

HATiP

HIV & AIDS Treatment in Practice

Issue 94 | 31 October 2007



In this issue:

Defining, and then watching for, treatment failure; **page 2**

- How does immunological or clinical failure correlate with virological failure?
- Kigali, Rwanda
- Dar es Salaam, Tanzania
- Malawi
- Table A: Relationship of clinical conditions to virological failure in Malawi ART programme
- Kampala, Uganda
- Cambodia
- MSF in Rwanda
- Table B. WHO clinical failure criteria as a test for virological failure - MSF/Rwanda
- Table C: Alternate viral load thresholds
- The Home Based AIDS Care Study
- Comparative cost effectiveness
- A word about algorithms and rationing viral load
- Making sense of it all
- References

Defining, and then watching for, treatment failure

How does immunological or clinical failure correlate with virological failure?

Several studies presented at the HIV Implementers Meeting in Kigali, Rwanda, and the International AIDS Society conference in Sydney asked whether recommended definitions of clinical and CD4 cell count failure can be used to detect virological failure. The answers varied a bit by site and definitions of failure used.

But first, it should be noted that the WHO guidelines never suggest that clinical and CD4 cell failure are surrogates for virological failure. Rather, they suggest that, in the absence of switching options, these may be better monitoring strategies to make certain that a person with HIV has received every last drop of possible effectiveness from the first line regimen before going onto the second line regimen. Whether this is the best time to switch for the patient's long-term outcome is another matter to debate.

Nevertheless, one consistent finding among many of the following studies is that often clinical and immunological failure criteria are met in people who have undetectable viral loads.

Kigali, Rwanda

At the HIV Implementers' Meeting, Dr Fabienne Shumbusho presented data from the Biryogo Social and Health Center in Kigali, Rwanda, which put 728 people on first line ART between July 2004 and October 2006. After an average of 20 months on treatment, 75 had signs of clinical failure or immunological failure (using the suggested WHO definitions but with monitoring every three months). Viral load is not available for routine use in Rwanda, but it can be used to confirm treatment failure before switching to the more expensive second line regimen, after assessment for good adherence. If the patient has low but detectable viral load (below 10,000 copies/ml) the test is repeated after three months.

Of the 75 "failing" patients, immunological failure was detected in 42; clinical failure in 22; and 11 had signs of both. However, viral load testing showed that only 4 (9.5%) of the patients with immunological failure, 1 (4.5%) of those with clinical failure, and 1 (9%) of the patients with both had a detectable viral load. "Overall, only 6 out of 75 persons — which is only 8% — were found with virological failure. These patients with virological failure were the only ones eligible for a change of ART regimen. The remaining 92% — this is 69 persons — had undetectable viral load and did not need to change treatment. They remained on their regimen," she said.

Note: there was one subject who had a low viral load who was given adherence support and who became undetectable three months later and did not switch regimens. Only four of the patients had viral loads over 10,000 copies/ml.

Dr Shumbusho also worried that, based upon failure rates reported in the literature, there might be as many as 47 patients out of the 728 with detectable viral loads, but who were not showing signs of clinical or immunological failure. However, this study did not actually screen for that.

Dar es Salaam, Tanzania

Dr Grace Magembe presented a similar analysis focusing on immunological failure from MDH, a collaborative programme

between Muhimbili University College of Health Sciences, Dar es Salaam City Council and Harvard School of Public Health in Tanzania.

At MDH clinics, ART regimen changes are usually based on immunological and clinical criteria, defined slightly differently from WHO suggestions. Notably, immunological failure is defined as a 30% drop in CD4 counts from peak value or a return to pre-ART baseline or lower (monitoring is every four months). Then, to make certain that a "failing" CD4 test was not simply a chance result, MDH performs a confirmatory test one week later, and those patients who are still suspected as failing are given a viral load test, which has only recently become available. If viral load is detectable after more than 24 weeks on treatment, the patient is switched to second line.

Over a follow up period between November 2004 and November 2006, 1657 patients on ART who had suspected immunological failure were given a repeat confirmatory test. Only 994 (60%) were confirmed. "Compared to a single CD4 test, a confirmatory CD4 count reduces the number of patients labelled as immunological failure — who would otherwise be put in a second line regimen — by 40%," said Dr Magembe.

Among the 994 confirmed to have immunological failure, only 179 (18%) of patients had detectable viral loads (confirmed virological failure). For the sake of this study, the researchers also tested those without confirmed CD4 failure, 46 (7%) out of 663 were found to have detectable viral load.

"This analysis shows that the use of immunological criteria alone cannot precisely predict patients with virological failure and may lead to significant misclassification of therapeutic responses," concluded Dr Magembe.

Malawi

Over 100,000 people have been started on first-line ART (d4T/3TC plus nevirapine or AZT/3TC plus efavirenz) in Malawi, one of the most resource poor countries in the world. However, it can be expected that eventually there will be a high demand for second line therapy, which at present in this country consists of AZT/3TC/tenofovir plus Kaletra. This regimen is considerably more complicated, with potential drug interactions, a higher pill burden and lab monitoring required (for safety reasons). It is also dramatically more expensive (\$150 per month versus \$13 a month paid for the first line regimen.)

"Therefore for us it's critical to determine the best definition for our setting of ARV failure to avoid unnecessary switches to second line treatment," said Dr Mina Hosseinipour, who presented the results of a two-centre prospective study comparing clinical and immunological definitions of ART failure to virological failure (> 400 copies/ml) at the IAS meeting in Sydney. One of these sites, the Lighthouse clinic in Lilongwe offers six-monthly CD4 testing, while the other site, the Queen Elizabeth Central Hospital in Blantyre, does not. In Malawi, the definition of clinical ART failure is a new or progressive worsening or recurrence of a WHO Stage IV condition, and the definition of immunological failure is a greater than 30% decline from peak or below pre-treatment values after six months on treatment and confirmed good adherence.

From the end of December 2005 until January 2007, 152 patients who met the Malawi definition of ART failure were identified. Of these 75% were immunologic failures, 21% met the clinical criteria and 6 (patients) met both clinical and immunological criteria for failure. The mean age was 39 years; 51% were women;

the mean CD4 cell count was 182 (range 1 to 927). The duration on ART at the time of the failure was 33 months (range 7 to 120).

Ninety out of 152 (59%) of the failure suspects were confirmed to have detectable viral loads: 68% of those with signs of clinical failure and 58% of those with signs of immunological. During the question and answer session, Dr Hosseinipour added that using a higher viral load threshold (above 10,000 copies/ml) wouldn't have affected these calculations much, because the majority had viral loads well over 10,000 copies/ml (mean of 37 000).

Dr Hosseinipour gave an interesting breakdown of which clinical conditions seemed to be associated with virological failure (see table). She noted that if KS were omitted as an endpoint (since it can confound the analysis because chemotherapy lowers CD4 cell counts and has been known to progress in people with HIV in Africa despite effective ART), clinical criteria would diagnose treatment failure correctly 75% of the time and immunological criteria 66% of the time.

Table A: Relationship of clinical conditions to virological failure in Malawi ART programme

| Stage 4 condition | N | % Confirmed Failures > 400 copies/ml |
|-------------------------|---|--------------------------------------|
| KS | 9 | 31% |
| Extrapulmonary TB | 8 | 75% |
| Wasting | 8 | 75% |
| Oesophageal candidiasis | 7 | 100% |
| Cryptococcal meningitis | 2 | 0% |

In a multivariate analysis, risk factors for confirmed virological failure included being on ART for more than three years, with an odds ratio of 8.8 [95% confidence interval of 2.9-26.7] and having a CD4 cell count below 200, with an odds ratio of 3.0 [1.2-7.8].

"We would recommend that viral load testing should probably be used to confirm ART failure prior to switching – particularly in immunologic cases and for KS patients. Immunologic ART failure is particularly vulnerable to misclassification," she said; suggesting that combinations of clinical criteria, the duration of ART and the absolute CD4 count should be looked at as a potentially better ART failure definition.

In the question and answer session, Dr Hosseinipour added that this switching approach would probably lead to long periods of viral replication in patients who fail virologically before showing clinical or immunological signs of failure. Thus, in a second part of this study, the researchers will be genotyping each of the failing patients to see the level of resistance that has developed, to see how well they respond to the second line treatment and whether it varies according to the burden resistance conferring mutations.

"Obviously it's not ideal," she said, "but we need to come up with the best solution for Malawi that's practical."

Kampala, Uganda

The Infectious Disease Institute in Kampala has been providing free ART to 4,200 people with HIV since September 2004. At the IAS meeting in Sydney, Dr Apollo Basenero presented an evaluation of the use of a consensus meeting (attended by doctors, nurses, counsellors and pharmacists who work at the clinic) to consider CD4, clinical and adherence assessments in patients suspected to

be failing their first-line regimen (failure is defined according to the WHO guidelines and adherence is measured through self-report and pill count). CD4 testing takes place every six months.

At the consensus meetings, the team decides whether to do a viral load test or employ other interventions like intensive adherence counselling before switching to second line treatment.

When patients are discussed in these meetings, they are classified into one of three categories:

- 1 **Clearly failing:** A switch to second line is recommended for patients with good adherence (>95%) but with proven clinical, immunologic, and if the viral load is present, virological failure.
- 2 **Poor adherence:** Those with clinical and/or immunologic failure; but with an adherence of <95%. Adherence counselling is emphasized during every subsequent clinic visit, and a CD4 cell count repeated after three or six months. If still failing, a switch is recommended.
- 3 **Inconclusive:** A viral load test is ordered in patients with immunologic failure if they are clinically stable and have good adherence.

Since August 2005 to March 2007, 100 patients with suspected failure have been discussed in the switch meeting: 20 patients were put in category 1; 26 patients in Category 2; and 54 patients were in category 3. Viral loads were performed in as many patients as possibly at the time of switching to confirm how many had detectable viral loads (over 400 copies/ml).

In category 1, viral load was detectable in 15 out of 15 patients who could be tested. In category 2, the adherence intervention appeared to improve CD4 cell counts (by 36-113 cells) in 16 (62%); however, the repeat CD4 cell tests suggested that 10 out of 26 (38%) were still failing. These were switched to second line and viral load was detectable in each case at switching.

But only 30 (56%) out of the 54 in category 3 had detectable viral loads, with a median of 93,686 and a range of 2,611 to 694,993 copies/ml, while the viral load was undetectable in 24 (44%).

The data suggest that a consensus meeting with experienced clinical staff may be better able to determine when clinical events and CD4 cell decline constitute true failure. However, for programmes that only want to switch to second line ART in patient with virological failure, relying on CD4 cell counts alone could be still misleading and would lead to premature switches.

Cambodia

Dr Sokkab An of Sihanouk Hospital Center in Phnom Penh presented data from his site on the sensitivity and specificity of WHO criteria (the 2003 version, in which WHO stage 3 events excluding TB are also considered clinical failure) for virological failure (a viral load >50 copies/ml) in patients on first line ART for more than 6 months, and for possible predictors of virological failure.

This was a cross-sectional study in which a viral load test was performed in 399 people on ART at the site, and the results were then matched to the patient's clinical and immunological status at that time, as well as to changes since going onto ART, and to other baseline characteristics.

About half of the subjects were men. At baseline, 47.6% and 36.8% were WHO stage III and IV respectively; the median age was 34; median body mass index 19; and median CD4 cell count was 52 (interquartile (IQR) range 13-131). All but one were on a nevirapine- or efavirenz-based regimen. The median time between

the start of ART and performance of the viral load test was 12.8 months (IQR 10.7-19.5).

Thirty-three out of the 399 (8.3%) of the participants were found to have detectable viral loads.

Clinical and immunological failure criteria were found to be very insensitive measures for detectable viral load (defined as the gold standard for treatment failure). Both clinical failure and CD4 tests picked up only about 18% of the cases of detectable viral load individually, and about 30% when combined. Since a significant number of clinical/immunological failures also had undetectable viral loads, the positive predictive value of CD4 or clinical failure for virological failure was quite low. However, the criteria were more specific (91% and 98% respectively, 89% when combined) with a negative predictive value of 93% (in other words, patients who did not show signs of CD4 or clinical failure were more likely to have suppressed viral loads).

In the multivariate analysis, predictors for virological failure included:

- male gender, with an odds ratio of 3.1 [95% CI, 1.3-7.3], $p=0.008$, possibly because of differences in health seeking behaviour or risk behaviour specific to males in this setting.
- stable or declining CD4 cell counts within the past six months versus an increase, with an odds ratio of 2.6 [1.2-5.6], $p=0.016$.
- a decline of more than 300 in the total lymphocyte count (TLC) within the past six months vs a smaller decrease, or increase, odds ratio 2.9 [1.4-6.3].

This last finding is interesting because studies have been reaching different conclusions about the utility of TLC as a substitute for CD4 cell responses to monitor treatment outcomes for years. Some recent studies suggest part of the problem may be wide variability in results from one setting to another (possibly indicating the need to establish local norms and quality control for haematology analysers) (Marshall, Mbanya).

But if these two factors (change in CD4 and TLC) are added to immunological and clinical failure to refine WHO failure criteria, the sensitivity for virological failure improves from 30% to 79%, but it still miscategorises some virological successes as failures, so the positive predictive value of this 'test' remains low. However, the negative predictive value increases slightly to 96%.

"This high negative predictive value could help identify patients that do not need a viral load to make a treatment decision," concluded Dr An.

MSF in Rwanda

Dr Johan Van Griensven of Médecins sans Frontières (MSF) presented a similar analysis at the HIV Implementers' Meeting, looking at the sensitivity/specificity and predictive values of immunological failure criteria for identifying virological failure at two MSF supported health centres in Rwanda. The ART programme at these sites began in late 2003 and has placed over 2,500 people on treatment. Viral loads were performed routinely after one year on ART.

This analysis was based on viral load results and immunological data from 863 adult patients in this programme. Baseline CD4 data were available for 604 patients; data on recent and peak CD4 levels were available for 862.

88% of the 863 subjects had undetectable viral loads (defined as < 40 copies/ml) after one year on treatment. Once again, the

sensitivity and positive predictive value of CD4 failure identifying people with a detectable viral load were quite low (See Table B).

Table B. WHO clinical failure criteria as a test for virological failure - MSF/Rwanda

| CD4 Failure Criterion | Sensitivity (%) | Specificity (%) | Positive Predictive Value (%) | Negative Predictive Value (%) | Acc (%) | AUC |
|-----------------------|-----------------|-----------------|-------------------------------|-------------------------------|---------|------|
| Fall below baseline | 12.3 | 90.8 | 15.5 | 88.3 | 81.3 | 0.52 |
| 50% fall from peak | 23.3 | 89.9 | 23.8 | 89.6 | 81.9 | 0.57 |
| Both | 27.4 | 81.9 | 17.2 | 89.1 | 75.3 | 0.55 |

"The main problem is the false positives and the false negatives," said Dr Van Griensven. "For the false positives, we have 16% — so these are patients with an undetectable viral load who we would define as having treatment failure based on the CD4 cell count. So if we were being strict, we would put these patients onto second line therapy — without any need. The other problem is false negatives. These patients have a detectable viral load, but we don't pick them up with the CD4 cell count measurements."

Dr Van Griensven then adjusted the definitions of failure to see whether that altered the calculations. For example, if a '30% fall from peak values' was used as the definition of failure (criteria which are used in Tanzania and Malawi), the sensitivity would increase somewhat to 38%. Likewise, a paper in last December's *JAIDS* suggested that using 'no increase in CD4 cell counts' as a definition of treatment failure would detect about a third of the people with virological failure (Moore), while a paper by Bisson et al in Botswana reported that using a CD4 cell counts rise of 'less than a 50 cells rise above baseline' as a criteria for failure captured as many as 61% of the virological failures.

Using the latter failure criterion in the MSF cohort, the sensitivity would increase to 33%, and to 56% when combined with the previously mentioned criteria. But all of this comes at the expense of specificity.

So then Dr Van Griensven looked at using different thresholds for virological failure. "Maybe we can use a threshold of more than 1,000 copies as treatment failure, at least this will eliminate virological blips [transient low but detectable viral load results which tend to resolve spontaneously on treatment] which is especially important if you only have single measurements." He also looked at 10,000 copies/ml, because of its proposed importance in making a treatment decision (as per the WHO guidelines).

At 1000 copies/ml, the performance of immunological criteria improved only slightly, but at 10,000 copies/ml, the improvement was more pronounced. See Table C.

Table C: Alternate viral load thresholds

| Threshold for viral load | WHO CD4 failure criteria | | Alternative CD4 failure criteria | |
|--------------------------|--------------------------|-----------------|----------------------------------|-----------------|
| | Sensitivity (%) | Specificity (%) | Sensitivity (%) | Specificity (%) |

| | | | | |
|----------------|------|------|------|------|
| 40 copies/ml | 27.4 | 81.9 | 56.2 | 65.2 |
| 1000 copies/ml | 28.6 | 81.5 | 59.5 | 64.2 |

| | | | | |
|-----------------|------|------|------|------|
| 10000 copies/ml | 47.6 | 81.8 | 80.9 | 64.1 |
|-----------------|------|------|------|------|

Overall, these data suggest again, that CD4 failure only poorly detects virological failure. However, Dr Van Griensven focused on the positives.

"I think we can still look for some good use for CD4 counts, and one is to use it or to consider it as a screening tool to select the patient where we need to do a viral load. If we would use WHO criteria in that way, we would do viral load measurements in 20% of the patients and would detect half the patients who need second-line therapy. If we modify the criteria a bit, we would do viral loads in 37% of the patients and we would detect 86% of the patients who need second line therapy," he said.

He also suggested that doing repeat measurements and incorporating other clinical and laboratory data could improve the detection of patients who may need to be switched on to second line.

"Over the long term, we need to increase access to viral load measurement, which is adapted to resource constrained settings and we need to go for development of alternative tests; at a reduced price with better logistics: with no need for refrigeration; ideally can be done on site; results available the same day and minimum training required. The same way you do a pregnancy test, HIV test or glucose measurements. Maybe also we don't need exact values — at least for these settings, but can use different thresholds," he concluded.

The Home Based AIDS Care Study

Despite the previous findings, the first data from a large prospective study comparing monitoring strategies, the Home Based AIDS Care (HBAC) study, suggests that when it comes to preventing morbidity and mortality, CD4 monitoring of people on ART can perform quite well — at least for the first three years on treatment and when adherence is excellent. However, clinical monitoring alone was associated with an increased rate of new AIDS-defining events and a trend towards increased mortality when compared to either CD4 cell count or virological monitoring.

Although the fewest clinical events occurred in the virological monitoring arm, the difference when compared to CD4 monitoring did not reach statistical significance— possibly because there were so few poor outcomes in this cohort during the course of the study. In other words, in this study, CD4 cell count monitoring resulted in equivalent clinical outcomes out the three years of follow-up. However, routine viral load testing did add considerably to the expense of monitoring (see below).

HATIP has previously made reference to this study, because of the exceptionally high standard of care offered to its participants despite being located in an extremely remote and under-resourced setting. The HBAC programme offers comprehensive holistic family-based care in the rural Tororo and Busia Districts in eastern Uganda, which includes a basic preventive package of care (described here: <http://www.aidsmap.com/cms1234609.asp>) and intensive community-based support, including weekly home visits by trained lay workers with delivery of ART and TB medications and routine symptom and adherence assessments (see

<http://www.aidsmap.com/cms1234974.asp>). With this model of care, clinical symptoms and adherence problems would probably be detected sooner than at facility-based clinics with quarterly or even less frequent clinical monitoring.

So Dr Mermin and colleagues conducted a randomised study evaluating whether the approach to monitoring would lead to differences in the clinical outcomes (severe morbidity, essentially new AIDS defining events, and mortality) in people on ART during three years of follow-up, as well as the cost-effectiveness of the different monitoring regimens.

- Arm A: quarterly viral loads, CD4 cell counts and clinical monitoring (through weekly home visits)
- Arm B: quarterly CD4 cell counts and clinical monitoring (through weekly home visits)
- Arm C: clinical monitoring alone, provided by weekly home visits (see the annex on HBAC symptom/illness screening)

The definition of virological failure in Arm A was somewhat flexible. A patient could be categorised as failing virologically if they had two consecutive detectable viral loads (over 500 copies/ml) after six months of therapy; but if the viral load was between 50 to 5000 copies/ml and the participant was clinically well, first-line therapy could be continued. If there were clinical events or a fall in CD4 cell counts, the bias was to continue the first-line regimen as long as their viral load was suppressed.

For Arm B, failure was defined as a persistently declining CD4 cell count measured on two separate occasions and/or clinical failure.

In Arm C, definitions for clinical failure included:

- Unintentional weight loss of greater than 10%
- Appearance of CDC Category C illness (see <http://www.aidsmap.com/cms1031859.asp>)
- Diarrhoea or fever for more than one month without correctable cause, and
- New or recurrent oral, oesophageal or vaginal candidiasis without a known cause

A clinical case conference, involving a multi-disciplinary team of physicians, nurses, pharmacists and counsellors, was held to report and discuss information regarding all deaths, hospitalizations, opportunistic illnesses, abnormal labs and decided whether changes in ART regimens were necessary. But in each arm, Dr Mermin stressed that, "the first response to an indication of treatment failure was adherence counselling and support."

The study enrolled 1,116 ART-naïve people with CD4 cell counts ≤ 250 cells or WHO clinical stage III or IV disease. A total of 1,094 people started the first-line regimen: stavudine/lamivudine (d4T/3TC) plus nevirapine (efavirenz was substituted for nevirapine in those with concomitant TB treatment). The second line regimen was lopinavir/ritonavir (*Kaletra*), didanosine (ddI) and tenofovir.

Baseline characteristics were well matched (the median age was 38, the CD4 cell count was between 127-131, and viral loads were over 200,000 copies/ml) with the exception of gender (67% of the participants in Arm C were women, compared to 75% in Arms B and C, $p=0.01$).

After a median follow-up of three years, there were:

- 126 deaths (11.2% of the total population), 47% of which occurred within the first three months of therapy
- 148 new AIDS-defining illnesses, 57% of which were in the first three months
- 61 (5.8%) of participants had two consecutive viral loads greater than 500 copies/ml after the first 6 months

- Only 28 (2.7%) of participants were ultimately switched to second line drugs.

In a Kaplan-Meier analysis of the time to an event of severe morbidity or mortality, the clinical monitoring arm had a significantly higher rate of morbid events and/or death than Arm A, whether using an intention-to-treat approach from the date of randomization ($p = 0.02$) or a per protocol approach starting three months after initiating ART ($p=0.004$). The per protocol analysis is particularly important because the first laboratory results after baseline were not drawn until three months after initiation of ART. In the per protocol analysis, the difference was also statistically significant in the comparison of Arm B to Arm C ($p=0.034$). There were no significant differences in time to death using the intent-to-treat or per protocol analysis.

After adjusting for age, gender, baseline CD4 cell count, viral load, body mass index and depression scores, the rate of first severe morbidity or mortality advances was roughly one and a half times higher in Arm C than in either Arms A or B, in the intent-to-treat analysis (see Table below), and twice as high in the per protocol analysis. There was no significant difference between Arms A or B.

"Although we have not yet analysed the adherence data, in Arms A & B, participants with either viral load or CD4 cell count measurements that indicated failure were counselled and improved medication adherence and this may be the reason that viral loads improved without the need for a change to a second-line regimen," said Dr Mermin. "However, in Arm C we did not detect drug failure as rapidly and people had elevated viral loads for a longer period of time."

Dr Mermin noted research from the Western Cape in South Africa (previously reported in HATIP) by Dr Catherine Orrel and colleagues, which has recently been published in *Antiviral Therapy*. "Using a system of early detection of virological breakthrough leading to implementation of a targeted adherence strategy, 23 (53%) of 43 people with initial VL >1,000 copies/ml subsequently achieved full virological suppression," wrote Orrel et al. They also stressed that the timing of failure detection was critical, because when the adherence intervention comes too late, it is less likely to result in viral suppression.

Of course, that study referred strictly to viral load, though the study by Basenero et al reported at IAS also found that adherence support offered on the basis of CD4 failure was able at least to stabilise CD4 cell counts in a large proportion of the 'failing' patients (at least temporarily). It's not clear whether the improvement in viral load in Arm B could really be considered significant (though at least it didn't get worse). But while CD4 cell counts and viral load might be increasingly discordant with time on ART, there is a chance that CD4 cell counts picked up participants who simply weren't taking their medication at all (as opposed to those who were erratically adherent).

For example, CD4 cell counts can be sustained in people on incompletely suppressive for years, but studies have shown that people on incompletely suppressive ART regimens who stop treatment altogether suffer a rapid increase in viral load and a rapid fall in CD4 cell counts (Deeks).

Note, however, that in HBAC, the immunological definition of failure is simply two declining CD4 cell counts — criteria that could have a different sensitivity for early detection of virological failure. Perhaps, given common variations in CD4 cell counts, transient dips in CD4 cell counts prompted more adherence support.

"Our first effort was always adherence monitoring... so we first counselled people and then we re-assessed with the laboratory and clinical monitoring to see if they got better," said Dr Mermin. "But for the arm with clinical monitoring only we couldn't wait — We would have to switch them. That's maybe why we had a greater number of people [in the clinical monitoring arm] who were switched to second line regimens after an AIDS-defining illness."

"Clinical criteria alone were poorly sensitive in detecting early failure, and poorly specific in restricting drug changes to those participants with an elevated viral load," said Dr Mermin. "Clinical monitoring alone was associated with an increased rate of new AIDS-defining events and a trend towards increased mortality. This supports efforts to build laboratory capacity and provide laboratory monitoring."

"But there was no benefit seen for adding quarterly viral load measurements to CD4 cell counts, supporting WHO guidelines and suggesting that priority should be given to a standing access to CD4 cell count measurements — at least in patients adhering well to ART," added Dr Mermin.

Comparative cost effectiveness

Given the extremely low event rate for Arms A and B after the first six months during the first few years on treatment, and the fact that the vast majority of people still have undetectable viral load, and the fact that even fewer people in Arm B have been switched second-line ART, it is not surprising that a subsequent analysis, presented by Dr Jim Kahn at the HIV Implementers meeting, suggested that routine viral load monitoring was not cost effective at all — at least within this time frame.

"Adding CD4 to clinical monitoring, at a cost of \$831-838 per DALY [disability adjusted life year saved or averted] is about as cost effective as putting another person on ART in Tororo [district in Uganda] which we've separately estimated at about \$600 per DALY. We also found that adding viral load to CD4 and clinical monitoring has a cost per DALY averted between \$3,600 and nearly \$12,000, which is 4 to 20 times higher than the cost per DALY averted for CD4. This analysis suggests that CD4 monitoring or starting a patient on ART are economically preferable to viral load monitoring, at least in the context of this trial," he said.

However, both Dr Mermin and Dr Kahn conceded that this could change after year three years on treatment. Again, a number of other studies suggest that this is when failure starts becoming much more common (as described by the study in Malawi). So the HBAC study will continue with Arm C re-randomised to Arms A & B.

"It's important for us to determine cost-effectiveness and long-term outcomes for follow-up greater than 3 years - in case over time routine viral load measurements might become beneficial, as an increasing number of people develop resistant virus," said Dr Mermin.

A word about algorithms and rationing viral load

While there may be some uncertainty about when to switch (what failure criteria should trigger a switch), most clinicians and people with HIV would like to have access to laboratory results to help guide those treatment decisions. But the cold and hard truth is that viral load testing simply is not available for routine use in most settings because of cost and logistics. Even where it is starting to be scaled up, its use may have to be rationed for some time to come.

In the meantime, programmes have to make the most of the monitoring tools that are currently locally available. One option is to use algorithms with clinical and basic laboratory indicators as

proposed by Colebunders et al and described in HATIP #69. There has since been some criticism of the algorithm, which a letter by Lawn et al to the Lancet says is not good enough to use as a replacement for virological monitoring. However, they said it could have some use as a screening algorithm.

Using this strategy to assess 24 patients with confirmed virological failure, 16 (67%) were identified as “possible treatment failure” but the remaining eight (33%) were not. The most useful criteria were related to adherence assessment; however, the team wrote that “this again represents a very substantial improvement over existing WHO recommendations.”

One clinical symptom stood out as a potential flag for early treatment failure: the development of prurigo/pruritic papular eruption (PPE), sometimes called itchy bump disease, (hyper pigmented raised bumps that bear a striking similarity to spider bites). This relatively common clinical condition appears to be highly associated with virological failure (Lawn, Orrell, Wood).

However, other clinical and laboratory criteria included in the model were of limited use probably because they are “dependent upon development of immunological failure, which only occurs some time after development of virological failure.”

Nevertheless, the team concluded “use of this or a modified algorithm in the field might require a substantial level of training of health-care personnel, but we agree with the authors that its use may reduce the overall demand for viral load testing where resources are scarce.”

As Dr Van Griensven suggested, incorporating CD4 criteria (as CD4 tests becomes more widely available) into these algorithms could also improve the sensitivity.

Making sense of it all

Again, different programmes may take home different messages from these studies, however a few things seem clear.

- Support adherence! If adherence is near perfect, during the first few years on ART, the majority of patients will do well, regardless of how the regimen’s effectiveness is monitored.
- Routine laboratory monitoring, whether by viral load or CD4 cell counts, may provide an early indication of whether the patient is not taking their treatment, allowing for intensive adherence interventions, which could preserve the first line regimen. Viral load is by far the most sensitive measure but even CD4 counts may detect patients who have stopped or never started taking their medications. The question for HIV programmes is whether it is cost effective (and whether the capacity exists or can soon be developed) to perform viral loads to detect and respond to poor adherence problems or to detect early failure. If adherence is generally good at a site, then testing everyone may not be the most cost-effective way to identify the few who are poorly adherent. The answer may be different at sites where patients get less time with a healthcare worker.
- At the very least, viral load may be needed to confirm which patients are truly failing — especially in light of the extremely high cost of second line therapy. This might be the foot in the door for viral load testing, then as capacity increases, and the logistics are worked out, or more inexpensive point-of-care assays are developed, viral load monitoring could become more routine with time.
- Even if routine viral load monitoring was available everywhere, it isn’t absolutely clear what threshold of viral load should be used to make the decision to switch the second-line treatment (we will

discuss this more in an upcoming article on antiretroviral resistance in resource limited settings).

- CD4 cell testing should be made available everywhere, because it really is an indispensable tool for making a number of treatment decisions including when to start, when to provide more support, whom to monitor more closely, and probably when to switch as well. “We may have to start thinking about a using combination of CD4 counts, viral loads and clinical criteria that might then in combination provide us with even better clinical outcomes,” said Dr Mermin in Kigali. Indeed, 25 years of clinical research have demonstrated that CD4 cell counts are the best predictor of a person with HIV’s risk of imminent clinical progression to severe illness or death.
- In settings that cannot yet afford or yet handle the logistics of routine viral load, it may be possible to refine screening algorithms so that they are more sensitive or specific for virological failure. Some clinical endpoints, in particular PPE, might be good early indicators of a virological failure before immunological decline, but these obviously won’t occur in every failing patient. CD4 cell counts could help, especially if paired with other markers of immune activation such as CD38.

Many experts believe that all these apparent contradictions and controversies will only be resolved by further study — and we fully support operational research and the sharing of best practices from resource limited settings.

But before launching into any large simple trials in thousands of people with HIV that will take years to complete, we need to ask ourselves: What are we randomising people to and why?

If we really want to know the best failure criteria to use for switching when there is only one other switching option available, then we should think about whether those studies would be necessary if third and fourth line ART regimens were available.

It is not acceptable that people in resource -limited settings have to settle for just two ART regimens. It took four to five years for triple drug combination therapy to start being made available in Africa. Why should it take 10 to 20 years to get access to the newer drugs being used in the West? The drugs exist.

What is the point of ‘universal’ treatment access if it only lasts three to six years?

Ultimately, the monitoring strategy that is chosen depends upon what policy makers believe the future holds. Its up to us as advocates to make sure that future is bright.

References

An S et al. Predictors of virologic failure in a Cambodian setting: findings from a cross-sectional study: findings from a cross-sectional study at Sihanouk Hospital Centre of Hope, Phnom Penh. Fourth International AIDS Society Conference on HIV Treatment and Pathogenesis, Sydney, abstract WeAb103, 2007.

Basanero A et al. Inadequacy of clinical and immunologic criteria in identifying virologic failure of 1st line ART: the Ugandan experience. Fourth International AIDS Society Conference on HIV Treatment and Pathogenesis, Sydney, abstract WeAb102, 2007.

Bisson, GP. Diagnostic accuracy of CD4 cell count increase for virologic response after initiating highly active antiretroviral therapy. *AIDS* 20: 1613–1619, 2006.

Calmy A et al. HIV Viral load monitoring in resource-limited regions: optional or necessary? *Clinical Infectious Diseases* 44:128–134, 2007

Colebunders R et al. A new model to monitor the virological efficacy of antiretroviral treatment in resource-poor countries. *Lancet Infect Dis* 6: 53–59, 2006.

Cozzi-Lepri A. Evolution of drug resistance in HIV-infect patients remaining on a virologically failing combination antiretroviral therapy regimen. *AIDS* 21: 721–732, 2007.

Deeks SG et al. Duration and predictors of CD4 T-cell gains in patients who continue combination therapy despite detectable plasma viremia *AIDS* 16: 201–207, 2002.

Gandhi T et al. Long-term immunologic and virologic responses in patients with highly resistant HIV infection who are treated with an incompletely suppressive antiretroviral regimen. *Clin Infect Dis* 45: 1085 – 1092, 2007.

Hosseinipour M et al. Validating clinical and immunological definitions of antiretroviral treatment failure in Malawi. Fourth International AIDS Society Conference on HIV Treatment and Pathogenesis, Sydney, abstract WeAb101, 2007.

Lawn SD, Myer L, Orrell C, Bekker LG, Wood R. Early mortality among adults accessing a community-based antiretroviral service in South Africa: implications for programme design. *AIDS* 19: 2141–2148, 2005.

Lawn SD, Orrell C, Wood R. Evaluating a model for monitoring the virological efficacy of antiretroviral treatment in resource-limited settings. *Lancet* Vol 6: 385–386, 2006.

Magembe, G.E. et al. Role of combined, repeat CD4 count and viral load tests in defining first line ARV failure in resource poor countries. HIV Implementers' Meeting, Kigali Rwanda, abstract 1197, 2007.

Marshall S Quality control of the total lymphocyte count parameter obtained from routine haematology analyzers, and its relevance in HIV management. *Afr J Med Med Sci.* 35(2):161–163, 2006.

Mbanya D Monitoring antiretroviral therapy in HIV/AIDS patients in resource-limited settings: CD4 counts or total lymphocyte counts? *Int J Infect Dis.* 11(2): 157–160, 2007. (Epub 2006 Jun 9)

Mermin J. Efficacy and cost-effectiveness of routine viral load, CD4 cell count, and clinical monitoring of adults taking antiretroviral therapy in Uganda. HIV Implementers's Meeting, Kigali Rwanda, 2007.

Moore DM et al. Performance of immunologic responses in predicting viral load suppression: implications for monitoring patients in resource-limited settings. *J Acquir Immune Defic Syndr* 43(4):436–439, 2006.

Munderi P et al. Discordance between virological/immunological and clinical outcomes: NORA efficacy 4th IAS Conference on HIV Pathogenesis, Treatment and Prevention, Sydney, 22–25 July 2007

Murri R, Lepri AC, Cicconi P, et al. Is moderate HIV viremia associated with a higher risk of clinical progression in HIV-infected people treated with highly active antiretroviral therapy: evidence from the Italian cohort of antiretroviral-naïve patients study. *J Acquir Immune Defic Syndr* 41:23–30, 2006.

Orrell C et al. Conservation of first-line antiretroviral treatment regimen where therapeutic options are limited. *Antiviral Therapy* 12: 83–88, 2007.

Shumbusho, F., Ndirabatware, B. Viral load testing before advancing to 2nd line or salvage ART regimens: observations in 75 Rwandan patients experiencing clinical and/or immunological failure. HIV Implementers' Meeting, Kigali Rwanda, abstract 356, 2007.

Van Griensven, J. et al. Validity of WHO-recommended Immunological Criteria for Treatment Failure. HIV Implementers' Meeting, Kigali Rwanda, abstract 1039, 2007.

WHO. Antiretroviral therapy for HIV infection in adults and adolescents: towards universal access. Recommendations for a public health approach. 2006
<http://www.who.int/entity/hiv/pub/guidelines/artadultguidelines.pdf>

about HATIP

A regular electronic newsletter for health care workers and community-based organisations on HIV treatment in resource-limited settings.

The newsletter is edited by Theo Smart (Cape Town) and Keith Alcorn, NAM's Senior Editor (London).

For further information please visit the HATIP section of aidsmap.com