteristics

Characteristics	n = 764*
Sex: male, n (%)	392 (51.3)
Age (yrs): median (range)	34 (18–68)
Weight (kg): mean (SD)	48.8 (9.3)
Height (cm): mean (SD)	160 (8)
Body mass index (score): mean (SD)	19.1 (3.1)
CD4 (cells/μL): median (IQR)	55 (16–161)
Total lymphocyte count (cells/μL): median (IQR)	1270 (900–1760)
Hemoglobin (g/dL): mean (SD)	11.2 (2.1)
ART experience: n (%)	59 (7.7)
WHO stage at initiation, n (%)	
Stage 1	4 (0.5)
Stage 2	40 (5.2)
Stage 3	301 (39.4)
Stage 4	417 (54.6)
Missing	2 (0.3)
Regimen at initiation, n (%)	
D4T-3TC-NVP	580 (75.9)
ZDV-3TC-NVP	45 (5.9)
D4T-3TC-EFV	111 (14.5)
ZDV-3TC-EFV	27 (3.5)
Other	1 (<0.2)

*All patients with at least 1 viral load measurement after month 6 and at least 12 months on ART.

3TC, lamivudine; D4T, stavudine; EFV, efavirenz; IQR, interquartile range; NVP, nevirapine; ZDV, zidovudine.

Wherever viral load assays are easily accessible and affordable, a more sensitive approach could be adopted. If viral load testing is performed in all patients with a score of ≥1, this would result in a better sensitivity (63.2%), but many more assays would be required (636 or 35.3% of visits).

DISCUSSION

Our research showed that a scoring system based only on clinical, immunological, and adherence data but without viral load testing was inadequate to predict first-line treatment failure. A threshold score of ≥2 had a sensitivity of 41.4% and a specificity of 92.6%, with a PPV of 22.1%. With a prevalence of failure fixed at the observed level of 4.8%, this means that for every 100 visits, 3 treatment failures will not be detected and 7 premature switches will occur. Therefore, we developed a 2-step algorithm based on the score followed by viral load testing in those with an intermediate risk. This algorithm reduced the false-positive rate to 0% and the overall misclassification to 3% (false-negatives), whereas a viral load was needed in only 9% of patient visits. These features compare favorably with the WHO criteria for the assessment of treatment failure. We believe that our approach is a feasible and effective strategy in LMICs. Nonetheless, the performance of our algorithm is likely to be site and time dependent. The median time of follow-up on ART in this Cambodian cohort was 18 months, and the number of patients failing first-line ART at 18 months was less than 10%. Although this low failure rate confirmed the high efficacy of first-line ART as described in other studies in LMIC,[17,27–30] it also contributed to the high negative predictive value of our scoring system. Failure rates and follow-up duration may vary though, and validation studies of the algorithm are therefore needed in other populations, including in cohorts with longer periods of follow-up.

Another limitation of our study is that the definition of virological failure was based on 1 viral load measurement only. A study in South Africa has shown that 53% of patients with a viral breakthrough returned to undetectable viral loads after a targeted adherence intervention.[31] Studies have shown that in LMICs, between 20% and 50% of patients with a detectable viral load have no major resistance mutations.[17,32] WHO recommends that patients should only be switched to second-line ART if a CD4 count decrease is confirmed by a repeat CD4 count and if the CD4 count is below 200 cells per microliter.[12] In our study, we did not repeat CD4 counts. When we restricted switching only to patients with CD4 count <200 cells per microliter, the WHO criteria (stringent) would still have a false-positive rate of 82.9%.

We used patient visits, not patients, as units of observation in our analysis. In certain patients, up to 3 viral load time points were available for analysis. While we could have restricted the analysis to 1 viral load time point per patient, we preferred to maximize the use of available information. Predictors for treatment failure were selected based on point estimates of effect size (LHR), which are unaffected by correlations between observations. Therefore, the dependence between observations for a same patient did not influence the scoring system.

Many other studies have looked at single markers of virological failure,[33–37] but few have looked at combinations of predictors.[17] In addition to the classical criteria based on CD4 count, we found that hemoglobin decrease, suboptimal adherence, prior ART exposure, and a new-onset papular pruritic eruption were associated with virological failure. A decrease of at least 1 g/dL of hemoglobin was (weakly) associated with virological failure. Other studies reported an increase in hemoglobin with ART.[36,37] A possible explanation for a decrease in hemoglobin at the moment of virological failure is that immune activation due to viral replication increases the circulating cytokines, which in turn suppress hematopoiesis. Adherence is clearly associated with treatment outcome.[38–42] However, there is no gold standard for measuring adherence. In our study, indications of low adherence by VAS were highly predictive of virological failure. A modified version of SMAQ was not predictive but observations were too few to assess this tool. Pharmacy refill data, which have proven to predict survival in a South African cohort, were not available in our study.[40] Prior ART exposure was identified as a risk factor for treatment failure, as was reported by previous studies.[17,28] A recurrent papular pruritic eruption while on ART was associated with treatment failure, as was also suggested by others.[18,43]

Changes in total lymphocyte count were not predictive for failure, which confirms the findings of other studies.[44–45]

TABLE 2. Number of Visits With Virological Failure (Defined as Viral Load >1000 Copies/mL) and Crude Likelihood Ratios for Predicting Virological Failure by Risk Factor

Risk Factor	No. Visits*	Visits With VL >1000 copies/mL, n (%)	Crude Likelihood Ratio (95% CI)
Total	1803	87 (4.8)	—
Sex			
Male	905	56 (6.2)	1.30 (1.05 to 1.61)
Female	898	31 (3.5)	0.71 (0.49 to 1.02)
ART exposure			
Yes	134	14 (10.5)	2.35 (1.16 to 4.58)
No	1,669	73 (4.4)	0.90 (0.80 to 1.02)
Age >35 yrs			
Yes	943	44 (4.7)	0.97 (0.73 to 1.27)
No	860	43 (5.0)	1.04 (0.78 to 1.38)
Time on ART ≥18 mo			
Yes	1203	53 (4.4)	0.91 (0.77 to 1.07)
No	600	34 (5.7)	1.19 (0.91 to 1.54)
More than 1 substitution/interruption			
Yes	195	10 (5.1)	1.10 (0.50 to 2.26)
No	1608	77 (4.8)	0.99 (0.90 to 1.09)
New papular pruritic eruption			
Yes	134	13 (9.7)	2.17 (1.18 to 3.80)
No	1669	74 (4.4)	0.91 (0.83 to 1.01)
WHO stage 3 clinical event in the last 6 mo			
Yes	287	16 (5.6)	1.19 (0.73 to 1.87)
No	1516	71 (4.7)	0.97 (0.87 to 1.08)
WHO stage 4 clinical event in the last 6 mo			
Yes	78	5 (6.4)	1.46 (0.48 to 3.78)
No	1725	82 (4.8)	0.98 (0.93 to 1.05)
Weight drop >1 kg			
Yes	311	17 (5.5)	1.16 (0.76 to 1.70)
No	1491	70 (4.7)	0.97 (0.88 to 1.07)
Hemoglobin drop ≥1 g/dL			
Yes	215	19 (8.8)	2.01 (1.35 to 2.91)
No	1499	61 (4.1)	0.86 (0.77 to 0.97)
Total lymphocyte count drop ≥300 cells/mm^3			
Yes	465	38 (8.2)	1.74 (1.38 to 2.18)
No	1239	45 (3.6)	0.74 (0.61 to 0.88)
CD4 below baseline			
Yes	37	7 (18.9)	4.79 (1.84 to 11.52)
No	1580	71 (4.5)	0.92 (0.86 to 1.00)
25% drop in CD4 from peak			
Yes	194	30 (15.5)	3.64 (2.55 to 5.19)
No	1559	54 (3.5)	0.71 (0.60 to 0.85)
50% drop in CD4 from peak			
Yes	32	11 (34.4)	10.51 (5.22 to 20.77)
No	1721	73 (4.2)	0.88 (0.81 to 0.95)
CD4 count <100 cells/μL after 12 mo			
Yes	90	16 (17.8)	4.35 (2.51 to 7.34)
No	1663	68 (4.1)	0.84 (0.76 to 0.94)
CD4 count ≤250 cells/μL			
Yes	744	25 (2.5)	1.71 (1.43 to 2.05)
No	1009	59 (7.9)	0.51 (0.34 to 0.75)
SMAQ (at least 1 positive answer)			
Yes	26	1 (3.9)	1.13 (0.11 to 5.74)
No	1697	83 (4.9)	1.00 (0.98 to 1.03)
VAS (<95% adherence)			
Yes	9	5 (55.6)	23.44 (6.75 to 87.77)
No	1260	57 (4.5)	0.92 (0.86 to 0.99)

*Data are missing when the total number of patient visits for a risk factor was less than 1803.

CI, confidence interval (calculated using robust standard errors); VL, viral load.

We found no association between weight loss and treatment failure. Moreover, clinical stage 3 and 4 conditions were not predictive of treatment failure. This is probably explained by the early detection of treatment failure, the use of cotrimoxazole prophylaxis, or because some stage 3 and 4 conditions were late-onset immune reconstitution inflammatory syndrome events. The most frequent stage 3 and 4 conditions were >10% weight loss, bacterial infections, and chronic genital herpes.

TABLE 3. Adjusted LHR of the Different Risk Factors and Corresponding Scores to Predict Viral Load >1000 Copies Per Milliliter

	Outcome: Viral Load >1000 Copies/mL		
Risk Factor	Crude LHR	Adjusted LHR	Score
Antiretroviral therapy exposure			
Yes	2.35	2.56	+1
No	0.90	0.89	0
New papular pruritic eruption			
Yes	2.17	2.29	+1
No	0.91	0.91	0
Hemoglobin drop ≥1 g/dL			
Yes	2.00	1.87	+1
No	0.86	0.88	0
CD4 below baseline			
Yes	4.79	2.15	+1
No	0.92	0.96	0
25% drop in CD4 from peak			
Yes	3.64	2.42	+1
No	0.71	0.79	0
50% drop in CD4 from peak			
Yes	10.51	2.35	+1
No	0.88	0.95	0
CD4 count <100 cells/μL from M12 on ART			
Yes	4.35	2.72	+1
No	0.84	0.89	0
VAS (<95% adherence)			
Yes	23.44	26.97	+3
No	0.92	0.91	0

LHR, likelihood ratio; M12, month 12.

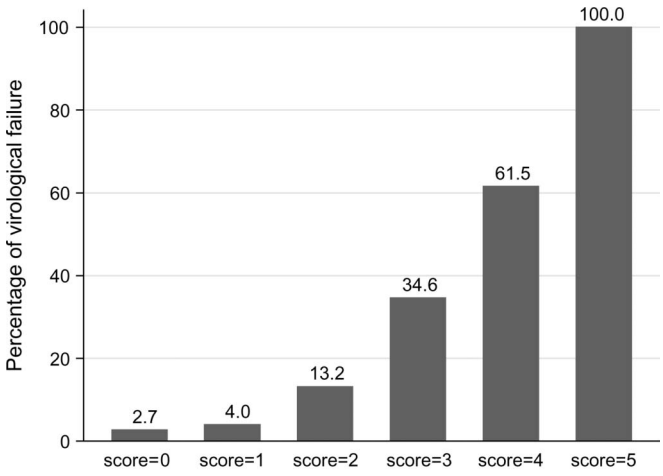


FIGURE 2. Percentage of virological failure by prediction score. No patients with scores >5 were observed in our sample.

What are the possible implications of our study for clinicians and policy makers? In LMICs, where treatment options are limited, patients should be able to benefit for as long as possible from first-line regimens. Although the sensitivity of our algorithm to detect treatment failure is low, the predictor score can easily be repeated every 6 months. Low sensitivity leads to delayed diagnosis of treatment failure. Late switch will lead to accumulation of resistance mutations, which is happening at all levels of viral load in patients who are kept on a failing regimen.[46–50] Data from Malawi suggest that when first-line ART failure diagnosis is based exclusively on clinical and immunological monitoring, extensive resistance to nucleoside and non–nucleoside reverse transcriptase inhibitors is present, impacting future treatment options.[51]

On the other hand, there are hardly any data from LMICs on the risk of disease progression or death in patients who had a late compared with an early treatment switch.[52] A study in Uganda showed that after a median follow-up of 3 years, there was no difference in the rate of AIDS-defining events or death in the group with viral load measurements compared with those without.[53] A recent computer-simulation model by Phillips et al[54] showed that the effect on survival of having access to viral load testing was negligibly small. Both studies suggest that late treatment switches do not have a large effect on survival rates, which leads to a conservative approach in terms of viral load testing. The benefits of targeted viral load testing however were not discussed in the article by Philips et al.[55] Moreover, using a new WHO stage 4 event as a criterion for switching, as proposed by Phillips et al, would result in our cohort in a 93.6% premature switching to second-line treatment. Taking into account the current high cost of second-line treatment, it would be cost-effective to try to avoid this with a system of targeted viral load testing.[56] In settings where resources are limited, we propose that clinicians and policy makers adopt such an approach based on a treatment failure risk assessment. Further operational research is needed to determine the performance of the algorithm in other settings to study the evolution of scores over time, the rapidity of changes in score, and the acceptability of the use of such an algorithm by clinicians in LMICs.

CONCLUSIONS

Clinical and immunological criteria have low diagnostic accuracy in identifying patients in need of second-line ART. Systematic assessment of viral load at regular intervals is a reliable method for detecting treatment failure early, but it is expensive and technically demanding. Cheaper and simple viral load assays are needed. Meanwhile, targeted viral load testing in a subgroup of patients with an intermediate risk of treatment failure in which the diagnostic benefit is greatest may be a feasible and effective strategy in LMICs. Our 2-step algorithm based on a predictor score

TABLE 4. Observed Diagnostic Accuracy of a Simple Scoring System and WHO Criteria to Identify Virological Failure

Score	Virological Failure Defined as Viral Load >1000 Copies/mL				
	n (%)	Sensitivity, % (95% CI)	Specificity, % (95% CI)	PPV*, % (95% CI)	NPV*, % (95% CI)
≥1	639 (35.4)	63.2 (51.4 to 73.7)	66.0 (63.2 to 68.6)	8.6 (6.2 to 11.8)	97.3 (95.9 to 98.2)
≥2	163 (9.0)	41.4 (31.2 to 52.4)	92.6 (91.2 to 93.8)	22.1 (15.7 to 30.1)	96.9 (95.7 to 97.8)
≥3	42 (2.3)	23.0 (15.4 to 32.9)	98.7 (98.1 to 99.2)	47.6 (32.6 to 63.0)	96.2 (94.9 to 97.2)
≥4	16 (0.9)	12.6 (6.8 to 22.3)	99.7 (99.3 to 99.9)	68.8 (41.7 to 87.1)	95.7 (94.4 to 96.8)
≥5	3 (0.2)	3.4 (1.1 to 10.1)	100	100	95.3 (93.9 to 96.4)
WHO criteria (strict)†	201 (11.1)	28.7 (19.9 to 39.5)	89.7 (87.9 to 91.4)	12.4 (8.1 to 18.6)	96.1 (94.8 to 97.1)
WHO criteria (lenient)‡	449 (24.9)	40.2 (30.3 to 51.0)	75.9 (73.4 to 78.2)	7.8 (5.4 to 11.1)	96.2 (94.7 to 97.2)

*PPV: percentage of virological failures in observations with score equal to or above the score cutoff. NPV: percentage of virological suppression in observations with score below the score cutoff.

†Defined as new or recurrent stage 4 conditions, a CD4 count below baseline, a CD4 decrease of 50% from the peak CD4 during treatment, or a CD4 count below 100 cells per microliter after 12 months of ART.

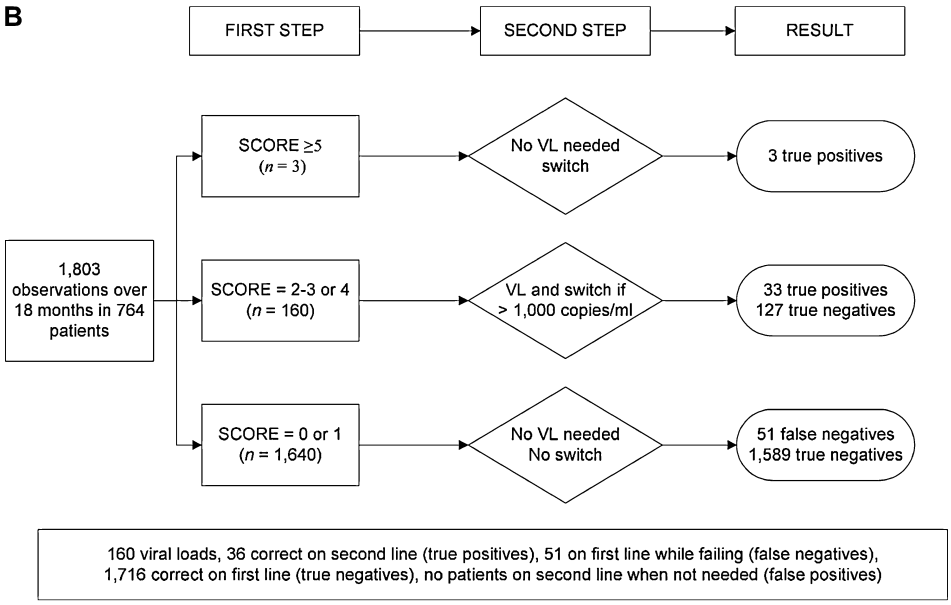
‡Defined as a new or recurrent stage 3 or 4 condition, a CD4 count below baseline, a CD4 decrease of 50% from the peak CD4 during treatment, or a CD4 count below 100 cells per microliter after 12 months of ART.

CI, confidence interval, calculated from logistic regression model with robust standard error adjusted for repeated observations on same patient. CIs for cutoffs with a specificity or PPV of 100% could not be calculated; NPV, negative predictive value.

A

Predictor	Score if condition present
VAS < 95%	3
CD4 below baseline	1
CD4>25% decrease from peak	1
CD4>50% decrease from peak*	1
CD4<100 after 12 months of ART	1
Prior ART exposure	1
Hemoglobin drop ≥ 1g/dl	1
New onset or recurrent papular pruritic eruption	1
Maximum SCORE (Sum)	10

*A patient with a >50% CD4 decrease will also score on the CD4>25% decrease and reaches a total score of 2 based on these two criteria.



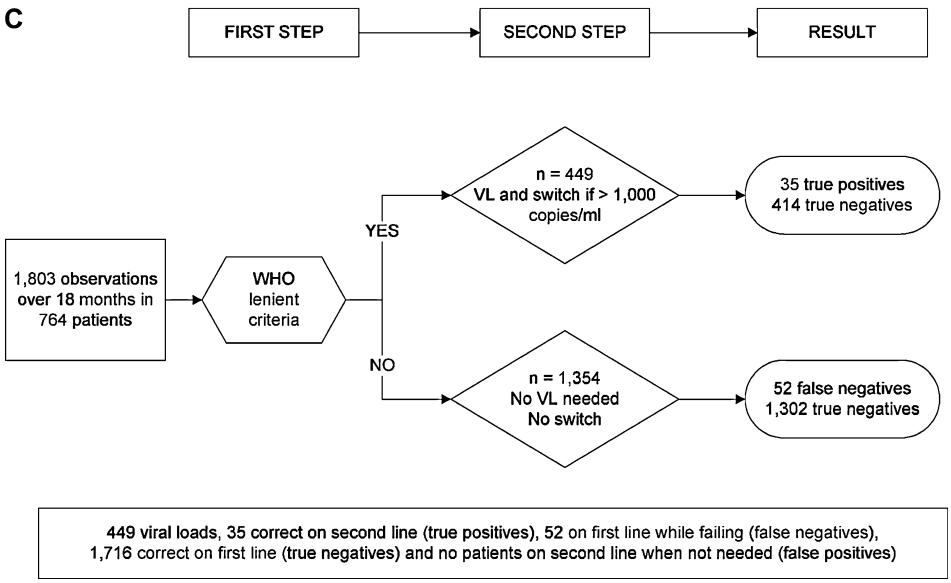


FIGURE 3. Scoring system and WHO lenient criteria as part of a 2-step algorithm using targeted viral load to confirm treatment failure. A, Scoring system leading to a total score for each patient in function of the number of predictors present. B, Algorithm allowing targeted viral load testing in all patients with a score of ≥2. C, Algorithm allowing targeted viral load testing in all patients who have WHO lenient criteria fulfilled. VL, viral load.

coupled with targeted viral load testing is now being implemented at our study site in Cambodia. This predictor score and algorithm require further evaluation and possibly adaptation for use in other settings.

ACKNOWLEDGMENTS

We would like to thank the patients and the doctors of the SHCH who participated in the study. We thank Marianne Mangelschots and Teav Syna for their excellent management

of viral load samples. We thank Anne Buvé for the opportunity to participate in the Europe AID project.

REFERENCES

1. WHO, UNAIDS, UNICEF. Towards universal access: Scaling up priority HIV/AIDS interventions in the health sector. Progress Report, June 2008. Available at: http://www.who.int/hiv/pub/towards_universal_access_report_2008.pdf. Accessed July 24, 2008.
2. UNAIDS. Report on the Global HIV/AIDS Epidemic 2008: Executive Summary. July 2008. Available at: http://data.unaids.org/pub/GlobalReport/2008/JC1511_GR08_ExecutiveSummary_en.pdf. Accessed December 12, 2008.
3. Ministry of Health Cambodia, NCHADS, Surveillance Unit. Report of a Consensus Workshop: HIV Estimates and Projections for Cambodia, 2006–2012. 2008. Available at: http://www.nchads.org/Events/28062007.php. Accessed December 2007.
4. DHHS panel on antiretroviral guidelines for adults and adolescents. Guidelines for the Use of Antiretroviral Agents in HIV-1-Infected Adults and Adolescents. December 2007. Available at: http://aidsinfo.nih.gov/contentfiles/AdultandAdolescentGL.pdf.
5. Gazzard B, Bernard AJ, Boffito M, et al. British HIV Association (BHIVA) guidelines for the treatment of HIV-infected adults with antiretroviral therapy (2006). *HIV Med*. 2006;7:487–503.
6. Kent DM, McGrath D, Ioannidis JP, et al. Suitable monitoring approaches to antiretroviral therapy in resource-poor settings: setting the research agenda. *Clin Infect Dis*. 2003;37:S13–S24.
7. WHO consultation on technical and operational recommendations for scale-up of laboratory services and monitoring HIV antiretroviral therapy in resource-limited settings. December 2004. Available at: http://www.who.int/hiv/pub/meetingreports/labmeetingreport.pdf. Accessed December 2008.
8. Calmy A, Ford N, Hirschel B, et al. HIV viral load monitoring in resource-limited regions: optional or necessary? *Clin Infect Dis*. 2007;44:128–134.
9. Schooley RT. Viral load testing in resource-limited settings. *Clin Infect Dis*. 2007;44:139–140.
10. Bagchi S, Kempf MC, Westfall AO, et al. Can routine clinical markers be used longitudinally to monitor antiretroviral therapy success in resource-limited settings? *Clin Infect Dis*. 2007;44:135–138.
11. Gilks CF, Crowley S, Ekpini R, et al. The WHO public-health approach to antiretroviral treatment against HIV in resource-limited settings. *Lancet*. 2006;368:505–510.
12. WHO. Antiretroviral therapy for HIV infection in adults and adolescents: recommendations for a public health approach. 2006 revision. Available at: http://www.who.int/hiv/pub/guidelines/adult/en/index.html. Accessed December 2008.
13. An S, Koole O, Haverkamp M, et al. Predictors of virological failure in a Cambodian setting: findings from a cross-sectional study at Sihanouk Hospital Centre of HOPE, Phnom Penh [WEAB103]. Presented at: 4th IAS Conference on HIV Pathogenesis and Treatment; July 22–25, 2007; Sydney, Australia.
14. Van Griensven J, Corthouts A, Atté EF, et al. Validity of WHO-recommended immunological criteria for treatment failure [1039]. Presented at: 2007 HIV/AIDS Implementers' Meeting; June 16–19, 2007; Kigali, Rwanda.
15. Hosseinipour M, van Oosterhout J, Weigel R, et al. Validating clinical and immunological definitions of antiretroviral treatment failure in Malawi [WEAB101]. Presented at: 4th IAS Conference on HIV Pathogenesis and Treatment; July 22–25, 2007; Sydney, Australia.
16. Colebunders R, Moses KR, Laurence J, et al. A new model to monitor the virological efficacy of antiretroviral treatment in resource-poor countries. *Lancet Infect Dis*. 2006;6:53–59.
17. Ferradini L, Laureillard D, Prak N, et al. Positive outcomes of HAART at 24 months in HIV-infected patients in Cambodia. *AIDS*. 2007;21:2293–2301.
18. Lawn SD, Orrell C, Wood R. Evaluating a model for monitoring the virological efficacy of antiretroviral treatment in resource-limited settings. *Lancet Infect Dis*. 2006;6:385–386.
19. WHO. Scaling up antiretroviral therapy in resource-limited settings: treatment guidelines for a public health approach, 2003 revision. 2003, Available at: http://www.who.int/3by5/publications/documents/arv_guidelines/en/. Accessed December 21, 2003.
20. Knobel H, Alonso J, Casado JL, et al. Validation of a simplified medication adherence questionnaire in a large cohort of HIV-infected patients: The GEEMA Study. *AIDS*. 2002;16:605–613.
21. Giordano TP, Guzman D, Clark R, et al. Measuring adherence to antiretroviral therapy in a diverse population using a visual analog scale. *HIV Clin Trials*. 2004;5:74–79.
22. Seymour DG, Green M, Vaz FG. Making better decisions: construction of clinical scoring systems by the Spiegelhalter-Knill-Jones approach. *BMJ*. 1990;300:223–226.
23. Spiegelhalter DJ. Statistical methodology for evaluating gastrointestinal symptoms. *Clin Gastroenterol*. 1985;14:489–515.
24. Berkley JA, Ross A, Mwangi I, et al. Prognostic indicators of early and late death in children admitted to district hospital in Kenya: Cohort study. *BMJ*. 2003;326:361.
25. Zhou XH, Obuchowski NA. *Statistical Methods in Diagnostic Medicine* (Wiley Series in Probability and Statistics). New York, NY: Wiley-Interscience; 2002.
26. Murphy MK, Black NA, Lamping DL, et al. Consensus development methods, and their use in clinical guideline development. *Health Technol Assess*. 1998;2:1–88.
27. Madec Y, Laureillard D, Pinoges L, et al. Response to highly active antiretroviral therapy among severely immuno-compromised HIV-infected patients in Cambodia. *AIDS*. 2007;21:351–359.
28. Calmy A, Pinoges L, Szumilin E, et al. Generic fixed-dose combination antiretroviral treatment in resource-poor settings: multicentric observational cohort. *AIDS*. 2006;20:1163–1169.
29. Ferradini L, Jeannin A, Pinoges L, et al. Scaling up of highly active antiretroviral therapy in a rural district of Malawi: an effectiveness assessment. *Lancet*. 2006;367:1335–1342.
30. Braitstein P, Brinkhof MW, Dabis F, et al. Mortality of HIV-1-infected patients in the first year of antiretroviral therapy: comparison between low-income and high-income countries. *Lancet*. 2006;367:817–824.
31. Orrell C, Harling G, Lawn SD, et al. Conservation of first-line antiretroviral treatment regimen where therapeutic options are limited. *Antivir Ther*. 2007;12:83–88.
32. Marconi VC, Sunpath H, Lu Z, et al. Prevalence of HIV-1 drug resistance after failure of a first highly active antiretroviral therapy regimen in KwaZulu Natal, South Africa. *Clin Infect Dis*. 2008;46:1589–1597.
33. Bisson GP, Gross R, Strom JB, et al. Diagnostic accuracy of CD4 cell count increase for virologic response after initiating highly active antiretroviral therapy. *AIDS*. 2006;20:1613–1619.
34. Meya D, Tibenderana H, John L, et al. Evaluation of clinical and laboratory parameters to predict viral response to ART in Uganda [531]. Presented at: 14th Conference on Retroviruses and Opportunistic Infections; February 25–28, 2007; Los Angeles, CA.
35. Moore DM, Mermin J, Awor A, et al. Performance of immunologic responses in predicting viral load suppression: implications for monitoring patients in resource-limited settings. *J Acquir Immune Defic Syndr*. 2006;43:436–439.
36. Berhane K, Karim R, Cohen MH, et al. Impact of highly active antiretroviral therapy on anemia and relationship between anemia and survival in a large cohort of HIV-infected women: women's interagency HIV study. *J Acquir Immune Defic Syndr*. 2004;37:1245–1252.
37. Moore RD, Forney D. Anemia in HIV-infected patients receiving highly active antiretroviral therapy. *J Acquir Immune Defic Syndr*. 2002;29:54–57.
38. Weiser S, Wolfe W, Bangsberg D, et al. Barriers to antiretroviral adherence for patients living with HIV infection and AIDS in Botswana. *J Acquir Immune Defic Syndr*. 2003;34:281–288.
39. Weidle PJ, Malamba S, Mwebaze R, et al. Assessment of a pilot antiretroviral drug therapy programme in Uganda: patients' response, survival, and drug resistance. *Lancet*. 2002;360:34–40.
40. Nachega JB, Hislop M, Dowdy DW, et al. Adherence to highly active antiretroviral therapy assessed by pharmacy claims predicts survival in HIV-infected South African adults. *J Acquir Immune Defic Syndr*. 2006;43:78–84.
41. Bisson GP, Gross R, Bellamy S, et al. Pharmacy refill adherence compared with CD4 count changes for monitoring HIV-infected adults on antiretroviral therapy. *PLoS Med*. 2008;5:e109.

42. Perno CF, Ceccherini-Silberstein F, De Luca A, et al. Virologic correlates of adherence to antiretroviral medications and therapeutic failure. *J Acquir Immune Defic Syndr*. 2002;31(Suppl 3):S118–S122.
43. Castelnuovo B, Byakwaga H, Menten J, et al. Can response of a pruritic papular eruption to antiretroviral therapy be used as a clinical parameter to monitor virological outcome? *AIDS*. 2008;22:269–273.
44. Florence E, Schrooten W, Van Esbroek M, et al. The role of non-viral load surrogate markers in HIV-positive patient monitoring during antiviral treatment. HAART monitoring without viral load. *Int J STD AIDS*. 2004;15:538–542.
45. Mahajan AP, Hogan JW, Snyder B, et al. Changes in total lymphocyte count as a surrogate for changes in cd4 count following initiation of HAART: implications for monitoring in resource-limited settings. *J Acquir Immune Defic Syndr*. 2004;36:567–575.
46. Kantor R, Shafer RW, Follansbee S, et al. Evolution of resistance to drugs in HIV-1-infected patients failing antiretroviral therapy. *AIDS*. 2004;18: 1503–1511.
47. Kristiansen TB, Pedersen AG, Eugen-Olsen J, et al. Genetic evolution of HIV in patients remaining on a stable HAART regimen despite insufficient viral suppression. *Scand J Infect Dis*. 2005;37:890–901.
48. Cozzi-Lepri A, Phillips AN, Ruiz L, et al. Evolution of drug resistance in HIV-infected patients remaining on a virologically failing combination antiretroviral therapy regimen. *AIDS*. 2007;21:721–732.
49. Napravnik S, Edwards D, Stewart P, et al. HIV-1 drug resistance evolution among patients on potent combination antiretroviral therapy with detectable viremia. *J Acquir Immune Defic Syndr*. 2005;40:34–40.
50. Riddler SA, Jiang H, Tenorio A, et al. A randomized study of antiviral medication switch at lower- versus higher-switch thresholds: AIDS Clinical Trials Group Study A5115. *Antivir Ther*. 2007;12:531–541.
51. Hosseinipour M, van Oosterhout JJ, Weigel R, et al. Resistance profile of patients failing first line ART in Malawi when using clinical and immunological monitoring [TUAB0105]. Presented at: XVII International AIDS conference, August 3–8, 2008, Mexico City, Mexico.
52. Humphreys EH, Hernandez LB, Rutherford GW. Antiretroviral regimens for patients with HIV who fail first-line antiretroviral therapy. *Cochrane Database Syst Rev*. 2007;CD006517.
53. Coutinho A, Mermin J, Ekwaru J, et al. Utility of routine viral load, CD4 cell count, and clinical monitoring among HIV-infected adults in Uganda: a randomized trial [125]. Presented at: 15th Conference on Retroviruses and Opportunistic Infections; February 3–6, 2008; Boston, MA.
54. Phillips AN, Pillay D, Miners AH, et al. 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.
55. Moore DM, Mermin J. Monitoring antiretroviral failure in resource-poor settings. *Lancet*. 2008;371:1396–1397.
56. Médecins Sans Frontières. Untangling the web of price reductions: a pricing guide for the purchase of ARVs for developing countries, 10th edition. 2007. Available at: http://www.accessmed-msf.org/fileadmin/user_upload/diseases/hiv-aids/Untangling_the_Web/UTW10_RSep_horizontal.pdf.