Monitoring HIV Antiretroviral Therapy in Resource-Limited Settings: Time to Avoid Costly Outcomes

Frederick K. Sawe¹ and James A. McIntyre²

¹HIV Program, Kenya Medical Research Institute/Walter Reed Project, Kericho, Kenya; and ²Perinatal HIV Research Unit, University of the Witwatersrand, Johannesburg, South Africa

(See the article by Kantor et al on pages 454–62)

There has been a dramatic and commendable rapid scale-up of human immunodeficiency virus (HIV) prevention and care services in resource-limited settings, including sub-Saharan Africa in the past 5 years; thousands of lives have been saved, and >3 million people are now receiving antiretroviral therapy (ART). Although much remains to be done to continue the scale-up of HIV prevention, care, and treatment, we must strive to hold onto gains we have achieved. More efficacious drugs, which are less toxic, are easier to take, and have greater genetic barriers to the emergence of resistance, as well as improved models on how to best deliver HIV services, are needed.

A critical area of need in resource-limited settings is improved access to appropriate laboratory tests for initiating and monitoring ART. To sustain the benefits of ART, maximal suppression of HIV is required. This is true both for the long-term prognosis of the individual and for potential community-level benefits, such as reduced risk of transmission. Currently, the only reliable way to ensure that treatment is achieving viral suppression is to measure HIV levels. Kantor et al [1] present additional evidence from western Kenya of the inadequacy of relying on immunological monitoring to predict virologic failure, in line with other recent results from similar settings [2, 3].

Although ART can sometimes be initiated on clinical criteria without CD4 cell count or viral load testing, these are necessary for optimal monitoring of ART. The World Health Organization recommends use of clinical and/or immunological monitoring (CD4 cell count) if viral load testing is not available [4]. This situation applies to most ART settings in resource-limited countries, where HIV load testing is most often unaffordable or unavailable.

An important goal in ART is ensuring maximum durability of current drug regimens, first through continued support of health care systems to ensure uninterrupted access to ART services and improved ART adherence support, but also through the identification of the early warning signs of potential virologic failure, before the development of multiple HIV drug resistance, which will limit the response to future ARTs. A number of studies have found that adherence to ART in Africa is high. Patients in Africa have been reported to take up to 90% of their prescribed antiretroviral [5, 6], and this has aided greatly in the scale-up of ART, and continued success of first-line regimens in resource-limited settings.

Access to second-line regimens remains a major barrier for many ART programs in resource-limited settings because of costs, infrastructure, technical capacity, and other logistic issues, despite the major price reductions in drug costs achieved through the interventions of groups such as UNITAID and the Clinton HIV AIDS Initiative. As the new ART programs mature in resource-limited settings, an increasing number of patients will experience treatment failure of first-line regimens, and the needs to switch to second-line ART will inevitably increase. With no third-line regimens available within “the public health approach” to providing ART in most resource-limited settings, a move to a second-line regimen represents the last chance for therapy for many people.

Viral load testing is the gold standard for monitoring patients on ART, where resources are available. The reliability of clinical and/or immunological monitoring strategies is currently the subject of raging debates about the best way to monitor patients on ART in resource-limited settings [7–9]. Cost-effectiveness models of the inclusion of viral load in different monitoring strategies are sensitive to the rates of
virologic failure and to rates of misclassification of failure [10, 11].

The sensitivity, specificity, and positive and negative predictive values of immunological monitoring are low. The sensitivity and specificity of clinical monitoring in detecting lack viral of suppression compared to immunological monitoring are even lower [2, 12]. This has lead to misclassification of the viral suppression status of many patients who are receiving ART in resource-limited settings. The consequences of these misclassifications have significant individual and public health implications that can erode the gains that have been made in ART in these settings.

Misclassification of virologic failure by clinical and/or immunologic monitoring results in unnecessary switches to second-line regimens that are more expensive, may be difficult to obtain, may be more difficult to take, have more adverse effects, limit future antiretroviral drug options, or be the last treatment option available. The major public health impact of this misclassification is the increased cost of ART programs. This misclassification has been found to be as high as 50% in some studies [13], reinforcing the necessity to confirm suspected viral failure that is based on clinical and/or immunological monitoring with a viral load test.

The other consequence of clinical and/or immunological monitoring misclassification is failing to detect patients who have early virologic failure. This means that they continue to receive, for too long, a failing drug regimen and, thus, may accumulate multiple HIV drug-resistance mutations. This could lead to patients completing eliminating their ART drug options by the time a decision is made to switch regimens, and this is further compounded by the fact that HIV resistance testing is even less available than viral load testing in resource-limited settings.

As access to ART continues to be scaled up in resource-limited settings, optimal monitoring strategies of the response to ART are needed that take into consideration local capacities. The quality of and access to CD4 cell count tests and viral load measurements vary in resource-limited settings, even where they are recommended in local treatment guidelines, because of inadequate resources. This is going to become even more challenging as treatment programs are rolled out from big hospitals in urban centers to primary health care facilities in the rural areas that are closer to the patients for initiation and continued care of stable patients on treatment. In the initial stages after commencement of ART, high-quality monitoring of response to therapy using viral load is preferable to detect and address issues of early suboptimal viral suppression to avoid increased morbidity, mortality, and development of HIV drug resistance. Patients who experience good viral suppression after the first few months or years need further monitoring to detect early virologic failure before they develop multiple drug-resistant viral mutations; the frequency of viral load monitoring for this group in resource-limited settings is something that would need to be determined [14]. Some studies from sub-Saharan Africa confirm that, when ART drug regimen switching is done early because virologic failure, with the help of viral load monitoring, the response to second-line therapy is excellent [15, 16]. Patients who experience complete viral suppression in the initial stages of treatment are more likely to maintain complete suppressed even if adherence drops off with time to <75%, emphasizing the importance of close monitoring of the early stages of treatment [17].

Viral load testing needs to be made more accessible in resource-limited settings to improve patient care. Although immediate and universal access may not be yet feasible in these ART programs, there is a need for a renewed international consensus on the issue and the inclusion of viral load monitoring in international guidelines, as a goal for programs and an incentive to the industry to supply lower cost technology. Innovative point-of-care technologies also need to be developed to increase access for all populations. Where viral load monitoring remains unavailable, enhanced and more nuanced clinical and immunological screening approaches, including consideration of trends in the CD4 cell percentage, may help avoid unnecessary regimen changes.

A growing body of evidence is demonstrating that clinical and immunological monitoring misclassifies virologic failure in almost one-half the cases and results in premature switches to second-line regimens. The cost of this, to programs and individuals, is immense. In 2008, Smith and Schooley [18] referred to managing ART without viral load monitoring as “running with scissors.” The emerging data on the high levels of unnecessary treatment switches in ART programs in resource-limited settings suggest that it is more akin to throwing these programs onto drawn swords. The time has come to work toward the progressive introduction of appropriate viral load monitoring technology in these programs with the same sense of urgency and commitment as the world approached ART access. To do less is to abandon the global success of ART to an early collapse.

Acknowledgments


References