The Kiel Institute for the World Economy
Duesternbrooker Weg 120
24105 Kiel (Germany)

Kiel Working Paper No. 1371

Optimal Sequencing of Antiretroviral Drug Cocktails under Uncertainty and Irreversibility

by

Mintewab Bezabih and Michael Stolpe

October 2007

Optimal Sequencing of Antiretroviral Drug Cocktails under Uncertainty and Irreversibility*

Mintewab Bezabih and Michael Stolpe

Abstract:

This paper develops a real options approach to the optimal sequencing of antiretroviral drug cocktails for HIV/AIDS patients in resource-poor settings. The analysis focuses on the implications of endogenous resistance mutations in the virus that reduce or eliminate the effectiveness of individual drugs within a cocktail when lack of laboratory equipment prevents these from being identified. Using a model with two drug cocktails, we show that the first-line therapy should be introduced later than in the case without resistance mutations and that the second-line therapy should be introduced earlier. We go on to discuss implications for comparative cost-effectiveness analyses.

Keywords: HIV/AIDS; Real option theory; Cost-effectiveness analysis; Combination therapy; Developing countries

JEL-Classification: D81, H51, I12

Corresponding author:
Michael Stolpe
Kiel Institute for the World Economy
24100 Kiel
Germany
Telephone: +49–431–8814 246
E-mail: michael.stolpe@ifw-kiel.de

"…… Everything will be OK if [the ART drugs] are available. I just pray I don't become resistant to them." taken from a BBC interview with a Kenyan HIV patient (BBC news, 2005)

---

* We thank Tessa Lennemann of the University of Frankfurt am Main for her advice on the medical background for this paper.

1 Introduction

This paper presents a stylized model of treating HIV/AIDS with antiretrovirals[1] (ART) to show how insights from the economic theory of real options can improve the cost-effectiveness in resource-poor settings. Our focus is on the optimal timing of first- and second-line treatment when the HI-virus develops resistance against one or several drugs within a cocktail that is administered to the patient. Resistant mutations are of particular concern in developing country settings, where the medical infrastructure often lacks basic resources, such as laboratory equipment and trained staff, to detect the emergence of drug-resistant virus strains in the course of treating an HIV/AIDS patient and monitor their evolution and spread in the population at large. In addition to rendering drugs ineffective in a particular patient, drug resistance may also diminish society's arsenal in the fight against HIV/AIDS, when resistant strains are transmitted to other people and the range of effective drugs is reduced.

In the absence of resistance testing that could identify the drugs or drug classes whose susceptibility is reduced, we propose to use the theory of sequential decision making under uncertainty to find flexible rules for the withdrawal of entire treatment drug cocktails before resistant mutations are likely to occur. The optimal sequencing rule makes efficient use of the limited available information, taking into account that the uncertain outcome of treatment may be an irreversible adverse event, such as the patient's death or the emergence of drug-resistant virus mutations, and maintaining flexibility in adjusting the treatment strategy to relevant new information. We believe our approach can improve on widespread current practice in which entire drug cocktails, including drugs that are still effective, are replaced in response to clinical, immunological or virological signs of treatment failure. This is not only wasteful in itself, but also often fails to withdraw the resistance-inducing drug early enough to minimize the risk of infecting others with resistant virus strains. This risk is further compounded when the general first-line regimen is based on drugs from only one class, such as triple NRTI, which is common practice in many resource-poor settings.

The social implications of more effective and efficient treatment strategies are of enormous importance. In sub-Saharan Africa, the HIV/AIDS epidemic is now the major stumbling block to economic and social progress. Philipson and Soares (2005) estimate the welfare loss from

[1] Antiretroviral therapy (ART) is any treatment that suppresses the replication of the HI-virus and significantly slows the natural progression of the infection (Wood et al. 2000).

HIV/AIDS in sub-Saharan Africa to be of the order of $ 800 billion, equivalent to the region's entire domestic production in one year. In 2006, almost 25 million, or two thirds of all persons worldwide, infected with HIV are living in sub-Saharan Africa; and up to 3.2 million adults and children are estimated to have been *newly* infected with HIV in 2006, more than in all other world regions combined (UNAIDS 2006). Many of the infected are at their prime working age, so that the treatment of HIV/AIDS does not only constitute an immediate public health, but also a general economic development priority.

Yet, efforts to scale up the diffusion of effective antiretroviral treatment of HIV/AIDS in sub-Saharan Africa have so far failed to reach the vast majority of infected people. Indeed as Freedman and Poku (2005) observe across the continent, poverty structures not only contour the pandemic, but also the likely outcome for an individual infected with HIV. The affordability of treatments and their cost-effectiveness have, therefore, rightly become a major concern of international health organizations, global pharmaceutical companies and national authorities seeking guidelines for clinical decision making and resource allocation in the fight against HIV/AIDS. Medical researchers, such as Goldie et al. (2006), have begun to evaluate the cost-effectiveness of alternative treatment strategies in resource-poor settings. But contributions from the field of global health economics have so far been largely limited to the issue of access.

In an ideal setting, where state-of-the-art monitoring of patients, including regular blood tests, is available, treatment guidelines for HIV/AIDS recommend using a cocktail of several different antiretroviral drugs that are administered simultaneously and should ideally come from at least two different drug classes. Indeed, combination antiretroviral therapy with a cocktail of three or more drugs from two or more drug classes has become the standard of care for patients with HIV infection in rich countries, such as Germany and the United States (Freedberg et al. 2001). The careful combination of treatment drugs with different pharmacodynamic characteristics and mechanisms of action is primarily intended to reduce the probability of resistant mutations, as any such mutations would be at a disadvantage in surviving and replicating unless they succeeded in becoming resistant against all three or more of the different drugs administered at the same time.

Nonetheless, any particular component of the drug cocktail may become ineffective when resistant mutations begin to replicate in the patient. Moreover, when resistant strains of the virus are transmitted to other people, the use of the *resistance-inducing* drug will be restricted to patients not yet infected with the resistant strains. Under constant state-of-the-art

monitoring, the risk of resistance mutations can be limited to a minimum by substituting a new drug for the resistance-inducing drug in the cocktail; ideally, the new drug should be from a different class to minimize the risk of cross-resistance. Under close monitoring of resistant mutations, the maximization of the patient's individual health and the minimization of risks to public health are two sides of the same coin. By substantially reducing the danger that resistant mutations spread, all drug cocktail-components can also retain their effectiveness against the virus in the population in the population at large for much longer. The challenge for resource-poor countries is to come close to this ideal even when constant monitoring and resistance-testing are unavailable.

The remainder of this paper proceeds as follows. Section 2 reviews the medical background to our problem. Section 3 discusses the relevant theory and develops the real option approach. Section 4 outlines potential policy implications, identifies promising directions for further research and concludes.

2 Medical Background

The pathogenesis (disease progression) of HIV/AIDS is well understood to be a function of the HIV RNA level in blood plasma (viral load) and the number of circulating CD4+ lymphocytes (CD4 count). CD4+ cells play a central role in the immune system: their number decreases when the HIV RNA level increases. Health states can be defined in terms of both CD4 counts and HIV RNA levels and disease progression can be modeled as the transition between health states. Clinically effective medication leads to an increase in the CD4 cell count and a decrease of the HIV RNA level, reducing the risk of opportunistic diseases and death and improving the patient's quality of life. The onset of serious opportunistic infections defining AIDS is best predicted by the number of circulating CD4+ lymphocytes (CD4 count) and the level of HIV RNA in blood plasma (viral load): Vulnerability of the patient increases markedly when CD4 lymphocyte levels drop below 200/μL. The viral load provides a measure of response to antiretroviral therapy. High levels predict future rates of decline in CD4 counts, even in asymptomatic patients.

Resistance to antiretroviral drugs may limit the power and duration of the response to the treatment (Little et al. 2002). And rapid viral replication provides many opportunities for the emergence of viral mutants resistant to antiretroviral drugs. However, combinations of drugs, ideally targeting two enzymes (HIV reverse transcriptase and protease), make the emergence of drug-resistant viral mutants less likely.

Changes in plasma HIV-RNA levels are used to measure the effects of a single or combinations of drugs. However, increasing levels may indicate noncompliance with drugs or resistant mutations of the virus. Inadequate compliance increases the likelihood of resistant mutations. Susceptibility testing is required to identify the drugs to which the virus mutants are less susceptible. Lazzarin et al. (2003) report incomplete viral suppression in every second patient treated with highly active antiretroviral therapy, so that patients need to switch to another combination of antiretroviral drugs. Cross-resistance within drug classes is widespread and often limits treatment options in third- or fourth-line treatment. In resource-poor settings, treatment options are exhausted more rapidly because the total number of available (or affordable) drugs is smaller to begin with and because individual drugs to which virus mutants are less susceptible cannot usually be identified.

Against this background, the World Health Organization (WHO) has issued a set of guidelines for resource-poor settings. The treatment guidelines issued in 2003, and updated in

2006, address the whole range of issues around the optimal timing of antiretroviral treatment, including initiation and replacement of drugs, the monitoring of disease progression and side-effects of the treatment. Specific recommendations for subgroups of patients facilitate the planning of national and international HIV care strategies in developing countries. However, the primary goal of the WHO guidelines is to establish standardized formularies for first- and second-line ART that entail 2 NRTIs + 1 NNRTI in the first-line and 1 PI class + 2 new NNRTI to minimize cross-resistance in the second-line treatment.

The WHO guidelines are motivated by practical questions, such as when to start, when to substitute for toxicity, when to switch for failure, and when to stop. They point out that treatment decisions must be based on clinical criteria alone when laboratory tests are unavailable. Studies investigating the optimal point of switching ART on the basis of clinical criteria, in the absence of CD4 cell counts, are hence urgently needed. New treatment strategies should seek to maximize the durability and efficacy of first-line therapy because second-line treatment is more expensive and often unavailable in resource-poor settings. In general, the optimal time for switching from first- to second-line therapy is determined by the *trade-off* between the potential loss of several months or years of survival benefit from any remaining first-line effectiveness and the loss in effectiveness of second-line therapy when it is started too late in a patient's pathogenesis. As a rule, the WHO recommends that the entire regimen or drug cocktail is changed in the event of first-line treatment failure.

Only very few studies have so far addressed the cost-effectiveness of HIV/AIDS treatment in resource-poor settings, defined as the additional costs of a medical strategy divided by the gain of additional life time in years or by the gain of quality-adjusted life time. The development of new medical strategies generally aims at improving the cost-effectiveness in treating a given disease. In the case of HIV/AIDS a substantial decrease in mortality and morbidity among patients was achieved by antiretroviral treatments after they became available in developed countries in 1996. However, antiretroviral medication is expensive and the majority of HIV/AIDS infected people in developing countries still do not have access to the medication (Freedberg et al. 2001; Goldie et al. 2006). Moreover, the effectiveness of any treatment of HIV/AIDS in developing countries is further limited by high co-infection rates with tropical diseases, by malnutrition and by limited opportunities for monitoring HIV/AIDS progression (Severe et al. 2005).

Implementation of cost-effectiveness analysis for ART is typically based on a mathematical simulation model of disease progression, using CD4 cell counts and HIV RNA levels as

predictors (Freedberg et al., 2001). The required empirical data include the direct costs (costs for CD4 cell counts and HIV RNA tests as well as drug costs) and some measure of health-related quality of life (quality-adjusted years of life gained by ART). Freedberg et al. (2001) develop a general model to evaluate the cost-effectiveness of treatment strategies of HIV/AIDS, which Goldie et al. (2006) subsequently modify to focus on developing countries. They develop and compare 22 strategies in which thresholds for initiating therapies are based either on clinical criteria alone or on both clinical criteria and CD4 cell counts, which they consider a crucial distinction for studying effective forms of treatment in developing countries. They also investigate the optimal timing of therapies and the appropriate combination of antiretroviral treatments and prophylaxis.

Freedberg et al. (2001) main findings are that three-drug therapy is always more cost-effective than two-drug therapy and that the timing of treatment start can have a substantial effect on the cost-effectiveness ratio. They point out that the optimal timing of treatment and of alternative strategies to decrease the risk of failure still need to be investigated. As for the developing country context, Goldie et al. (2006) find that treatments combining antiretroviral and prophylactic drugs, such as trimethoprim-sulfamethoxazole, are more cost-effective than antiretroviral strategies alone. Moreover, strategies based on CD4 cell counts are more effective than strategies based on clinical criteria only. Where CD4 cell counts are unavailable, the initiation of therapy is recommended after one or two severe opportunistic infections.

The impact of resistance mutation on the efficacy of antiretroviral treatment was first studied in a developed country setting by Little et al. (2002) and to our knowledge has not been the main focus of developing country studies so far. In any case, identification of mutations that are not natural polymorphisms and can hence be considered evidence of transmitted drug-resistance is not easy, so that overestimation of the true prevalence of transmitted drug-resistance is likely. Needless to say, the prevalence of multi drug-resistant HIV has important implications for the use and the management of antiretroviral treatments.

3 The Model

We present a model of treatment processes with ART drug cocktails that examines the effect of a possible onset of resistant mutations on the optimal timing and switching between these processes. For tractability reasons, our analysis considers a simple case with only two ART drug cocktails. Each treatment is supposed to generate protection to white blood cells (WBC) by blocking the replication of the HI virus or by hampering the copying of the genetic codes of the virus into white blood cells (Wood et al. 2000). We refer to the level of (healthy) white blood cells generated at each point in time as k, suppressing time subscripts for ease of exposition. Accordingly, k_1 represents WBC during the first treatment and k_2 represents the WBC during the second treatment and the generation of WBC under two sequential treatments takes the form

$$k = \begin{cases} k_1 = \alpha k dt + \gamma_1 q_1 k dt + \sigma k dz \\ k_2 = \alpha k dt + \gamma_1 q_1 k dt + \sigma k dz \end{cases} \tag{1}$$

where α is a constant, representing expected natural growth rate of WBC in the respective treatment stages. It should be noted that the natural growth rate of WBC in the two stages is assumed to be constant, which is reasonable since the treatment is not supposed to alter the intrinsic growth rates of WBC. k_1 represents the level of WBCs during the first-line treatment and k_2 represents the level of WBCs during the second-line treatment. The contribution of anti-retroviral Therapy drugs in the first stage is given by $\gamma_1 q_1$, where q_1 is the total amount of the drugs in the first stage and γ_1 represents the conversion factor, the number of healthy WBC per unit of medicine, or simply the 'effectiveness' of the medicine. Similarly, γ_1 and q_1 represent the conversion factor and the total cost of medicine in the second stage respectively. σ instantaneous standard deviation of k which stands for the spread of the levels of WBC around the mean in the respective stages; and $dz(t)$ is the increment to a standard Gauss-Wiener process.

The decision to enter into the second-line treatment depends on the level of WBC, k. If k is below the threshold level that triggers the second-line treatment, the introduction of the second stage treatment is called for. Otherwise, the treatment will not commence.

Following Goldie et al. (2006), we measure the benefit from the treatment in terms of the gains in life expectancy every year. $L(k)$ denotes the statistical value of life corresponding to the level of WBC *k*, which is given by:

$$L(k) = \mu k^2 \tag{2}$$

where μ is a conversion factor that measures the contribution of WBC to life expectancy.

Since going for the treatment costs money, the net benefit from the treatment is the increase in life expectancy net of the costs of the treatment, *sqk*, where *s* is the monetary cost of the treatment per unit of the combination therapy and *q* is the quantity (dosage) of combination therapy. The net benefit is given by:

$$Net\ benefit = \mu k^2 - sqk \tag{3}$$

Our objective is to assess the need for withdrawing an incumbent treatment and replacing it with a new treatment in a condition where a resistant mutation occurs. To that effect, we analyse two situations. In the first situation, we consider the generation of WBC protection in the absence of resistant mutations. In the second situation, we consider the generation of WBC protection when there is a possible development of resistant mutations. Our analysis of these two cases determines the levels of investment where one treatment should be withdrawn and be replaced by the next treatment. We thus obtain treatment regimes – for the case with and without the resistant mutations – that are characterized by their threshold levels of white blood cells, *k*, which trigger the need to switch from first- to second-line treatment.

To establish an analytical benchmark, we first analyze the case in which investment in the treatment takes place under uncertainty and the irreversibility of death, but without the risk that treatment induces resistant mutations by the virus. We assume that the medical practitioner acts as a perfect agent and maximizes the objective function of the patient. In stage one, the practitioner selects a cocktail of drugs and a level of treatment that maximizes the patient's statistical value of life and makes a commitment to using the first-line treatment. Then, in stage two, the practitioner decides to withdraw the first-line treatment and replaces it with a new drug cocktail as the second-line treatment.

3.1 Sequencing treatments under uncertainty and irreversibility: the case of no resistant mutations

In this case, investment in the treatment is characterized by uncertainty and irreversibility. However, there is no risk that a treatment leads to a build-up of resistant mutations by the virus. Under this condition, the generation of white blood cells (WBC) under two sequential treatments with switching at a threshold value of *k* takes the form:

$$dk = \begin{cases} dk_1 = \alpha k dt + \gamma_1 q_1 k dt + \sigma k dz \ if \ k \geq k_1 \\ dk_2 = \alpha k dt + \gamma_2 q_2 k dt + \sigma k dz \ if \ k < k_2 \end{cases} \quad (4)$$

where q_1 is the quantity (dosage) of combination therapy in the first-line treatment and γ_1 is the effectiveness of the therapy in the multiplication of healthy WBC. Similarly, q_2 is the amount of combination therapy in the second-line treatment, and γ_2 is the effectiveness of the therapy in the multiplication of healthy WBC.

Given this, our objective involves finding the threshold levels of WBC in the first and second line treatments which mark the switch from one line treatment to the next. Following Dixit and Pindyck (1994) and Bar-Ilan and Strage (1998), our problem is solved using backward induction where the decision rule for the second-line treatment is solved first and the decision rule for the first-line treatment problem is obtained based on the solution to the second line problem.

Second-line treatment

Using equations (3) and (4), and assuming a discount at rate ρ, the value of the option to invest in the second-line treatment, *W(k)*, is given by:

$$\mu k^2 - s_2 q_2 k + (\alpha + \gamma_2 q_2) k W'(k) - \rho W(k) + \frac{1}{2}\sigma^2 k^2 W''(k) = 0 \quad (5)$$

where $W'(k)$ and $W''(k)$ are the first and second derivatives of the option to invest, $W(k)$,

subject to the boundary conditions:

$$W(0) = 0 \quad (6)$$

$$W(k_2^*) = Gk_2^{2^*\beta_2} + \frac{1}{\rho}(\mu k_2^{2^*} - s_2 q_2 k_2^*) \tag{7}$$

$$W'(k_2^*) = \beta_2 G\theta_2^{\beta_2 - 1} + \frac{2\mu\theta_2}{\rho} - \frac{s_2 q_2}{\rho} \tag{8}$$

With the boundary conditions in equations (6) to (8), the solution to equation (5) is:

$$W(k) = \begin{cases} Ak^{\beta_1} & if\ \mu k < s_2 q_2 \\ Gk^{\beta_2} + \frac{1}{\rho}(\mu k^2 - s_2 q_2 k) & if\ \mu k > s_2 q_2 \end{cases} \tag{9}$$

The constants β_1 and β_2 are solutions to the fundamental differential equations

$$\beta_1 = \frac{1}{2}(1 - m + [(1-m)^2 + 4r]^{1/2}) \text{ and} \tag{10}$$

$$\beta_2 = \frac{1}{2}(1 - m - [(1-m)^2 + 4r]^{1/2}) \tag{11}$$

with $m = 2(\alpha + \gamma_2 q_2)/\sigma^2$, $r = 2\rho/\sigma^2$.

Similarly, the constants A and G are determined from the continuity of $W(k)$ and $W'(k)$ at $\mu k = s_2 q_2$, where

$$\theta_2 = s_2 q_2 \tag{12}$$

$$G = \mu\theta_2^{2-\beta_2}\left(\frac{\beta_1 - 2)_1}{\rho(\beta_2 - \beta_1)}\right) - sq\theta_2^{1-\beta_2}\left(\frac{\beta_1 - 1}{\rho(\beta_2 - \beta_1)}\right) \tag{13}$$

The solutions to these constants are given by:

$$G = \left(\frac{\mu\theta_2^{2-\beta_2}(\beta_1 - 2) - sq\theta_2^{1-\beta_2}(\beta_1 - 1)}{\rho(\beta_2 - \beta_1)}\right) \tag{14}$$

$$A = \frac{\mu\theta_2^{2-\beta_1}}{\rho\beta_1}\left(\frac{\beta_2(\beta_1 - 2)}{(\beta_2 - \beta_1)} - 2\right) + \frac{sq\theta_2^{1-\beta_1}}{\rho\beta_1}\left(\frac{\beta_2(\beta_1 - 2)}{(\beta_2 - \beta_1)} - 1\right) \tag{15}$$

Substituting the solutions for G in equation (14) and A in equation (15) into (5), gives:

$$0 = Gk_2^{\beta_2-\beta_1} + \frac{\mu}{\rho}k_2^2 - \frac{sq}{\rho}k_2 \tag{16}$$

Equation (16) gives the solution to k_2 in terms of the critical parameters. However, because of the nature of the equation, the highly non-linear relationships between k_2 and the other parameters, it is not possible to find a closed form solution. We present a numerical solution after laying out the solution procedure for the first-line treatment.

First-line treatment

As in the second line treatment, the value of the option to invest in the first line treatment, *F(k),* is given by:

$$(\mu k^2 - s_1 q_1 k) + (\alpha + \gamma_1 q_1)kF'(k) - \rho F(k) + \frac{1}{2}\sigma^2 k^2 F''(k) = 0 \tag{17}$$

subject to the boundary conditions:

$$F(0) = 0 \tag{18}$$

$$F(k_1^*) = W(k_1^*) - \mu k_1^* - sqk_1 \tag{19}$$

$$F'(k_1^*) = W'(k_1^*) \tag{20}$$

where:

$$F(k_1) = Dk^{\beta_1} \tag{21}$$

Using the boundary conditions (18) to (21), the solution to the constant D in equation (21) is:

$$D = \frac{\beta_2}{\beta_1}Gk_1^{\beta_2-\beta_1} + \frac{2\mu}{\beta_1\rho}k_1^{2-\beta_1} - \frac{s_1 q_1}{\beta_1\rho}k^{1-\beta_1} \tag{22}$$

Using the solution for G in equation (13), the final solution for D becomes:

$$D = \frac{\beta_2}{\beta_1}Gk_1^{\beta_2-\beta_1} + \frac{2\mu k_1^{2-\beta_1}}{\beta_1\rho} - \frac{s_2 q_2 k_1^{1-\beta_1}}{\beta_1\rho} + \frac{2\mu k_1^{2-\beta_1}}{\beta_1\rho} - \frac{s_1 q_1 k_1^{1-\beta_1}}{\beta_1\rho} \tag{23}$$

which is equivalent to:

$$D = \frac{\beta_2}{\beta_1} G k_1^{\beta_2 - \beta_1} + \frac{4\mu k_1^{2-\beta_1}}{\beta_1 \rho} - \frac{k_1^{1-\beta_1}}{\beta_1 \rho}(s_2 q_2 + s_1 q_1) \tag{24}$$

The expression for D in equation (24) and the boundary condition (19) give the solution to k1.

$$\rho G k_1^{\beta_2}\left(\frac{\beta_2 - \beta_1}{\beta_1}\right) + \mu k_1^2 (4 - 2\beta_1) - k_1 (s_2 q_2 + s_1 q_1)(1 - \beta_1) = 0 \tag{25}$$

Equation (16) gives the solution to k_2 in terms of the critical parameters and equation (25) gives the solution for k_1 in terms of the critical parameters. However, because of the nature of the equation the highly non-linear relationships between k_1, k_2 and the other parameters, it is not possible to find a closed form solution. A numerical solution is presented after laying out the solution procedure for the first line treatment.

In Table 1, the parameters corresponding to the baseline scenario are used to generate the benchmark figures for the threshold WBC values in the first and second line treatments. In order to see the effect of the different parameters on the threshold WBC, changes were made to the base line parameters. Accordingly, the second row in Table 1 stands for the threshold WBC in the first line treatment corresponding to the different parameters. Similarly, the third row in Table 1 contains the threshold WBC in the second line treatment.

From the baseline scenario (the first column in Table 1), we see that the threshold WBC for the first line treatment is set at 1.05 units. It should be noted that the baseline parameters are arbitrary figures and hence the figure corresponding to the threshold WBC is arbitrary. That is why we do not have any specification of units of WBC. As we argued above, the comparative statics is done by changing the parameter values and studying the direction of change in the WBC.

Still focusing on the threshold WBC for first line treatment, a decrease in the individual discount rate leads to a lower threshold WBC value. As with any other investment, this implies that lower discount rate increases the individual gets impatience and leads to an earlier introduction of the treatment. On the other hand, the threshold WBC decreases when the intrinsic growth rate decreases. Increase in the amount of combination therapy drug (i.e. effectively increased available treatment) leads to a lower threshold WBC, which implies that

when the amount of the drugs is higher, it is more efficient to introduce the treatment earlier. Increase in gamma, which measures the effectiveness of the therapy (the number of healthy WBC created per unit of therapy), means that the effectiveness of the therapy is enhanced. The simulation shows that an increase in gamma shows that there is a need to introduce the treatment earlier to gain a lot from the treatment.

The threshold WBC in the second treatment is significantly more sensitive to changes in parameters compared to the threshold WBC in the first line treatment. One example is a 10% reduction in the intrinsic growth rate which leads to a rise in the threshold WBC from 67 units to 177. In the case of the first line treatment, an equal (10%) change in the intrinsic growth rate, only leads to a 0.02 units change. What is similar in the values of threshold WBC in the first and second line treatments is that the direction of changes as response to shifts in parameters are the same as those of the first line treatment.

Table 1: Numerical Solutions for the Threshold WBC Values in the First- and Second-line Treatments (the Case of No Resistant Mutations)

		Baseline scenario	σ=0.05 ρ=0.1 α=0.05 γ=0.05 μ=1 q=1 s=0.05 a=0.6 b=1.1 δ=0.006				
Threshold WBC	baseline	decrease ρ=0.001	decrease α=0.0001	decrease μ=0.1	increase q=10	increase s=0.5	increase γ=0.005
First stage	1.05	1.002	1.07	1.073	1.009	1.031	1.006
Second stage	65.89	8.36	177.01	1.6858	1.27	269	9.49

Proposition 1: The trigger level of WBC needed to induce the first treatment, k_1^*, is a function of the natural growth rate of WBC, the degree of uncertainty in the growth pattern, the discount rate, and the costs of the available first- and second-line treatments. Similar factors determine the optimal timing of second-line treatment. However, the cost of first-line treatment does not affect the second line trigger level in terms of WBC.

3.2 Sequencing treatments under uncertainty and irreversibility: the case of resistant mutations

The case we consider here is similar to that in 2.1, but in addition, any particular treatment now leads to a build-up of resistant mutations by the virus. The possible development of resistant mutations along the treatment process means that the effectiveness of the treatment drugs declines as a given treatment progresses and the generation of WBC protection will be slower even under two sequential treatments. The rate at which the progress is hindered by the onset of resistant mutation is denoted by γ. This transforms the dynamics of WBC in the two stages in equation (4) into:

$$dk = \begin{cases} dk_1 = \alpha k dt + (\delta_1 - \gamma_1) q_1 k dt + \sigma k dz \ if \ k \geq k_1 \\ dk_2 = \alpha k dt + (\delta_2 - \gamma_2) q_2 k dt + \sigma k dz \ if \ k < k_2 \end{cases} \tag{26}$$

Following the same solution procedure as in section 2.1, the solution to the constant G is:

$$G = \mu \theta_2^{2-\beta_2} \left(\frac{\beta_1 - 2)_1}{\rho(\beta_2 - \beta_1)}\right) - sq\theta_2^{1-\beta_2} \left(\frac{\beta_1 - 1}{\rho(\beta_2 - \beta_1)}\right) \tag{27}$$

Similarly, the solution for the trigger level of WBC during the second-line treatment is given by:

$$\rho\beta_2 G k_1^{\beta_2 - \beta_1} - \rho\beta_1\beta_2 G k_1^{\beta_2 - 1} + 2\mu k_1^{2-\beta_1} - 4\beta_1 \mu k_1^2 - s_1 q_1 k^{1-\beta_1} + \beta_1 (s_2 q_2 + s_1 q_1) k_1 = 0 \tag{28}$$

Table 2: Numerical Solutions for the Threshold WBC Values in the First- and Second-line Treatments (the Case of Resistant Mutations)

			σ=0.05 ρ=0.1 α=0.01 γ=0.05 μ=1 q=1 s=0.05 δ=0.006						
Threshold WBC	Baseline	decrease	decrease	decrease	increase	increase	increase	increase	increase
	σ =0.05	ρ=0.001	α=0.0001	μ=0.1	q=10	s=0.5	γ=0.05	a=6	b=
First stage	1.11	0.98	1.15	1.164	1.018	1.064	1.011	1.111	1.111
Second stage	105.37	8.37	531	2.39	1.28	6.19	9.51	99.68	619.48

The following discussion compares the results for the first and second threshold levels of WBC in the two scenarios with and without resistant mutations. Again, due to the difficulty of finding a closed form solution, we compare the results based on the simulations corresponding to changes in each parameter. Increases in a and b seem to have minute effects on increasing the threshold WBC from 1.05 units to 1.053 units.

For the first line treatment, decrease in σ, the instantaneous standard deviation of change in WBC over time, which measures the stochastic nature of the pattern of growth in WBC, leads to a higher level of WBC when resistant mutation is taken into consideration. This implies that in the case of resistant mutations, the reduction in uncertainty delays the need to introduce the treatment. This is intuitive since any increase in the standard deviation reduces the level of uncertainty.

On the other hand, for the first line treatment, reduction of the discount rate in a no resistant mutations case leads to a higher threshold WBC than in the case of resistant mutation (see the values of k_1 corresponding to σ in Tables 1 and 2). An increase in the discount rate implies that the value of healthy WBC is higher in closer periods than at later points in time. Since the incidence of resistant mutations reduces the values of WBC, an earlier introduction is called for.

Proposition 2: The possible development of resistant mutations implies a delay in the introduction of the first-line treatment and an earlier introduction of the second-line treatment.

4 Discussion and Conclusion

The paper has investigated the extent to which uncertainty and irreversibility in ART determine the optimal duration between its onset and withdrawal of a given drug cocktail. Our premise was that in a setting where regular monitoring by means of blood tests is not affordable as part of combination treatment, a conventional procedure of ART[2] carries the risk of the development of resistant mutations that renders the treatment ineffective. Our analysis focused on a more sophisticated treatment strategy that involves a series of ART cocktails similar to the conventional one, but that ensures the withdrawal of a drug cocktail before the onset of the corresponding resistant mutations, and the replacement with another drug cocktail. This treatment strategy ensures that the withdrawn drugs in the initial cocktail are not ineffective and the resistant mutations are not developed and hence even with continued spread of the virus, resistant strains are not transmitted.

Our analysis highlights the implications of two important constraints in poor country settings:

- lack of monitoring options, and
- more limited treatment options in developing settings.

Each individual option is hence relatively more valuable than in a rich country, but also at greater risk of becoming obsolete. Moreover, each drug component also tends to be used on a lower scale per patient, resulting in lower cost-effectiveness. We argue that the optimal approach to enhancing the cost-effectiveness of using existing treatment options in a resource-poor setting – in the absence of monitoring the development of resistance mutations directly – is must rely on a real option approach. Rigorous application of economic theory hence can contribute important novel insights and help to substantially increase the cost-effectiveness of HIV/AIDS treatment in resource-poor settings and thus in those countries where the disease is most prevalent.

We believe that our economic model does not only provide a normative framework for optimal treatment strategies in resource-poor settings, but that it can also serve as a guide to empirical assessments of cost-effectiveness at the level of individual patients and of society as a whole. Our paper thus complements recent empirical studies on the economics of pharmaceutical drugs in the age of resistance, such as Arrow et al. (2004), and helps to

[2] By conventional ART, we are referring to a procedure where a given drug cocktail is used until the time it is rendered ineffective due to the onset of a resistant mutation and only then replaced with a different drug.

improve the state-of-the-art in the treatment of HIV/AIDS in resource-poor settings, as documented in Goldie et al. (2006).[3]

We do not deny that the first-best treatment strategy presupposes the availability and affordability of the full set of treatment drugs that can be combined into an effective cocktail for the treatment of HIV/AIDS. However, many of the newer and more potent drugs are typically under patent protection and therefore much more expensive than the generic drugs on which treatment must be based in many resource-poor settings. Despite recent progress in lowering the prices of many relevant drugs, as documented in Schwartländer et al. (2006), the high overall costs still render first-best combination therapies prohibitive in the African context, underlining the need to consider alternative, more easily affordable treatment strategies.

With this goal in mind, the World Health Organization (WHO, 2001) has issued a set of guidelines that recommend a series of ART drug cocktails, which entail withdrawing all drug components in the administered drug cocktail upon the development of resistant mutations, without first identifying which of the components is rendered ineffective by the resistance mutations. Scaling-up the implementation of these WHO guidelines, which is currently under way in sub-Saharan Africa, may result in a significantly increased number of individuals eligible for treatment (Badri et al. 2004), but it may also exacerbate the spread of drug-resistant strains of the virus, thus creating a potentially large liability for the future.

We believe that our model yields important new insights how the implementation of the WHO recommendations can be improved in a dynamically efficient way. The main objective of this paper was to find the optimal sequencing rule for antiretroviral (ART) drug cocktails in a setting where there is uncertainty and irreversibility in the form of resistant mutations. The study identified the threshold level of a given drug cocktail that corresponds to the time of withdrawal and replacement by another treatment. Furthermore, we identified the ex ante cost-effective sequencing rule for ART drug cocktails. We derived these results by means of a real-options-approach that takes into account uncertainty and irreversibilities in the application of ART. The model was solved using a stochastic dynamic optimisation approach

---

[3] Kenneth Arrow, in his pioneering article on "Uncertainty and the Welfare Economics of Medical Care," which appeared in the American Economic Review in 1963, first pointed out the pervasive role of uncertainty in the production of medical care. Sue Goldie's research is applying the modern tools of decision science to evaluate the clinical benefits, public health impact, and cost-effectiveness of alternative prevention and treatment strategies against viruses that are major current public health problems.

that provides the optimal treatment rule at every point in time. We thus obtained a chain of rules that determines the optimal sequencing of ART drug cocktails. The findings of this analysis are expected to help steer treatment strategies for HIV/AIDS patients in resource-poor settings.

References

Arrow, K.J., C.B. Panosian and H. Gelband (2004). Saving Lives, Buying Time: Economics of Malaria Drugs in an Age of Resistance. Committee on the Economics of Antimalarial Drugs, Board on Global Health. Washington, D.C.: The National Academic Press.

Badri, M., L.-G. Bekker, C. Orrell, J. Pitt, F. Cilliers and R. Wood (2004). Initiating Highly Active Antiretroviral Therapy in Sub-Saharan Africa: An Assessment of the Revised World Health Organization Scaling-up Guidelines. *Aids* 18: 1159–1168.

Bar-Ilan, A. and W.C. Strange (1998). A Model of Sequential Investment. *Journal of Economic Dynamics and Control* 22 (3): 437–463.

Dixit, A.K. and R.S. Pindyck (1994). Investment under Uncertainty. Princeton, N.J.: Princeton University Press.

Freedberg, K.A., E. Losina, M.C. Weinstein, A.D. Paltiel, C.J. Cohen, G.R. Seage, D.E. Craven, H. Zhang, A.D. Kimmel and S.J. Goldie (2001). The Cost-effectiveness of Combination Antiretroviral Therapy for HIV Disease. *The New England Journal of Medicine* 344: 824–831.

Freedman, J. and N. Poku (2005). The Socioeconomic Context of Africa's Vulnerability to HIV/AIDS. *Review of International Studies* 31 (4): 665–686.

Goldie, S.J., Y. Yazdanpanah, E. Losina, M.C. Weinstein, X. Anglaret, R.P. Walensky, H.E. Hsu, A. Kimmel, C. Holmes, J.E. Kaplan and K.A. Freedberg (2006). Cost-Effectiveness of HIV Treatment in Resource-Poor Settings – The Case of Côte d'Ivoire. *The New England Journal of Medicine* 355: 1141–1153.

Lazzarin, A., B. Clotet, D. Cooper, J. Reynes, K. Arastéh, M. Nelson, C. Katlama, H.-J. Stellbrink, J.-F. Delfraissy, J. Lange, L. Huson, R. DeMasi, C. Wat, J. Delehanty, C. Drobnes and M. Salgo (2003). Efficacy of Enfuvirtide in Patients Infected with Drug-Resistant HIV-1 in Europe and Australia. *New England Journal of Medicine* 348 (22): 2186–2195.

Little, S., S. Holte, J.-P. Routy, E.S. Daar, M. Markowitz, A.C. Collier, R.A. Koup, J.W. Mellors, E. Connick, B. Conway, M. Kilby, L. Wang, J.W. Whitecomb, N.S. Hellmann and D.D. Richman (2002). Antiretroviral-Drug Resistance among Patients Recently Infected with HIV. *New England Journal of Medicine* 347 (6): 385–394.

Philipson, T.J., and R.R. Soares (2005). The Economic Costs of AIDS in Sub-Saharan Africa: A Reassessment. In: G. López-Casasnovas, B. Rivera, and L. Currais (eds.), Health and Economic Growth: Findings and Policy Implications. MIT Press, Cambridge MA.

Schwartländer, B., I. Grubb and J. Perriens (2006). The 10-year Struggle to Provide Antiretroviral Treatment to People with HIV in the Developing World. *The Lancet* 368 (9534): 541–546.

Severe, P., P. Leger, M. Charles, F. Noel, G. Bonhomme, G. Bois, E. George, S. Kenel-Pierre, P. Wright, R. Gulick, W.D. Johnson, J.W. Pape and D.W. Fitzgerald (2005). Antiretroviral Therapy in a Thousand Patients with AIDS in Haiti. *New England Journal of Medicine* 353 (22): 2325–2334.

UNAIDS (2006). UNAIDS/WHO AIDS Epidemic Update: December 2006.

WHO (2001). The World Health Report. Geneva.

Wood, E., P. Braitstein, J.S.G. Montaner, M.T. Schechter, M.W. Tyndall, M.V. O'Shaughnessy and R.S. Hogg (2000). Extent to Which Low-Level Use of Antiretroviral Treatment could Curb the AIDS Epidemic in Sub-Saharan Africa. *The Lancet* 355 (9221): 2095–2100.

Appendix

Antiretroviral Drug Classes

Category	**Examples** with different types of possible adverse effects	**Mechanism of action**
Nucleoside or nucleotide revers-transcriptase inhibitors (NRTI)	Zidovudine Lamivudine Stavudine Didanosine Abacavir Ienofovir	
Nonnucleoside reverse-transcriptase inhibitors (NNRTI)	Efivarenz Nevirapine	
Protease inhibitors (PI)	Amprenavir Atazanavir Indinavir Lopinavir Nelfinavir Ritonavir Saquinavir	Interfering with activity of HIV protease

Figure 1: First-line and Second-line ART (adapted from Freedberg et al. 2001)

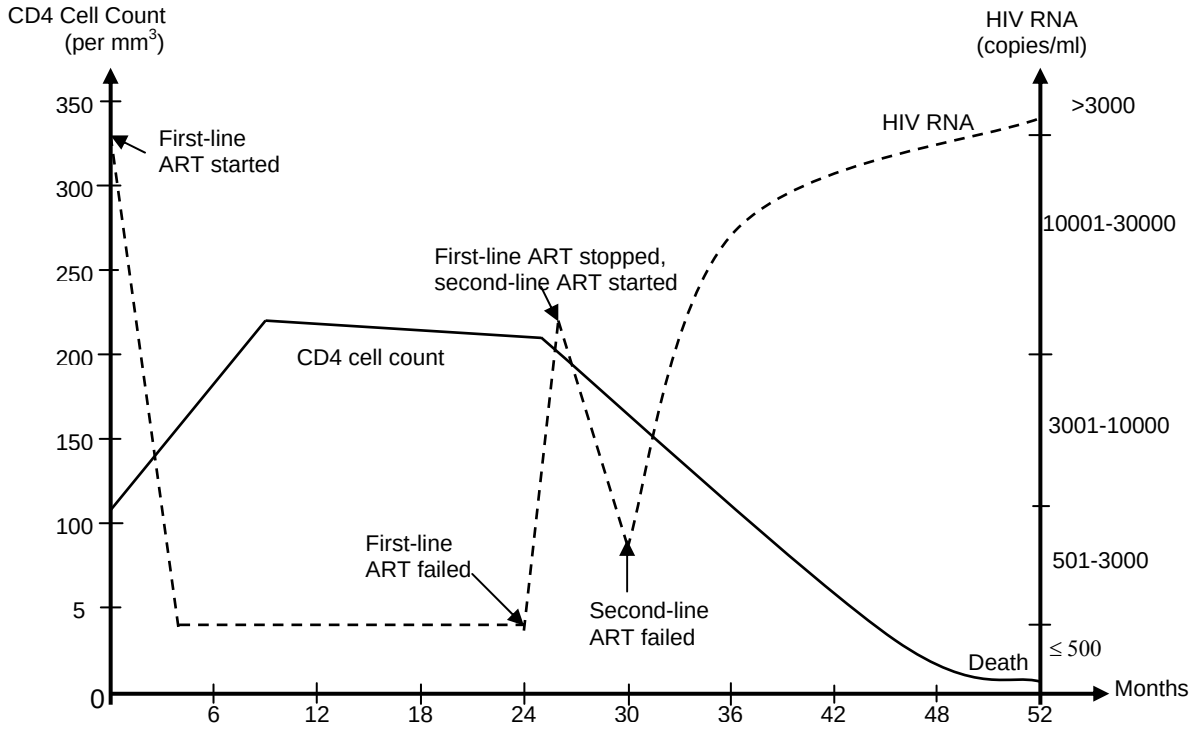

Figure 2: Cost-effectiveness of Treatment Strategies (adapted from Goldie et al. 2006)