
THE COST OF TREATMENT FOR HIV/AIDS INFECTED ADULTS: A STUDY OF PUBLIC SECTOR PROVISION IN SOUTH AFRICA

Christopher Browne¹

1. Introduction

With just over 5 million of a total of 46 million South Africans HIV-positive in 2004 (Dorrington *et al.* 2004), there are many reasons to study “The costs of HIV/AIDS”. What do we mean by ‘costs’; who among the infected population is most worthy of scholarly attention; and in what focussed areas will research effort bear most fruit? With the largest roll-out of antiretroviral (ARV) therapy in the world just a year old, South Africa’s researchers are in a unique position to conduct primary studies that draw on this experience, to learn more about the composition of resources actually consumed at different types of institution, unit costs of those resources at today’s prices, bottlenecks and constraints in service provision, equity and access issues and the ability of hospital services to provide a full range of health care to HIV and non-HIV patients alike. Analysis can be either from a provider budget perspective (e.g. province, hospital) or wider societal one. The former emphasises direct / financial costs, whereas the latter measures economic costs. Empirical studies help us to establish affordability as well as guide policy makers towards productive efficiency, in terms of the least-cost method and place of care delivery.

The question, which this paper addresses, is: “what do we know about the costs of treating HIV/AIDS at the average patient level that can help predict total programme costs?” The question of whether ARV treatments are affordable in national budget terms has, in the popular media, almost pre-empted the more central economic issues of whether they are a cost-effective (CE) use of resources and whether human and physical resources exist for their implementation. A discourse of “unaffordability” (Nattrass 2004) has claimed that since ARV therapy is expensive, because of the high cost of antiretroviral drugs, this form of treatment will continue to be beyond the reach of low-income countries’ budgets. The conclusion is then drawn that

¹ School of Economics, University of Kwa-Zulu Natal, Durban, South Africa. email: brownec@ukzn.ac.za. I am indebted to Geoff Harris, my supervisor, for his comments on a draft of this paper. This work has been assisted by a grant from the University Research Fund. Input from Enhancing Care Initiative, KZN, is also acknowledged.

prevention offers better use of scarce resources than treatment (see e.g. Creese *et al.* 2002). On the other hand in South Africa, it is largely irrelevant to conduct sophisticated C.E. studies that aim to illuminate how the marginal dollar (or Rand) should best be spent: the scale of this communicable disease dwarfs all others, both infectious and otherwise, and has already made it a priority for fairly obvious reasons, largely socio-political. It therefore seems more appropriate to ask, as was done in advocating full access to developed country standards of care (notably drugs) for prevention of mother-to-child HIV-transmission: can we afford not to provide universal access for adults to latest standards of care – known as ‘Highly Active Antiretroviral Therapy’ (HAART). If it turns out that HAART is a dominant strategy, producing more effectiveness while at lower cost per patient, refined cost-utility analysis may become unimportant.

Following Drummond (1997), the method is to ask how the cost of an intervention compares to the cost of “doing nothing” / the next best thing? HAART implies categories of costs incurred, but for each cost category there are also counterpart savings, or ‘offsets’, as compared to a defined status quo, which may have been underestimated. Drawing out the implications of a ‘counterfactual’ status quo is problematic, however, as it involves specifying disease events and associated resource use that cannot always be fully observed.

One reason for this is that owing to a paucity of local data, resource utilization has been frequently modelled using data from dissimilar settings and studies. Furthermore, few studies analyse costs from a societal perspective. Private travel costs, costs of lost time, and the costs associated with lost opportunities to work, engage in informal activities, go to school etc. are frequently ignored, as are society-borne costs of orphan care and state provision of grants of various kinds. This is a serious omission as AIDS is a disease that undermines the very fabric of society, especially amongst the young. The net effect will be to understate the net cost/benefit of HAART as direct financial costs dominate the calculation.

The rest of the paper is as follows: Section 2 describes a framework for costing health care, Section 3 describes some of the conceptual sources of uncertainty in cost estimation, Section 4 sets out some results from a few major South African studies, Section 5 selectively reviews the international literature, while Section 6 concludes.

2. A Framework for Economic Analysis

The Harvard AIDS Institute's Enhancing Care Initiative (ECI) identifies the costs of HIV/AIDS at four levels:

- (1) personal medical care deriving from AIDS;
- (2) non-personal costs, such as information, education and research;
- (3) cost of lost output and income due to excess morbidity and mortality;
- (4) psychological costs caused by the epidemic.

AIDS is a costly disease because technologies to diagnose and treat cases and opportunistic infections are expensive. Moreover, infected people are often in the most economically productive period of their lives, causing an important loss of income. This four-way taxonomy can be further disaggregated: personal medical care costs consist of all the costs of identifying, treating and caring for People Living with HIV/AIDS (PLWHA). The non-personal costs constitute important public goods, which must be spread over programmes. The cost of lost output and income applies to individuals, households, industries and countries. The importance of the first and second levels is that they make up the *direct* costs of the epidemic, diverting scarce resources away from other priorities. The third and fourth levels are rarely measured, at least in Africa, owing to the difficulty of quantifying the indirect (opportunity) costs of being unable to work and, even more so, the costs of pain and suffering to sufferers and unpaid care by their relatives.

In Africa estimates of average direct costs of care range from 1.1 to 2.8 times the per capita GNP, and are “often overestimated because they are based on unrepresentative samples of urban, hospitalised patients, whereas most PLWHA have very limited access to specialised care” (ECI, 2000). The challenge of economic analysis is to take into consideration a representative sample of the general population living with HIV/AIDS in terms of residence, gender and age. They conclude: “it could be useful to obtain the direct costs for different areas of a country, allowing for comparisons between rural and urban populations. A more in-depth analysis of direct costs would be useful, in particular the proportion of average annual costs for different components of HIV treatment: early diagnosis, treatment, nursing care, counselling, home based care and social support”. Knowledge of the relative proportions for HAART and no-HAART could prove

especially useful. Also, identifying the average cost of hospital inpatient care as a proportion of average total medical cost (per annum) could suggest opportunities for savings by a re-allocation of functions between inpatient and ambulatory care.

3. Estimation of Costs

Projecting the cost of ART in South Africa has recently occupied leading clinical epidemiologists and health economists. Andrew Boule and Susan Cleary (2004) have investigated the interrelation between numbers potentially in need of ART, the costs per unit of provision and clinical reality. In terms of ‘need’, most models equate this to Stage 4 (AIDS). Uncertainty exists because while the median CD4 count of stage 4 is typically around 120, the median CD4 count of those starting ART in their work was typically below 50, at least in the early stages of programmes. The question of ‘need’ is not a neutral one, as there is debate about the proper percentage of those HIV positive that should be starting HAART. Figures vary from 10% to 25% and there are also constitutional obligations on government to ensure that very poor rural populations are not disadvantaged relative to urban ones.

On the supply side, the Department of Health has set specific targets of ARV provision for those of the HIV positive population in need. About 40,000 out of 54,000 targeted had begun treatment as of March 2005, which was originally the target for March 2004. Demand will partly be driven by the simple clinical eligibility criteria: WHO stage 4 AIDS-defining illnesses and / or symptomatic with $CD4 < 200$ and training in compliance to the drug regimen. The cost implications are discussed in the next section. Boule and Cleary comment that uncertainty around uptake completely dwarfs any uncertainty around unit costs. However, to gauge demand in South Africa for HIV-related care is a separate, very difficult task. Thus the intended focus here is on unit costs of the ‘typical’ patient.

Further uncertainties exist around numbers presently in chronic HIV care. The numbers in care not on ART are around three times those on ART, at Khayelitsha (a township on the outskirts of Cape Town). The Cape Town model assumes three in care for every patient started, but this assumption needs testing outside of the Cape Town setting. Alternative estimates are found in the GOALS model, which simply puts lifetime costs onto all patients dying, or models which

base incident opportunistic infections (OIs) only on HIV prevalence (Boulle and Cleary, 2004). Costing can be economic or financial. Financial costing allows planners to see the budgetary and household resources required, while economic costing includes services provided where the price does not reflect the true resource cost and allows ‘exceptional resources’ to be identified. An example might be the valuation of Home-Based Care-givers’ time. Difficulties were identified in the Cape Town study in the measurement of human resources and capital spending; costing of personnel must recognise that quality of care is not homogeneous and if we seek to gain efficiencies in terms of staff, the quality of care may suffer. As it is, there are grave shortages of staff across the board. Economic costing includes the cost of space and capital in per unit costs and may underestimate short-term requirements. In the Cape Town model, a programme – level set of cost assumptions accounts for financial costs of infrastructure (Cleary *et al.* 2004).

The model of care selected by government also affects costs. In choosing to accredit more expensive hospital sites to provide ART, outpatient costs are correspondingly higher. A major challenge, discussed later, is how the legislated model of HIV care for those not on ART will affect costing – the opportunity cost of ART. The present model of adherence support (treatment buddies, supporters and counsellors) affects costs but is itself subject to uncertainty and potential change.

In conclusion, cost estimates involve many assumptions and uncertainties. Their main value is as a planning tool, notably as a way to comprehensively consider all aspects of a programme of ART. As we turn our attention to local studies, we can summarise economists’ present interests as follows:

- What factors are likely to govern the public uptake of ART?
 - What are the socio-economic characteristics of those accessing ART?
 - In particular, are the very poor, ill-educated or very sick presenting?
- What are the net costs per patient after accounting for cost savings and realignment in the composition of the cost ‘pie’ for the typical individual?
 - Are there significant urban-rural differences, which might help us predict likely numbers taking treatment?

4. National Studies

This sub-section examines the findings of three major South African studies.

4.1 DoH/ Treasury (2003)

A Joint Health and Treasury Task Team examined costs per patient to implement a “simple, robust and affordable model of care, suitable for use in South African public health facilities by non-specialist Medical Officers and nurses,” according to a guideline developed by a panel of South African clinical experts, using WHO guidelines (DoH/Treasury, 2003:Appendix 2, p2). These unit cost estimates were then used in a population model to estimate costs of differing levels of access to ART. Their Technical Team reports total costs of four alternative treatment options: “No ART”, but comprehensive provision of the legislated standard of primary care; then “No ART” plus phase up over three years to 2008 to provide ARVs for 20, 50 and 100% respectively of all new AIDS cases. The costs of “No ART” were calculated per the government’s Standard Treatment Guidelines, which means the “full current package” of No-ARV treatment of opportunistic infections (including TB) and their prophylaxis, and palliative/terminal care, for all those not receiving ART; full costs of ART – drugs, laboratory monitoring and service delivery – for those receiving it, plus nutritional support for 50% of these patients.

The costs of standard non-ARV care can be expected to increase as numbers of AIDS cases rise, leading the Technical Team to assert that “introduction of ARV therapy can certainly be expected to lead to significant decreases in the utilization of No-ARV therapy, and hence will reduce costs in this area.” It recognizes that the need for standard care and treatment for those developing OIs while on ART, or failing therapy, remains. And obviously, “those people who do not have access to ART would continue to require standard treatment as before” (p.20).

However, the Technical Team does not use empirical data for No-ARV costs and relies on the GOALS SA model to estimate them. This model allows for a 50% reduction in the need for treatment of opportunistic infections and for a lesser reduction in the end-care/palliative care load for those on ART. The Technical Team then estimates a total cost for universal access by public sector patients to a comprehensive No-ARV care package. Two conclusions are worth highlighting: first, that introducing high levels of access to ART “would lead to a significant

Draft

reduction in the costs of No-ARV treatment” – about R2 billion per year by 2008 assuming 100% coverage were attained that year. Second, No-ARV treatment continues to involve significant cost, at around R6 billion per year for the bulk of the decade.

Achieving 50% coverage of ART by 2008 is estimated to cost between R4 and 4.9 billion, in 2003 prices. The 50% coverage scenario is felt to be more likely and is close to that which exists *de facto* in Brazil, where no limit on access exists. If self-limiting behaviour of this kind is applicable in South Africa, 50% coverage may actually be a good proxy for ‘universal’ access.

Combining the costs of No-ARV, ARV and nutrition components produces total programme costs of AIDS treatment and care.

Table 1: Total Treatment Costs by Scenario (Million Rand per year)

Scenario	2003	2005	2008	2010
No ARV	5,439	6,287	6,739	6,708
50% ARV coverage	5,529	6,985	10,536	12,941

Source: DoH/Treasury, (2003, Appendix p.26)

Additional costs imposed on the state by ART shows as the difference between the two rows. What is noticeable is that the No-ARV costs rise only modestly as the decade continues, whereas ARV costs rise considerably. In the technical team’s presentation to Ministers on 9th May 2003, despite significant funding increases devoted to strengthening the treatment package since 2002, it was admitted that quality of existing health care still needed improvement, and work on nutrition interventions was still “in progress”. Much of the burden of palliative care was also being borne by communities themselves. The No ART counterfactual therefore could be much more costly if this is taken into account. Similarly, it is counter-intuitive for the ARV costs to rise by so much when we suspect that much cost-saving occurs as the large numbers presently infected do not regress into expensive end-care (Stage 4) treatment patterns but rather are returned to an earlier clinical stage. Patients are admittedly surviving longer on ART, but it seems worth exploring these costs further.

Achieving 50% (100%) coverage of ART by 2008 would involve a net additional cost of between R2.8 (6.7) billion. and R3.8 (8.9) billion. relative to *full coverage* of No-ARV treatment in that same year. This seems cheap by comparison with the potential benefits, which include both measurable items such as significant reduction in mortality and morbidity; the decrease in hospitalisation; the decreased burden of disability grants; the decrease in orphans and their support requirements; and the significant production/income losses averted at the individual, family, business and national level.

4.2 Cleary *et al.* (2004)

A pioneering full cost-effectiveness evaluation, in the MSF pilot clinic setting of Khyelitsha, near Cape Town, is that by Cleary *et al.*, 2004. ART is economically efficient versus no-ART: the incremental cost-effectiveness ratio is \$1 684 /QALY against \$ 1 901 /QALY, but slightly more per Life Year saved. The incremental cost-effectiveness ratio is robust to increases of the order of 20% in inpatient costs. Lifetime Costs are \$ 7 520 and \$ 2 472 respectively; QALYs gained are 4.30 and 1.30 respectively. We examine four aspects of their analysis:

Assessing Chronic HIV Care: Their Markov modeling originally used estimates of the incidence of each individual opportunistic infection (OI) in order to gain precise estimates of resource use. The study however demonstrated the difficulty of separating care episodes into disease episodes; it made more sense to cost the care episodes as a whole. Costing OI episodes individually would tend to overestimate drug costs and position drug costs as the major cost driver. Their hospital work also suggests that inpatient care costs for HIV-medical patients and medical patients generally are similar.

Programme-level costs: These concern coordinating and implementation overhead costs that must be allocated to each treatment centre at the district, provincial and national level; they consist of monitoring and evaluation activities; resistance testing; staff training; technical support from consultants; adherence support at the community level, not linked to the patient – e.g. the so-called ‘buddy’ system of observed therapy and education to encourage adherence; and various forms of social and food security necessary to make the programme work.

Their Cape Town model provides three ways to deal with programme-level costs, but according to authors Boulle and Cleary they are all equally problematic: one may take a percentage of non-

Draft

drugs, non-laboratory costs, or use a cost per person year, or finally use an arbitrary fixed amount. Whatever method is chosen can be subjected to criticism. The extent of their importance can be assessed by further study, perhaps pointing to a method.

Drug prices: “The cost-effectiveness of ART improves markedly as the prices of antiretrovirals drop.” (p.44). In order to illustrate the magnitude of drug price reductions, we may refer to UNAIDS 2004 reports from Uganda, where the prices of a first-line regimen were examined between 1998 and 2003. In June 1998 the cost per year was \$12 000 per patient, which by September had fallen to \$7 000 with the launch of special access pricing for ART in developing countries by major drug companies. In June 2000 prices plummeted to \$1 200. These were reduced to about \$350 by the competition of generic with patent-drug producers, by June 2001.

Study Design Issues: A major challenge in the methodology is to produce comparable service utilization estimates. The authors rightly point out the inherent problems in comparing patients who receive ART to those who do not, chiefly the potential incomparability of the chosen 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 the incomparability of cohorts. This design does however produce 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 utilization 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” (p.43).

Methodologically, there is a trade-off: one can either accept the potential difficulty of measuring the resource/service use for two cohorts, which may not match in important respects that influences their resource use in a non-random manner. Then one derives cost, effectiveness and quality of life data from a single cohort and makes a separate estimate of later-stage utilization to match up that cohort’s patients against patients with similar care levels, at least for hospitalisation, as best one can. Alternatively, one may risk the mis-match of cohorts at the outset, but follow the *actual* resource use for the selected non-ARV patients as they progress

through the resource-intensive phases of Stage 4 AIDS, instead of having to impute resource use from another source.

Service utilization remains a key area of uncertainty that a new study could address. The quantity of inpatient days for HIV+ people at various stages of their illness has been modelled but rarely observed empirically. Cleary *et al.* observe that “many attempts at modelling costs have relied on a single set of inpatient utilization results estimated from an outpatient cohort followed up in Johannesburg at Baragwanath hospital (Karstaedt, Lee *et al.* 1996). This study found approximately 23 inpatient days per year for patients with Stage 4 disease.” The Cleary study, which derives hospital utilization data from the same cohort as primary care cost and effectiveness data, finds an average of only 10 days inpatient for the year of patients’ death, less than half the 1996 Baragwanath estimate. Lifetime costs do increase, but the extent to which this matters for provincial and national budgeting (an affordability question) is by no means obvious or clear, suggesting the need for further empirical study.

4.3 Geffen et al / Treatment Action Campaign (2003)

Geffen *et al.* (2003) calculate the cost of concurrent prevention and treatment intervention, using a three-scenario technique. Employing the ASSA 2000 model, they estimate costs over a period of years and the impact on GNP spending of a universal package of public health HIV interventions. In its most expensive year, 2015, the package would cost about 20.3 billion Rand (about 1.74 % of GNP), involving a 50% increase in public spending on health, from 3.7% in 2000 to 5.4% of GNP in 2015. Total spending on health in 2000 was 8.8% of GNP, the significant difference being private health spending. This caused the *Lancet*, in an editorial, to question whether the country can afford to follow the “Brazil model” of a publicly - funded comprehensive health-care response to AIDS (Lancet 2003). Reflecting the Brazilian findings, people on HAART are theorized to experience fewer opportunistic infections - e.g. TB, PCP, Pneumonia, Kaposi’s Sarcoma) – translating into a saving of hospital costs. They estimate these averted costs from various primary and secondary source, in particular health professionals working in a range of health facilities.

Resource usage is not explicitly described in the paper, which is a deficiency, since price data, including very crucially that of ARV medication, was collected between September 2001 and

October 2002. The data is therefore very outdated, depending on prevailing exchange rate and pharmaceutical market conditions. Scenarios that depict the importation of vastly cheaper generic substitutes, or their local production, are not modelled.

The study discusses the importance of staff costs, which, with drugs, are key cost drivers and notes that these are typically based on other studies rather than empirically measured. A new study could refine these estimates by, for example, observing hours per day actually spent by key personnel with subjects. The study also notes that counselling costs vary between provinces, according to the authors. Very importantly, their calculations paid attention to the reduction in state obligations due to lower incidence of OIs and reduced orphan grant spending. These are seen as direct spin-off's of treating people with HAART and draws into focus the cost of *not* responding with HAART (Decosas, 2003). There would also be savings in palliative care averted or at least postponed, but these are not properly explored.

The reality is that treatment and prevention are a unified conceptual whole and that effectively addressing AIDS involves a “comprehensive health care package”, implying a wide variety of costs of provision (or not providing) and justifying a broad conception of costs. They do this by means of a scenario approach, which follows Johnson and Dorrington (2002). Four distinct health interventions are modelled:

- Voluntary counselling and testing (VCT)
- Prevention of mother-to-child transmission (pMTCT)
- Better management of STIs
- Highly active antiretroviral therapy (HAART)

These interventions may be applied separately, or in conjunction with one another in a general equilibrium sense. HAART is antiretroviral therapy using three or more drug combinations, available to adults with AIDS- defining symptoms or whose CD4 count is below 200. VCT, pMTCT and STI management constitute ‘prevention’ since they primarily reduce transmission of HIV. HAART is ‘treatment’ because it primarily serves to improve the health of PLWHA. Counselling also affects treatment costs while HAART has prevention aspects; the latter are beyond the scope of this paper. Three different scenarios are modelled. Scenario One is meant to

Draft

depict the 2003 default situation in the public sector, when none of the new interventions were implemented and therefore no added direct costs were incurred. Note also that there is one significant AIDS-related cost, namely the treatment of OIs of PLWHA: “hospitalisation costs rise over time under Scenario One as HIV morbidity in the general population increases” (p3).

The scenarios are:

- Scenario One (S1): Treatment of OI’s only.
- Scenario Two (S2): Government adds the three prevention interventions to treatment of OI’s.
- Scenario Three (S3): Adds HAART to prevention measures and treatment of OI’s.

A cautionary note is that the ASSA 2000 model used is calibrated for a base of 2000 and assumes that the interventions are gradually phased in from then. Johnson and Dorrington (2002) assumed that 80% of the population would have access to these interventions by 2005, a figure that is far from the case. Turning to the all-important question of savings, the study finds that S3 has the greatest benefits in terms of reduced mortality, reduced morbidity and reduction of new infections. It also has “a wide range of socio-economic benefits – reduced sick leave; higher productivity; fewer days take off for employees to attend funerals; lower employee replacement costs; preservation of human capital; fewer orphans” (Nattrass, 2002). It is important to note that the authors deliberately do not attempt to measure these socio-economic benefits. There are three points to be made in connection with their analysis of costs:

- From a narrow DOH viewpoint the cost of S3 is substantial at the peak
- S3 involves both costs *and savings* for the public health sector
- Shedding light on the nature and magnitude of the savings enables us to reach a *net cost* calculation of S3.

We now turn to consideration of the financial savings implied by S3. The logic is that when HAART is introduced the consequent improvement in morbidity translates into a fall in utilisation of hospital resources and lower hospitalisation costs. However, the savings to be potentially realised are subject to two sources of error. First, the model used (ASSA 2000) simply *estimates* numbers of people in each stage of the disease. Secondly, the in- and out-patient hospitalisation unit costs for people in each clinical stage have been taken from Kinghorn *et al.* (1996) for Baragwanath Hospital, Soweto, a large urban hospital. The authors thereby rely

Draft

on inflation-adjusted rather than contemporary observed data, from a single institution. This data is now nine years old. The unit cost estimates are shown in Table 3 below:

Table 3 : Hospitalisation costs per patient with HIV

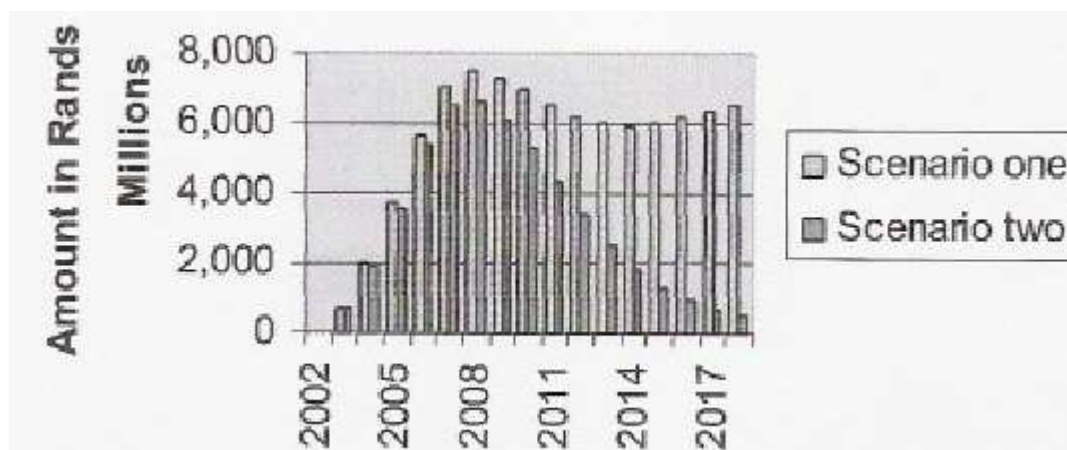
Stage	Cost per patient per year (Rands)
Adult stage 1	1 378
Adult stage 2	1 378
Adult stage 3	6 572
Adult stage 4 (Aids)	18 020
Adult on HAART who has become healthy again	1 378

Source : Geffen et al (2003) citing Kinghorn et al. (1996)

As shown, an assumption is made that people whose immune systems have been reconstituted through HAART have the same costs as those in S1. A further point concerns the demand for, and cost of, hospitalisation in the absence of HAART. The Treatment Action Campaign (TAC) make sophisticated arguments to show the extent of *de facto* rationing occurring in the health sector, by comparing the health budget and the projected hospitalisation costs under S1. In 2002, hospitalisation costs were already 71% of the national health budget, of R34 bn. Costs under S1 are therefore forecast to overtake the available budget without some form of intervention (S2 or S3). The TAC project actual in-and-out patient AIDS hospitalisation needs for 2000 at R13.6 bn against only R3.6bn spent on in-and-out patient hospitalisations in 2000, calculated by the government (Department of Health, 2001). How do we reconcile this information?

The DoH's figures may be under-estimates: they cite a figure of 24% of all hospital admissions being for AIDS-related illness, which seems low in relation to other estimates (e.g. Colvin, 2000). Recall that Cleary finds estimates for individual in-patient costs made in 1996 by Kinghorn to be twice as large as theirs in 2004. Either way, the study's inferences need testing: either a substantial number of patients are voluntarily not presenting with their HIV-related illness at hospitals (what obstacles might hamper their attempts to get treated?) or a high level of rationing exists. A benefit of a detailed empirical study will be to quantify the extent of such deficits, refining the estimates and testing assumptions.

Figure 1 below reveals how much extra spending is involved in providing non-drug care to those not on HAART (or prevention only) compared to those receiving HAART at the phased in levels, (again caveats apply to assumed coverage percentages - from a 2002 baseline):

Figure 1: The Potential Savings of Scenario 3 over its Alternatives

Source: Geffen 2003 (p26)

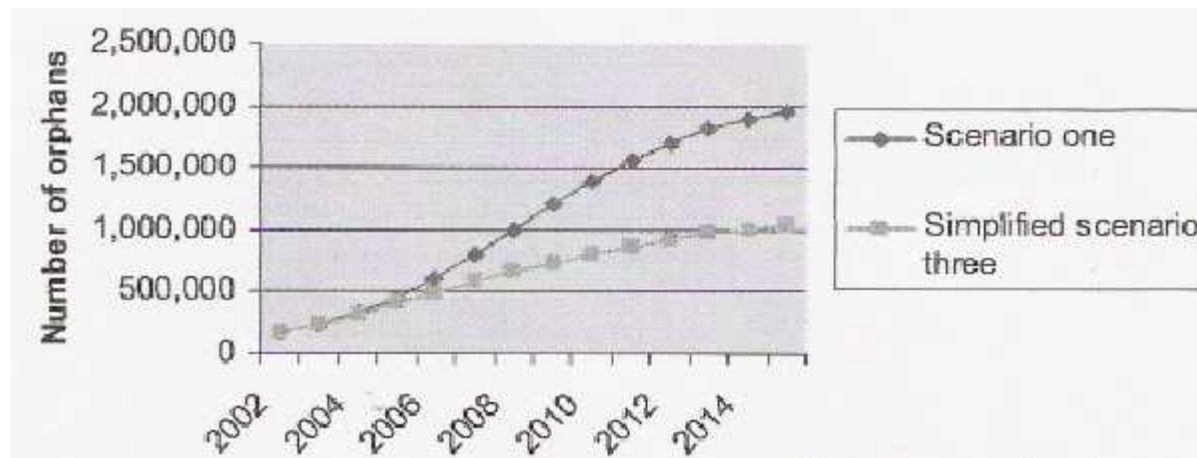
These estimates are based on the assumption that government is already meeting its legal and policy obligations, i.e. that no rationing of health care is occurring. However, for the reasons given above, Geffen comments that “the level of rationing that is probably occurring implies that current government health care policy objectives are not being realised. It is probable that many patients who use the public sector are getting sub-optimal treatment irrespective of their status” (2003, p26). The opinion is backed up by evidence from other writers that the health infrastructure is under “enormous pressure” (South African Health Review, 2002; Mowatt and Quinlan, 2003). The latter found that the burden of HIV/AIDS is felt more acutely at in-patient than at outpatient level. They also highlight that HIV positive patients re-present more frequently than HIV negative patients and this may be a significant cost factor, even though treatment costs for HIV-positive patients are not an average higher than for others.

High levels of rationing seem to be the norm under all three scenarios, it is conjectured, “unless per capita health expenditure is increased to compensate for the shortfall.” Per capita health spending is currently under R700 (about 60 pounds) per annum. The importance of S3, with added direct costs but involving some significant savings, is that the extra spending should relieve pressure on the public health system; this would follow from the finding of reduced demand by patients on HAART compared to their no HAART counterparts. The consequent reduction in rationing should have a range of positive consequences in terms of quantity and quality of care. Again, future work could investigate this.

Draft

Significant savings can potentially be realised if the ballooning number of orphans in South Africa can be arrested by appropriate measures. Under a simplified S3 (VCT + HAART only), 927 000 fewer orphans are expected in 2015 compared to S1. Foster care grants and subsidies to children's homes imply that of at least R1 billion would be saved under S3, using the most conservative estimates and Geffen believes that the saving will be much higher. The subsidy for children's homes in Kwa-Zulu Natal is R1000 per child per month. To put this figure in context, the monthly cost of first-line ARV drugs is between R 350 and R900 per patient per month at 2003 prices, depending on whether generic or patented medicines are bought. Drug prices have been the subject of consistent focus by commentators, but perhaps the issue of orphan care costs needs more emphasis (See Figure 2).

Figure 2 : Number of Double-orphans under S1 and a simplified version of S3



Source: Geffen 2003 (p28)

Further research is also warranted into the net effect of S3 on levels of disability grant spending. Rules on qualifying for disability grant are changing frequently. Currently, the qualification for a disability grant is a CD4 below 200, or if your doctor certifies persistent sickness, so that the patient is unlikely to be able to work. Disability grant recipients must take a blood test once a year. If their CD4 is restored above 200 they may continue to receive grant, but only until the next test. If HAART (S3) restores immune function and can push people who were initially below 200 back above the qualifying threshold, for a period of perhaps 6 – 8 years, the disability grant would no longer be necessary and obligations on the state can be drastically reduced. To illustrate, if 500,000 people were eligible for a disability grant and half were restored to health, the budgetary savings would be R2.34 billion per annum. The total direct costs of S3 do not include savings from the costs of hospitalisation and orphans averted as described above. The

Draft

overall estimate of costs is thus a prudent one. One difficulty faced in reaching conclusions about savings is that we have yet to observe, or properly model, the *total* lifetime costs of a patient who is on HAART. It is not clear whether these will be actually *reduced* by providing HAART, although they certainly will be *postponed* several years. Whereas benefits accrue mostly now, these deferred costs will then appear smaller in today's terms once future cost streams are discounted to present values.

Markov modelling can assist us in this task. However, a persistent difficulty in measurement is that whether a patient is on HAART or not, almost all their health care consumption occurs at the end of their life. As end points cannot be observed, new research must focus on annual costs at the patient level, for patients that can actually be observed now, and then modelled using appropriate intermediate markers. Transition probabilities must be taken from the best available local data. The samples will have to be carefully segmented to reflect appropriate proportions of patients at each disease stage.

From a macroeconomic standpoint, the work of TAC overcomes one of the obstacles to government prevarication, that of budgetary feasibility. Using the projections contained in Geffen's report, which admittedly need revision as they start from a 2002 baseline, one sees that placing 1,147,000 adult patients on HAART in 2007 at a cost of R11bn, amounts to 1% of GNP (excluding savings discussed above). "The WHO (2002: 6) comments that it is feasible, on average, for low and middle-income countries to increase budgetary outlays for health by 1% of GNP by 2007 and 2% of GNP by 2015 compared with current levels, although this may be optimistic given intense competing demands for scarce public resources." It should be recalled that this calculation is made on the narrowest of budgetary perspectives. If the S3 described is modified by changes in the price of ARV medicines, the effect can be put in budget perspective: reducing from R355 to R300 (current best prices assumed) for first line, and from R611 to 450 for second line regimen drops the percentage of GNP from 1.74 to 1.44% in 2015, with 2.3 million adults on treatment in that peak year of consumption, and, again, is well within the 2% range of additional spending recommended by the WHO. This translates to a fall in the cost of adult HAART of R4 billion in that year.

The study has several main conclusions:

- S1 is very costly, both in terms of human lives (no attempt is made to value them) and the resources consumed in hospital from the treatment of opportunistic infections.
- Even if S3 were not selected (i.e. continue with bulk of population not receiving HAART) the state's spending on health and social welfare per capita must increase substantially; rationing attests to the present inadequacy of the health budget. The DoH plan to spend R7-billion a year on managing AIDS in hospitals by 2004. If 24% of all hospital admissions in 2000 were AIDS-related, but deaths are forecast to peak only in 2010, this justifies the TAC's chairman in his comment that unless ARVs are introduced, the massive demand for care from people with AIDS would "destroy the healthcare system."
- As to the reliability of the estimates, there remain uncertainties: although the Lancet (editorial, April 2003) said that Geffen took great care in his estimates, including costs of training and infrastructure, the cost of ARVs remains a "major determining variable". Also, estimates may be realistic, even conservative, but "the total cost projection is only as good as the epidemiological and impact projections to which the unit costs are applied. No model has so far proven accurate in predicting the future profile of HIV epidemics."

5. International Studies

Many studies suggest that in the short term the costs of HAART are partially or fully offset by a decrease in costs associated with inpatient care. For instance, in the USA, Bozette *et al.*'s, large HISCUS survey showed that overall expenditures declined by 10% per annum, 1996-98, when drugs accounted for 56% of the total, against 30% for inpatient care. In 1998 expensive protease inhibitors had just begun to be used in triple therapy, and a month's regimen cost \$784 / patient.

In the Italian ICONA study (70 centres, 1997-2002) the pattern found in Bozette *et al* was confirmed. Average inpatient costs fell from 345 euros per year to 205 euro per year. ARV-related cost increased 34% per person year, but this was more than outweighed by a fall of 88% for inpatient HIV costs. HAART was credited with permanently modifying health resources consumption, moving it from hospital to patient's homes. Dr Des Martin from the HIV

Draft

Clinician's Society is quoted citing research done in John Hopkins University, showing that patients on ARVs reduced their hospital costs by an average of 46%.

Sendi et al, 2001 reviewed published studies evaluating the cost effectiveness of HAART, find nearly all published articles reporting cost and C/E from a third-party payer perspective. Finding immediate economic benefits of preventing disease progressing in patients with moderate to severe immunodeficiency (25 to 250 CD4), the clear suggestion is that there is potential to produce savings in non-drug health-care costs, which may partly or fully offset ARV drug costs, producing cost-effectiveness ratios which compare favourably to other interventions (for example, Freedberg *et al.* 2001, at \$23 000/QALY, at mean CD4 = 87 starting HAART).

Sendi found that although OIs and associated inpatient costs cannot be excluded in the long term, postponement of costs still represents a favorable outcome. This is not only because society has a positive time preference; postponing the occurrence of OIs will raise societal welfare by a prolonged period of better quality of life for infected individuals. Choice of discount rate matters, since a higher rate "reduces the impact of costs resulting from opportunistic infections that are postponed." Studies agree that all future costs should be brought into account, medical and non-medical. The study examined job productivity in HIV-infected patients and found that "in developed countries with a low unemployment rate, the discounted value of savings caused by increased productivity in earlier years exceeds the discounted value of later increases in costs resulting from morbidity" (2001, p709). Under these conditions HAART can be a very efficient treatment strategy, producing cost savings overall. Productivity losses are notoriously hard to quantify.

One such attempt has been made in the sugar industry in Kwa-Zulu Natal by Morris and Cheevers (2000). They assessed the impact of HIV on a cohort of HIV-infected sugar mill employees and the cost this involved for their employers. For 23 workers who took ill-health retirement due to HIV between 1991 and 1998, the cost of HIV treatment increased heavily in the last two years of the employee's tenure. The cost of lost productivity was R 100 for each day lost and the total cost per worker was R 9 544 annually. These costs were divided roughly equally into replacement worker costs, productivity losses and absenteeism. Similar findings

have been found in studies conducted in Zimbabwe and Kenya, which highlight the significant amount of economic activity lost due to HIV infection, and the increasing burden on the economy as the epidemic matures. More up-to-date studies, of other sectors, would be welcome.

6. Conclusion

With 10% of the world's known HIV population living in South Africa, there is an urgent need to study the cost implications of addressing the epidemic, which is concentrated largely among the young, heterosexuals and females, but which affects all race groups and all income brackets. With competing demands on the public purse, notably from education and housing, the argument for an expanded state response, recognised already by some major firms, must come initially through an appeal to budgetary sense. Between 500 000 and 1 000 000 citizens urgently need HAART to restore some economic and social functioning, and at least buy time for them to prepare their households for strategies to cope, and sustain their livelihoods for them. Significant private costs are already borne by these households for what has been described as 'futile' (non-HAART) primary care. Evidence of rationed health services abounds. Many of the anticipated extra costs to the state will not be additional economic costs, as elements of the care burden will in part simply be transferred from households to the state. There is thus a need first to conduct detailed empirical studies, to establish the relative efficiencies of hospital, clinic and home-based care. From Baragwanath hospital one writer confirms "given the divergence in estimates of cost, operational projects that accurately assess cost, benefits and feasibility of provision of comprehensive AIDS care, including ARV therapy, are necessary. Any cost-benefit exercise should take into account the impact of HAART on family members, and decreased hospital admissions for the patient being treated" (Martinson *et al.* 2003). If it turns out that the latest ARV therapies are cost-saving, on a wide view of costs, then spending more on treatment may actually release more resources for further prevention. Economics can provide some of the answers to a problem that is largely behavioural — but ultimately the solutions are in political and spiritual leadership.

References

- Actuarial Society of South Africa. *AIDS and the demographic model of the Actuarial Society of South Africa* (ASSA 2000). ASSA 2002.
- Bozette SA, Joyce G, McCaffrey DF, *et al.* *Expenditures for the Care of HIV-infected Patients in the Era of Highly Active Antiretroviral Therapy*. N Engl J Med, 2001. **344**(11): 817-823.
- Cleary S, Boulle A, McIntyre D, Coetzee D. *Cost-Effectiveness of Antiretroviral Treatment for HIV-positive Adults in a South African Township*. Health Systems Trust: Durban, 2004: 1-67. www.hst.org.za
- Colvin M, Dawood S, Kleinschmidt I, Mullick S, Lallo U. (2000) *Prevalence of HIV and HIV-related diseases on the adult medical wards of a tertiary hospital in Durban, South Africa*. Medical Research Council, Durban, South Africa.
- Creese A, Floyd K, Alban A, Guinness L, *Cost-effectiveness of HIV/AIDS Interventions in Africa: A Systematic Review of the Evidence*. Lancet, 2002. **359**:1635-42.
- Decosas J, from druginfo@healthlink.org.za, accessed 27/2/2003
- Department of Health, *An Enhanced Response to HIV/AIDS and Tuberculosis in the Public Sector: Key Components and Funding Requirements, 2002/2-2004/5*. Pretoria: Department of Health. 2001.
- Dorrington RE, Bradshaw D, Johnson L, Budlender D, *The Demographic Impact of HIV/AIDS in South Africa. National Indicators for 2004*. Cape Town: Centre for Actuarial Research, South African Medical Research Council and Actuarial Society of South Africa. 2004
- Drummond M, *et al.* *Methods for the Economic Evaluation of Health Care Programmes*. OUP: 1997.
- ECI 2000: Accessed 26/03/2002
www.eci.harvard.edu/research/framework/analysis/perspectives/perspectives_econ
- Editorial, Lancet 2003, 361(9364). *HIV Prevention and Treatment in South Africa: Affordable and Desirable*.
www.thelancet.com/journal/vol361/iss9364/full/llan.361.9364.editorial_and_review
- Freedberg KA, Losina E, Weinstein MC, *et al.* *The Cost-effectiveness of combination antiretroviral therapy for HIV disease*. N Engl J Med, 2001. **344**(11): 824-31.

Draft

Geffen N, Natrass N, Raubenheimer C, *The Cost of HIV Prevention and Treatment Interventions in South Africa. Centre for Social Science Research Working Paper No. 28*, University of Cape Town: Cape Town, 2003.

Ijumba P, Ntuli A, Barron P, eds. *South African Health Review, 2002*: 2003

Johnson L, Dorrington R, *The demographic and epidemiological impact of HIV/AIDS treatment and prevention programmes: an evaluation based on the ASSA 2000 model. A paper presented at the 2002 Demographic Association of Southern Africa Conference*. Cape Town 2002.

Joint Health and Treasury Task Team (JHTTT). *Summary 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*, released 1 August 2003.

Karstaedt AS, Lee TCM, Kinghorn AW, *et al.*, *Care of HIV infected adults at Bragwanath Hospital, Soweto: Part 2. Management and costs of inpatients*: S Afr Med J, 1996. **86**(11): 1490-1493

Kinghorn AW, Lee TCM, Karstaedt AS, Khounane B, Schneider H, *Care of HIV-infected Adults at Baragwanath Hospital, Soweto: Part 1. Clinical Management and Costs of Outpatient Care*. S Afr Med J, 1996. **86**(11): 1484-1489.

Martinson *et al.* *South African Health Review, 2003*: Perinatal HIV Research Unit, Baragwanath Hospital.

Merito M, *et al.* *Economic evaluation of HIV treatments: The I.C.O.N.A. cohort study*. Health Policy 2005: in press

Mowatt R and Quinlan T, *A Rapid Appraisal of the cost of HIV/AIDS Treatment to the KZN Public Health Services*: HEARD 2003

Natrass N, *The Moral Economy of AIDS in South Africa* : Cambridge : 2004

Sendi *et al.* *Highly active antiretroviral therapy: pharmacoeconomic issues in the management of HIV infection*. Pharmacoeconomics 2001: **19**(7): 709-13.

World Health Organisation *World Health Report*: WHO 2002.