COST-EFFECTIVENESS OF ANTIRETROVIRAL TREATMENT FOR HIV-POSITIVE ADULTS IN A SOUTH AFRICAN TOWNSHIP

Report to Health Systems Trust

November 2003

Susan Cleary [1], Andrew Boulle [2], Di McIntyre[1] and David Coetzee[2]

[1] Health Economics Unit,

[2] Infectious Diseases Epidemiology Unit

School of Public Health and Family Medicine, University of Cape Town

Acknowledgements

Many people gave very generously of their time during the process of this research. We would particularly like to acknowledge the staff and patients of the HIV clinics, as well as the staff of Tygerberg hospital. The following people deserve particular credit: Emi Maclean, Marta Darder, Eric Goemaere, Toby Kasper, Clas Rehnberg, Kurt Maart, Robin Wood and Gary Maartens.

This research was supported by a grant from the Health Systems Trust.

Contents

Tables

Figures

Executive Summary

Background to this report

This report is the product of research into the cost-effectiveness and cost-utility of treatment for HIV-positive adults. The research was conducted in three HIV-dedicated clinics in Khayelitsha, a township on the outskirts of Cape Town. The formal objectives of the research were:

- To describe the costs of providing antiretroviral treatment (ART) at primary care clinics
- To describe the costs of providing care for HIV in the absence of ART at primary care clinics
- To determine the associated health care costs for patients with HIV referred to other levels of care
- To describe the effectiveness of ART and no-ART (i.e. treatment in the absence of ART) in terms of life years (LYs) gained and quality adjusted life years gained (QALYs)
- To describe the cost-effectiveness and cost-utility of ART and no-ART in this setting
- To describe the incremental cost-effectiveness and cost-utility of ART
- To describe the incremental cost-effectiveness of different starting times of ART

Introduction to Economic Evaluation and Markov Modelling

Economic evaluations use commonly accepted methodology to establish the costs and the outcomes of different courses of action, in order to provide clarity to the decision-making process. In economic terms, this type of research aims to establish technical and allocative efficiency. Technical efficiency refers to “doing it the right way” and allocative efficiency refers to “doing the right thing”. However, this type of analysis cannot establish economic feasibility. Even if an intervention were technically efficient, it may still be unaffordable. Instead, economic evaluations attempt to clarify the resource implications of interventions (the lifetime cost) as well as the gain (in life expectancy and quality of life) in order to allow society to have an informed choice in decision-making.

There are a number of different forms of the economic evaluation. This research uses both cost-utility and cost-effectiveness analyses. The cost-utility analysis is a specialised form of the cost-effectiveness analysis (often the terms cost-effectiveness and cost-utility are used interchangeably in the literature).

The cost-effectiveness analysis uses an outcome measure that has only one dimension (the Life Year, for instance). It can be used to compare interventions that lead to outcomes that have similar quality of life, but perhaps different effectiveness in terms of life expectancy. Although this is not a strictly appropriate form of analysis for comparing ART to no ART, it is nevertheless useful from a budgeting point of view because it can calculate the annual cost.

The cost-utility analysis is a far more appropriate form of economic evaluation when comparing ART to no ART. It uses a multi-dimensional outcome measure, and can capture the different effects of ART and no ART in terms of both quantity and quality of life. A common outcome measure that combines both quantity and quality of life is the Quality Adjusted Life Year (or QALY). Put simply, the QALY adjusts the average

gain in life expectancy from an intervention with a factor indicating the value (or utility) derived from the extension in life expectancy.

Given that the cost-utility analysis is merely a specialised form of the cost-effectiveness analysis, for simplicity this report uses the terms cost-effectiveness and cost-utility interchangeably.

On the cost side, economic evaluations include all recurrent and capital costs required to deliver an intervention. Costing can be done from a number of different perspectives (such as society, or the health system) and depending on the perspective chosen, different categories of costs are included. For instance, if a societal perspective were chosen, the costing would include costs to the health system and costs to the patient (e.g. patient waiting time and travel costs) but if a health system perspective were chosen, these patient costs would be excluded.

One of the key difficulties in predicting the cost-effectiveness or cost-utility of ART is that data are not yet available for the full course of patients' lifetimes on ART in the public sector in South Africa, although there is much less uncertainty for patients who are not on ART. This is a common problem in the economic evaluation of long-term interventions and chronic diseases, and has lead to the widespread adoption of Markov modelling, which is a technique that allows current data to be extrapolated forward into future health states in order to predict future costs and future effects. Although this does imply a degree of uncertainty in the results, sensitivity analyses can be very useful in clarifying the degree of variability in the estimates.

Key Findings

This research considers the costs, and effects of treating opportunistic and HIV-related infections either with antiretroviral therapy, or without, for HIV-positive adults with CD4 cell counts of less than 200 cells/ μl[1]. The two interventions (ART or no ART) are considered from the point of view of the public health system using a cost-utility and a cost-effectiveness framework.

This research is contributing three key findings to the current state of knowledge in this area. Firstly, it provides an indication of the relative efficiency of ART compared to no ART in a setting that is similar to future ART rollout sites in South Africa. Secondly, this research is able to give a better indication of the costs of providing ART over a patient's lifetime than is currently available. Thirdly, it is able to give a solid indication of the current costs of treating opportunistic and HIV-related infections for patients who are not on ART. The latter two pieces of information are essential for budgeting for the ART rollout adequately, whilst the former gives an indication of the relative efficiency of ART versus no ART in similar settings.

The costing of ART and no ART includes all recurrent costs required to deliver ART, to treat opportunistic and HIV-related infections, to encourage adherence and to minimize transmission of the virus (including viral load, CD4 count and other laboratory testing, co-trimoxazole prophylaxis, ongoing palliative care, extensive counselling of patients, referrals for tuberculosis treatment and inpatient care, nutritional supplementation, and the provision of male and female condoms). The capital costs associated with infrastructure, medical equipment, furniture and staff training were also included (annuatized using a real discount rate).

[1] Patients can only access ART once their CD4 cell counts are less than 200 on this project, unless they are classified as WHO Stage IV

In terms of the overall efficiency of the intervention, this research has calculated that ART costs R13 754 per QALY and no ART costs R14 189. The incremental cost per QALY gained on ART is R13 621. This result indicates that ART is efficient in economic terms, and ought to be pursued if economically feasible and desirable to society.

The lifetime cost of treating a patient on ART was calculated to be just over R93 000, and off ART the lifetime cost was on average just under R24 000 (for patients with CD4 counts less than 200 cells/µl). The average life expectancy is 8.33 years on ART, and 2.27 years for patients not on ART. In other words, ART leads to an average gain in life expectancy of 6.06 years. This translates into 6.79 QALYs on ART or 1.59 QALYs for no ART. Patients reported higher Health Related Quality of Life (HRQoL) on ART than off ART.

When the lifetime cost is broken down into its key cost components, over 50% of the cost of the ART option relates to the cost of the antiretrovirals. Although the prices of some ARVs have fallen recently (and this research uses the latest prices as of October 2003) second-line ARV regimens are still expensive, as is efavirenz, which is an important component of the first-line regimen in this research. Laboratory tests at the clinic level for the ART option are also relatively expensive, and account for 9.2% of the lifetime cost. For non-ART, the most important component is the cost of inpatient care.

Recommendations

These findings have a number of immediate policy implications.

- The current focus on reducing the cost of antiretroviral drugs is warranted, as on the whole, ARVs still account for nearly 50% of the lifetime cost on ART. This is particularly important for the drugs that remain relatively expensive (such as Efavirenz, ddI and Kaletra). Although personnel costs are not a major cost driver, recruiting and training sufficient human resources to deliver ART will still be a major challenge.

- More emphasis should be placed on reducing the cost of HIV RNA (viral load) testing. There should also be clarification of the role of this test in the provision of ART in South Africa.

- The clinical results on which this study is based are a clear demonstration of the potential for the intervention to extend life, improve quality of life, and delay many of the individual and societal consequences associated with premature mortality.

Abbreviations

3TC: Lamivudine
AIDS: Acquired Immunodeficiency Syndrome
ART: Antiretroviral treatment
ARVs: Antiretroviral drugs
AZT: Zidovudine
CD4: CD4+ Lymphocyte Cells
d4T/3TC/NVP: Triomune
d4T: Stavudine
ddI: Didanosine
EFV: Efavirenz
FL: First-line antiretroviral regimen
HAART: Highly Active Antiretroviral Therapy
HIV: Human Immunodeficiency Virus
HIV+: HIV-positive
HRQoL: Health Related Quality of Life
LFT: Liver Function Test
LYs: Life Years
MSF: Medecins sans Frontieres
NGO: Non-governmental Organisation
No ART: Treatment for HIV-positive adults in the absence of antiretroviral treatment
NVP: Nevirapine
PAWC: Provincial Administration of the Western Cape
PDE: Patient Day Equivalent
QALYs: Quality Adjusted Life Years
SL: Second-line antiretroviral regimen
TB: Tuberculosis
VL: Viral Load (HIV RNA) test

1 Introduction and Background

This project aims to establish the costs and effectiveness of antiretroviral therapy (ART) for HIV positive adults in a resource-constrained public-sector setting. The research compares ART to the current status quo for HIV-positive adults who are dependent on the public sector for care in South Africa – i.e the treatment of opportunistic and HIV-related infections and events (e.g. wasting) in the absence of ART. This research is clearly important in the developing country context, where the HIV epidemic is expected to have a dramatic impact on life expectancy and to lead to early mortality for a large proportion of the population (Dorrington, Bourne et al. 2001).

Although ART has been shown to be effective in poor settings (Laurent, Diakhate et al. 2002; Weidle, Malamba et al. 2002), and poor people have demonstrated their potential to be adherent to therapy (Orrell, Bekker et al. 2001; Orrell, Bangsberg et al. 2003), there has been very little primary research into the cost-effectiveness of treatment (to date, none has been published from Africa). Limited budgets in poor countries imply that resources should be put to their most cost-effective use - a lack of evidence on the cost-effectiveness of treatment for the most severe health crisis in poor countries is therefore a serious shortcoming.

The purpose of this research is to fill this gap. Specifically, a comparison is made between the costs and effects of treating opportunistic and HIV-related infections either with ART or without ART for clients with CD4+ lymphocyte counts below 200 cells/µl (a clinical prerequisite to qualify for treatment). The perspective of this evaluation is that of the public health system.

Primary outcomes (i.e. actual survival time) should ideally be used as the measure of effectiveness in a study such as this. However, this would imply that results would be delayed until the entire cohort had died. Given that the level of the epidemic requires decisions to be taken sooner rather than later with the best available information, this research uses Markov Modelling to anticipate future effectiveness by extrapolating from the first three years of the ART programme. Until primary outcomes on the effectiveness of ART are available, results based on Markov modelling can provide valuable insight into the potential costs and effects of ART.

2 Literature review

2.1 International experience: Economic Evaluation of ART

Literature on the cost-effectiveness of ART was reviewed by conducting Medline, AIDSline, Cochrane Collection and ad hoc bibliography searches.

In the United States and Europe, a number of economic evaluations of adult antiretroviral treatment have been conducted. However, only one study was found that evaluated ART versus no ART (Freedberg, Losina et al. 2001). The remainder evaluate monotherapy versus combination therapy, different starting times of therapy or different drug combinations [For instance, Oddone and Cowper et al (1993) evaluate different starting times of monotherapy; Miners and Saber et al (2001) evaluate HAART versus two Nucleoside Reverse Transcriptase Inhibitors (NRTI's) and Messori and Becagli et al (1997) evaluate monotherapy]. Many of these studies model costs and outcomes using Markov modelling. Frequently, cost data is taken from the AIDS Cost and Services Utilisation Survery (Hellinger 1993; Shapiro, Morton

et al. 1999; Freedberg, Losina et al. 2001). This is a large survey that was conducted at 26 sites in 10 different cities in the United States. Based on reports from the respondents in the survey, an estimate was made about the monthly costs of health care at different stages of HIV disease.

2.2 South African experience: Economic Evaluation

In South Africa and in Africa in general, there are no published studies based on existing programmes examining the cost-effectiveness of ART versus no ART and there are no published Markov models of HIV.

While no economic evaluations based on primary research have been conducted, a number of spreadsheet models have estimated the lifetime cost of ART (AbtAssociates 2000; Boulle, Kenyon et al. 2002; Marseille, Hofmann et al. 2002; Geffen, Nattrass et al. 2003). One study (Boulle, Kenyon et al. 2002) examined the cost effectiveness of the additional expenditure as a result of ART, but did not include non-ART costs in the primary cost-effectiveness measure. Furthermore, the National Departments of Health and Treasury have developed a cost model of ART (2003).

In addition, a small number of studies have undertaken primary costing of inpatient and outpatient care for HIV-positive people who are not on ART (Karstaedt, Lee et al. 1996; Kinghorn, Lee et al. 1996; Govender, McIntyre et al. 2000; Haile 2000)

3 Aims and objectives

3.1 Aim

To describe the cost-effectiveness of providing ART in a poor public health setting in terms of the economic cost per life year gained and quality adjusted life year gained and to describe the lifetime costs of ART and no ART.

3.2 Objectives

- To describe the costs of providing ART at primary care clinics
- To describe the costs of providing care for HIV in the absence of ART at primary care clinics
- To determine the associated health care costs for patients with HIV referred to other levels of care
- To describe the effectiveness of ART and no-ART in terms of life years (LYs) gained and quality adjusted life years gained (QALYs)
- To describe the cost-effectiveness and cost-utility of ART and no-ART in this setting
- To describe the incremental cost-effectiveness and cost-utility of ART
- To describe the cost-effectiveness of different starting times of ART

4 Setting

This research was conducted in an antiretroviral pilot situated in a township on the outskirts of Cape Town in the Western Cape Province of South Africa. Life in this township is characterised by high levels of unemployment (approximately 40% under the broad definition (Nattrass 2002)), high crime rates, a shortage of basic services and only limited formal housing.

In April 2000, the Western Cape provincial government in collaboration with Médecins sans Frontières (MSF) launched three dedicated HIV clinics within the existent community health centres in the township. The dedicated HIV clinics provide treatment, counselling and coordinated support groups for HIV-positive people. From 2001, the service was extended to include ART. Referrals from the HIV-dedicated clinics are mainly to GF Jooste secondary hospital and to Tygerberg Academic Hospital Complex for inpatient care, and to tuberculosis (TB) clinics in Khayelitsha.

This is the oldest public sector pilot in South Africa where ART is offered at the clinic level. Although there have been older treatment programmes in the country, these have typically been clinical trials delivered from the hospital setting (and are therefore likely to have higher costs and a different level of effectiveness). The results from this research therefore provide a valuable opportunity to learn about the costs and benefits of ART in a setting that is similar to what will be encountered when ART is scaled up.

5 Study design and methodology

The most appropriate form of economic evaluation in this context is the cost-utility analysis[2]. This type of economic evaluation takes both quantity and quality of life into account, which is important because it is clear that both quantity and quality of life are different for patients on ART relative to patients not on ART. A common outcome measure that combines both quantity and quality of life is the QALY. We have also presented the costs per LY (i.e. cost-effectiveness results) because this is a more useful format for budgeting purposes. We have used Markov Modelling to calculate the cost per LY or QALY gained for each treatment option. The model also calculates total costs, total LYs and total QALYs.

5.1 General aspects of Markov Modelling

In the economic evaluation of ART, one is concerned with the costs over a patient's entire lifetime on ART, and also the effectiveness of the intervention in terms of the average life expectancy and the quality of this life. However, this presents an immediate problem because ART has not been available for long enough for primary outcome data to be available.

This difficulty has lead to the widespread adoption of Markov modelling techniques in the economic evaluation of long-term interventions such as ART. This is a very useful technique, as it allows available data to be used in such a way as to generate estimates of life expectancy and costs for the average patient. While the information that is generated from Markov modelling can never be as accurate as primary outcome and cost data, sensitivity analyses performed on the key assumptions of the model can give a fairly concrete indication of the overall reliability of the results.

Markov models depict the natural history of a disease as an evolving sequence of mutually exclusive health states. In other words, any patient in the model can only be in one state at a time. Each health state (or Markov state) is defined to capture important clinical events. All patients assigned to a given state incur similar economic costs and enjoy a similar quality of life and life expectancy. In order for a Markov process to terminate, an "absorbing" state is required to absorb all patients after a

[2] Cost-utility analysis is a specialised form of cost-effectiveness, and in the literature, there is often no distinction made between the two. This report therefore uses these two terms interchangeably.

sufficient number of cycles has passed. In this case, the absorbing state is the "Dead" state.

The net probability of making a transition from one Markov state to another during a single cycle is called a transition probability. Transition probabilities are calculated from transition rates – i.e. the number of occurrences of an event for a given number of patients per unit of time.

The time horizon of the model is divided into equal cycles called Markov cycles where the length of the cycle is chosen to represent a clinically meaningful time interval.

Evaluation of the Markov process (i.e. solving the model through either matrix algebra, Markov cohort simulation, or Monte Carlo simulation) yields the average number of cycles of survival and their associated costs. For a basic introduction to Markov modelling, refer to Sonnenberg and Beck (1993).

5.2 A brief description of antiretroviral therapy

Antiretroviral drugs cripple enzymes that are crucial in the replication of HIV. When a patient with HIV starts taking these drugs, the concentration of HIV (viral load) drops rapidly. Whereas a person with HIV who is not on ART might have an HIV concentration of 100,000 copies per millilitre of blood, this can be reduced to below the level of detection by current technology within three to four months. On stopping ART, HIV replication resumes and the viral load rapidly reverts to what it was prior to ART. In the absence of ART the circulating virus progressively weakens the immune system, as reflected by the fall in the CD4+ lymphocyte count.

When taking one or two antiretroviral drugs, the virus mutates to escape the drugs (resistance) after some time. The rate at which this happens is however vastly reduced when taking three drugs, although resistance eventually does occur, at which point new antiretrovirals need to be used. For various reasons only certain antiretrovirals can be combined with others, and resistance to some also confers resistance to others the patient has never taken. The net result is that currently there are only two combinations of three drugs (regimens) that can be used sequentially and be effective in suppressing viral replication. These are often termed first-line and second-line regimens. When resistance develops to the second-line regimen, theoretically the increased viral replication will lead to a progressive loss in CD4+ lymphocytes and eventual progression to AIDS and then death. Many patients will remain on the regimen in spite of failure as there is some evidence that the resistant virus is less fit than that in the absence of ART. Some however will stop the treatment once failure occurs due to side effects. Not much is known about this period of treatment due to the relatively short time that ART has been available.

An analogy often used to summarise concepts in HIV is that of a train (the patient) moving along its tracks towards a barrier (death). The speed at which it is moving is likened to the viral load, and the distance from the barrier is the CD4+ lymphocyte count. ART removes the forward momentum, allowing the train to roll slowly backwards away from the barrier (CD4+ lymphocytes re-accumulating). On stopping ART, the train regains its forward momentum.

5.3 A Markov model for HIV/AIDS

Two separate Markov models were constructed for this setting – one for the ART option and one for the no ART option.

The first step in the modelling exercise was to identify the appropriate Markov states. Each state is defined to have a homogenous outcome (survival and quality of life in this case) and level of costs associated with health care.

The initial choice of states was based on analyses of cohort data (Hogg, Yip et al. 2001; Egger, May et al. 2002) that have shown three CD4 count strata to have significantly different survival levels[3]. In this manner, the Markov states were designed to capture a similar risk of death. The next step was to divide these CD4 count based strata into smaller states in order to capture differences in costs.

Once the states had been identified based on indications from the literature and primary data, the results were presented to the clinical staff working in the HIV-dedicated clinics, and based on their advice, further adjustments were made.

An associated process was the definition of the Markov cycle length. This is the amount of time that any cohort member spends in a state before transitions are allowed to new states. We chose a Markov cycle length of 3 months.

Thereafter, the various movements between the states were defined, and the corresponding probability of each movement was calculated (these are known as transition probabilities – see full details from 5.3.1 to 5.3.7).

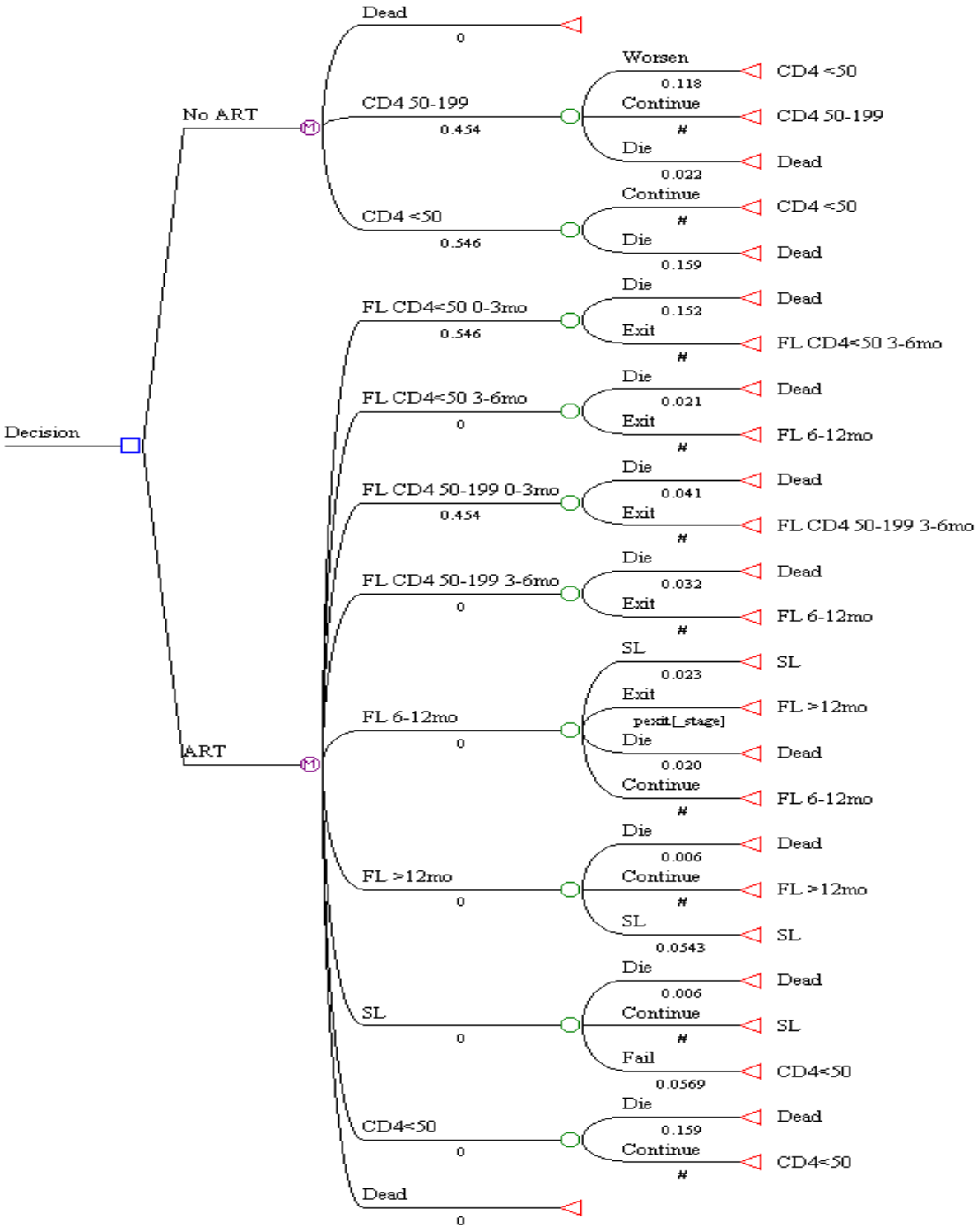
Figure 1 shows a slightly simplified version of the Markov models for ART and no ART (excluding options that are used in sensitivity analysis). After the Decision node, cohort members are randomised into one of the two Markov nodes, labelled ART or no ART in the diagram. The lines to the right of the Markov nodes depict the full set of Markov states in each model. Once patients have been randomised into ART or no ART, all patients start in a Markov state reflecting CD4<50 or CD4 50-199, depending on the baseline CD4 counts in the ART pilot. The lines emanating from the Markov states describe the possible transitions in the model, and the transition probability for each possible movement.

For instance, a patient on ART who has CD4<50 will start in “FL CD4<50 0-3 months”. Once in this state, he or she has a probability of 0.152 of dying, otherwise he or she moves to “FL CD4<50 3-6 months”. Once in this state, he or she has a probability of 0.021 of dying, otherwise he or she continues to “FL 6-12 months”. The process continues until over 99% of the cohort has been absorbed into the absorbing state (“Dead”).

[3] These are CD4<50, CD4 50-199 and CD4>200

Figure 1: The structure of the Markov Model.

At the Decision node, HIV+ patients with CD4<200 cells/µl can either remain off ART or go on ART. Thereafter, patients start in either CD4 50-199 or CD4<50 according to the baseline CD4 counts of patients in the pilot who went on ART. Patients transition through the model as shown by the transition probabilities on the diagram. The model runs until over 99% of the cohort has been absorbed into the "Dead" state.



5.3.1 Initial CD4+ lymphocyte count categories

The relative proportion of patients initiating treatment in each CD4 category were determined from the data of those starting ART in the first 2 years of the pilot. For patients on ART, 54.6% of patients started in “FL CD4<50 0-3 months” (i.e. their baseline CD4 count was less than 50 cells/µl) and the remainder started in “FL CD4 50-199 0-3 months”. This split was maintained for the no-ART group, with 54.6% of patients starting in “CD4<50” and the remainder in “CD4 50-199”. Although this might seem an unintuitive starting point for the no ART group, it is necessary to ensure that the ART and no ART models are comparable with each other.

Although the clinical criterion for accessing ART is CD4<200, data indicate that patients have tended to start treatment with relatively low CD4 counts. Informal reports suggest that the reason that patients present late for treatment is partly because patients present late for treatment and partly because patients with the lowest CD4 counts are prioritised for treatment first. If this intuition is true, it is likely that patients will continue to access treatment with low CD4 counts until the public sector ART rollout has become more established. For this reason, this analysis has utilised these starting CD4 counts when evaluating ART versus no ART, as this is likely to give a realistic picture of the expected costs and effectiveness of the intervention in future ART rollout sites. However, results are also presented indicating the relative cost-effectiveness of different starting times of ART (see 6.10).

5.3.2 Overall survival probabilities for ART and no ART

The transition probabilities between the Markov states determine the speed at which the cohort moves through the model towards the absorbing state (“Dead”) and are the basis of the calculation of effectiveness (life-years gained) in each treatment option.

For the ART model, transition probabilities were calculated from survival data for the first 21 months on ART at the HIV-dedicated clinics.

Actual survival probabilities were utilised for the first year on ART in the model, stratified by starting CD4+ lymphocyte count. Thereafter, a constant mortality rate was utilised irrespective of initial CD4+ lymphocyte count. This mortality rate was generated by averaging the observed survival experience in the first 18 months for all patients combined, and extrapolating forwards at a constant rate.

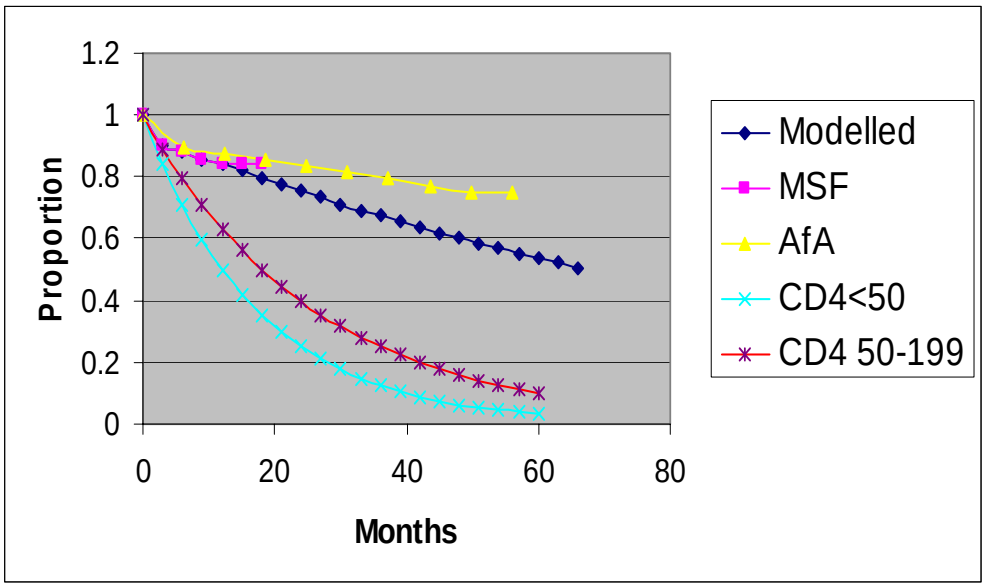
The survival curve that was generated in this manner (Figure 2) was validated by comparing it with the survival data from ‘Aid for AIDS’ (a South African cohort of people on antiretrovirals in the private health care sector which has data up to 54 months). As mortality in the first few months on ART is higher than subsequently, averaging the mortality in the early period on ART to arrive at subsequent mortality amounted to a conservative prediction of the longer-term survival benefit of ART in this programme. In other words, we have assumed a lower survival rate than has been displayed by the “Aid for AIDS” programme.

The probability of dying converged between initial CD4+ lymphocyte states between 6 and 12 months on ART. This allowed us to merge the two CD4 count defined groups into one group.

For the no-ART model, a similar process was followed, using survival data from published research in the Cape Town area (Maartens, Wood et al. 1997). Data on

the risk of transitioning from one CD4 defined state to another was also extrapolated from this source.

Figure 2: Survival curves



5.3.3 AIDS versus non-AIDS deaths for the ART group

Having determined the overall probability of survival for patients on ART over a given period, it was necessary to make an assumption about whether patients would die while on antiretroviral treatment (or while in one of the states in the model where ARVs are offered), or whether patients would fail treatment and move into an off treatment state before dying. Although this assumption would not effect the overall life expectancy assumptions, it would have implications for the costs.

The literature suggests that it is reasonable to anticipate that up to a quarter of HIV-related mortality occurs in individuals prior to failing treatment, and the early experience at the MSF clinics validated the use of this assumption. Based on real data, mortality in the first year on ART is assumed to occur without cycling through additional Markov states. Once immune reconstitution is assumed to be substantial (after one year on treatment), the above schema (whereby 75% of patients die by moving through the Markov states) is applied.

5.3.4 Treatment failure on ART and transition to death

The exact manner in which patients failing ART are likely to transition to death is unlikely to be known for some years to come. Many patients failing ART will remain on their failing regimen due to the fact that the resistant virus has altered virulence and there remains some clinical benefit to remaining on failing regimens. The exception is those patients who do not tolerate the failing regimen.

Rather than defining an additional failing state for which costs and transition probabilities are unknown, the assumption was made that if patients exit the second-line regimen directly into the CD4+ lymphocyte count category below 50 cells/µl, this would provide a reasonable averaged division between time on ART and off-ART once failure occurs prior to death, and would also provide a fair assumption of

averaged hospitalisation after virological failure. In other words, the assumption is that patients stay on treatment despite virological failure. In terms of costs, this is a fairly conservative assumption, because in effect, patients are assumed to stay in a state with a higher overall cost level than the off-ART states.

5.3.5 Probability of transitioning to second-line treatment for the ART group

The final set of probabilities required to build the effectiveness component of the Markov model for the ART group, were the probabilities of initiating second-line treatment. For the first year on ART, real data were utilised. Thereafter, the apportionment of survival benefit between first and second-line regimens was determined by a ratio of 60% on first-line and 40% on second-line. Primary data on the time spent on each regimen in this setting are not available, and this split was based on expert opinion[4]. No one began second-line treatment in the first six months on ART, and this is likely to be a rare occurrence in a population where most patients have never taken ART before.

5.3.6 Defaulting treatment

In the main model, patients are assumed not to default treatment. This is justified because the rate of defaulting treatment is exceptionally low in this programme. Furthermore, the conservative estimates on survival compensate for this, as does the modelled time off ART described above. The impact of differing probabilities of defaulting treatment are explored in sensitivity analysis.

5.3.7 Arriving at final transition probabilities

The final transition probabilities were determined by the following sequenced calculations:

ART Group:

- The overall mortality was determined from the real and extrapolated survival distributions described above.
- The direct mortality from each state was calculated:
    - Real data were used for all of the states in the first year
    - Thereafter one quarter of the mortality was assumed to occur prior to failure and moving into CD4<50 and this was uniformly applied to the first-line and second-line states
- The transitions from first-line to second-line regimens and from the second-line regimen to CD4 less than 50 cells/µl was determined empirically[5] to satisfy two conditions:
    - That the resultant mortality combining all three transition probabilities (first-line to second-line, second-line to CD4 < 50 cells/µl, and CD4 < 50 cells/µl to death) accounted for the remaining mortality from the survival distributions (excluding the direct mortality from each state)

[4] Personal communications Prof. G Maartens and Prof. R Wood

[5] Mean survival times were calculated from the exponential distributions given by Markov states with constant transition probabilities, the resultant mean apportioned between states, and an exponential transition probability recalculated.

    - That prior to entering the final state before death, 60% of the time would be spent on the first-line regimen, and 40% on the second-line regimen
- The real mortality rate for those who had failed treatment and moved into the "CD4<50" state was determined from secondary data (Maartens, Wood et al. 1997) and applied directly to this state

For the No-ART group:

- The real mortality rate for those with CD4< 50 cells/µl and for those with CD4 50-199 was determined from secondary data and applied directly to these two state (Maartens, Wood et al. 1997)
- The transition from CD4 50-199 cells/µl to CD4 <50 cells/µl was determined empirically in order for the resultant mortality to concur with that described in the secondary data for the group with CD4+ lymphocyte counts between 50 and 199 cells/µl.

All possible transition probabilities together with an explanation of assumptions are found in the following tables.

Table 1: Transition probabilities per cycle in the first year on ART

First-line 0-6 months CD4 < 50 cells/ul

Probablilty Dying Markov Cycle	Prob	Comment
0	0.152	Directly from ART pilot data until end 2002, 289 patients, median f/u 6.5 months
1	0.021	

First-line 0-6 months CD4 50-199 cells/ul

Probablilty Dying Markov Cycle	Prob	Comment
0	0.041	Directly from ART pilot data until end 2002, 289 patients, median f/u 6.5 months
1	0.032	

First-line 6-12 months all patients

Probability Second-Line Markov Cycle	Prob	Comment
All	0.023	Directly from ART pilot data until end 2002, 289 patients, median f/u 6.5 months

Probablilty Dying Markov Cycle	Prob	Comment
All	0.020	Directly from ART pilot data until end 2002, 289 patients, median f/u 6.5 months

Table 2: Transition probabilities per cycle beyond the first year on ART

First-line 12 months onwards all patients

Probability Second-Line Markov Cycle	Prob	Comment
All	0.0543	From ART pilot data extrapolated

Probablilty Dying Markov Cycle	Prob	Comment
All	0.006	25% of deaths assumed to happen on first-line or second-line, the rest via CD4<50

Second-line 12 months onwards all patients

Probability Failing and Moving to CD4<50 Markov Cycle	Prob	Comment
All	0.0569	Derived experimentally to yield same net survival as if mortality were equal in all states

Probablilty Dying Markov Cycle	Prob	Comment
All	0.006	25% of deaths assumed to happen on first-line or second-line, the rest via CD4<50

CD4 < 50 for patients failed treatment

Probablilty Dying Markov Cycle	Prob	Comment
All	0.159	From Maartens (1997)

Table 3: Transition probabilities for no ART

CD4 < 50

Probablilty Dying Markov Cycle	Prob	Comment
All	0.159	From Maartens (1997)

CD4 50-199

Probablilty Dying Markov Cycle	Prob	Comment
All	0.022	

Probability moving to CD4 < 50 Markov Cycle	Prob	
All	0.118	Derived experimentally to yield same net survival as if mortality was equal in all states

5.3.8 Software implementation of Markov models

The Markov model was created and evaluated using Decision Analysis by TreeAge (DATA™) software version 4.0. Cohort simulation was used to solve the model. This calculates the expected value of the process by multiplying the percentage of the

cohort in a Markov state by the incremental value (cost, effect and utility) assigned to that state, and summing these products over all states and all Markov cycles.

5.4 Costing methodology for Markov states

5.4.1 General approaches

The costing of ART and no-ART includes all direct costs that accrue to HIV positive people with a CD4<200. Both recurrent and capital costs are included from the perspective of the health system. Costs associated with clinic visits, district or secondary level inpatient care, tertiary level inpatient care and tuberculosis treatment are included, but the costs of hospital outpatient department visits and home-based care are excluded because of insufficient data. However, these costs would be similar for both ART and no-ART patients, and their exclusion is unlikely to bias the cost-effectiveness results.

The primary difference between the costing approach developed in this research and that of previous approaches (Freedberg, Losina et al. 2001) is the method used for costing opportunistic infections. Previous research has estimated the incidence of OIs per Markov state and combined this with the average cost of treating each OI.

We have maintained this approach for costing Tuberculosis (which is the most important OI in this setting) but have found this approach to be unworkable for the costing of other OIs for a number of reasons. Firstly, this approach would require data on the incidence of OIs that is currently unavailable (with the exception of Tuberculosis). Secondly, this would also require very large samples to adequately calculate the average cost of treating each OI at each level of care and stage of disease. If one considers this problem at the inpatient level, it becomes even more complex. Experience in this setting has shown that patients tend to be admitted for inpatient care with a number of different OIs together, and it is impossible to unpack the relative contribution of each OI to the total cost of a period of hospitalisation. In addition, costing OIs separately from one another and basing costs on the incidence of OIs combined with an average cost of each episode could lead to a large overestimation in costs, especially in settings where it is clear that not all patients are receiving the care that they require for each episode of morbidity.

Instead, this costing has focussed on establishing the total utilisation of different types of services by the cohort in each Markov state. These are combined with other patient-specific items that are provided on an ongoing basis to the cohort. Different categories of services utilised by the cohort include:

- Ongoing clinic consultations
- Tertiary level inpatient care (Cleary and Committee 2002)
- Secondary/district level inpatient care (Haile 2000)
- Tuberculosis treatment (Sinanovic, Floyd et al. 2000)

Utilisation data are combined with the average cost of each type of service to generate costs for each Markov state for these services.

Patient-specific items include:

- Laboratory testing and imaging
- Primary and secondary prophylaxis (co-trimoxazole/dapsone and fluconazole)
- Antiretrovirals (where appropriate)

All primary costing used 2002 prices, January - December 2002 expenditure and January – December 2002 visit statistics. At the inpatient level, utilisation statistics and expenditure exactly matched the data collection period (April to August 2002).

Because ART is a long-term intervention, many costs are incurred over time. To avoid the need to inflate costs, a real interest rate has been used for annuatizing capital items (Walker and Kumaranayake 2002). According to the South African Reserve Bank: "A real interest rate is calculated (in the simplest way) by deducting the inflation rate from the nominal interest rate" (http://www.reservebank.co.za/ accessed 1 April 2003)

For the purposes here, CPIX[6] is used to proxy the inflation rate, and the repurchase rate (money market interest rate) of the Reserve Bank is used to proxy the interest rate. Using the averages of these two rates for 2002, the real interest rate is calculated to be 1.311 – we have used annuatization factors of 2% from Drummond et al (1987).

Euro and US$ denominated amounts were converted to Rand using the average interbank rate over the last 4 years (so as not to overestimate foreign currency amounts owing to volatility of the Rand).

The following table gives details of these exchange rates.

Table 4: Exchange Rates

	US$: Rand	Euro: Rand
Average (1461 days):	8.34755	7.99429
High:	13.845	12.475
Low:	5.935	6.0341

5.4.2 Costing Samples

Estimates of quantities of service utilisation (visits, inpatient days and TB treatment) and patient-specific items were taken from the cohort of patients who went onto ART (n=288) by the end of 2002 stratified by CD4 cell count (i.e. CD4< 50 or CD4 50-199). The period before going on ART informs health service utilisation for the non-ART option, and the period after baseline provides information about utilisation for the ART option. This was facilitated by the clinics not offering ART in the first year of their operation, providing a sufficient at risk period from which non-ART costs could be calculated for patients who eventually received ART. The quality of data and record keeping for these patients provides more detail than would have been available from patients who never received ART at all. The latter group of patients would also have been qualitatively different from the group that did receive ART on the basis of their not being eligible for treatment.

In other words, the cost for patients in "FL CD4<50 0-3 months" are calculated from the group of patients who went onto ART with a baseline CD4 count below 50 cells/µl. Costs for this Markov state would include all utilisation of services (i.e. quantities of inpatient days / clinic visits / lab tests / medicines etc) including baseline laboratory testing (although this happens before month 0).

[6] CPIX is the consumer price index excluding interest rates on mortgage bonds

There were a number of methods used to allocate costs to the appropriate Markov state and cycle. For instance, certain costs were calculated per visit, others per inpatient day, and others from actual utilisation by cohort members (e.g. ARVs and lab tests). If costs were calculated per visit or inpatient day, quantities of visits or inpatient days were estimated from cohort members falling into the relevant Markov state and cycle. Each of these will be discussed separately in the following sections.

5.4.3 The cost of ongoing clinic consultations

The cost per visit has been calculated for a representative public sector setting and for the pilot project scenario. The former gives an indication of the resource requirements for the scaling up of ART in the public sector. The latter gives an indication of the actual costs of the pilot. This is justified because it is clear that costs in the pilot are likely to be higher than in the public sector for a number of reasons.

Firstly, the pilot does not achieve economies of scale in developing adherence tools and the training of staff in treatment of HIV-related illnesses and in HAART (which would be done at provincial or national level). Secondly, some office staff and doctors in the pilot are expatriates and therefore earn foreign denominated salaries that are higher than those paid in South Africa. Thirdly, the pilot project undertakes more extensive monitoring of patients (for research purposes) than would be required in the public sector. A full explanation of these differences is found in Appendix A.

The following items have been allocated per visit:

- Overheads (utilities, security, office supplies, condoms, nutritional supplements etc – see appendix for full list)
- Non-clinical staff (counsellors, monitoring data enterers, office staff, pharmacists, cleaners etc)
- Clinical staff (doctors and nurses)
- Medicines (prescribed at visits for opportunistic and HIV-related infections)
- Capital costs (medical equipment, other electronic equipment, furniture, buildings and staff training)

A brief description of each of these will be given below and full details are given in the Appendix.

5.4.3.1 Overheads
Overhead costs were split using the relevant visit head counts, and results were fed back to staff to double-check their validity.

5.4.3.2 Non-clinical staff
All non-clinical staff were interviewed to ascertain:

1. The proportion of their time relating to treatment
2. The proportion of this time relating to ART or non-ART

The resultant costs were allocated to visits using the relevant visit head counts.

5.4.3.3 Clinical staff
Further effort was put into estimating the appropriate split of clinical staff time for an ART or non-ART visit. Clinical staff in 2002 included 3 full-time doctors, 3 full-time nurses and sessional doctors for approximately 5 hours per day.

Researchers observed and timed 54 consultations with ART clients and 94 consultations with non-ART clients. It was found that, on average, doctors and nurses spent slightly longer with an ART client than a non-ART client. The average cost per ART and non-ART visit is presented below for the public sector model and the pilot. The costs for the pilot are higher owing to the difference between salaries offered in the public sector and salaries paid to expatriates, although the difference is fairly insignificant.

Table 5: Average cost per visit for clinical staff

	ART Visit	***Non-ART Visit***
No. visits in sample	54	94
Average time per visit (minutes)	18.53	17.06
Clinical staff cost per visit in the pilot	58.61	54.11
Clinical staff cost per visit in the public sector	50.91	46.88

5.4.3.4 Medicines

Medicines prescribed (excluding antiretrovirals and medicines used for primary and secondary prophylaxis) were extracted from the folders of a sub-sample of 60 patients who had been on ART for at least one year as of December 2002 (with the period before baseline informing the costs of the non-ART option as described above). Prices for medicines were sourced from the 2002 Provincial Tender Price list. The costs per consultation are allocated to visits, stratified by ART and non-ART, and for non-ART, stratified by CD4 count.

5.4.3.5 Capital costs

2002 replacement values were obtained for all furniture, medical and electronic equipment, buildings and the costs of staff training (all clinical staff and counsellors receive training – see Appendix) and annuatized using the real interest rate in 2002 to calculate an annual economic cost. This was divided by the relevant visit statistics in 2002 to get a capital cost per visit.

5.4.4 The cost of tertiary and secondary level inpatient care

5.4.4.1 Utilisation of inpatient care

Utilisation data from this research indicate that patients access ongoing inpatient care for episodes of morbidity, but that this utilisation becomes more concentrated in the period preceding death.

An important refinement in the costing of HIV care introduced by this research has been the treatment of the costs of hospitalisation in the period prior to death. Previous analyses have usually spread these costs across Markov states, which leads to an artificial determination of a constant rate of hospitalisation across time, when in reality this utilisation is highly concentrated around the time of death. Alternatively, it can lead to the assumption that patients on ART who die prior to treatment failure (which is approximately one-quarter of patients) receive no inpatient care prior to death, which our utilisation data has shown to be an inaccurate simplification.

Thus, the costs associated with hospitalisation in the six months prior to death have been modelled as transition costs in the model, and accrue to any cohort-member when transitioning to the *"dead"* state. This transition cost was applied over and above ongoing hospitalisation associated with each Markov state.

For the frequency of hospitalisation unrelated to death, the clinical cohort (excluding patients who died) was utilised to derive estimates of the number of hospital days per patient per Markov state and cycle.

These estimates of the utilisation of inpatient care are combined with the cost per inpatient day to arrive at an overall cost relating to hospitalisation per Markov state, and per transition cost.

The process of arriving at a cost per inpatient day is described in the following sections.

5.4.4.2 Cost per inpatient day

Primary costing was undertaken for tertiary level inpatient care for HIV+ people at Tygerberg Academic Hospital (Cleary and Tygerberg Steering Committee, 2002). Methodology employed in this cost analysis replicated a study (Haile, 2000) undertaken at G.F. Jooste Hospital. Thus, estimates were available for the cost of inpatient care for HIV+ people at both the district and the tertiary levels. These costs are combined with quantities of inpatient days from the clinical cohort in applicable Markov states and before death.

The cost per inpatient day was calculated separately for patient-specific costs, recurrent overhead costs (hotel costs) and capital costs.

To estimate patient-specific costs, doctors and nurses were asked to fill out cost data sheets (see the appendix for an example of the data collection sheet) for each HIV+ patient under their care on a daily basis. Data was collected in the last week of April, the last week of May and the last week of August 2002 for an overall data collection period of 3 weeks.

Clinical suspicion and/or laboratory results were used to identify patients who had HIV/AIDS. HIV-positive laboratory identification was accepted without a confirmation test in some circumstances. The clinical suspicion HIV-positive identification (using WHO criteria) was accepted for patients who refused a laboratory test.

Table 6 gives information on the costing sample at Tygerberg.

Table 6: Tygerberg sample

No. Patients	**61**
Inpatient Days:	**243**
Venue:	
Ordinary Hospital Bed	220
Intensive Care Unit	2
High Care Unit	10
Unknown	11
Department:	
Medicine	143
Surgery	62
Gynaecology	24
Obstetrics	1
Oncology	8
Unknown	5

The following patient-specific costs were collected and priced as follows:

1. Laboratory tests were priced using SAIMR billing tariffs for 2002.
2. Imaging was priced using the Uniform Patient Fee Schedule (UPFS) 2002 tariffs. A Facility Level 3 tariff (for Tygerberg Hospital) was combined with specific charges for the different categories of imaging.
3. Medication was priced according to the Provincial Tender price during the sample period.
4. Social worker, clinician and physiotherapy costs were calculated from the global cost per staff category, averaged by the number of staff in each category. This was applied to estimates of time spent with patients from the cost data sheets.
5. It was not possible to estimate patient-specific costs for nurses or pharmacists; these are included under overhead costs

For 96 out of 243 inpatient days, doctors did not indicate their time spent with patients on their cost data sheets. For these days, the average doctor cost of R56.15 (based on the time of those who did fill out their timesheets) was used to arrive at an estimated doctor cost. It was found that this assumption did not prejudice the results. It was reported by the doctors that 8 patients received physiotherapy and 19 patients received counselling from a Social Worker. These consultations last 30 minutes at Tygerberg.

Collecting patient-specific costs by asking doctors and nurses to fill out cost data sheets is fraught with difficulties. For this reason, some time was spent ensuring the validity of these estimates (see section 6.1.1 on page 29). It was found that the final results were highly reliable.

Recurrent overhead costs were calculated by subtracting capital and patient-specific expenditure from total expenditure at each facility during the relevant period. This was allocated to inpatient days by dividing expenditure by the patient day equivalent (PDE). The PDE is a weighted combination of the number of inpatient days and outpatient visits in the relevant period, and provides a useful means of allocating recurrent expenditure between inpatient days and outpatient visits. Recurrent expenditure per inpatient day is likely to be higher relative to an outpatient visit (because inpatients receive hotel services for example), so this implies that expenditure should be weighted more for inpatients relative to outpatients.

The PDE is calculated as follows:

PDE = (Inpatient Days*appropriate weight) + (outpatient visits)

The “appropriate weight” was calculated as the average ratio of the cost per outpatient visit to the cost per inpatient day in the medicine and surgery departments at Groote Schuur hospital from April 2002 to January 2003 (where a cost-centre accounting system allows this calculation to be made). This calculation estimated that an inpatient day was approximately 3.77 times more expensive than an outpatient visit. This factor was used to weight recurrent overheads and capital costs at Tygerberg and Jooste.

A model produced by Rod Bennet at the National Department of Health (Bennet personal communication) was used to estimate building and equipment capital costs. The model generates estimates based on the level of the facility, the number of beds

and the Building Price Index. This model was adapted to fit the number of beds at Tygerberg and Jooste, using the Building Price Index for 2002.

5.4.5 The cost of tuberculosis treatment

The cost of tuberculosis treatment is allocated to every new tuberculosis case in the cohort. This cost has been sourced from secondary data inflated to the 2002 level (Sinanovic, Floyd et al. 2000) from an appropriate setting (Guguletu).

It is not clear what type of tuberculosis treatment patients from the HIV clinics access, or whether they are new or retreatment TB cases, therefore an average cost was calculated from the costs of TB treatment for new cases, retreatment cases and community or clinic-based supervision (see Table 7), including all relevant recurrent and capital costs. The incidence of new tuberculosis cases per Markov state was derived directly from the clinical cohort. This was multiplied by the cost of TB treatment to obtain a cost for TB treatment in each Markov state. Although tuberculosis treatment spans more than one quarter, the entire cost is captured against a single quarter since incidence is being utilised rather than the proportion on tuberculosis treatment. This approach may overestimate the cost of tuberculosis treatment because a small proportion of patients will die without completing the full treatment course, and will have been assigned full treatment costs in the model.

Table 7: Average cost of Tuberculosis treatment

	Clinic-based supervision	Community-based supervision	Average
New patients	3,747.69	1,717.17	2,732.43
Retreatment patients	5,282.57	3,047.69	4,165.13
	Cost per Markov State		3,448.78

Source: Sinanovic, Floyd et al, 2000

5.4.6 The cost of patient-specific items

A number of items were allocated on a patient-specific basis. These include laboratory testing, chronic medication (fluconazole and bactrim / dapsone), ARVs and radiology. Utilisation of these items was estimated from the clients of the HIV clinics in each Markov state per 3 month period, and was priced as follows:

- Laboratory Tests: SAIMR tariffs for 2002
- Chronic Medication: prices for Bactrim / Dapsone and patented Fluconazole from Provincial Tender Price list in 2002; price of Biozole™ (generic version of fluconazole imported by MSF) was used as the basis for the fluconazole costing except if patented prices were used for ARVs.[7]
- Imaging: priced according to the relevant Uniform Patient Fee Schedule (UPFS) 2002 tariff using a Facility Level 1 tariff.
- ARVs: International US$ estimates of generic and patented manufacturer's prices (MSF 2002) updated until October 2003 with the latest price changes

[7] Although Fluconazole is available as a donation from Pfizer, this is an economic costing so includes the costs of donated items. However, if generic ARV prices were used, the Biozole price was used.

The number of laboratory tests performed at the clinics (which are guided by a standardised protocol) was used as the basis for calculating the patient specific laboratory costs.

The proportion of patients in each Markov state on prophylactic fluconazole, co-trimoxazole (bactrim) or dapsone was calculated directly from the cohort. The clinical protocol at the clinics stipulates that all patients with CD4+ lymphocyte counts below 200 cells/µl are on co-trimoxazole one tablet daily, which is double when below 100 cells/µl. For this reason, those with CD4+ lymphocyte counts of between 50 and 200 cells/µl were presumed to be on an average of 1.5 co-trimoxazole tablets per day if the database indicated that they were receiving prophylaxis. Secondary fluconazole prophylaxis was assumed at 200mg per day.

The number of radiology requests are routinely captured into the clinical database and formed the basis of calculating the mean number of investigations per patient per state per period.

The proportion of patients on each antiretroviral by Markov state was calculated from the clinical cohort. In building alternative models utilising alternative regimens, uniform utilisation of a single regimen per state was assumed. This was justified by the very low rate of treatment substitution for reasons of intolerance.

This costing has used a wide variety of prices for ARV drugs. The following price sources were used:

- *Untangling the Web of Price Reductions* (MSF 2002) was used for the prices of generics from Cipla, Ranbaxy and Hetero and for patented prices from Boehringer Ingelheim, Abbott, Merck and Bristol Myers Squibb.
- GlaxoSmithKline patented prices were sourced from their latest press release as of end October 2003 (GlaxoSmithKline 2003)
- Press releases were the source of prices for drugs procured through the Clinton Foundation HIV/AIDS Initiative (a new innovative programme that procures drugs from Cipla, Aspen Pharmacare, Ranbaxy and Matrix at very competitive prices) (Schoofs 2003)
- Far Maguinhos in Brazil was the source for the current MSF regime[8]

If prices were “Free on Board” or “Cost Insurance Freight”, a 30% mark-up was included to cover the costs of warehousing, distribution etc. While there is no research done in South Africa to justify this assumption, it is in line with assumptions used in the costing of antiretroviral treatment developed by the National Department of Health and National Treasury (2003).

5.5 Quality of Life and Utilities

Data on the Health Related Quality of Life (HRQoL) was assessed by implementing the EUROQOL EQ-5D instrument at baseline, 1, 3, 6 and 12 months on antiretrovirals in the same setting (Jelsma, MacLean et al. 2003). The EQ-5D provides a measure of overall health-related quality of life based on five descriptive questions with three levels of answers and a rating scale. Utility values between 0 (death) and 1 (full health) for the different combinations of possible answers in the descriptive part have been established in the general population in the United Kingdom using the time trade-off method. These values were used to calculate utilities for each patient who participated in the survey at the clinics. While it would be

[8] Marta Darder, Personal Communication

preferable to calculate utilities using health state valuations from a general South African population survey, this information is unavailable.

The assessment of quality of life in Khayelitsha used a similar before and after study design, and provided quality of life data that could be linked to Markov States. Baseline quality of life was used for patients not on ART. This is likely to overstate the overall quality of life of patients not on ART, and therefore once again biases the results away from ART in general.

5.6 Choice of sensitivity analyses

The choice of which items to vary for sensitivity analysis was informed by those items for which there was insufficient data for certainty, and by those items that constituted the major costs, or for which programme choices could lead to significant cost differences. On the effectiveness side, this necessitated examining different assumptions about overall survival, and about retention of patients on treatment programmes. On the cost side, variations in antiretroviral prices, hospitalisation costs and usage of viral load testing were modelled.

Full details of the sensitivity analysis scenarios are included with the results below.

6 Results

6.1 Survival estimates and time in each Markov state

The mean survival for all patients from the time of initiating ART to the time of death was 8.33 years. Although the survival distribution is not strictly exponential due to the real mortality rates being utilised in the first year on treatment, this approximates roughly to a median survival of 7.25 years.

The mean survival time for patients off ART in the same CD4 count strata as those on ART was 2.27 years.

The mean amount of time spent in each Markov state (in years) is shown below.

Table 8: Survival time by Markov state (years)

		ART	Non_ART
FL	ART first 3 months	0.25	
	ART second 3 month	0.23	
	ART second 6 months	0.43	
	All patients thereafter	3.32	
SL	Second regimen	3.01	
Non-ART	CD4<200	0	0.81
	CD4<50	1.08	1.46
	Total	8.33	2.27

6.2 Costs of ongoing clinic consultations

This section describes the utilisation and costs for items that are allocated on a per visit basis.

6.2.1 Utilisation

Patients initiating ART had approximately 7.8 visits in the first 3-month period. This quickly diminished as time on treatment increased to between 3.9 and 3.26 visits per cycle. Patients not on ART had approximately 3.25 visits per cycle.

The following costs are applied to each visit in the relevant Markov state.

6.2.2 Capital costs, overheads and staff

Table 9 presents the costs of capital, overheads and staff per visit in the public sector model and the pilot. As explained, the actual costs of the pilot have been calculated, and assumed costs for a representative public sector model have been calculated.

Table 9: Overheads, staff and capital costs per visit

	Public Sector Model		MSF Pilot	
	ART Visit	Non-ART Visit	ART Visit	Non-ART Visit
Recurrent Costs				
Recurrent Overheads	19.87	19.87	19.87	19.87
Clinical Staff	50.91	46.88	58.61	54.11
Counselling Staff	21.22	1.06	20.63	1.03
Counselling Coordination	-	-	16.46	0.82
Monitoring	9.78	4.66	20.28	9.66
Pharmacist	1.44	1.44	1.44	1.44
Office Staff	7.56	7.56	10.93	10.93
Cleaners	1.51	1.51	1.60	1.60
Admin and translation	4.43	4.43	4.30	4.30
Capital Costs				
Clinical Staff Training	1.35	0.17	1.59	0.21
Buildings, Furniture and Equipment	6.78	6.78	6.78	6.78
Total	***124.85***	***94.36***	***162.49***	***110.75***

The differences between the pilot costing and the public sector model relate to the following:

- Clinical staff: Some pilot doctors are expatriates, paid in Euros. For the public sector model, appropriate South African public sector salaries have been used.
- Counseling Staff and Counseling Coordination: the pilot employs an extra counseling coordinator (who provides support to the existent counseling coordinators from NGOs in the area and provides training for the counselors). For the public sector model, the Lifeline Coordinator (NGO in the area that coordinates counselors) was included in the costing as part of the overall cost of counseling staff (which is why this cost is slightly higher in the public sector model.)
- Monitoring: Some monitoring of patients is for research purposes; this would not be necessary in a public sector rollout.
- Office Staff: In the public sector model, expatriate office staff salaries were replaced with South African salaries for these positions.
- Cleaners: in the public sector model, the cost of cleaning the MSF office is excluded.
- Clinical staff training: This is higher in the MSF Pilot because of expatriate salaries.

Further details are given in Appendix A.

6.2.3 Medicines

The non-ARV medicine costs (ad-hoc curative prescribing for HIV-related conditions) were comprised chiefly of antibiotics, antifungals and dermatologicals (Table 10). Although the number of patients on anticonvulsants is few, many are on sodium valproate (safest to use with antiretrovirals) which is disproportionately expensive.

Table 10: Curative medicine cost by category

Category	% of cost
Simple antibiotics	24.4
Systemic and mucosal antifungals	21.2
Dermatologicals including antifungal creams	17.4
Acyclovir	9.9
Simple analgesics	7.9
Anticonvulsants	6.5
Vitamins and other supplements	2.8
Amitriptilline	1.5
Promethazine	1.1

The average cost per visit for these medicines was R15.11 for those patients not yet on ART and with CD4 counts below 50 cells/µl, and R13.31 for those with CD4+ lymphocyte counts between 50 and 199 cells/µl. The average cost per visit for patients on ART was R5.17. While the average medicine cost in the first three months on ART was still quite high, these prescriptions were spread over more than 7 visits, so the cost per visit was low. After the first period on ART, the medicine cost per period fell rapidly, but was spread over fewer visits.

6.3 Costs of inpatient care

6.3.1 Cost per inpatient day

The following tables provide a detailed breakdown of the patient specific costs at Tygerberg and Jooste. The Jooste study did not collect patient-specific data for doctors, social workers and physiotherapists and neither study collected patient specific costs for nursing staff. The relatively large cost for imaging at Tygerberg is owing to the expense associated with CT scans for a number of patients as well as a large number of chest X-rays.

Table 11: Tygerberg patient specific cost per inpatient day

	Total Cost	Average Cost	Proportion
Medication	7,605.43	31.30	11.9%
Lab tests	10,863.19	44.70	17.0%
Imaging	30,729.00	126.46	48.2%
Doctor	13,644.63	56.15	21.4%
Counselling	686.66	2.83	1.1%
Physiotherapy	236.88	0.97	0.4%
	63,765.79	262.41	100.0%

(n=243 inpatient days)

Table 12: Jooste patient specific cost per inpatient day

	Average Cost	Proportion
Medication	69.15	35.8%
Lab tests	90.66	46.9%
Imaging	29.09	15.0%
Diagnostic Procedures	2.59	1.3%
Treatment Procedures etc	1.82	0.9%
	193.32	100.0%

(n=367 inpatient days; 2002 prices) (Haile, 2000)

Intuitively, the patient specific results seem slightly underestimated at Tygerberg, or overestimated at Jooste. One would expect care to be more expensive at the Tertiary level than at the Secondary level, but these results are indicating that the cost per inpatient day is similar when clinical personnel are excluded from Tygerberg. Potential reasons for this could be underreporting at Tygerberg owing to the cost data sheet collection method. Another reason could be that rationing of services has increased since the earlier days of the HIV epidemic. Refer to section 6.1.1 for details about the validity of these estimates.

Table 13 contains details of the recurrent overhead costs per inpatient day. For Jooste, expenditure data from September 1996 to September 1997 were inflated to the 2002 level using the consumer price index (CPIX) and the PDE assumptions were adjusted. Tygerberg expenditure and PDE data were for the period April to August 2002.

Table 13: Recurrent overhead cost per inpatient day

	Tygerberg	Jooste
Inpatient Days	156,588	42,454
Outpatient visits	221,166	25,963
Casualty Visits	-	2,857
PDE	215,196.99	50,091.30
Overhead Expenditure	196,596,442.00	40,522,631.53
Cost per Inpatient Day	913.57	808.98

The capital cost per inpatient day is presented below using PDE data for calendar year 2002 for both Jooste and Tygerberg.

Table 14: Capital cost per inpatient day

	Building Replacement Value	Annual Economic Cost	PDE	Cost per Inpatient Day
Tygerberg	1,770,596,461.02	56,346,073.05	561,413.28	100.36
Jooste	97,045,844.27	3,088,310.83	76,755.47	40.24
	Equipment Replacement Value	**Annual Economic Cost**	**PDE**	**Cost per Inpatient Day**
Tygerberg	885,298,230.51	120,851,577.44	561,413.28	215.26
Jooste	29,113,753.28	3,974,302.54	76,755.47	51.78

Combining patient-specific, recurrent and capital costs gives the following cost per inpatient day:

Table 15: Summary of the cost per inpatient day

Tygerberg	
Patient-specific Cost	262.41
Overhead Cost	913.57
Capital Cost	315.62
Total	***R 1,491.60***
Jooste	
Patient-specific Cost	193.32
Overhead Cost	808.98
Capital Cost	92.02
Total	***R 1,094.32***

It was found that 71% of all referrals for inpatient care from the Khayelitsha cohort were to the District level, and the remainder (29%) were to the Tertiary level. This weighting applied to the costs for each inpatient day gave an average cost per inpatient day of R1,209.53.

6.3.2 Validity of recurrent cost per inpatient day estimates

Costing of HIV at the inpatient level is clearly an imprecise science, given the wide variety of diseases encountered in any one patient, and the marked difference in diseases between patients.

There were a number of difficulties in establishing patient-specific costs at Tygerberg. Firstly, ethical considerations prevented us from doing a retrospective record review of patient's charts. Therefore, the costing relied on overworked clinicians and nurses filling out cost data sheets and time sheets, and there is a strong likelihood that some cost data were not entered onto the sheets. For instance, over one-third of clinicians did not fill out the part of the cost data sheet relating to the time they spent with each patient. A further limitation was that the one-week data collection period did not always capture data for each patient's full length of stay. This could also lead to an underestimation of costs.

A sound way of verifying these estimates is to calculate the average cost per inpatient day by performing straight step-down costing. This involved calculating the total recurrent expenditure (i.e. excluding capital expenditure) in the facility and allocating it to inpatient days using the PDE as before. The resultant estimate suggests a recurrent cost of R1208 per inpatient day at Tygerberg, which is remarkably similar to the result we obtained of R1176. At Jooste, straight step-down costing lead to a difference of under R4.00.

In other words, the results at Tygerberg and Jooste are likely to be accurate on average.

Although the results at Jooste were accurate at the time of the study, a futher problem could arise from using the CPIX to inflate the results. Medical inflation is not necessarily the same as consumer price inflation. To verify the rate of inflation, a random sample of 24 medicines was selected and the 1996/97 price was compared

to the 2002 price. Although some prices were lower in 2002 than in 1996/97, on average the prices in 2002 were 2.04 times higher, which is higher than the CPIX inflator over this period (1.47). In other words, drug prices in this sample have increased at a higher rate than consumer price inflation, which on the whole indicates that inflating the Jooste results by using the CPIX might lead to an underestimation.

Although this verification of the costs per inpatient day indicates that the estimates are relatively accurate, we have nevertheless varied these costs in sensitivity analysis. This should more than account for any over or underestimation.

6.3.3 Cost of hospitalisation per Markov state and Transition costs

To calculate the ongoing cost of hospitalisation in Markov states, the cost per inpatient day was combined with estimates of utilisation of inpatient services from the cohort of survivors who went on ART in the clinics. A separate calculation was made of the number of days spent in hospital prior to dying from a cohort of patients who died. These were modelled as transition costs for patients transitioning to dead from the various Markov states. Both categories of inpatient costs are summarised in Table 16.

Table 16: Transition costs and ongoing hospitalisation costs in Markov states

Markov State	Ongoing IPD		IPD prior to death	
	# / qtr	Cost	#	Cost
No ART, CD4 < 50, all quarters	0.66	R 795.97	7.13	R 8,622.79
No ART, CD4 50-199, all quarters	0.46	R 560.19	5.28	R 6,383.64
0 - 3 months on ART, CD4 < 50	0.52	R 632.05	4.00	R 4,838.12
3 - 6 months on ART CD4 < 50	0.44	R 531.35	4.00	R 4,838.12
0 - 3 months on ART, CD4 50 - 199	0.19	R 228.38	4.00	R 4,838.12
3 - 6 months on ART CD4 50 - 199	0.09	R 112.46	4.00	R 4,838.12
6 - 12 months on ART	0.28	R 338.52	4.00	R 4,838.12
Beyond 12 months on ART	0.11	R 127.02	3/4 at 4.00, 1/4 at 7.13	R 5,784.58

6.4 Patient-specific cost items

This section describes the utilisation and costs of patient specific items in each Markov state and cycle.

6.4.1 Laboratory and imaging costs

The following table shows the average utilization of laboratory tests and imaging by the cohort.

Table 17: Laboratory tests per quarter

	Pre-ART	Months on ART			
		3	6	12	18
Viral Load		1.5	0.9	0.5	0.4
CD4	1.3	1.0	0.6	0.5	0.3
ALT		2.1	0.7	0.4	0.2
AST		1.2	0.3	0.1	0.1
FBC	1.3	3.2	1.1	0.5	0.4
Differential	1.2	3.1	1.0	0.5	0.4
Lipase		0.4	0.3	0.2	0.2
Amylase		0.4	0.1	0.1	0.1
Sputum	0.6	0.2	0.1	0.1	0.0
Full LFT's		0.4	0.0	0.0	0.0
RPR	1.1	0.6	0.1	0.1	0.1
Xray	0.1	0.2	0.1	0.0	0.0

Although some baseline viral loads are done prior to the commencement of ART, these are captured in the 0-3 months on ART period, as they apply only to patients who ended up on ART. The current clinical protocol with regard to laboratory testing at the clinics for patients on ART is provided for reference (Table 18).

Table 18: Current clinical protocol for laboratory testing

Reg-imen	Visits / procedures / tests	M -0.5	M 0	M 0.5	M 1	M 2	M 3	M 6	M 12	M 18	M 24
All	Informed consent	X									
	Consultations		weekly		bi-weekly		monthly		1-3 monthly		
	Treatment assistant	-	X	-	-	-	X	If VL detectable			
	CD4 cell count	X	-	-	-	-	-	X	X	X	X
	HIV RNA	X	-	-	-	-	X	X	X	X	X
AZT	FBC & Differential	X	-	-	X	X	X	X	X	X	X
D4T	FBC & Differential	X	Clinical reasons only								
EFV	Alanine Transferase (ALT)	X	Clinical reasons only								
NVP	Alanine Transferase (ALT)	X	-	X	X	X	-	X	X	X	X
PI	Cholesterol/Triglycerides	X	-	-	-	-	-	-	X	-	X

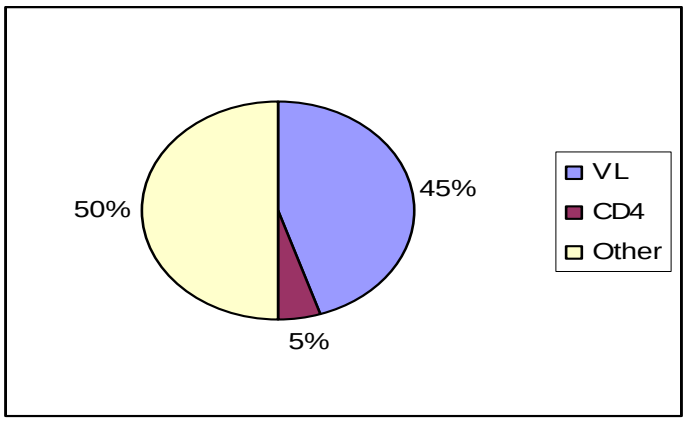
The utilization and costs by stage for laboratory tests are presented for key Markov states below.

Table 19: Laboratory costs and utilisation in key Markov states

		ART 0-3 months		ART 3-6 months		ART 6-12months		ART >12 months		No ART	
Lab Tests	Price	Q	Total	Q	Total	Q	Total	Q	Total	Q	Total
VL	450.00	1.50	673.91	0.92	415.38	0.50	225.00	0.50	225.00	-	
CD4	88.90	1.02	91.05	0.56	50.15	0.48	42.46	0.48	42.46	0.34	29.95
CD8	160.00	-	-	-	-	-	-	-	-	-	-
FBC	28.00	3.24	90.76	1.06	29.79	0.54	15.25	0.38	10.71	0.32	9.00
DIff	20.00	3.15	63.00	1.03	20.51	0.53	10.60	0.38	7.65	0.31	6.12
ALAT	25.40	2.09	53.01	0.69	17.42	0.40	10.05	0.24	5.98	-	-
ASAT	25.40	1.18	30.06	0.30	7.65	0.14	3.60	0.12	2.99	-	-
GGT	25.40	0.41	10.31	0.03	0.81	0.03	0.76	-	-	-	-
Amylase	23.50	0.42	9.76	0.10	2.41	0.13	2.98	0.06	1.38	-	-
Bilirubin	18.80	-	-	-	-	-	-	-	-	-	-
Choles-terol	20.03	0.00	0.10	0.01	0.13	-	-	0.03	0.59	-	-
Glucose	16.45	0.00	0.08	0.01	0.11	-	-	0.03	0.48	-	-
Creat	16.45	-	-	-	-	-	-	-	-	-	-
RPR	16.45	0.56	9.14	0.10	1.58	0.14	2.33	0.12	1.94	0.27	4.45
Sputum	17.80	0.18	3.27	0.06	1.14	0.06	1.06	-	-	0.14	2.48
Lipase	25.40	0.43	11.04	0.25	6.35	0.22	5.50	0.18	4.48	-	-
Xray	61.00	0.19	11.49	0.08	4.69	0.02	1.37	0.03	1.79	0.02	1.50
Cost per quarter			**1,056.98**		**558.14**		**320.96**		**305.44**		**53.50**

Table 19 indicates the large cost of laboratory testing for patients on antiretrovirals versus patients not on ART, especially in the first 3 months (R1,057 versus R54).

Figure 3: Contribution of Viral Loads to the cost of lab testing on ART



Viral load testing is the most important driver of the cost of lab testing on ART, at approximately 45% of the lab costs across Markov states.

6.4.2 Prophylactic medication

Generally, primary and secondary prophylactic medicine was not an important cost driver
(

	Price per Qtr	ART CD4<50		ART 6-12months		No ART CD4 50-199		No ART CD4<50	
		Q	Total	Q	Total	Q	Total	Q	Total
Fluconazole	305.01	0.05	16.37	0.05	16.37	0.08	24.68	0.08	24.68
Co-trimoxazole	7.80 - 11.69	0.94	11.08	0.94	7.58	0.94	7.48	0.94	11.32
Dapsone	101.48	0.03	2.61	-	-	0.01	0.78	0.02	2.41
Cost per quarter			**30.05**		**23.95**		**32.94**		**38.41**

). Despite high prices for both Fluconazole and Dapsone, the proportion of patients prescribed these medicines was low, so the overall cost was low. Note that the table presents results using the generic price of Fluconazole. The patented price for Fluconazole per period is R2,110.37.

Table 20: Prophylactic medication costs

	Price per Qtr	ART CD4<50		ART 6-12months		No ART CD4 50-199		No ART CD4<50	
		Q	Total	Q	Total	Q	Total	Q	Total
Fluconazole	305.01	0.05	16.37	0.05	16.37	0.08	24.68	0.08	24.68
Co-trimoxazole	7.80 - 11.69	0.94	11.08	0.94	7.58	0.94	7.48	0.94	11.32
Dapsone	101.48	0.03	2.61	-	-	0.01	0.78	0.02	2.41
Cost per quarter			**30.05**		**23.95**		**32.94**		**38.41**

6.4.3 Antiretroviral medicine costs

ARV prices vary greatly depending on the choice of regimen and the manufacturer. The baseline costing uses prices of WHO pre-approved ARVs or MCC registered ARVs. Other scenarios use “best-offer” prices –reported to be the lowest prices for ARVs currently available - and the prices of patented ARVs, some of which are only available to the public sector or to NGO’s (MSF, 2002). The MSF Pilot costing uses the ARVs currently sourced by MSF from Far Maguinhos in Brazil. Details of these prices are presented in the following tables.

Table 21: Cheapest WHO pre-approved or MCC registered ARVs

First Line	***Strength***	***Daily Dose***	***Cost per unit $***	***Rand***	***Annual Cost***	***Plus 30% markup***	***Total FL Cost***	***Manufac-turer***
Triomune (3TC+d4T+NVP)	150 + 40 + 200	2	$0.19	R 1.59	R 1,157.81	R 1,505.15	**R 1,505.15**	Clinton Foundation
OR							**OR**	
3TC+D4T	150 + 40	2	$0.17	R 1.43	R 1,042.02	R 1,354.63		Ranbaxy
EFV	600	1	$0.95	R 7.93	R 2,894.51	R 3,762.87	**R 5,117.50**	Merck
Second Line							***Total SL Cost***	
AZT	300	2	$0.25	R 2.05	R 1,499.05	R 1,948.77		Ranbaxy
ddI	100	4	$0.21	R 1.77	R 2,583.73	R 3,358.85		BMS
Kaletra (LPV/r)	133.3 + 33.3	6	$0.23	R 1.90	R 4,168.10	R 5,418.53	**R 10,726.15**	Abbott

Table 22: "Best-offer" ARVs

First Line	*Strength*	*Daily Dose*	*Cost per unit $*	*Rand*	*Annual Cost*	*Plus 30% markup*	*Total FL Cost*	*Manufac-turer*
Triomune (3TC+d4T+NVP)	150 + 40 + 200	2	$0.19	R 1.59	R 1,157.81	R 1,505.15	**R 1,505.15**	Clinton Foundation
OR							**OR**	
3TC+D4T	150 + 40	2	$0.17	R 1.43	R 1,042.02	R 1,354.63		Ranbaxy
EFV	600	1	$0.95	R 7.93	R 2,894.51	R 3,762.87	**R 5,117.50**	Merck
Second Line							***Total SL Cost***	
AZT	300	2	$0.19	R 1.60	R 1,169.99	R 1,520.99		Aurobindo
ddI	100	4	$0.13	R 1.06	R 1,547.80	R 2,012.14		Hetero
Kaletra (LPV/r)	133.3 + 33.3	6	$0.23	R 1.90	R 4,168.10	R 5,418.53	**R 8,951.66**	Abbott

Table 23: Patented ARVs

First Line	*Strength*	*Daily Dose*	*Cost per unit $*	*Rand*	*Annual Cost*	*Plus 30% markup*	*Total FL Cost*	*Manufac-turer*
Combivir (AZT + 3TC)	300 + 150	2	$0.33	R 2.71	R 1,980.46	R 2,574.59	**R 7,327.69**	Glaxo
NVP	200	2	$0.60	R 5.01	R 3,656.23	R 4,753.09		B-I
OR							**OR**	
EFV	600	1	$0.95	R 7.93	R 2,894.51	R 3,762.87	**R 6,337.46**	Merck
Second Line							***Total SL Cost***	
d4T	40	2	$0.08	R 0.63	R 457.03	R 594.14		BMS
ddI	100	4	$0.21	R 1.77	R 2,583.73	R 3,358.85		BMS
Kaletra (LPV/r)	133.3 + 33.3	6	$0.23	R 1.90	R 4,168.10	R 5,418.53	**R 9,371.52**	Abbott

Table 24: Current MSF Regime

First Line	*Strength*	*Daily Dose*	*Cost per unit $*	*Rand*	*Annual Cost*	*Plus 30% markup*	*Total FL Cost*	*Manufac-turer*
Combivir (AZT + 3TC)	300 + 150	2	$0.29	R 2.41	R 1,757.02	R 2,284.13	**R 4,277.79**	Far Maguinhos
NVP	200	2	$0.25	R 2.10	R 1,533.58	R 1,993.66		Far Maguinhos
OR							**OR**	
EFV	600	1	$0.95	R 7.93	R 2,894.51	R 3,762.87	**R 6,046.99**	Merck

6.5 Costs of tuberculosis care

The incidence of tuberculosis in the cohort was between 0.13 and 0.17 per annum for patients on treatment and between 0.36 and 0.56 per annum for patients off treatment. The cost per episode of tuberculosis treatment (Table 7) multiplied by the

incidence in each quarter yielded the cost of tuberculosis per Markov state (Table 25).

Table 25: Incidence and average cost of TB treatment in key Markov states

Markov state	Incidence per cycle	Average cost per cycle
ART CD4<50	0.04	**150.20**
Other ART States	0.03	**115.14**
No ART CD4 50-199	0.09	**325.00**
No ART CD4<50	0.14	**487.05**

6.6 Utility estimates for each Markov State

The following table shows the Health State Value that was obtained through implementing the EQ-5D instrument to patients at the HIV clinics. These values are combined with life expectancy estimates to calculate QALYs. In other words, for a patient at Baseline, one year of life is worth 0.7 QALYs.

Time on ART	Baseline	3 Months	6 Months	12 Months
Sample size	118	116	113	48
Overall health state value	**0.7**	**0.79**	**0.81**	**0.84**

Source: (Jelsma, MacLean et al. 2003)

6.7 Costs, average survival and utility per Markov state

Table 26 shows the average utilisation of services and the average cost for each Markov state over a three month period, the average transition cost (if a patient were to die from the particular Markov state), the average time spent in each state and the utility value. Transition costs are incurred by any patient moving to the “Dead” state in the model. Patients who died off ART with a CD4<50/µl incurred the highest amount of time in hospital prior to death.

Table 26: Costs, survival and utility per Markov state; transition costs from Markov states

Costs in Rands

ART

	CD4<50 0-3 months		CD4>50 0-3 months		CD4<50 3-6 months		CD4>50 3-6months	
	Quantity	*Total*	*Quantity*	*Total*	*Quantity*	*Total*	*Quantity*	*Total*
Clinic Visits	7.80	978.64	7.80	978.64	3.90	492.56	3.90	492.56
Hospitalisation	0.52	632.05	0.19	228.38	0.44	531.35	0.09	112.46
TB Treatment	0.04	150.20	0.03	115.14	0.04	150.20	0.03	115.14
Lab Tests		1,056.98		1,056.98		558.14		558.14
Chronic Meds		30.05		23.95		30.05		23.95
3TC + d4T + NVP	0.39	144.74	0.39	144.74	0.39	144.74	0.39	144.74
3TC + d4T + EFV	0.61	769.73	0.61	769.73	0.61	769.73	0.61	769.73
Cost per State		**R 3,762.39**		**R 3,317.55**		**R 2,676.78**		**R 2,216.71**
Transition Cost	4.00	**R 4,838.12**	4.00	**R 4,838.12**	4.00	**R 4,838.12**	4.00	**R 4,838.12**
Utilities EQ-5D [1]		0.70		0.70		0.79		0.79
Ave Survival		0.14		0.11		0.13		0.10

	FL 6-12months		FL >12 months		SL		No ART CD4<50	
	Quantity	*Total*	*Quantity*	*Total*	*Quantity*	*Total*	*Quantity*	*Total*
Clinic Visits	3.66	461.71	3.26	412.77	3.26	412.77	3.25	321.51
Hospitalisation	0.28	338.52	0.11	127.02	0.11	127.02	0.46	795.20
TB Treatment	0.03	115.14	0.03	115.14	0.03	115.14	0.09	487.05
Lab Tests		320.96		305.44		305.44		53.50
Chronic Meds		23.95		23.95		23.95		38.41
3TC + d4T + NVP	0.39	144.74	0.39	144.74	-			
3TC + d4T + EFV	0.61	769.73	0.61	769.73	-			
AZT + ddI + Kaletra					1.00	2,644.80		
Cost per State		**R 2,174.74**		**R 1,898.79**		**R 3,629.12**		**R 1,695.67**
Transition Cost	4.00	**R 4,838.12**	4.00	**R 4,838.12**	4.00	**R 4,838.12**	7.13	**R 8,622.79**
Utilities EQ-5D [1]		0.81		0.84		0.84		0.70
Ave Survival [2]		0.43		3.32		3.01		1.08

No ART

	CD4>50		CD4<50	
	Quantity	*Total*	*Quantity*	*Total*
Clinic Visits	3.25	319.71	3.25	321.51
Hospitalisation	0.46	559.66	0.66	795.20
TB Treatment	0.09	325.00	0.14	487.05
Lab Tests		53.50		53.50
Chronic Meds		32.94		38.41
Cost per State		**R 1,290.81**		**R 1,695.67**
Transition Cost	5.28	**R 6,383.64**	7.13	**R 8,622.79**
Utilities EQ-5D [1]		0.70		0.70
Ave Survival [2]		0.81		1.46

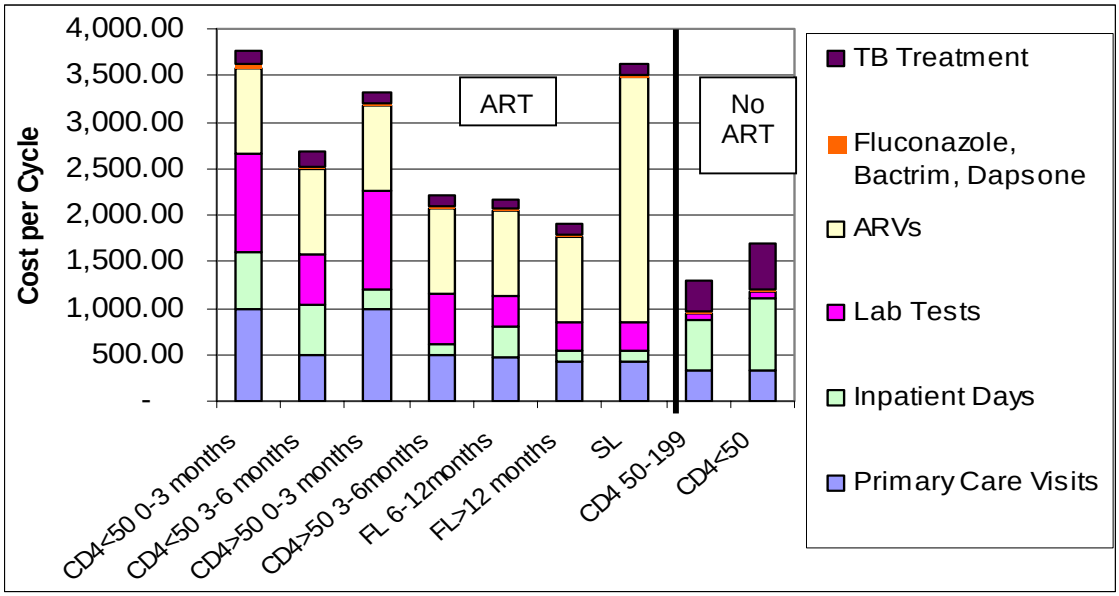
All costs in 2002 prices

[1] Utilities from Jelsma et al, 2003

[2] No ART Survival derived from Maartens et al, 1997

The cost results (excluding transition costs) are shown graphically in Figure 4.

Figure 4: Average cost per Markov state



This figure depicts the importance of the prices of ARVs, especially for second-line. Lab tests are important costs at baseline, but decrease in importance once patients are on treatment for longer than 3 months. Inpatient costs and TB treatment costs decrease as time on treatment increases, but are important costs off ART.

6.8 Cost-effectiveness of ART versus no ART

The Markov model was evaluated using cohort simulation until over 99% of the cohort had died. As shown in Table 27, ART is efficient when compared to no ART, as indicated by the lower cost per QALY for ART relative to no ART. However, if unadjusted LYs are considered, ART is slightly less cost-effective. This difference illustrates the importance of considering the cost per QALY when comparing ART to no ART as it is clear that the interventions are different in terms of both the average life expectancy and the quality of life.

Table 27: Cost-effectiveness results in the baseline scenario

	Lifetime Cost	LY's	QALY's	Cost per LY	Cost per QALY	Incremental Cost per QALY gained
ART	R 93,370.44	8.33	6.79	R 11,208.94	R 13,751.17	R 13,620.04
No ART	R 22,546.24	2.27	1.59	R 9,932.26	R 14,180.03	

All costs in 2002 prices

6.9 Lifetime costs of ART and no ART

The lifetime cost on ART is approximately R93,000 and for no ART it is approximately R23,000. Patients on ART have an average life expectancy of 8.33

years (6.79 QALYs) and patients off ART have an average life expectancy of 2.27 years (1.59 QALYs).

The following figures show a detailed breakdown of costs over the average life expectancy of patients, either with transition costs included under inpatient days in each Markov state so as to show the full distribution of costs (this is a function of the relevant transition cost, the probability of dying from each state, and the average time spent in each state) or without transition costs.

Figure 5: Distribution of costs across time on ART including transition costs

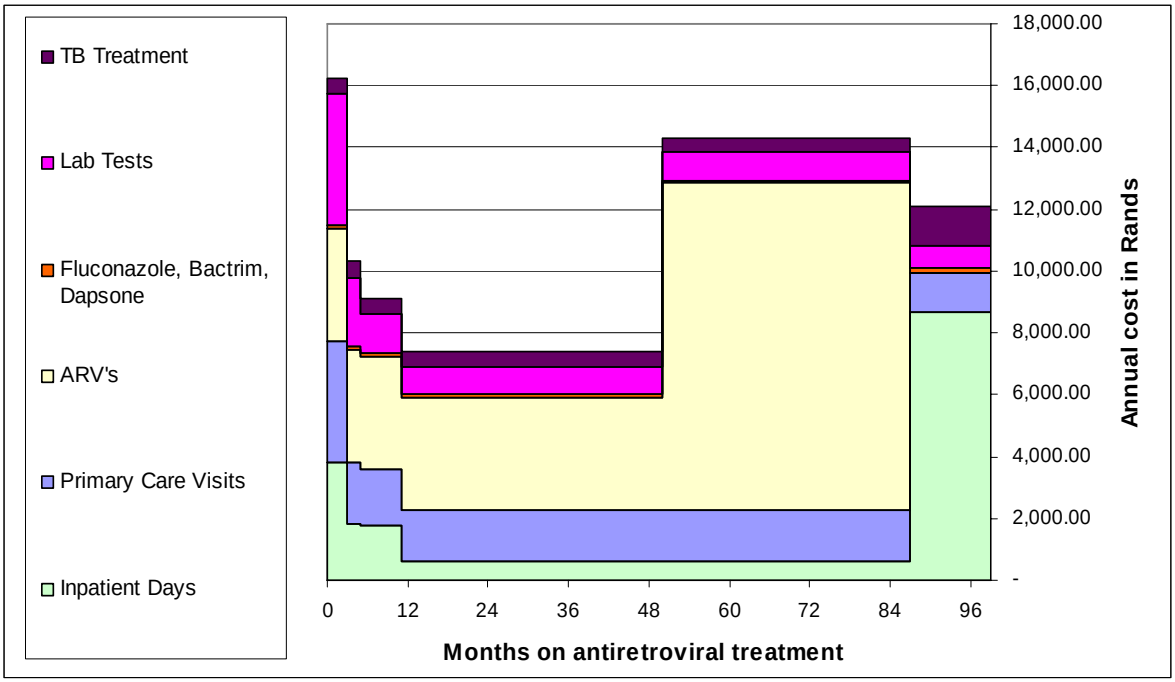


Figure 6: Distribution of costs on ART excluding transition costs

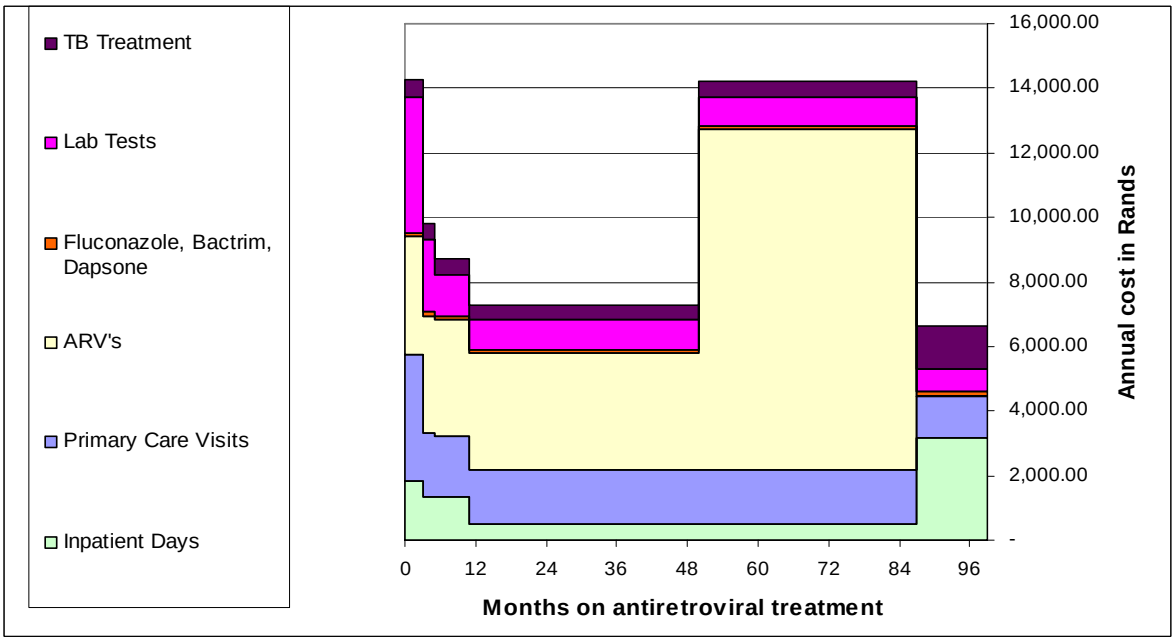


Figure 7: Distribution of costs off ART including transition costs

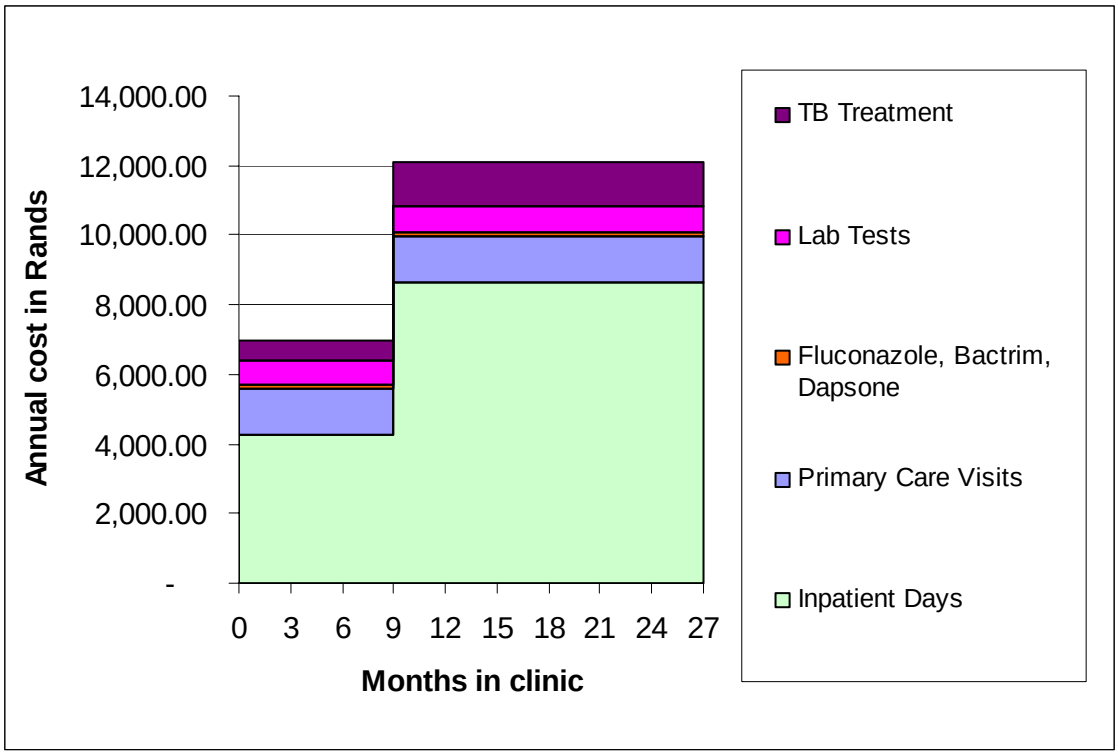


Figure 8: Distribution of costs off ART excluding transition costs

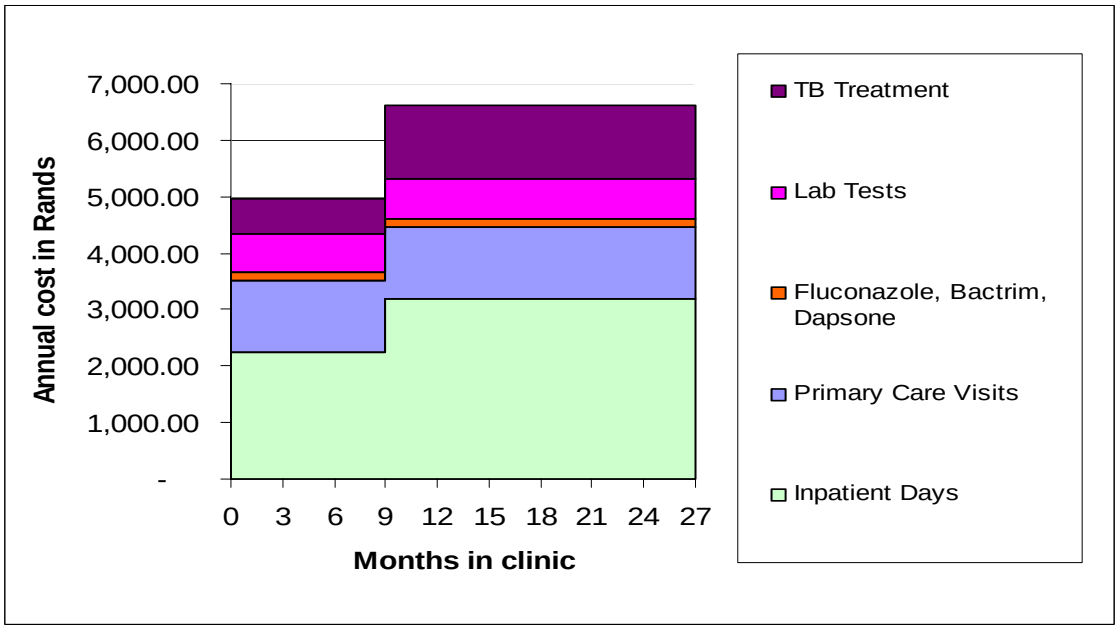


Figure 5 and Figure 6 highlight the following points:

- ARVs make up a large proportion of the cost of the ART option
- Second Line ARVs are particularly important cost drivers
- Lab tests are a large cost item at baseline, but decrease in importance over time
- Inpatient costs are important at baseline (indicating morbidity and mortality due to the low CD4 counts at baseline in this cohort) but decrease in importance until

cohort members go off treatment when they become the most important cost driver

- The costs associated with clinic visits are relatively unimportant. In all states other than the first 3 months on treatment (where there are almost 8 visits per client), the costs of ARVs (approx R257 per month for First Line and R880 for Second Line) are more than double the costs of clinic visits (approx R162 per month).

Figure 7 and Figure 8 clearly show that the most important cost driver off ART is inpatient care.

A slightly different way of expressing the lifetime costs of ART and No-ART is to divide the service-related costs into their different components. In other words, for each type of service, the contribution of salaries, medicines, capital, labs and imaging, and other overheads can be calculated. When each of these is combined with similar categories from other types of services, the lifetime cost from personnel / medicines etc can be calculated. Thus, in the following table, we have grouped costs in this manner to show the *approximate* contribution of salaries, other overheads, capital, medicines, lab tests and imaging, and antiretrovirals, to the total lifetime cost in the baseline scenario.

Table 28: Contribution of different service components to Lifetime costs

	ART	Lifetime Cost	Non-ART	Lifetime Cost
Hospital Personnel	13.07%		47.08%	
HIV Clinic Personnel	11.24%		8.00%	
TB Clinic and NGO Personnel	2.82%		5.83%	
Total personnel	27.13%	25,332.26	60.92%	13,734.14
Hospital Medicines	1.02%		3.66%	
Clinic Medicines	1.53%		3.21%	
TB Clinic Medicines	0.76%		1.57%	
Total Medicines	3.31%	3,087.06	8.45%	1,904.73
Hospital Labs + Imaging	2.04%		7.34%	
Clinic Labs + Imaging	9.20%		7.06%	
TB Labs + Imaging	0.38%		0.79%	
Total Labs + Imaging	11.62%	10,847.09	15.18%	3,421.94
Other Hospital Overheads	0.75%		2.72%	
Other HIV Clinic Overheads	2.31%		2.35%	
Other TB Clinic and NGO Overheads	0.77%		1.59%	
Total Other Overheads	3.83%	3,575.01	6.66%	1,501.53
Hospital Capital	2.00%		7.20%	
HIV Clinic Capital	0.94%		0.82%	
TB Clinic and NGO Capital	0.38%		0.79%	
Total Capital	3.32%	3,099.31	8.82%	1,988.58
Antiretrovirals	50.80%	47,429.72	0.00%	-
Lifetime Cost		***93,370.44***		***R 22,546.24***

As is shown, personnel account for approximately 61% of the total cost of No-ART, but only 27% of ART. If the contribution of antiretrovirals is removed, personnel make up 55% of the lifetime cost, and if antiretrovirals and lab tests are removed, personnel make up about 72% of the lifetime cost of ART.

6.10 Cost-effectiveness of different starting times of ART

Although commenting on the relative cost-effectiveness of different starting times of ART was not initially an objective of this research, it has become an increasingly important issue now that the planning of the ART rollout has commenced. To clarify the importance in terms of efficiency of different starting times of ART, we have assumed that patients start ART either with a CD4<50 or a CD4 50-199.
In Table 29 we present results of this analysis.

Table 29: Cost-effectiveness of different starting times of ART

	Lifetime Cost	LYs	Cost per LY	Incremental Cost per LY
ART Started with CD4<50	R 89,421.63	7.92	R 11,296.82	
ART Started with CD4 50-199	R 98,119.27	8.82	R 11,118.92	R 9,569.52

All costs in 2002 prices

As expected, it is more effective and more cost-effective for patients to start ART with CD4 50-199 than with CD4<50.

Although patients will probably start ART with low CD4 counts during the initial phases of the rollout, hopefully in time it will be possible to enrol patients earlier, in order to take advantage of the extra gains in effectiveness and in efficiency.

6.11 Sensitivity Analysis

Based on the breakdown of the lifetime costs of ART and no ART, the most important cost drivers were identified as inpatient care, the prices of ARVs and whether or not Viral Loads were performed. These were thoroughly varied in sensitivity analysis. The following scenarios were run:

1. Public Model with WHO pre-qualified ARVs or MCC Registered ARVs (baseline scenario)
2. Public Model with best-offer ARVs
3. Public Model using patented drug prices for ARVs and Fluconazole
4. MSF Pilot Model using actual costs of the pilot
5. 15% price reduction on all ARVs from baseline scenario
6. 30% price reduction on all ARVs from baseline scenario
7. 45% price reduction on all ARVs from baseline scenario
8. 10% price reduction on First Line ARVs, 45% off Second Line ARVs from baseline scenario
9. 10% price reduction on First Line ARVs, 30% off Second Line ARVs from baseline
10. 15% price reduction on First Line ARVs, 45% off Second Line ARVs from baseline
11. No hospitalisations in any Markov States or transition costs
12. 20% increase in hospitalisations in Markov States and transition costs from baseline scenario
13. 20% decrease in hospitalisations in Markov States and transition costs from baseline scenario
14. All patients dying have the same Transition Cost (an average of baseline Transition Costs) irrespective of the Markov state from which they transition to the Dead state
15. No Viral Load testing

The following sensitivity analyses were performed on the effectiveness assumptions:

16. An additional 4% of patients defaulting treatment per year
17. Uniform 15% decreased mortality
18. Uniform 15% increased mortality
19. 50% 50% regimen split in survival benefit between regimens

20. 70% 30% regimen split in survival benefit between regimens

In varying the cost items, we have varied one item at a time, holding all others constant in order to untangle each item's impact on the cost. However, for some of these scenarios, there ought to be an impact on the overall effectiveness (for instance, in the scenario where we remove all hospitalisations, it is unlikely that the number of QALYs would stay constant). Although this scenario is unrealistic, it allows a comparison between our results and some of the theoretical costing models that have excluded hospitalisations.

Table 30: Sensitivity analysis results

Scenario	Total Cost No ART	Total Cost ART	Incre-mental cost	QALY's No ART	QALY's ART	QALY's Gained	CUR No ART	CUR ART	ICUR	% Change from Baseline ICUR
Baseline	22,546	93,370	70,824	1.59	6.79	5.20	14,189	13,754	13,621	
Best offer ARVs	22,546	88,080	65,534	1.59	6.79	5.20	14,189	12,975	12,604	-7.5%
Patented ARVs	23,873	105,333	81,460	1.59	6.79	5.20	15,024	15,516	15,667	15.0%
MSF Pilot Costing	23,052	100,037	76,985	1.59	6.79	5.20	14,507	14,736	14,806	8.7%
15% reduction on all ARVs	22,546	86,256	63,710	1.59	6.79	5.20	14,189	12,706	12,253	-10.0%
30% reduction on all ARVs	22,546	79,142	56,595	1.59	6.79	5.20	14,189	11,658	10,885	-20.1%
45% reduction on all ARVs	22,546	72,027	49,481	1.59	6.79	5.20	14,189	10,610	9,516	-30.1%
10% 45% Reduction by regimen	22,546	77,435	54,888	1.59	6.79	5.20	14,189	11,407	10,556	-22.5%
10% 30% Reduction by Regimen	22,546	82,232	59,685	1.59	6.79	5.20	14,189	12,113	11,479	-15.7%
15% 45% Reduction by Regimen	22,546	76,662	54,116	1.59	6.79	5.20	14,189	11,293	10,408	-23.6%
No Hospitalisation	7,628	77,944	70,315	1.59	6.79	5.20	4,801	11,482	13,523	-0.7%
Increased Hospitalisation	25,537	96,460	70,922	1.59	6.79	5.20	16,071	14,209	13,640	0.1%
Less Hospitalisation	19,568	90,288	70,720	1.59	6.79	5.20	12,315	13,300	13,601	-0.1%
Average Transition Costs	19,521	91,362	71,842	1.59	6.79	5.20	12,285	13,458	13,817	1.4%
No Viral Loads	22,546	86,228	63,682	1.59	6.79	5.20	14,189	12,702	12,248	-10.1%
Defaulting at 4% per year	22,546	79,263	56,716	1.59	5.81	4.22	14,189	13,636	13,427	-1.4%
Decreased mortality 15%	22,546	105,330	82,784	1.59	7.56	5.97	14,189	13,932	13,863	1.8%
Increased mortality 15%	22,546	86,069	63,523	1.59	6.31	4.72	14,189	13,634	13,448	-1.3%
50% 50% regimen split in survival benefit	22,546	97,645	75,098	1.59	6.79	5.20	14,189	14,384	14,443	6.0%
70% 30% regimen split in survival benefit	22,546	87,541	64,995	1.59	6.79	5.20	14,189	12,895	12,500	-8.2%

In this table, the last column is especially informative, as it shows how the incremental cost-utility ratio changes relative to the baseline results.

The sensitivity analysis brings the following points to light:

ARVs:

- Lower priced ARVs have a large impact on the incremental cost-utility ratio.

- Using patented ARVs is relatively cost-ineffective (15% increase in ICUR from baseline).
- 10% reduction on FL and 45% reduction on SL gives a similar result to 30% reduction overall; if FL price reductions have bottomed-out, SL price reductions will still have a positive effect on total cost and cost-utility results.

Hospitalisations and Transition costs:

- If no time is spent in hospital, the total costs of ART and no-ART are reduced by approximately the same amount each (it probably does not make sense to comment on the cost-effectiveness here as it is unrealistic to assume that Life Years would remain unaltered).
- The ICUR is robust to changes in hospitalisation assumptions.

Viral Load Testing:

- Viral Loads have a relatively large impact on the ICUR (-10%).

Effectiveness:

- Varying assumptions as to the survival benefit of ART or the numbers of patients defaulting had a minimal effect on the incremental cost utility of the intervention (-1% to +1%).
- Assuming more time on the first-line regimen improved the incremental cost-utility of ART by 8%, reflecting the differential in the regimen costs

General points:

- In 14 out of 20 sensitivity analyses, ART is more efficient than no-ART
- Results for the MSF Pilot are similar to the baseline scenario, indicating that higher salaries, more monitoring, more counselling and diseconomies of scale for the development of tools and training are not important cost drivers.
- Comparing all of the sensitivity analyses, it is only the scenario in which more expensive patented medicines are accessed that increases the incremental cost-utility by over 9% (15% increase)
- Clearly, the results are the most sensitive to changes in ARV prices and usage of Viral Loads
- Given that inpatient time is necessary care for an HIV+ person, the clear emphasis for policy makers should be on price reductions of ARVs.

7 Discussion

7.1 Study design

One of the difficulties in comparing patients who receive ART to those who do not is the potential incomparability of cohorts. The study design for the costing in this case was to use patients as their own control, partly facilitated by the delay in the introduction of ART into a service that was already providing dedicated HIV care. This protects to some extent against incomparability of cohorts. This design does however introduce a different bias in that there is a survivor effect – those patients who ended up on treatment did not experience the period of increased morbidity and associated costs prior to death. The study was able to address this to some extent with respect to hospitalisation, where separate estimates of utilisation related to this period were made from a group of patients who had a similar level of care in the same service, but who died before being able to access ART.

7.2 Effectiveness estimates

Anticipating the survival benefits of ART has been one of the key uncertainties to date in exercises that have considered the cost-effectiveness of the intervention. The main reason for this uncertainty is that the intervention has not existed in a standardised manner for long enough to ascertain what the benefit will be. Making assumptions of constant effects over time is fraught with difficulty. The limited number of antiretrovirals, and the inevitable emergence of resistance and toxicity, has lead many experts to be cautious as to the sustained benefit of the intervention. Nevertheless, the increase in AIDS-related mortality that was anticipated in those countries where ART has been the standard of care since 1996, has not materialised (Mocroft, Ledergerber et al. 2003), suggesting an impact of advances in technology together with a more complex relationship between viral resistance and clinical benefit than had previously been anticipated.

This exercise has attempted to deal with this uncertainty by being conservative in the use of current data that form the basis of extrapolation into the future. By averaging the early mortality experience, and applying this uniformly over time, the model has in effect modelled the worst-case scenario, as mortality is always likely to be highest in the early period on treatment in this setting. The resultant mean and median survival on treatment are slightly longer than those used in South African modelling exercises to date (previous studies have utilised a median survival or survival benefit of between 4.5 and 5 years (Boulle, Kenyon et al. 2002; 2003; Geffen, Nattrass et al. 2003)

Although it might appear that this uncertainty would compromise attempts to comment on cost-effectiveness, when looking at the primary end-point of incremental cost-effectiveness of ART compared to treatment without ART, the estimates are remarkably robust in sensitivity analyses that vary the survival estimates.

7.3 Markov states and transition probabilities

The evidence supporting the Markov states utilised is solid for both the ART and the no ART models, however there were two key areas of uncertainty surrounding transition probabilities for the ART model.

The first area regards the overall time spent on first-line versus second-line regimens. It is essential to include regimen changes, as currently there is a large price differential in antiretroviral medicines between regimens. Our baseline scenario assumes 60% of time on first-line and 40% on second-line, which has been varied in sensitivity analysis. As expected, assuming 50% of time on ART is on first-line and 50% is on second-line worsens the ICUR of ART by 6%. Alternatively, assuming 70% of time on first-line and 30% on second-line improves the ICUR by 8.2%.

The second area of uncertainty regards the manner in which patients failing treatment incur costs and transition to death. The decision was taken in this study to assume a transition from the *second-line* state to the CD4+ lymphocyte count state of less than 50 cells/µl. An alternative approach is to anticipate CD4+ lymphocyte count gains and losses based on primary data (Freedberg, Losina et al. 2001). However, the relationship between the CD4+ lymphocyte count and survival once on ART is complex, and attempts to utilise this approach have dramatically underestimated the benefit of the intervention. In addition, it is very unlikely that patients would cease treatment even if virologically failing before a significant deterioration in immune status.

A potential area of overestimation in costs comes about because we have not made a half cycle correction in the model (i.e. all patients are assumed to die at the end of each Markov cycle and incur full costs for that cycle). Although it would have been possible to apply this correction to the effectiveness component of the model, it was not possible to do the same for the costs due to varying costs in different Markov states. With the selection of a short cycle period (3-months) the maximum error in effectiveness introduced by this is 1.5 months. Over 8.3 years, this is unlikely to have affected the cost-effectiveness estimates, given the robustness of the results to variations in survival in sensitivity analysis.

7.4 Service utilisation

A further key area of uncertainty to date has been the quantity of inpatient days for HIV+ people at various stages of their illness. Many attempts at modelling costs have relied on a single set of inpatient utilisation results estimated during the mid-1990's from an outpatient cohort followed up in Johannesburg (Karstaedt, Lee et al. 1996). This study found approximately 23 inpatient days per year for patients with Stage IV disease.

A strength of our study is that hospital utilisation data have been derived from the same cohort as the primary care cost and effectiveness data. Our results indicate an average of only 10 days in hospital for the year when a patient dies, which is considerably lower than Karstaedt et al (1996). However, increasing inpatient costs by 20% in sensitivity analysis (for both ART and no ART) indicated very little change in the ICUR (although the lifetime costs increased).

7.5 Cost-effectiveness estimates

The overall results of research in this setting indicate that ART is economically efficient relative to no ART. ART costs less than no-ART per QALY, but slightly more per LY. The cost-effectiveness of ART improves markedly as the prices of antiretrovirals drop.

7.6 Lifetime and annual cost estimates

The estimate of the lifetime cost of ART represents one of the first attempts to provide the economic costs of ART that is fully inclusive of all levels of care. Encouragingly, the results that have been obtained in the No Hospitalisation scenario for ART (R9359 per year) are similar to the theoretical costs produced by other South African studies which have not included hospitalisations. The recent study commissioned by the Treatment Action Campaign calculated a cost per year of almost R9,000 if viral loads were included (Geffen, Nattrass et al. 2003).

Clearly, at about 50% of the lifetime cost, the most important factor in the cost-effectiveness of ART is the prices of ARVs are the key cost-driver of the ART intervention (about 50% of the lifetime cost). Fortunately, various recent changes have opened far more opportunities for the procurement of generic antiretrovirals.

Firstly, South African generic manufacturer Aspen Pharmacare has developed the capacity to manufacture generic versions of ddI, 3TC, AZT, combivir and nevirapine (through voluntary licenses) at a price of approximately $1.00 per day for a first-line regimen (Meldrum and Smart 2003). Aspen has also recently entered into an agreement with the Clinton Foundation to deliver Triomune (a fixed dose combination of 3TC, d4T and Nevirapine) at $0.38 per day (Schoofs 2003).

A further welcome change in the market for antiretrovirals has been the decision by GlaxoSmithKline (2003) to extend its voluntary licenses (on AZT, 3TC and Combivir) to include sale to the private sector and to all countries in Sub-Saharan Africa.

However, there are a number of essential ARVs that are still relatively expensive. These include Merck's Efavirenz (an essential component of first-line regimens), Abbott's Kaletra and Bristol Myers Squibb's ddI, both of which are essential second-line drugs. If the prices of these drugs could be reduced in the coming years, the impact on the lifetime cost of ART could be highly favourable.

A further key cost driver for ART is viral load testing. This research has applied the same viral load testing schedule irrespective of regimen, whereas many experts would be happy to omit viral load tests when patients are on second-line. Further, WHO guidelines for delivering ART in developing countries accept that viral load testing may not be feasible in many settings (WHO 2002).

In this analysis, excluding viral loads reduced the incremental cost-effectiveness estimate by 10%. In previous studies based on theoretical costing the omission of viral loads has reduced the cost of care by between 14 and 21%, and in one study reduced the incremental cost-effectiveness by 45% (Boulle, Kenyon et al. 2003). The latter study was working off a much more basic and less costly baseline cost, hence the larger impact of viral load tests on costs.

7.7 Strengths and limitations

This study is a valuable contribution to existing work on the lifetime cost of ART and no ART, and presents the first cost-effectiveness results derived from a public sector clinic-based treatment programme. Cost, effectiveness and quality of life data have all been derived from a single cohort. A further strength is that this research has been conducted in a setting that is closer to the envisaged public sector service model than any other at present.

However, a general limitation of many aspects of this study is that the follow-up duration on ART has been insufficient to capture the full benefit of ART with respect to service costs (although there is much less uncertainty regarding the no ART

costing in this setting). In other words, we have used real data from the first 18 months on treatment to estimate health service utilisation, and have applied these estimates to Markov states for which no real cost data exist. The following is a list of examples of potential cost overestimates:

- Utilisation of just over one clinic visit per month throughout ART despite evidence to suggest that patients might visit less frequently once established on their regimens
- Assumption that patients remain on second-line when treatment has failed
- Insufficient follow-up time to fully capture the reduction in the incidence of tuberculosis
- Insufficient follow-up time to fully capture the reduction in morbidity requiring inpatient care

Although there is less uncertainty surrounding the no ART costing, some care should be taken in applying the lifetime costs calculated in this setting to other settings. It is likely that a variety of factors could affect the demand and supply of health services for HIV-positive patients who are not on ART in other settings.

Finally, this analysis does not include the utilisation of specialised forms of inpatient services, such as tuberculosis hospitals and hospices. Data for these admissions were insufficient for their inclusion. A further omission is the failure to adequately capture outpatient visits at hospitals. Both of these omissions are likely to have biased the study against the cost-effectiveness of ART.

8 Conclusion

Cost-effectiveness analyses use widely accepted methodology to establish which of two or more competing interventions can give the maximum output (e.g. QALYs) for a given level of input (health system resources valued in terms of their economic cost). In other words, it establishes technical efficiency (*doing it the right way*). When compared to other interventions in the health sector, it can also establish allocative efficiency (*doing the right thing*).

That said, cost-effectiveness / cost-utility cannot determine economic feasibility. Even if an intervention were economically efficient, it would not necessarily be affordable. Affordability is an important issue in this case because of the sheer number of people that need ART.

This research is contributing three key findings to the current state of knowledge in this area. Firstly, it provides an indication of the relative efficiency of ART compared to no ART in a setting that is similar to future ART rollout sites in South Africa. Secondly, this research is able to give a better indication of the costs of providing ART over a patient's lifetime than is currently available. Thirdly, it is able to give a solid indication of the current costs of treating opportunistic and HIV-related infections for patients who are not on ART. The latter two pieces of information are essential for budgeting for the ART rollout adequately, whilst the former gives an indication of the relative efficiency of ART versus no ART in similar settings.

ART has been shown to be cheaper per QALY and to lead to enhanced life expectancy. In other words, the intervention is efficient in economic terms, and ought to be pursued if economically feasible and desirable to society. Furthermore, ART can prevent some of the devastation associated with the early mortality of breadwinners and caregivers that is currently being felt across the continent, and

therefore offers immense benefits of the sort that are typically excluded from this type of analysis.

These finding have a number of immediate policy implications.

- The current focus on reducing the cost of antiretroviral drugs is warranted, as on the whole, ARVs still account for nearly 50% of the lifetime cost on ART. This is particularly important for the drugs that remain relatively expensive (such as Efavirenz, ddI and Kaletra). Although personnel costs are not a major cost driver, recruiting and training sufficient human resources to deliver ART will still be a major challenge.

- More emphasis should be placed on reducing the cost of HIV RNA (viral load) testing. Viral load testing makes up almost 50% of the cost of laboratory testing for ART. There should also be clarification of the role of this test in the provision of ART in South Africa.

- The clinical results on which this study is based are a clear demonstration of the potential for the intervention to extend life, and delay many of the individual and societal consequences associated with premature mortality

We hope that these results will assist planners of the ART rollout in the public sector in South Africa to anticipate the costs of the services and to implement ART in the most efficient manner.

References

-- (2003). Full report of the joint Health and Treasury task team charged with examining treatment options to supplement comprehensive care for HIV/AIDS in the public health sector, Department of Health, National Treasury.

-- (2003). GlaxoSmithKline takes further action to help the world's poorest fight HIV/AIDS. London, GlaxoSmithKline.

AbtAssociates (2000). Projected impacts of the HIV/AIDS epidemic on the South African health sector. Johannesburg, Abt Associates.

Boulle, A., Kenyon, C., et al. (2003). A review of antiretroviral costing models in South Africa. Economics of AIDS and Access to HIV/AIDS Care in Developing Countries. Issues and Challenges. J. P. Moatti, B. Coriat, Y. Souteyrandet al. Paris, ANRS**:** 293 -310.

Boulle, A., Kenyon, C., et al. (2002). "Exploring the costs of a limited public sector antiretroviral treatment programme in South Africa." S.Afr.Med.J. **92**(10): 811-817.

Cleary, S. and Committee, T. A. H. S. (2002). Prospective Cost Analysis of HIV/AIDS treatment and prophylaxis at Tygerberg Academic Hospital Complex. Cape Town, Health Economics Unit, University of Cape Town.

Dorrington, R., Bourne, D., et al. (2001). The impact of HIV/AIDS on adult mortality in South Africa. Cape Town, Medical Research Council.

Drummond, M. F., Stoddart, G. L., et al. (1987). Methods for the Economic Evaluation of Health Care Programmes. Oxford, Oxford University Press.

Egger, M., May, M., et al. (2002). "Prognosis of HIV-1-infected patients starting highly active antiretroviral therapy: a collaborative analysis of prospective studies." Lancet **360**(9327): 119-129.

Freedberg, K. A., Losina, E., et al. (2001). "The cost effectiveness of combination antiretroviral therapy for HIV disease." N Engl J Med **344**(11): 824-31.

Geffen, N., Nattrass, N., et al. (2003). The Cost of HIV Prevention and Treatment Interventions in South Africa. Cape Town, Centre for Social Science Research, University of Cape Town.

GlaxoSmithKline (2003). GlaxoSmithKline takes further action to help the world's poorest fight HIV/AIDS, GlaxoSmithKline.

Govender, V., McIntyre, D., et al. (2000). The Costs and Perceived Quality of Care for People Living with HIV/AIDS in the Western Cape Province in South Africa, Partnerships for Health Reform

Haile, B. (2000). The Costs of Adult Inpatient Care for HIV Disease at GF Jooste Hospital. Health Economics Unit. Cape Town, University of Cape Town.

Hellinger, F. J. (1993). "The Lifetime Cost of Treating a Person with HIV." JAMA **270**(4): 474-478.

Hogg, R. S., Yip, B., et al. (2001). "Rates of disease progression by baseline CD4 cell count and viral load after initiating triple-drug therapy." JAMA **286**(20): 2568-77.

Jelsma, J., MacLean, E., et al. (2003). An investigation into the Health Related Quality of Life of individuals living with HIV who are receiving Highly Active Anti-retroviral Therapy (HAART). Cape Town, University of Cape Town.

Karstaedt, A. S., Lee, T. C. M., et al. (1996). "Care of HIV-infected adults at Baragwanath Hospital, Soweto: Part II. management and costs of inpatients." S Afr Med J **86**(11): 1490-1493.

Kinghorn, A. W., Lee, T. C. M., et al. (1996). "Care of HIV-infected adults at Baragwanath Hospital, Soweto: Part I. Clinical management and costs of outpatient care." S Afr Med J **86** (11): 1484-1489.

Laurent, C., Diakhate, N., et al. (2002). "The Senegalese government's highly active antiretroviral therapy initiative: an 18-month follow-up study." AIDS **16**(10): 1363-1370.

Maartens, G., Wood, R., et al. (1997). "Independent epidemics of heterosexual and homosexual HIV infection in South Africa--survival differences." QJM **90**(7): 449-54.

Maartens, G., Wood, R., et al. (1997). "Independent epidemics of heterosexual and homosexual HIV infections in South Africa-survival differences." Quarterly Journal of Medicine **90**: 449-454.

Marseille, E., Hofmann, P., et al. (2002). "HIV Prevention before HAART in sub-Saharan Africa." The Lancet **359**: 1851-1856.

Meldrum, J. and Smart, T. (2003). South African HIV treatment to depend on generic drugs, AIDSmap. **2003**.

Messori, A., Becagli, P., et al. (1997). "Advanced HIV infection treated with zidovudine monotherapy: lifetime values of absolute cost-effectiveness as a pharmacoeconomic reference for future studies evaluating antiretroviral combination treatments. The Osservatorio SIFO sui Farmaci." Ann.Pharmacother. **31**(12): 1447-1454.

Miners, A. H., Sabin, C. A., et al. (2001). "Assessing the cost-effectiveness of HAART for adults with HIV in England." HIV Med **2**(1): 52-8.

Mocroft, A., Ledergerber, B., et al. (2003). "Decline in the AIDS and death rates in the EuroSIDA study: an observational study." Lancet **362**(9377): 22-9.

MSF (2002). Untangling the Web of Price Reductions, Medecins Sans Frontieres.

Nattrass, N. (2002). Unemployment, Employment and Labour Force Participation in Khayelitsha/Mitchell's Plain. Cape Town, Centre for Social Science Research.

Oddone, E. Z., Cowper, P., et al. (1993). "Cost effectiveness of early zidovudine treatment of HIV infected patients." British Medical Journal **307**: 1322-1325.

Orrell, C., Bangsberg, D. R., et al. (2003). "Adherence is not a barrier to successful antiretroviral therapy in South Africa." AIDS **17**(9): 1369-1375.

Orrell, C., Bekker, L. G., et al. (2001). "Adherence to antiretroviral therapy--achievable in the South African context?" S Afr Med J **91**(6): 483-4.

Schoofs, M. (2003). Clinton Program Would Help Poor Nations Get AIDS Drugs. Wall Street Journal. New York.

Shapiro, M. F., Morton, S. C., et al. (1999). "Variations in the care of HIV-infected adults in the United States: results from the HIV Cost and Services Utilization Study." JAMA **281**(24): 2305-2315.

Sinanovic, E., Floyd, K., et al. (2000). Cost and cost-effectiveness of community-based and clinic-based supervision of tuberculosis treatment in Cape Town, South Africa. Cape Town, Health Economics Unit, University of Cape Town.

Sonnenberg, F. A. and Beck, J. R. (1993). "Markov models in medical decision making: a practical guide." Med.Decis.Making **13**(4): 322-338.

Walker, D. and Kumaranayake, L. (2002). "How to do (or not to do).... Allowing for differential timing in cost analyses: discounting and annualization." Health Policy and Planning **17**(1): 112-118.

Weidle, P. J., Malamba, S., et al. (2002). "Assessment of a pilot antiretroviral drug therapy programme in Uganda: patients' response, survival, and drug resistance." Lancet **360**(9326): 34-40.

WHO (2002). Scaling up antiretroviral therapy in resource limited settings: guidelines for a public health approach. Geneva, World Health Organisation.

Currency conversions from:
http://www.oanda.com/convert/fxhistory using the Interbank Rate for both the Euro and the Dollar exchange rate.

Appendix A: Costing Assumptions for the Clinics

1 Staff

1.1. Lay Counsellors

1.1.1 MSF Pilot

There was one counsellor per clinic up until June 2002, thereafter there were 2 per clinic due to increasing client load.

Four out of Five Counsellors are employed and overseen by Lifeline (a local NGO). The remaining counsellor is employed by MSF. MSF also employs a counselling coordinator to further oversee the counsellors in the pilot (and to provide support to other counsellors and counsellor coordinators in Khayelitsha). In effect, this means there is a degree of double counting because counsellors receive supervision via Lifeline and also in-house.

Approximately 50% of a lay counsellor's time is spent doing counselling, the remaining time is spent retrieving or filing client folders, translating between doctors and clients and dividing bulk-bought medicines into visit dosages. 80% of the counselling time is spent with ART clients and the remainder with non-ART clients (based on interviews with the counselling coordinator Leticia Mdani, from research conducted by Dr Taryn Young and an interview with the monitoring coordinator Katherine Hildebrand).

The counselling programme is divided between ongoing one-on-one sessions and support groups. All new referrals (there are 4-5 per clinic per day) to the clinics receive one-on-one counselling around the following topics, depending on the needs of the new client:

- Explanation of expected management in the clinics
- Explanation of clinic procedures
- Post-test counselling if required
- Disclosure
- Nutrition
- Safe sexual practices
- Opportunistic infections
- Family planning
- Referral to other sectors e.g. for Disability Grant / non-ARV support groups run at clinics in Khayelitsha

Once a client has been identified by the medical staff to be potentially eligible for ART, there are a number of counselling sessions:

Session 1:

- Explaining treatment
- Explaining the need for disclosure to a treatment assistant
- Explaining the expectations of the clinic staff around adherence
- Explaining the consent form
- Organising a home visit

Once the above is concluded satisfactorily, the selection committee makes a decision about the suitability of the candidate.

Session 2:
- Telling the selection committee results – candidates that were unsuccessful are explained the reasons for this

Candidates who are not approved are encouraged to seek counselling when they need it and to attend support groups at their local Khayelitsha clinic.

For approved candidates:

Session 3:
- Explanation of the informed consent form

Session 4:
- Topics from Session 1 to 3 are re-explained to the client and the treatment assistant who is included for the first time

Session 5:
- The consent form is signed between the doctor, the client and the treatment assistant
- Explanation of the pill box, the tick sheet, the daily schedule and the self-report sheet (for adverse events and emotional and physical problems)

Once enrolled in the ARV programme, patients visit once per week for the first 0.5 months, every two weeks until the end of the second month and then on a monthly basis to see medical staff and counsellors. Topics discussed with the counsellors include:

- A review of the pill box and tick sheet and an assessment of how treatment is taken
- A review of information on ART and side-effects
- Family planning
- Safe sex
- Opportunistic infections
- Nutrition
- Alcohol

Furthermore, if a patient's viral load tests (done every 6 months) show an increased viral load, the client is referred to the counsellors in order to:
- Explore how the patient has been taking the treatment
- Check for any concomitant medication
- Check if the client has seen any healthcare worker outside the MSF clinic
- Re-explain the adherence measures

Clients are also free to make appointments with the counsellors for more one-on-one counselling if needed.

Each clinic runs an ARV support group once per week for two hours– all clients are expected to attend one per month, but often attend more regularly. The first hour is spent discussing how treatment is taken, packing pillboxes, talking about side effects and emphasizing adherence measures. The second hour is spent discussing a topic identified by the counsellors or the group.

1.1.2 Public Sector
The main difference for the public sector costing is that it is assumed that there is no in-house counselling coordination, but instead this is carried out by the local NGO.

1.2 Pharmacists

1.2.1 MSF Pilot / Public Sector
Clients receive medication from the MSF doctor (including antiretrovirals, cotrimoxazole, multivitamins and certain antibiotics / antifungals etc) as well as from the adjacent public sector clinic pharmacy if appropriate. This cost is likely to be heavier for non-ART clients, but there is no way to know how to weight this split, and the cost is small.

1.3 Monitoring

1.3.1 MSF Pilot
The monitoring team enters client data from each visit into a database. The team consists of two full-time data enterers - 90% of their time is spent entering data for clients at the HIV clinics, the remainder is spent entering data for MTCT and other unrelated work. Besides the data enterers, MSF employs a monitoring coordinator to oversee the data enterers, design the database, audit the data etc. 40% of her time is spent on this work.

From Jan to end March 2002, the monitoring team entered all clients in the database. From March to end December 2002, all clients at Site B were entered, but only ART clients at Site C and MM. These visit totals are used to proportion the salaries of the three members of the monitoring team to ART and non-ART visits.

1.3.2 Public Sector
In the public sector costing, a monitoring coordinator would be situated at the district level; this cost is therefore excluded from the monitoring cost.

1.4 Office Staff

1.4.1 MSF Pilot
Based on interviews with MSF office staff, it was found that four people are involved in administration, coordination and management of the clinic part-time. Their jobs roughly correspond to financial administration, logistics and management (head of mission). The proportion of their time spent on these activities was costed at their actual salaries for this component.

1.4.2 Public Sector
All of these positions were maintained, but expatriate salaries were converted to an equivalent local level based on salaries in PAWC.

1.5 Cleaning

1.5.1 Public Sector and MSF Pilot
The office cleaner is proportioned to “clinic-related” using the split described below, and then shared evenly by ARVs and non-ARVs. The clinic cleaners are shared evenly by ARVs and non-ARVs.

2 Recurrent Overheads

Two types of overhead expenditure are incurred. Some expenditure is part of the Community Health Services Organisation (CHSO) expenditure for Khayelitsha CHC, Nolungile CHC and Michael M CHC. This includes electricity, security and some other general items. This expenditure is allocated using the CHSO visit headcounts for those clinics.

The remaining overhead expenditure is MSF expenditure. However, MSF undertakes many activities including research and advocacy. In order to allocate MSF overhead expenditure, all office staff were interviewed to determine the proportion of their time spent on “clinic-related” work, and time on clinic-work relative to time on other work was used to allocate non-staff overhead expenditure. A list of staff and clinic work time is provided below for 2002:

Researcher / administrator = 0.1 * 7 months
Head of mission = 0.25 * 12 months
Access to medicines campaign coordinator= 0.0 * 12 months
Administrator = 0.0 * 6 months
Monitoring coordinator = 0.4 * 12 months
Monitoring data enterer = 0.9 * 12 months
Monitoring data enterer = 0.9 * 12 months
Adherence researcher = 0.0 * 12 months
Counselling coordinator = 0.8 * 12 months
Receptionist for TAC / MSF = 0.0 * 12 months
Administrator / finance clerk = 0.5 * 12 months
Logistician = 0.5 * 12 months
Research field worker = 0.0 * 9 months
Research field worker = 0.0 * 9 months
Research field worker = 0.0 * 2 month
Research field worker = 0.0 * 2 month
Mobile exhibition assistant = 0.0 * 8 months
Mobile exhibition assistant = 0.0 * 8 months

Unrelated work = 0.67
Related work = 0.33

Table 31: Overhead cost per visit

Description	*Cost per Visit*	*Proportioned by…*
Electricity, water, security, laundry, telephone etc	4.32	COT South total PHC Visits
MSF operation running costs (printing, stationery, maintenance etc)	10.99	"Clinic-related" time of MSF office staff and total MSF visits
Medical and nutritional (condoms, nutritional supplements, gloves, speculae etc)	2.55	Total MSF visits
Logistic and sanitation (maintenance supplies etc)	0.51	"Clinic-related" time of MSF office staff and total MSF visits
Transport-Freight-Storage (vehicle running costs)	1.50	Assumed one clinic-related car, proportioned by total MSF visits
Total Overhead Cost	19.87	

3 Capital Costs

3.1 Adherence

The various adherence tools (pill boxes, tick sheets etc) took approximately 4 months to develop by one person.

This kind of tools development would theoretically be done at the provincial or national level and therefore have far greater economies of scale. It is unrealistic to assume that one person would have to do this for each treatment project that is started. This cost is not included in the public sector costing.

3.2 Staff Training

Counsellors, Nurses and Doctors all receive relevant training.

3.2.1 Counselors

Counsellors are required to complete 3 days of training on HAART / adherence. Training is done by MSF staff, and the individual cost to a counsellor is made up of his or her salary plus the cost of the relevant MSF staff member's time. Assumed 6 counsellors trained at one time. Included only salary and indemnity. The cost of this goes to ARV visit.

3.2.2 Nurse training

Nurses are trained in treatment of HIV-related diseases and HAART over 5 days by more experienced MSF staff. Half is on general HIV medicine, the other half on HAART, so half is shared by all visits, and the other half goes to an ARV visit. Assumed 3 nurses were trained together.

3.2.3 Doctor training

Doctors are trained in treatment of HIV-related diseases and HAART over 5 days by more experienced MSF staff. Assumed 4 doctors were trained together. Half is on general HIV medicine, the other half on HAART, so half is shared by all visits, and the other half goes to an ARV visit.

Training costs consist of the cost of the trainer's time (training was undertaken by existent staff – MSF Pilot used expat salaries, Public sector used average public sector doctor salary), cost of the trainees' time (actual cost of all doctors and all nurses at 2002 prices) plus the cost of stationery and books. It was assumed that all clinical staff received training material and a copy of "Oxford Handbook of HIV Medicine" (Wilson).

Table 32: Training capital cost per visit

	MSF Pilot	Public Model
Clinical staff		
Total Economic Cost	35,864.71	29,844.21
Annual Cost	7,608.93	6,331.64
Non-ART Visit Cost	0.21	0.17
ART Visit Cost	1.43	1.19
Counsellors		
Total Economic Cost	2,313.29	2,313.29
Annual Cost	490.78	490.78
ART Visit Cost	0.16	0.16
Total per ART Visit	**1.59**	**1.35**
Total per non-ART Visit	**0.21**	**0.17**

3.3 Building, Furniture and Equipment Capital costs

The split of staff time between "clinic-related" or "not-clinic related" work was used to allocate building and furniture capital costs for the office. For office equipment, it would be unrealistic to allocate in this manner because of the large amount of electronic equipment used by MSF for research. Instead, it is assumed that office staff require 1 computer, 2 laptops (for monitoring), one printer and one fax machine. Rental for photocopier comes under recurrent expenditure and has been proportioned by "clinic-related" staff time.

For the clinics, capital items were proportioned equally by all visits.

Capital items are annuatised by depreciating the item based on its 2002 replacement cost, its assumed useful life and a realistic discount factor to reflect the opportunity cost of capital.

Table 33: Infrastructure capital cost

Description	Replacement Value	Years of Useful Life	Annual Economic Cost r = 2%	**Cost per Visit r = 2%**
Clinics' Medical Equipment	59,143.20	8	8,073.61	0.44
Clinics' Electronic Equipment	23,997.00	5	5,091.12	0.27
Clinics' Furniture	86,816.73	8	11,851.30	0.64
Clinic Buildings	518,450.00	10 prefab; 50 other	52,275.15	2.82
Office Electronic Equipment	52,265.30	5	27,038.89	1.46
Office Furniture	26,469.37	8	3,769.69	0.20
Office Building	126,373.50	50	4,021.61	0.22
Vehicle	64,000.00	5	13,578.02	0.73
Total	957,515.10		**125,699.40**	***6.78***

Appendix B: Data Collection Sheet for patient-specific inpatient costs at Tygerberg (developed by Dr Kurt Maart)

-1-

DATA SHEET FOR THE COSTING OF HIV/AIDS AT TYGERBERG ACADEMIC HOSPITAL COMPLEX

1. **Folder number** ☐ (patient will remain anonymous)
2. **Date** ……../..……/……... **Age:** ☐ **Gender:** **F** ☐ **M** ☐
3. **Department :** **Paeds** ☐ **Medicine** ☐ **Gynaecology** ☐ **Obstetrics** ☐ **Surgery** ☐
4. **In-patient day** ☐ **(e.g. Day 1, Day 2, etc)**
5. **Ordinary hospital bed** ☐ **/ ICU** ☐ **/ High Care Unit** ☐
6. **HIV/AIDS related disease diagnosis and complications**

DIAGNOSIS	
a)	c)
b)	d)

7. **Describe the patient according to W.H.O. Clinical Staging Criteria** **1** ☐ **2** ☐ **3** ☐ **4** ☐
8. **Procedures done only in relation to HIV/AIDS illness or complication (please tick ✔)**

   Pleural tap ☐ Lumber puncture ☐ Biopsy ☐ Neb mask ☐ IV drip line ☐ Oxygen (FMO2) ☐

   Bronchoscopy ☐ ECG ☐ Sigmoidoscopy ☐ Naso-gastic tube ☐ ICD inserted ☐

   Other: ☐ ***Specify*:……..**
9. **Consumables used only in relation to HIV/AIDS illness or complication (the total amount used daily for each item)** ***please check consumable list***

   a)……d)……..

   b)……e)……..

   c)……f)……..…

10. **Laboratory service, and blood product used in relation to HIV/AIDS illness or complication. Please check laboratory service.**

    a)……e)……..

    b)……f)……..

    c)……g)……..…

    d)……h)……..

-2-

11. Radiology
Only in relation to HIV/AIDS illness or complication

Sonar	CT	X-Ray	Nuclear Med.

Specify : ……

12. Theatre : Only in relation to HIV/AIDS illness or complication

Yes ☐ No ☐

Description : ……..……. **Hours :** ☐

13. Doctor: Note down approximately how much time was spend with the patient examining, doing procedures and other tasks:

Less than15 min ☐ 15 min ☐ 30 min ☐
45min ☐ 60 min ☐ more than 60 min ☐

14. Received counselling from Social Worker: Only in relation to HIV/AIDS illness or complication

Yes ☐ No ☐

15. Received physiotherapy: Only in relation to HIV/AIDS illness or complication

Yes ☐ No ☐

16. Miscellaneous : Only in relation to HIV/AIDS illness or complication

Specify : ……

17. Any prophylactic treatment given? Yes ☐ **No** ☐

Example: Vaccinations or INH Specify:……..

18. Is the patient on TB treatment? Yes ☐ **No** ☐ **(***If yes please specify under 20)*

19. Is the patient on Anti-retroviral treatment? Yes ☐ **No** ☐ **(**If yes p*lease specify under 20***)**

20. Medication : Only in relation to HIV/AIDS illness or complications
Please tick where appropriate

a) ☐ **Discharge medication. Please note *total amount* as dispensed by our Pharmacy. Applicable only at last day of stay**

b) ☐ **Inpatient medication. Please note *total amount per day* as dispensed by our Pharmacy.**

Name of medication	*Dosage*	*PO/IV/IMI*	*Total amount*

21. | **Name of treating doctor:** | **Signature:** |