Development and validation of systems for rational use of viral load testing in adults receiving first-line ART in sub-Saharan Africa

Michael Abouyannis[a], Joris Menten[b], Agnes Kiragga[c], Lutgarde Lynen[b], Gavin Robertson[c], Barbara Castelnuovo[c], Yukari C. Manabe[c,d], Steven J. Reynolds[d,e] and Lesley Roberts[a]

Background: World Health Organization (WHO) immunological and clinical criteria for monitoring first-line antiretroviral treatment (ART) offer low accuracy for predicting viral failure. Targeting viral load assays to those at high risk has been recommended and a system to do this has been developed in Cambodia. Systems for use in sub-Saharan African populations were evaluated.

Methods: A new Ugandan-based scoring system for targeting viral load assays was developed from data from the first 4 years of a Ugandan cohort ($N = 559$) receiving first-line ART. The accuracy of this, the Cambodian system and the WHO criteria to predict viral failure, through targeting viral load assays, were compared in a separate population of 496 Ugandans.

Results: The new Ugandan scoring system included CD4 cell count, mean cell volume, adherence, and HIV-associated clinical events as predictors of viral failure. In the validation population, the Ugandan system undertook viral load assays in 61 (12.3%) cases offering 20.5% sensitivity and 100% positive predictive value (PPV) to predict viral failure. The Cambodian system undertook viral load assays in 33 (6.7%) cases producing 23.1% sensitivity and 90.0% PPV. WHO criteria recommended viral load assays in 72 (14.5%) cases offering 30.8% sensitivity and 100% PPV.

Conclusion: Locally developed algorithms based on clinical and immunological criteria may offer little additional accuracy over WHO criteria for targeting viral load assays. When possible, confirming viral load before switching therapy is recommended. Scoring systems are more flexible than WHO criteria in allowing ART providers to choose the proportion of the population that undergo targeted viral load testing.

AIDS 2011, **25**:1627–1635

Keywords: antiretroviral therapy, HIV, resource-constrained settings, treatment failure, Uganda, viral load

Introduction

Access to antiretroviral therapy (ART) in sub-Saharan Africa has dramatically improved recently with approximately 42% coverage of those requiring ART being reached in Uganda between December 2007 and 2008 [1]. With time, a proportion of these will fail therapy and become eligible for second-line treatment. Reported

[a]University of Birmingham, Birmingham, UK, [b]Institute of Tropical Medicine, Antwerp, Belgium, [c]Infectious Diseases Institute, Kampala, Uganda, [d]Department of Medicine, Johns Hopkins School of Medicine, Baltimore, and [e]National Institutes of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland, USA.

Correspondence to Dr Lesley Roberts, Primary Care Clinical Sciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK.

Tel: +44 121 414 3356; e-mail: l.m.roberts@bham.ac.uk
Received: 27 May 2010; revised: 24 May 2011; accepted: 7 June 2011.

DOI:10.1097/QAD.0b013e328349a414

rates of use of second-line therapies have been low [1–4] and research suggests that difficulties in diagnosing treatment failure may have contributed to this [5,6].

The viral load test is the gold standard for monitoring ART [7,8]. Currently, this test is expensive and technically demanding [9], prohibiting its use in resource-constrained settings [10,11]. Instead, the World Health Organization (WHO) recommends monitoring clinical and immunological criteria [12,13]. However, consistent evidence has mounted to suggest that these criteria have low sensitivity and positive predictive value (PPV) for identifying viral failure [14–20]. Late switches may lead to the development of drug resistance [21,22], opportunistic infections and increased mortality [23]. Inappropriate switching to second-line treatment incurs increased costs [24].

Systems based on additional clinical and immunological markers of viral failure, such as adherence, have shown capacity to improve accuracy over WHO criteria [16,25–27]. Although these systems have not offered an alternative to viral load monitoring, their use as tools for targeting viral load testing has been advocated [16,27]. A recent Cambodian-based study developed a scoring system (Table 1) for targeting viral loads with greater accuracy than WHO clinical and immunologic criteria [27]. The development of such a system for use in a sub-Saharan African population has yet to be fully explored.

In this study, a new Ugandan-based scoring system for targeting viral load tests was developed. In a separate population, this system was tested alongside the Cambodian scoring system and the WHO criteria to identify the most useful for targeting viral load tests.

Table 1. Scoring systems for predicting risk of viral failure and targeting viral load assays.

	Score
Cambodian scoring system predictors	
Current CD4 cell count below baseline	+1
CD4 cell count >25% decrease from peak	+1
CD4 cell count >50% decrease from peak	+1
Current CD4 cell count <100 after 12 months of ART	+1
Haemoglobin drop ≥1 g/dl	+1
Percentage adherence (VAS) ≤95%	+3
Prior ART exposure	+1
Prurigo	+1
Ugandan scoring system predictors	
Current CD4 cell count ≤200 cells/μl	+2
Current MCV ≤95 fl	+2
Percentage adherence (VAS) ≤90%	+1
Herpes zoster (varicella zoster virus; shingles)	+2
Severe bacterial pneumonia	+2
Prurigo	+1
Oropharyngeal candidiasis (thrush)	+1
Extrapulmonary tuberculosis	+1

Scoring systems in which the risk of viral failure is calculated by summing the scores associated with each present predictor. MCV, mean corpuscular volume; VAS, visual analogue scale.

Methods

Study setting

The study was undertaken at the Infectious Diseases Institute (IDI) at Mulago Hospital and Makerere University in Kampala, Uganda [28]. The IDI clinic provides free care to HIV-positive patients and is one of the largest urban HIV treatment centres in Uganda.

Derivation of the new Ugandan scoring system

The Ugandan scoring system was developed using data from a prospective observational cohort undertaken at the IDI which will be referred to as the derivation population. Between April 2004 and April 2005, 559 consecutive patients initiating their first course of ART were enrolled if they fulfilled the following eligibility criteria: confirmed HIV-1 infection, regular attendance (at least two clinic visits within the previous 6 months), and provision of written informed consent. Participants were initiated on zidovudine or stavudine and lamivudine, and either nevirapine or efavirenz according to WHO 2003 and Uganda Ministry of Health treatment guidelines [29,30]. Participants were provided with daily co-trimoxazole prophylaxis, or dapsone, if they were allergic to co-trimoxazole.

At baseline and at 6-monthly intervals study physicians completed standardized data collection forms, and laboratory measurements were recorded. As previously described [31], data included demographics, HIV-related clinical events, adherence [participant reported visual analogue scale (VAS)], complete blood cell counts, CD4 lymphocyte testing, and HIV plasma RNA measurement. Study physicians were not blinded to the viral load test result.

Six-monthly visits between the 12-month and the 4-year interval were used to develop the scoring system, and the visit was the unit of observation. Visits in which participants were not taking ART due to side-effects, contraindications, or complications were excluded from the analysis, as were those following a switch to second-line therapy. Viral failure was defined as a single episode of HIV plasma RNA measurement over 1000 copies/ml to maintain consistency with the study from which the Cambodian scoring system was developed [27].

The following potential predictors were evaluated: [Demographics] sex, age; [Clinical] new or recurrent WHO stage 3 and stage 4 conditions [13], body mass index (BMI), ART regimen, time on ART, previous exposure to ART, VAS adherence; [Laboratory] absolute CD4 cell count, change in absolute CD4 cell count from pretherapy baseline, percentage CD4 cell count, change in percentage CD4 cell count from pretherapy baseline, absolute lymphocyte count, haemoglobin, haematocrit, mean corpuscular volume (MCV) and mean corpuscular haemoglobin (MCH).

The Spiegelhalter and Knill-Jones method, adapted by Berkley *et al.* and Lynen *et al.* for use in Cambodia, was used to construct the new Ugandan scoring system [27,32–34]. Continuous variables were dichotomized at the point with the highest sum of sensitivity and specificity. These points were rounded to simplify use. Positive and negative likelihood ratios were calculated for each variable to predict viral failure (>1000 copies/ml). Those with a likelihood ratio at least 2.0 or 0.5 or less were retained and adjusted for confounding using a multivariate logistic regression analysis. Variables with adjusted likelihood ratios at least 1.5 or 0.67 or less were selected for use in the scoring system. The natural log of the likelihood ratio was rounded to the nearest integer to give a risk score for each predictor. Positive scores favoured viral failure, whereas negative scores argued against it. Constants were added to negative scores to simplify use.

External validation of the Cambodian scoring system, Ugandan scoring system and WHO criteria for targeting viral load assays

The three systems were validated using data from a separate cross-sectional population which will be referred to as the validation population. Participants were sampled from the IDI adult clinic between February and June of 2006. Approximately 10 patients per day were selected from the reception using a list of random numbers and a total of 500 were enrolled. The following eligibility criteria applied: confirmed HIV-1 infection, aged over 18 years, established on first-line non-nucleoside reverse transcriptase inhibitor (NNRTI)-based ART for more than 6 months, not viral load monitored as routine clinic practice, and not suffering acute illness. All data collection was undertaken at enrolment and this included the following: demographics, HIV-related clinical events, adherence (including use of VAS), previous and a current complete blood cell count, previous and a current CD4 lymphocyte count, and HIV RNA in plasma measurement [16].

The Cambodian and Ugandan scoring systems were applied and their total scores were calculated for each participant (missing data were assigned scores of 0). The Cambodian system diagnosed treatment failure in those with a score above 4 and excluded it in those with a score below 2. For those with a Cambodian score of 2–4 a viral load assay was recommended. This was the approach recommended by researchers on deriving the Cambodian scoring system [27]. The cut-off at which viral load testing was undertaken for the Ugandan system was chosen based on accuracies in the derivation population.

WHO clinical and immunologic criteria were applied and those with WHO-defined failure underwent confirmatory viral load testing. Immunologic failure was defined as any of CD4 persistently (≥6 months) below 100 cells/μl, a CD4 cell count at least 50% less than the on-treatment peak value, or a CD4 cell count below the pretherapy baseline value. Immunologic failure was excluded in visits with CD4 cell counts above 200 cells/μl as guidance recommends switching only if the CD4 cell count is below this [12]. Clinical failure was defined as the onset or recurrence of any WHO stage 4 disease. IDI research interviewers were also asked if they felt the participant was failing virologically and the accuracy of their judgement was described.

Accuracy characteristics for each scoring system and for WHO criteria were presented in terms of sensitivity, specificity, PPV and NPV with accompanying 95% confidence intervals (CIs). CIs from the derivation population were calculated using robust standard errors to account for intrapatient clustering. Wilson CIs were used in the validation population.

Stata 10.0 software was used for the statistical analysis [StataCorp (Stata Statistical Software), Release 10, StataCorp, College Station, Texas].

Ethical approvals

Informed consent was obtained from all participants of both included populations. Approval was received from the following prior to commencing this research: Makerere University Research and Ethics Committee (Kampala Uganda); the Infectious Diseases Institute Scientific Review Committee (Kampala Uganda); the BMedSc Medicine in Society Internal Ethics Committee (University of Birmingham, UK); the HIV/AIDS Research Committee (ARC, Kampala, Uganda) and the National Institute of Allergy and Infectious Diseases Institutional Review Board (Bethesda, Maryland, USA). Both studies were also approved by the Uganda National Council for Science and Technology (UNCST, Kampala, Uganda).

Results

Derivation of the new Ugandan scoring system

Of the 559 participants that were enrolled into the derivation population, 443 contributed a total of 2668 visits to the analysis. Inclusions and exclusions of visits from the analysis and frequencies of viral failure at each visit interval have been detailed in Table 2. Forty-eight participants were lost to follow-up and 105 died. From 1 year onward, 92 (20.8%) participants suffered one or more episodes of viral failure to produce 130 (4.9%) visits in which this was present. There were 61 visits with missing viral load data.

The characteristics of the original cohort ($N = 559$) have been previously described [35] and the characteristics of the 443 individuals who contributed data to this study were as follows. At baseline, 30.2% were men, median age was 35 years [interquartile range (IQR) 30–42 years],

Table 2. Visits included and excluded from the analysis and frequencies of viral failure at each visit interval (derivation population).

Event	Visit interval (months)							Total
	12	18	24	30	36	42	48	
No. visits included in analysis	439	406	388	373	357	355	350	2668
No. visits with viral failure (>1000 copies/ml)	36	26	15	13	13	13	14	130
No. visits not attended	18	4	1	6	7	4	12	52
No. visits viral load missing	3	4	7	10	18	14	5	61
No. visits not on ART	3	5	4	0	2	2	0	16
No. visits lost through death	81	8	5	4	2	4	1	105
No. visits lost to follow-up	12	22	7	2	2	1	2	48
No. visits lost through switching to second-line drugs	3	14	7	5	1	4	5	39

mean weight was 55.4 kg (SD 10.2 kg) and mean height was 164.3 cm (SD 9.1 cm). Most had advanced HIV disease with median CD4 cell count 102 cells/μl (IQR 31–168 cells/μl) and a high proportion with WHO stage III and stage IV disease (stage I, 0.5%; stage II, 11.7%; stage III, 57.8%; and stage IV, 30.0%). Stavudine, lamivudine and nevirapine was the initial regimen for 72.5% of participants, 27.3% were started on zidovudine, lamivudine and efavirenz, and 0.2% began with zidovudine, lamivudine and nevirapine. Twenty participants (4.5%) were ART-experienced. Median MCV was 85.4 fl (IQR 80.9–89.3 fl).

Of the potential predictors listed in the methods, eight demonstrated adjusted likelihood ratios at least 1.5 or 0.67 or less for detecting viral failure (>1000 copies/ml) (Table 3) and were used in the Ugandan scoring system. These included CD4 cell count, MCV, patient reported adherence, herpes zoster, prurigo, severe bacterial pneumonia, oropharyngeal candidiasis and extrapulmonary tuberculosis.

When applied to the derivation population, Ugandan system total scores ranged from 0 to 7. Table 4a illustrates accuracy characteristics of the Ugandan scoring system to predict viral failure (>1000 copies/ml) at each score cut-off in the derivation population. A two-step algorithm for diagnosing treatment failure was developed. Step 1 was to calculate the total risk score. Step 2 was to diagnose those with a score less than or equal to 2 as not failing, and to undergo confirmatory viral load testing in those with a score of 3 or higher. This cut-off was chosen because in the derivation population this offered acceptable accuracy whilst relying on a limited number of viral load assays – based on author judgement. This resulted in 194 (7.3%) visits being eligible for testing. In the derivation population sensitivity was 40% (95% CI 29.9–51.1%) and PPV was 100% for predicting viral failure. This accuracy was superior to the Cambodian and WHO criteria (data not presented).

Validation of the Cambodian scoring system, Ugandan scoring system and WHO criteria for targeting viral load assays

Of the 500 participants enrolled into the validation population, 496 had complete questionnaires and viral load results. The characteristics of this population have been previously described and were as follows [16].

Table 3. Frequencies, crude likelihood ratios with 95% confidence intervals, adjusted likelihood ratios and risk scores of parameters used in the Ugandan scoring system (derivation population).

Predictor of viral failure	No. of visits[a]	No. of visits predictor positive	Predictor present	Crude LRs (95% CI)	Adjusted LRs	Score if present
CD4 cell count ≤200cells/μ\l	2661	604	Yes	2.64 (2.22–3.13)	2.61	+2[b]
			No	0.57 (0.47–0.68)	0.57	0[b]
MCV ≤95 fl	2657	713	Yes	2.19 (1.85–2.59)	2.23	+2[b]
			No	0.60 (0.49–0.72)	0.59	0[b]
Patient reported adherence (VAS) ≤90%	2668	30	Yes	6.19 (2.77–13.81)	3.82	+1
			No	0.95 (0.91–0.99)	0.96	0
Herpes zoster	2668	6	Yes	5.29 (0.88–31.88)	6.29	+2
			No	0.99 (0.97–1.01)	0.99	0
Prurigo	2668	27	Yes	4.74 (1.90–11.84)	3.48	+1
			No	0.97 (0.93–1.00)	0.97	0
Severe bacterial pneumonia	2668	6	Yes	10.77 (2.32–50.03)	4.79	+2
			No	0.98 (0.96–1.01)	0.99	0
Oropharyngeal candidiasis	2668	5	Yes	13.84 (2.76–69.5)	1.92	+1
			No	0.98 (0.96–1.01)	1.00	0
Extrapulmonary TB	2668	3	Yes	11.63 (1.55–87.37)	2.08	+1
			No	0.99 (0.97–1.01)	1.00	0

CI, confidence interval (calculated using robust standard errors); LRs, likelihood ratios; MCV, mean corpuscular volume; VAS, visual analogue scale.
[a]Missing data is 2668 minus no. of visits.
[b]Constant of +1 added to avoid a negative score.

Table 4. Accuracy of Cambodian scoring system, Ugandan scoring system and WHO criteria to predict viral failure (derivation and validation populations).

Definition of failure	N[a]	N[a] with viral failure	Accuracy characteristics % (95% CI)[d] Sensitivity	Specificity	PPV	NPV
(a) Derivation population						
Ugandan scoring system						
Ugandan Score 0≤	2668	130	100	–	4.9 (3.9–6.1)	–
Ugandan Score + 1≤	1183	98	75.4 (64.9–83.5)	57.2 (53.6–60.9)	8.3 (6.4–10.7)	97.8 (96.7–98.6)
Ugandan Score + 2≤	1160	97	74.6 (64.1–82.9)	58.1 (54.4–61.8)	8.4 (6.4–10.8)	97.8 (96.7–98.6)
Ugandan Score + 3≤	194	52	40 (29.9–51.1)	94.4 (93–95.5)	26.8 (19.7–35.4)	96.8 (95.8–97.6)
Ugandan Score + 4≤	168	49	37.7 (27.6–49)	95.3 (94–96.3)	29.2 (21.3–38.6)	96.8 (95.7–97.6)
Ugandan Score + 5≤	15	10	7.7 (4–14.3)	99.8 (99.5–99.9)	66.7 (38.8–86.3)	95.5 (94.3–96.4)
Ugandan Score + 6≤	2	2	1.5 (0.4–6)	100	100	95.2 (94–96.2)
Ugandan Score + 7≤	1	1	0.8 (0.1–5.3)	100	100	95.2 (93.9–96.2)
(b) Validation population						
Cambodian scoring system						
Cambodian score 0≤	496	39	100 (91.0–100)	0.0 (0.0–0.8)	7.9 (5.8–10.6)	–
Cambodian score + 1≤	106	20	51.3 (36.2–66.1)	81.2 (77.3–84.5)	18.9 (12.6–27.4)	95.1 (92.5–96.9)
Cambodian score + 2≤	34	9	23.1 (12.6–38.3)	94.5 (92.0–96.3)	26.5 (14.6–43.1)	93.5 (90.9–95.4)
Cambodian score + 3≤	16	6	15.4 (7.2–29.7)	97.8 (96.0–98.8)	37.5 (18.5–61.4)	93.1 (90.5–95.1)
Cambodian score + 4≤	5	2	5.1 (1.4–16.9)	99.3 (98.1–99.8)	40.0 (11.8–76.9)	92.5 (89.8–94.5)
Cambodian score + 5≤	1	0	0.0 (0.0–9.0)	99.8 (98.8–100)	0.0 (0.0–79.3)	92.1 (89.4–94.2)
Ugandan scoring system						
Ugandan score 0≤	496	39	100 (91.0–100)	0.0 (0.0–0.8)	7.9 (5.8–10.6)	–
Ugandan score + 1≤	269	28	71.8 (56.2–83.5)	47.3 (42.7–51.8)	10.4 (7.30–14.6)	95.2 (91.5–97.3)
Ugandan Score + 2≤	261	27	69.2 (53.6–81.4)	48.8 (44.2–53.4)	10.3 (7.20–14.6)	94.9 (91.3–97.1)
Ugandan score + 3≤	61	8	20.5 (10.8–35.5)	88.4 (85.1–91.0)	13.1 (6.8–23.8)	92.9 (90.0–94.9)
Ugandan score + 4≤	54	7	17.9 (9.0–32.7)	89.7 (86.6–92.2)	13.0 (6.4–24.4)	92.8 (90.0–94.8)
Ugandan score + 5≤	4	1	2.6 (0.5–13.2)	99.3 (98.1–99.8)	25.0 (4.6–69.9)	92.3 (89.6–94.3)
WHO criteria						
WHO immunologic[b]	61	10	25.6 (14.6–41.1)	88.8 (85.6–91.4)	16.4 (9.2–27.6)	93.3 (90.6–95.3)
WHO immunologic and clinical[c]	72	12	30.8 (18.6–46.4)	86.9 (83.5–89.7)	16.7 (9.8–26.9)	93.6 (90.9–95.6)

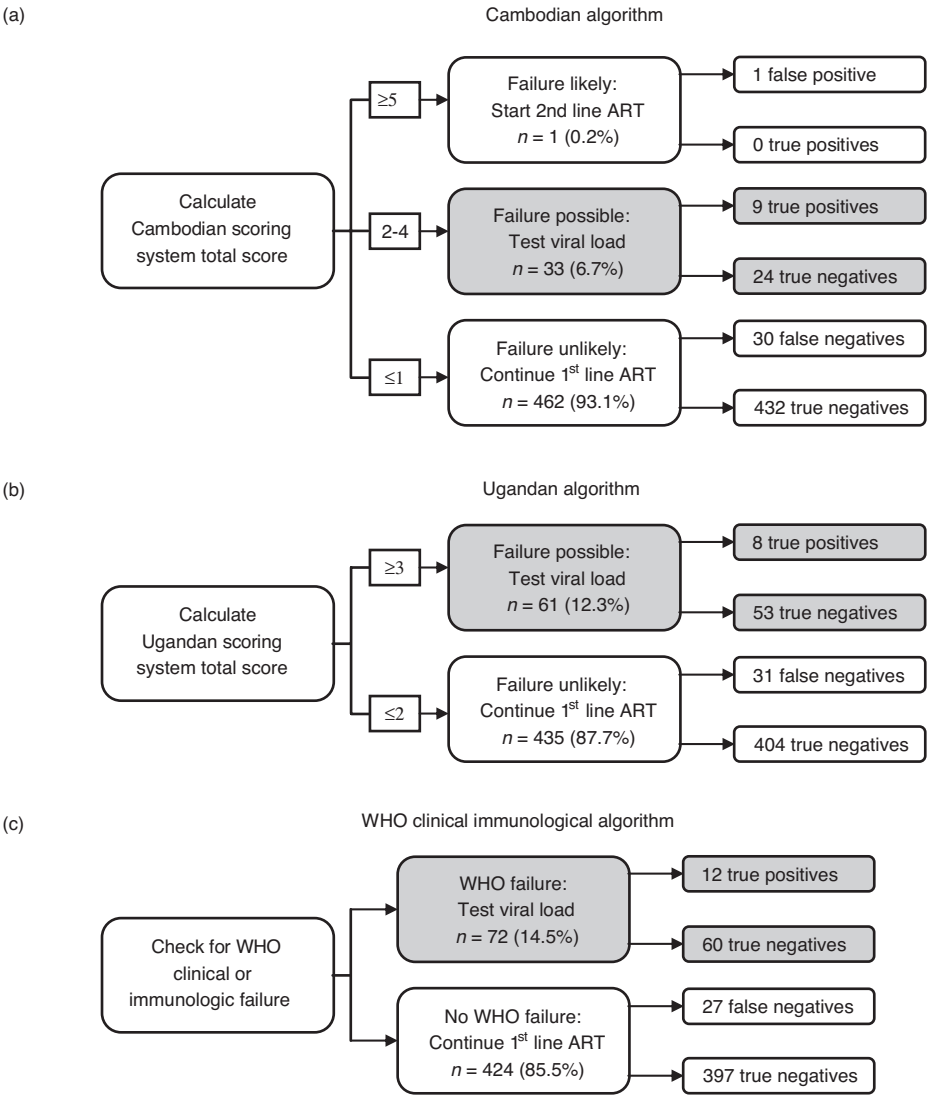
[a]*N*, number of visits in (a) derivation population and patients in (b) validation population for which the WHO, Cambodian or Ugandan system definitions of failure were fulfilled. NPV, negative predictive value; PPV, positive predictive value.
[b]Defined as, CD4 cell count below 200 cells/μl and any of: CD4 persistently (≥6 months) below 100 cells/μl, a CD4 cell count ≥50% less than the on-treatment peak value, or a CD4 cell count below the pretherapy baseline.
[c]Defined as, CD4 cell count below 200 cells/μl in the presence of a new or recurrent WHO stage 4 disease, or any immunologic criteria.
[d]CI, confidence interval. Generated using logistic regression model with robust standard errors adjusted for repeated observations of the same participant in the derivation population. Wilson confidence intervals were used in the validation population.

Median age was 38 years (IQR 34–44 years), 37.3% were men and median duration of ART was 13 months (IQR 10–16 months). Before starting ART the median CD4 cell count was 90 cells/μl (IQR 35–156 cells/μl) and 11 participants (2.2%) suffered new or recurrent WHO stage 3 and 4 opportunistic infection whilst on ART. Thirty-nine (7.9%) participants had viral failure (>1000 copies/ml).

The Cambodian and Ugandan scoring systems (Table 1) were applied and total scores ranged from 0 to 5. The accuracy characteristics of these systems to predict viral failure at each score cut-off have been described in Table 4b. Accuracy of the Ugandan system was reduced in the validation population with the cut-off of greater than or equal to three offering sensitivity of 20.5% (95% CI 10.8–35.5%) and PPV of 13.1% (95% CI 6.8–23.8%) – each less than half compared with accuracies in the derivation population, although confidence intervals did overlap. The Cambodian system in the validation population was superior with a cut-off of greater than or equal to two offering sensitivity of 23.1% (95% CI 12.6–38.3%) and PPV of 26.5% (95% CI 14.6–43.1%).

Upon application of WHO criteria immunologic criteria were fulfilled in 66 cases. Immunologic failure was present in 61 cases (5 fulfilling immunologic criteria had CD4 cell counts above 200 cells/μl) and 72 had combined clinical or immunologic failure. The combined immunologic and clinical criteria performed marginally better than the immunologic criteria alone offering sensitivity of 30.8% (95% CI 18.6–46.4%) and a PPV of 16.7% (95% CI 9.8–26.9%).

Each of the systems were evaluated in algorithms for targeting viral load assays which have been depicted in Fig. 1a–c with their accuracies summarized in Fig. 1d. The Cambodian algorithm allocated viral load assays in 33 (6.7%) cases and produced a sensitivity of 23.1% (95% CI 12.6–38.3%) and a PPV of 90.0% (95% CI 59.6–98.2%). The PPV was not 100% as one case with a score of 5 was falsely diagnosed as treatment failure. The Ugandan algorithm relied on more viral load assays with 61 (12.3%)



(d)

With targeted viral load testing	No. viral load tested (%)	Sensitivity	Specificity	PPV	NPV
Cambodian system (Fig 2a)	33 (6.7)	23.1 (12.6–38.3)	99.8 (98.8–100)	90.0 (59.6–98.2)	93.8 (91.3–95.6)
Ugandan system (Fig 2b)	61 (12.3)	20.5 (10.8–35.5)	100	100	93.6 (91.1–95.5)
WHO immunologic and clinical** (Fig 2c)	72 (14.5)	30.8 (18.6–46.4)	100	100	94.4 (92.0–96.1)

Fig. 1. Performance of Cambodian scoring system, Ugandan scoring system and WHO clinical and immunologic algorithms for targeting viral load assays. (a) Cambodian algorithm, (b) Ugandan algorithm and (c) WHO clinical immunological algorithm. ART, antiretroviral therapy; WHO, World Health Organization. (d) Accuracy of Cambodian, Ugandan and WHO two-step algorithms to predict viral failure (validation population).

being tested. This offered lower sensitivity at 20.5% (95% CI 10.8–35.5%) and the PPV was 100%. WHO clinical and immunologic criteria relied on a similar number of assays to the Ugandan system at 72 (14.5%) but offered the greatest sensitivity at 30.8 (95% CI 18.6–46.4%) and a PPV of 100%.

The validation population study interviewers judged 45 (9.1%) to be failing virologically, and 11 of these truly were – therefore, sensitivity was 28.2% (11/39 cases).

Discussion

This is the first study to develop and externally validate a system for targeted viral load assays in a sub-Saharan African population. Through confirmatory viral load testing inappropriate use of second-line therapy is avoided. Unfortunately, sensitivities of each algorithm were low, limiting their usefulness. The locally developed Ugandan scoring system offered no additional value over WHO criteria in its accuracy to predict viral failure.

The three systems evaluated in this study relied on viral load testing to confirm treatment failure and this significantly improved their PPV compared with clinical or immunologic criteria alone. In the validation population, WHO clinical and immunologic criteria diagnosed treatment failure in 72 cases – 60 of these were false. This low PPV is not atypical of those previously reported [14,16–18]. The concept of undertaking confirmatory viral load testing has been recommended by various researchers [15,16,27,36] and has recently been advocated by the WHO [37]. This is important as second-line regimens are expensive [5] and avoiding their inappropriate use with confirmatory viral load testing has been shown to be highly cost-effective [24].

There has been little research evaluating systems for targeted viral load assays. The new Ugandan system performed poorly requiring viral load assays in 12.3% of cases and only identifying 20.5% of cases of viral failure. Study limitations may have contributed to this. Participants were enrolled in an approved observational study and were likely to have received closer medical attention than is available at a typical resource-constrained ART provider. Separate episodes of viral failure from a single participant were included in the analysis as the system was designed for use in populations who may suffer prolonged treatment failure. Unlike in Cambodia [27], clinicians were not blinded to the viral load result and this may have influenced their reporting of symptoms or adherence, which may have spuriously increased PPV. Furthermore they may have switched patients at an earlier stage before clinical or immunologic failure became evident, which has been shown to occur in viral load monitored populations [38]. The definition of viral failure was based on a single measurement greater than 1000 copies/ml, whereas two consecutive episodes are regarded as a more reliable definition [7,18]. The symptom prurigo was not defined prospectively in either the derivation or validation population, which may make reporting of this unreliable.

In this study, the WHO immunologic and clinical criteria offered the greatest sensitivity at 30.8% (95% CI 18.6–46.4%). The Cambodian system identified fewer positive cases with sensitivity of 23.1% (95% CI 12.6–38.3%), but did rely on significantly fewer viral load assays: 33 (6.7%) compared with 72 (14.5%) for WHO criteria. Use of interviewer's opinion to target viral load assays offered similar accuracy to the Cambodian and WHO algorithms. This highlights the value of clinical judgement in diagnosing treatment failure. However, it should be noted that study interviewers had the advantage of undertaking a structured and detailed clinical history through completing questionnaires which may have inflated their diagnostic ability.

The Cambodian system was the only one not to offer a PPV of 100%. In cases with a score of 5 or more switching was recommended without confirmatory viral load testing. This resulted in false-positives suggesting that it would be more appropriate to always confirm viral failure before switching. A notable advantage of using scoring systems is that the cut-off at which to allocate viral load assays is flexible. As depicted in Table 4, lower cut-offs can be used to increase sensitivity at the cost of relying on more viral load assays. For example, if those with a Cambodian score of 1 or greater were tested, 21.4% would require an assay and sensitivity would be increased to 51.3% (95% CI 36.2–66.1%). ART providers could therefore choose a cut-off after weighing the risks associated with false-negative results against the costs of providing viral load assays. This may become more relevant as costs of viral load facilities reduce and access subsequently increases [9].

Although targeted viral load testing may represent the best available option for resource-constrained ART providers, the risks of not routinely monitoring viral load are not fully known and may be significant. In the absence of regular viral load testing, populations receiving first-line therapy develop significantly more resistance mutations [21,39]. These have been identified with high prevalence in resource-constrained settings [22,40] and may limit options for second-line therapy [40]. In addition, discordant responses, including incomplete viral suppression in the presence of immunologic response, are known to occur in resource-limited settings [41] and have been associated with increased mortality [42]. However, evidence to suggest routine viral load monitoring confers a significant reduction in mortality is limited [43]. A computer simulation model predicted little survival benefit with access to viral load monitoring [44], and results from three randomized trials comparing monitoring strategies in resource-limited settings have shown nonsignificant differences in rates of AIDS defining events and deaths [45–47]. On the basis of current research derived from clinical trial settings achieving high adherence rates, regular viral load testing does not appear cost-effective [44,48], although the longer-term benefits outside of research settings are not fully understood.

Conclusion

The study suggests that locally developed algorithms based on clinical and immunological criteria offer little additional value as tools for targeting viral load assays when compared with WHO criteria or clinical judgement. When possible, confirming viral load before switching to second-line therapy is recommended as clinical and immunological criteria have poor PPV for identifying viral failure, and inappropriate switches may be costly. One advantage of using a scoring system, particularly as costs of viral load assays decrease, is that ART providers can choose a lower threshold for testing; increasing sensitivity over WHO criteria through more

frequent viral load monitoring. The availability of a low cost point of care viral load assay is therefore eagerly awaited.

Acknowledgements

Conflicts of interest

The study was undertaken in part-fulfilment of an intercalated BmedSc and M.A. was financially supported by the College of Medical and Dental Sciences, University of Birmingham in this work. Funding was also provided by the Infectious Diseases Institute, Kampala, Uganda and by the Division of Intramural Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health. All other authors are employees of the institutions listed above and undertook activity as part of their core role or as an additional activity.

Author contributions: M.A. conceived the study and the design was developed by all authors with supervision from L.R. A.K., G.R., B.C., Y.M. and S.R. were responsible for original data collection and management. Data analysis was undertaken by M.A. under the guidance of J.M. The study was drafted by M.A. and all authors have contributed to critical revision of the manuscript. All authors approved the final version of the manuscript. L.R. is the guarantor of this study.

References

1. WHO 2009 – WHO, UNAIDS, UNICEF. **Towards universal access: scaling up priority HIV/AIDS interventions in the health sector: progress report, September 2009.** http://www.who.int/hiv/pub/2009progressreport/en/. [Accessed 19 February 2010]
2. Pujades-Rodríguez M, O'Brien D, Humblet P, Calmy A. **Second-line antiretroviral therapy in resource-limited settings: the experience of Medecins Sans Frontieres.** *AIDS* 2008; **22**: 1305–1312.
3. Collini P, Schwab U, Sarfo S, Obeng-Baah J, Norman B, Chadwick D, *et al*. **Sustained immunological responses to highly active antiretroviral therapy at 36 months in a Ghanaian HIV cohort.** *Clin Infect Dis* 2009; **48**:988–991.
4. Keiser O. **Switching to second line ART, and mortality in resource-limited settings: collaborative analysis of treatment programs in Africa, Asia and Latin America.** *Presented at 16th Conference on Retroviruses and Opportunistic Infections*; 2009; Montreal, Canada.
5. Lynen L, Van Griensven J, Elliott J. **Monitoring for treatment failure in patients on first-line antiretroviral treatment in resource-constrained settings.** *Curr Opin HIV AIDS* 2010; **5**:1–5.
6. Harries AD, Zachariah R, van Oosterhout JJ, Reid SD, Hosseinipour MC, Arendt V, *et al*. **Diagnosis and management of antiretroviral-therapy failure in resource-limited settings in sub-Saharan Africa: challenges and perspectives.** *Lancet Infect Dis* 2010; **10**:60–65.
7. Gazzard BG, BHIVA Treatment Guidelines Writing Group. **British HIV Association guidelines for the treatment of HIV-1 infected adults with antiretroviral therapy.** *HIV Med* 2008; **9**:563–608.
8. Hammer SM, Saag MS, Schechter M, Montaner JSG, Schooley RT, Jacobsen DM, *et al*. **Treatment for adult HIV infection: 2006 recommendations of the International AIDS Society–USA panel.** *JAMA* 2006; **296**:827–843.
9. Fiscus SA, Cheng B, Crowe SM, Dementer L, Jennings C, Miller V, *et al*. **HIV-1 viral load assays for resource limited settings.** *PLoS Med* 2006; **3**:e417.
10. Calmy A, Ford N, Hirschel B, Reynolds SJ, Lynen L, Goemaere E, *et al*. **HIV viral load monitoring in resource-limited regions: optional or necessary?** *Clin Infect Dis* 2007; **44**:128–134.
11. Schooley RT. **Viral load testing in resource-limited settings.** *Clin Infect Dis* 2007; **44**:139–140.
12. World Health Organization. *Antiretroviral therapy for HIV infection in adults and adolescents in resource-limited settings: recommendations for a public health approach (2006 revision)*. Geneva: WHO; 2006. http://www.who.int/hiv/pub/arv/adult/en/index.html. [Accessed 19 February 2010]
13. World Health Organization. *WHO case definitions of HIV for surveillance and revised clinical staging and immunological classification of HIV-related disease in adults and children*. Geneva: WHO; 2007. www.who.int/hiv/pub/guidelines/HIV-staging150307.pdf. [Accessed 19 February 2010]
14. Mee P, Fielding KL, Charalambous S, Churchyard GJ, Grant AD. **Evaluation of the WHO criteria for antiretroviral treatment failure among adults in South Africa.** *AIDS* 2008; **22**:1971–1977.
15. Moore DM, Awor A, Downing R, Kaplan J, Montaner JS, Hancock J, *et al*. **CD4+ T-cell count monitoring does not accurately identify HIV-infected adults with virologic failure receiving antiretroviral therapy.** *J Acquir Immune Defic Syndr* 2008; **49**:477–484.
16. Meya D, Spacek LA, Tibenderana H, John L, Namugga I, Magero S, *et al*. **Development and evaluation of a clinical algorithm to monitor patients on antiretrovirals in resource-limited settings using adherence, clinical and CD4 cell count criteria.** *J Int AIDS Soc* 2009; **12**:3.
17. Keiser O, Macphail P, Boulle A, Wood R, Schechter M, Dabis F, *et al*. **Accuracy of WHO CD4 cell count criteria for virological failure of antiretroviral therapy.** *Trop Med Int Health* 2009; **14**:1220–1225.
18. Reynolds SJ, Nakigozi G, Newell K, Ndyanabo A, Galiwongo R, Boaz I, *et al*. **Failure of immunologic criteria to appropriately identify antiretroviral treatment failure in Uganda.** *AIDS* 2009; **23**:697–700.
19. Castelnuovo B, Kiragga A, Schaefer P, Kambugu A, Manabe Y. **High rate of misclassification of treatment failure based on WHO immunological criteria in resource limited settings in Uganda.** *AIDS* 2009; **23**:1295–1296.
20. Rawizza H, Chaplin B, Rao T. **Comparison of Immunologic and Virologic Failure Criteria in a Large ART Program in Nigeria.** *Presented at 17th Conference on Retrovirus and opportunistic Infections*; 2010; San Francisco, California, USA.
21. Gupta R, Hill A, Sawyer AW, Cozzi-Lepri A, von Wyl V, Yerly S, *et al*. **Virological monitoring and resistance to first-line highly active antiretroviral therapy in adults infected with HIV-1 treated under WHO guidelines: a systematic review and meta-analysis.** *Lancet Infect Dis* 2009; **9**:409–417.
22. Hosseinipour MC, van Oosterhout JJG, Weigel R, Phiri S, Kamwendo D, Parkin N, *et al*. **The public health approach to identify antiretroviral failure: high-level nucleoside reverse transcriptase inhibitor resistance among Malawians failing first-line antiretroviral therapy.** *AIDS* 2009; **23**:1127–1134.
23. Petersen ML, van der Laan MJ, Napravnik S, Eron JJ, Moore RG, Deeks SG. **Long-term consequences of the delay between virologic failure of highly active antiretroviral therapy and regimen modification.** *AIDS* 2008; **22**:2097–2106.
24. Colby DJ. **Cost-effectiveness of viral load testing for suspected first-line antiretroviral drug therapy failure in Vietnam.** *Presented at 2009 HIV/AIDS Implementer's meeting*; 2009; Windhoek, Namibia.
25. Castelnuovo B, Byakwaga H, Menten J, Schaefer P, Kamya M, Colebunders R. **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.
26. Bisson GP, Gross R, Bellamy S, Chittams J, Hislop M, Regensberg L, *et al*. **Pharmacy refill adherence compared with CD4 count changes for monitoring HIV-infected adults on antiretroviral therapy.** *PLoS Med* 2008; **5**:e109.

27. Lynen L, An S, Koole O, Thai S, Ros S, De Munter P, *et al*. **An Algorithm to Optimize Viral Load Testing in HIV-Positive Patients With Suspected First-Line Antiretroviral Therapy Failure in Cambodia.** *J Acquir Immune Defic Syndr* 2009; **52**:40–48.
28. The Infectious Diseases Institute, Mulago Hospital, Makerere University in Kampala, Uganda. http://www.idi.ac.ug/.
29. World Health Organization. *Scaling up antiretroviral therapy in resource-limited settings: treatment guidelines for a public health approach (2003 revision)*. Geneva: WHO; 2004. www.who.int/hiv/pub/prev_care/en/arvrevision2003en.pdf. [Accessed 19 February 2010]
30. Uganda Ministry of Health. **National antiretroviral treatment and care guidelines for adults and children.** *Uganda Ministry of Health;* 2003. www.aidsuganda.org/pdf/ARV_Clinical_Guidelines_Final_draft.pdf. [Accessed 19 February 2010]
31. Kamya MR, Mayanja-Kizza H, Kambugu A, Bakeera-Kitaka S, Semitala F, Mwebaze-Songa P, *et al*. **Predictors of long-term viral failure among Ugandan children and adults treated with antiretroviral therapy.** *J Acquir Immune Defic Syndr* 2007; **46**:187–193.
32. 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.
33. Spiegelhalter DJ. **Statistical methodology for evaluating gastrointestinal symptoms.** *Clin Gastroenterol* 1985; **14**:489–515.
34. Berkley JA, Ross A, Mwangi I, Osier FHA, Mohammed M, Shebbe M, *et al*. **Prognostic indicators of early and late death in children admitted to district hospital in Kenya: cohort study.** *BMJ* 2003; **326**:361.
35. Nakanjako D, Kiragga A, Ibrahim F, Castelnuovo B, Kamya MR, Easterbrook PJ. **Sub-optimal CD4 reconstitution despite viral suppression in an urban cohort on Antiretroviral Therapy (ART) in sub-Saharan Africa: frequency and clinical significance.** *AIDS Res Ther* 2008; **5**:23.
36. Kantor R, Diero L, DeLong A, Kamle L, Muyonga S, Mambo F, *et al*. **Misclassification of first-line antiretroviral treatment failure based on immunological monitoring of HIV infection in resource-limited settings.** *Clin Infect Dis* 2009; **49**:454–462.
37. World Health Organization. **Rapid advice: antiretroviral therapy for HIV infection in adults and adolescents.** *WHO*; 2009. http://www.who.int/hiv/pub/arv/rapid_advice_art.pdf. [Accessed 12 March 2010]
38. Keiser O, Tweya, Boulle A, Braitstein P, Schechter M, Brinkhof M, *et al*. **Switching to second-line antiretroviral therapy in resource-limited settings: comparison of programmes with and without viral load monitoring.** *AIDS* 2009; **23**:1867–1874.
39. Reynolds S, Sendagire H, Newell K, Castelnuovo B, Kiragga A, Namugga I, *et al*. **VLM reduces the rate of accumulated genotypic resistance to commonly used ART in Uganda.** *Presented at 18th Conference on Retroviruses and Opportunistic Infections;* 2011; Boston, Massachusetts, USA.
40. Sungkanuparph S, Manosuthi W, Kiertiburanakul S, Piyavong B, Chumpathat N, Chantratita W. **Options for a second line antiretroviral regimen for HIV type 1-infected patients whose initial regimen of a fixed-dose combination of stavudine, lamivudine, and nevirapine fails.** *Clin Infect Dis* 2007; **44**: 447–452.
41. Tuboi SH, Brinkhof MWG, Egger M, Stone RA, Braitstein P, Nash D, *et al*. **Discordant responses to potent antiretroviral treatment in previously naïve HIV-1-infected adults initiating treatment in resource-constrained countries. The Antiretroviral Therapy in Low-Income Countries (ART-LINC) Collaboration.** *J Acquir Immune Defic Syndr* 2007; **45**:52–59.
42. Tuboi SH, Pacheco AG, Harrison LH, Stone RA, May M, Brinkof MW, *et al*. **Mortality associated with discordant responses to antiretroviral therapy in resource-constrained settings.** *J Acquir Immune Defic Syndr* 2010; **53**:70–77.
43. Chang LW, Harris J, Humphreys E. **Optimal monitoring strategies for guiding when to switch first-line antiretroviral therapy regimens for treatment failure in adults and adolescents living with HIV in low-resource settings.** *Cochrane Database Syst Rev 2010*:CD008494. doi: 10.1002/14651858.CD008494.
44. 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.
45. Coutinho A, Mermin J, Ekwaru J, Were W, Bunnell R, Kaharuza F, *et al.* **Utility of routine viral load, CD4 cell count, and clinical monitoring among HIV-infected adults in Uganda: a randomized trial.** *Presented at 15th Conference on Retroviruses and Opportunistic Infections*; 2008; Boston, Massachusetts, USA.
46. Jourdain G, Ngo-Giang-Huong N, Le Coeur S, Traisaithit P, Barbier S, Techapornroong M, *et al.* **PHPT-3: a randomized clinical trial comparing CD4 vs. viral load ART monitoring/switching strategies in Thailand.** *Presented at 18th Conference on Retroviruses and Opportunistic Infections;* 2011; Boston, Massachusetts, USA.
47. Kouanfack C, Laurent C, Laborde-Balen G, Aghokeng A, Tchatchueng-Mbougua J, Boyer S, *et al.* **HIV viral load, CD4 cell count, and clinical monitoring vs. clinical monitoring alone for ART in rural hospitals in Cameroon: Stratall ANRS 12110/ESTHER Trial, a randomized noninferiority trial.** *Presented at 18th Conference on Retroviruses and Opportunistic Infections*; 2011; Boston, Massachusetts, USA.
48. Bendavid E, Young SD, Katzenstein DA, Bayoumi A, 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.