The Southern African HIV Clinicians Society has appointed Dr Francesca Conradie as its new President. She is an HIV and TB advisor at Right to Care, the country’s largest non-governmental ARV treatment organisation. Other than adult HIV, the organisation also runs programmes for paediatric HIV, prevention of mother-to-child transmission, TB, cervical cancer, medical male circumcision, and HIV counselling and testing. Dr Conradie is well known for her commitment to combatting HIV and TB. At the Society, Dr Conradie takes over from outgoing President, Professor Francois Venter, saying, ‘Francois did a fantastic job. I have big shoes to fill. In keeping with Professor Venter’s legacy, I would like to underscore that the Southern African HIV Clinicians Society will continue its commitment to help the South African Department of Health to achieve the goals of the National Strategic Plan for HIV, STIs and TB for 2012 to 2016.’ Some further comments appear on page 5 of this issue.
CONTENTS

DEBATE
17 Why it’s time to say goodbye to stavudine ... everywhere
20 Low-dose stavudine trials: a public health priority for developing countries

ORIGINAL ARTICLES

REVIEW
28 Safety, strength and simplicity of efavirenz in pregnancy

CASE STUDY
34 Chronic genital ulcer disease with subsequent development of methicillin-resistant *Staphylococcus aureus* (MRSA) urethritis and bacteremia in an HIV-seropositive person – a case observation

CPD QUESTIONNAIRE
Loose insert
This issue of the Journal continues discussions started in the December 2011 issue. Professor Kuhn responds to the concerns raised by Saloojee about the national policy promoting breastfeeding in HIV-exposed infants; she discusses the scientific rationale for the policy revisions, and presents a population perspective rooted in overall child morbidity and mortality. For better or worse, infant feeding is likely to remain an important topic for years to come. Both viewpoints need careful consideration, and subsequent discussions in these pages hopefully can be informed by experiences from implementing the new infant feeding policies in different settings around the country.

Another matter is low-dose stavudine in HIV management. In the last Journal, Innes et al. presented a rationale for a trial investigating low-dose stavudine's impact on therapeutic and toxicity endpoints. Given increasingly scarce resources for ART in much of sub-Saharan Africa, this proposal has intuitive appeal. We now publish a strong response by Andrieux-Meyer et al., who argue that further research into stavudine's use is untenable — with a clear rebuttal from Venter et al. There are important nuances — the use of stavudine in adults v. children; follow-up duration of a trial investigating long-term outcomes; and the changing costs of tenofovir and other more expensive medications — that evade oversimplified judgments.

More generally, these two ongoing debates raise important questions about what we know and how well we know it. Although we aim to practice ‘evidence-based medicine’, the evidence base for many policies and decisions may be surprisingly thin and malleable. The same body of evidence can lead to opposing interpretations, as seen in the debates on infant feeding and low-dose stavudine. The challenge and talent of skilled clinicians and good policymakers is to make sensible decisions in the face of flawed evidence. Fundamental to this is perceiving the likelihood of misjudging — and in turn the ability to acknowledge opposing viewpoints and the importance of continually trying to improve the evidence base on which our decisions are based. These challenges re-emerge constantly, and again in this issue of the Journal. They have been entwined in the theory and practice of medicine for millennia, as recorded in one of Hippocrates’ aphorisms on the art of medicine, from around 400 BC: ‘Life is short; the art is long; opportunity fleeting; experiment fallible; judgment difficult.’

The Journal presents other exciting pieces, including an important critique on the role of efavirenz in pregnancy from Pillay and Black, where clinical judgment has greatly outpaced policymaking. Johnson presents a model-based analysis of ART initiation across the country, and suggests that the scope of the ART roll-out approaches the targets set by the NSF for 2007 - 2011. This is a major accomplishment that underscores the ability of the public health system to achieve ambitious goals, given adequate capacity and resources. In addition, an opinion piece by Kenyon and colleagues calls into question the widely held belief that poverty alone drives the sexual transmission of HIV across South Africa (a contentious assertion that may give rise to more debate), and Katusiime presents an interesting case study on chronic genital ulcer disease in the context of HIV infection.

Good reading!

Landon Myer
School of Public Health & Family Medicine
University of Cape Town
Landon.Myer@uct.ac.za

The first Southern African HIV Clinicians Society meeting that I attended was in or around 2002, at the Pharmaceutical Society in Glenhove Road. Professor Gary Maartens spoke on isoniazid preventive therapy, and the room was full (it was the last meeting in that too-small venue). The initial function of the Society was to help a group of private doctors to better manage HIV infections.

In the dark years, it seemed unlikely that ART would ever be affordable and available in either the public or private sectors. I was working in a public service clinic, and all we could do was treatment and prevention of opportunistic infections. ART was for a select few with money or taking part in research trials. Then, on 1 April 2004, the first patients accessed therapy from the government programme. For months after that, I went and opened the pharmacy cupboards — just to look at the medicines. The atmosphere at the clinic changed; while people arrived very ill, many got better. Informal support groups formed. I remember celebrating the first 1 000 patients on treatment at our clinic. Today, over 15 000 people receive treatment there.

South Africa now has the largest ART programme in the world, with some 1.5 million on treatment. The DoH, under the leadership of Dr Aaron Motsoaledi, commenced the largest-ever HIV testing campaign last year, and over 15 million South African were tested for HIV. There has been a reduction in mother-to-child transmission to 3.5%.

So do we have it all sorted out, and is there no need for a Southern African HIV Clinicians Society? What are the present challenges? What role will I play as President? I have always seen the function of the Society as pushing the boundaries and leading the way in getting the best possible care to HIV-infected South Africans. We must ensure that our guidelines for all aspects of HIV care and prevention are challenged and aligned with international guidelines. Research in South Africa is of the highest standard. Our researchers have been involved in many of the latest breakthroughs in HIV, including early treatment for infants (CHER), the use of treatment as prevention (HPTN052) and microbicides (CAPRISA 004). As soon as any research breakthroughs are made, we in the Society need to assist the DoH to implement them. And TB must receive more attention. South Africa’s TB incidence is high — second only to Swaziland’s — and we rank fourth-highest in the world in multidrug-resistant TB incidence. This huge increase in TB has been driven largely by HIV infection. It seems a great pity to have made such massive progress in HIV treatment, and then lose our people to TB.

FRANCESCA CONRADIE
President
Southern African HIV Clinicians Society
Johannesburg
OPINION

MATERNAL AND INFANT HEALTH IS PROTECTED BY ANTIRETROVIRAL DRUG STRATEGIES THAT PRESERVE BREASTFEEDING BY HIV-POSITIVE WOMEN

Louise Kuhn, PhD
Gertrude H Sergievsky Center, College of Physicians and Surgeons; and Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, USA

The South African Department of Health is justified in withdrawing support for free infant formula. By so doing, it recognises that any intervention that might detract from breast feeding poses a serious threat to infant survival. Since evidence is now strong that antiretroviral drugs used during lactation prevent transmission of infection from a seropositive mother, strategies that promote breastfeeding can now be recommended for enhancing the health of mothers and infants.

The Tshwane Declaration of Support for Breastfeeding in South Africa was recently championed by the national Department of Health as a concrete step to improving maternal and child health in the country. Saloojee, Gray and McIntyre (in the December 2011 edition of this journal) state they are not opposed to this declaration, and welcome the greater support for application of baby-friendly principles in the health services, strengthening community-based programmes to support breastfeeding, and stricter legislation to protect the rights of breastfeeding mothers.

They objected to only one item concerning the withdrawal of free formula for HIV-positive women, and lamented that there has been hardly any response from clinicians, health professionals or civil society groups to this decision. Aside from their objections, the overwhelming response to the Tshwane Declaration from clinicians, health professionals and civil society groups has been enthusiastic support. Moreover, the Tshwane Declaration itself was a culmination of more than two years of consultation between the Department of Health and clinicians, health professionals, civil society groups, including activists and women living with HIV, and Saloojee et al., themselves.

The latter authors state that the evidence base for withdrawal of formula is inadequate. In this paper, I present the extensive evidence base supporting the new South African government policies. The evidence is strong that provision of free infant formula is dangerous and that antiretroviral drugs (ARVs) work. I also discuss whether withdrawal of free formula could be considered unconstitutional – a very important accusation, and one which requires thoughtful consideration.

BREASTFEEDING SAVES LIVES
Saloojee et al. assert that to withdraw support for free formula is a luxury that South Africa can ill afford unless there is 'substantial evidence that the strategy is either ineffective or results in major harm' [emphasis added]. While these are reasonable criteria on which to make any decision about public health, it is extraordinary that they appear to disregard the overwhelming evidence from around the African continent, including countries in southern Africa, that formula feeding is associated with significantly higher mortality in young infants. It is precisely the accumulation of substantial evidence that provision of infant formula is either futile or results in major harm that informed the international recommendations released by the World Health Organization to guide national ministries of health.

Human breast milk is exquisitely regulated, containing not only nutrients but also immunologically active components to protect newborns against disease and support the maturation of their own immune system. Medical research dating back to the Middle Ages identified that orphans and abandoned infants would die unless human breast milk were provided. An experiment was undertaken in the 1970s by formula manufacturers, confident in their ‘modern’ product, who began marketing it in African countries. Provision of infant formula correlated with infant deaths. Fortunately, these deaths also sparked effective pro-breastfeeding advocacy that has helped to shape global public health initiatives. There are extensive biological, clinical, epidemiological and programmatic data indicating that infant formula results in major harm to infants and their mothers. Consequently, it is a falsehood to say that evidence showing the major harm associated with provision of formula is ‘simply lacking’. There is overwhelming evidence of the harmful effects of formula feeding in the general population in southern Africa and elsewhere for decades. Until recently, there was indeed a lack of evidence of any comparable effect among HIV-infected mothers and their exposed but uninfected infants.

Yet there is now substantial evidence in better- and less-resourced settings, including the better-resourced settings of South Africa and Botswana, that formula feeding results in elevated death rates among children who would otherwise be HIV-uninfected and alive. The serious threat to infant survival is the most important justification for the withdrawal of Department of Health support for infant formula.
IS HIV IN SOUTH AFRICA THE EXCEPTION?

Saloojee et al. do not appear to be aware of this expanse of biological, clinical, epidemiological and programmatic research. It seems that their position can only be held if they subscribe to two types of exceptionalism: (i) HIV exceptionalism and (ii) South African exceptionalism.

Postnatal transmission of HIV through breastfeeding is indeed a special case that requires cautious and courageous consideration of appropriate infant feeding policy. HIV transmission can occur throughout the period of breastfeeding, therefore complete abstention from breastfeeding will obviously not permit any transmission to occur via this route. Abstention from breastfeeding will not, however, prevent intrauterine or intrapartum transmission. In the absence of interventions, most (~70%) infants born to and breastfed by HIV-positive mothers will remain uninfected. When HIV-positive women avoid breastfeeding with the goal of preventing the proportion of vertical transmission attributable to breastfeeding, they place their infants at risk of malnutrition, pneumonia and diarrhoeal morbidity and mortality as well as increasing the child’s risk of developmental and cognitive delays. This is the nub of the dilemma, and provision of free formula is not a solution; instead, it’s part of the dilemma that the HIV epidemic has made us face.

In the era prior to the demonstration that ARVs used during lactation can provide a constructive solution to the infant-feeding dilemma, two wish-fulfillment strategies were used instead: Either deny that HIV is transmitted through breastfeeding or deny that there are substantial risks of death and other serious outcomes from formula feeding. Study after study clearly showed that HIV is transmitted to infants through breastfeeding. Denial of the dangers of formula became the more popular position. In 2000, the WHO recommended that HIV-infected women provide formula feeds to their infants as a means of preventing HIV infection. This guidance was based on the premise and intention that public health programmes could be set up that would eliminate the risks associated with handing out formula feeds. This strategy was very powerful because it was able to mobilise resources to buy formula for programmes and for research. But the strategy set aside the special case that requires cautious and courageous consideration of appropriate infant feeding policy. HIV transmission can occur throughout the period of breastfeeding, therefore complete abstention from breastfeeding will obviously not permit any transmission to occur via this route. Abstention from breastfeeding will not, however, prevent intrauterine or intrapartum transmission. In the absence of interventions, most (~70%) infants born to and breastfed by HIV-positive mothers will remain uninfected. When HIV-positive women avoid breastfeeding with the goal of preventing the proportion of vertical transmission attributable to breastfeeding, they place their infants at risk of malnutrition, pneumonia and diarrhoeal morbidity and mortality as well as increasing the child’s risk of developmental and cognitive delays. This is the nub of the dilemma, and provision of free formula is not a solution; instead, it’s part of the dilemma that the HIV epidemic has made us face.

For better or worse, several groups, including my own, bought into this notion and conducted studies to test whether complete avoidance of breastfeeding, or shortening the duration of breastfeeding, would have adverse consequences for infants born to HIV-positive mothers. Sadly, they did. These well-conducted, rigorous research studies with results that have been reviewed by peer scientists prior to publication in leading medical journals were conducted in a wide range of settings in Africa, including better-resourced settings such as Botswana and South Africa. For example, in a clinical trial in urban Botswana where women were randomised either to formula from birth or breastfeeding for 6 months, a doubled risk of death was observed among uninfected infants born to HIV-infected mothers. In this study, participants were carefully screened to ensure all had access to clean water and adequate sanitation, formula was provided free, counselling and support around formula feeding was extensive, and there was a well-functioning health service safety net. In another example in a well-resourced area in rural Uganda, with a sophisticated health service, women were counselled about infant feeding options following AFASS (affordable, feasible, acceptable, sustainable and safe — the acronym summarising the criteria that were proposed at that time as the requirement for formula feeding to be the better choice) and a 6-times greater risk of infant mortality was observed among women who selected formula feeding because they felt it was “AFASS” for themselves. There are several other studies, including numerous studies from South Africa. The consistency of the findings across diverse settings, across different study designs and with established biological processes makes it highly unlikely that the dangers of formula can be explained away as part of the vagaries of clinical research methodology. The findings of these studies, in conjunction with research findings demonstrating the efficacy of ARVs to significantly reduce the risk of HIV infection through breastfeeding, iteratively led the WHO to revise its recommendations from a position of recommending formula feeds as the default feeding practice for HIV-infected mothers, to recommending breastfeeding with ARVs.

Saloojee et al. dismiss this large body of research with the claim that it comes from settings with much higher rates of infant morbidity and mortality than those observed in South Africa (or those parts of South Africa with more resources). This is not true i.r.o. Botswana and other countries, such as Zambia and Malawi, that are more economically disadvantaged, and manifests a confusion between an absolute and a relative risk. An absolute risk quantifies the likelihood that an event will occur in a group; e.g. the risk of dying is 40 per 1000. A relative risk compares two groups: group A has an absolute risk of 40 per 1000 and group B has an absolute risk of 80 per 1000, therefore the relative risk of group B v. group A is double. Even in countries with very low absolute rates of infant mortality, such as the USA, UK and the Netherlands, formula increases mortality; i.e. the relative risk is elevated. But in countries with higher absolute infant mortality rates, the same relative risks translate into a larger absolute number of infant deaths. Moreover, synergy occurs: in populations with high absolute mortality rates, relative risks of death owing to formula are also higher; for example, water contamination, lack of access to adequate sanitation and poor health service infrastructure exacerbate the dangers of formula. But economic disadvantage does not create the biological disadvantage of formula. There is no threshold below which formula no longer causes harm. Breastfeeding saves lives in all countries – South Africa is no exception.

Saloojee et al. misquote three studies as evidence that replacement feeding can be safely accomplished. These studies do report equivalent or better outcomes with replacement feeding; however, the outcome reported is HIV-free survival, and at a time when ARVs were not available to prevent HIV infection through breastfeeding. HIV-free survival is a composite endpoint defined by the absence of either infant HIV infection or infant death. As a public health indicator, it is useful as it reminds us that there is little point in saving infants from HIV if they are only going to die of other causes. However, consideration of only HIV-free survival does not provide proof of safety of formula feeds. Equivalent HIV-free survival means that the number of HIV infections averted has been cancelled by the number of additional uninfected deaths caused.

Saloojee et al. dismiss this large body of research with the claim that it comes from settings with much higher rates of infant morbidity and mortality than those observed in South Africa (or those parts of South Africa with more resources). This is not true i.r.o. Botswana and other countries, such as Zambia and Malawi, that are more economically disadvantaged, and manifests a confusion between an absolute and a relative risk. An absolute risk quantifies the likelihood that an event will occur in a group; e.g. the risk of dying is 40 per 1000. A relative risk compares two groups: group A has an absolute risk of 40 per 1000 and group B has an absolute risk of 80 per 1000, therefore the relative risk of group B v. group A is double. Even in countries with very low absolute rates of infant mortality, such as the USA, UK and the Netherlands, formula increases mortality; i.e. the relative risk is elevated. But in countries with higher absolute infant mortality rates, the same relative risks translate into a larger absolute number of infant deaths. Moreover, synergy occurs: in populations with high absolute mortality rates, relative risks of death owing to formula are also higher; for example, water contamination, lack of access to adequate sanitation and poor health service infrastructure exacerbate the dangers of formula. But economic disadvantage does not create the biological disadvantage of formula. There is no threshold below which formula no longer causes harm. Breastfeeding saves lives in all countries – South Africa is no exception.

Saloojee et al. misquote three studies as evidence that replacement feeding can be safely accomplished. These studies do report equivalent or better outcomes with replacement feeding; however, the outcome reported is HIV-free survival, and at a time when ARVs were not available to prevent HIV infection through breastfeeding. HIV-free survival is a composite endpoint defined by the absence of either infant HIV infection or infant death. As a public health indicator, it is useful as it reminds us that there is little point in saving infants from HIV if they are only going to die of other causes. However, consideration of only HIV-free survival does not provide proof of safety of formula feeds. Equivalent HIV-free survival means that the number of HIV infections averted has been cancelled by the number of additional uninfected deaths caused.
Odimune

Tenofovir 300 mg
Emtricitabine 200 mg
Efavirenz 600 mg

3-in-1
ONCE DAILY

Recommended as preferred 1st line regimen by national and international guidelines$^2,3$

“Simplify, simplify.”
Henry David Thoreau
Philosopher

REFERENCE:
1. SEP excl. VAT
2. Antiretroviral therapy for HIV infection in adults and adolescents (WHO 2010 Guidelines)
3. Guidelines for the Use of Antiretroviral Agents in HIV-1-Infected Adults and Adolescents (DHHS, December 09)

Cipla Medpro (Pty) Ltd. Reg. No. 1995/004182/07, Rosen Heights, Pasita Street, Rosen Park, Bellville, 7530. Tel (021) 943 4200 E-mail: medicalpa@ciplamedpro.co.za Website: www.cipla.co.za
These stark statistics do not resolve the question of whether breastfeeding alone or formula feeding is the better feeding practice for HIV-exposed infants. New data from research studies undertaken in over 6 countries, including South Africa, completely transform the context in which the dilemma of infant feeding by HIV-infected mothers should be considered. ARV intervention can be used during breastfeeding to reduce the risk of transmission. None of the studies referred to above used ARVs during breastfeeding; the Kenyan study was done before even short-course perinatal interventions became available. These studies are uninformative for the current era when ARVs are available to prevent transmission through breastfeeding. Saloojee and colleagues remain locked into an evidence base and paradigm that does not recognise the potency of ARVs and the opportunity they present to improve the health and survival of HIV-exposed infants.

SOUTH AFRICA HAS PERSISTING INEQUITIES IN HEALTH AND WEALTH

As an argument in support of free formula, Saloojee et al. remind us that South Africa is not a single homogenous country. This is absolutely true. South Africa is a country with gross disparities in wealth, health and living conditions. This is not an argument for the government to support formula for the better-off. To the contrary, new government policies can serve to reduce inequities and provide highly effective interventions to everyone and not just a number of favoured groups; the new national policies proactively consider the needs of the poor first – as public health policies should.

In Saloojee et al.’s view, the government should provide free formula for women in the wealthier provinces, such as Gauteng and Western Cape, where women are sufficiently well-off to meet AFASS criteria but not so much as to purchase formula themselves. Yet this contradicts the first ‘A’ of AFASS, which is ‘affordable’ – a point which they seem to ignore. But it’s not actually the A which is most relevant – it’s the S for safety. Back to the first point: formula feeding is dangerous.

STUDY THE NUMBERS – THEY MATTER

It is therefore, reasonable to ask why HIV-positive women in the United States are required to formula feed. It is imperative to pay close attention to the actual absolute and relative risks in the South African context. Saloojee et al. in their protest to the Mail & Guardian claim that, since infant mortality may be as low as 25 per 1 000 in some better-off parts of South Africa, this figure is below the ‘accepted’ threshold where formula feeding can be considered ‘safe’. The basis for this claim is mathematical modelling, conducted by several different groups (including myself) in the 1990s, calculating the competing risks of HIV transmission associated with breastfeeding v. the increased risk of uninfected child deaths owing to abstention from breastfeeding. A ‘safe’ threshold is the point at which the number of HIV infections averted by formula is exactly equivalent to the number of deaths caused by formula feeding – hardly a basis for a resounding endorsement of formula. Nevertheless, the primary limitation of the models used by Saloojee et al. is that they ignore the new opportunities provided by ARV strategies. ARVs, when used throughout lactation, significantly reduce the risk of HIV transmission via lactation. If one applies the new rates of HIV transmission observed when ARVs are given, the infant mortality rate has to fall to below 10 per 1 000 before the increased number of deaths caused by formula feeding is counterbalanced by the number of HIV infections averted. Only when the infant mortality rate is <8 per 1 000 live births, is formula able to save one child per 1 000. If transmission rates are lower than assumed in the model, and are as low as observed in clinical trials, such as the trial in Botswana, or risks associated with formula feeding are higher than observed in clinical trial settings, as is likely to occur in practice, infant mortality rates need to be even lower before formula can be considered a desirable option (Fig. 1). The infant mortality rate is nowhere near this level in any of the populations affected by HIV in South Africa, even in the wealthier provinces. New government policies take into account the newest up-to-date data, in contrast to the complaints made by Saloojee et al. that rely on out-of-date data and arguments that exclude the availability of ARVs.

ANTIRETROVIRAL DRUGS PROVIDE A SOLUTION TO THE INFANT FEEDING DILEMMA

Many commentators are blithely optimistic about the safety of formula, yet this optimism does not extend to ARV strategies. Regarding the ARV strategies to prevent mother-to-child HIV transmission, they state these strategies are unproven, based on inference, and with many ‘unanswered questions’. This is surprising, since these authors have been at the forefront of testing ARV drug strategies and have published data showing the efficacy and safety of drug interventions and have been highly active in supporting their successful roll-out in Gauteng and elsewhere in South Africa. South African researchers have a stellar record in implementing ARV-based programmes including demonstrating the capacity of the routine health services to provide effective ARV strategies for pregnant HIV-positive women. This is not to say that ARV programmes are easy to implement and that they may fall short. But pessimism, and claiming that failure to implement perfect programmes will have ‘drastic consequences’, damages mobilisation of resources and the will to implement these programmes. ARV drug strategies are highly effective in reducing mother-to-child HIV transmission through all routes, including breastfeeding, and save women’s lives. Programmes to implement these strategies should be supported, not disparaged.

My major concern about this pessimism is that it implies that formula is a better option than ARV drugs. This is deeply
disturbing because formula does nothing to prevent mother-to-child transmission that can occur during pregnancy and delivery. Formula does nothing to improve maternal health. Even if formula is provided, ART treatment and prophylactic regimens remain vital. Based on current criteria, a large proportion of pregnant HIV-positive women meet criteria for ART treatment. They need this treatment as a matter of urgency for their own survival and well-being, and this treatment needs to be lifelong. To my mind, it is problematic to argue against ART therapy because formula is more cost-effective in preventing HIV transmission in a select group of mothers and children. Women who meet criteria for treatment are responsible for a large proportion of the infant infections (>80% of postnatal infections). Therefore, purely implementing standard adult guidelines for provision of ART therapy for pregnant women who require it based on their own health status would address the majority of postnatal HIV infections and would also reduce maternal deaths. Choices of infant nevirapine prophylaxis (option A), or therapeutic regimens that are stopped after the cessation of breastfeeding (option B), are available to address the remaining small proportion, but apply only to those asymptomatic women with high CD4 counts. The focus of public health interventions needs to be on reaching the women who need treatment for their own health and who are also most likely to transmit.

HIV IS NOT THE ONLY DISEASE FROM WHICH CHILDREN NEED TO BE PROTECTED

Saloojee et al. argue that ‘[a]n HIV-free generation can never be achieved while breastfeeding continues.’ This is true. But this statement could be more properly rephrased ‘An HIV-free generation can never be achieved while pregnancy continues.’ Current ART drug regimens do not result in zero transmission even in formula-fed populations. When ART drug regimens were started early in pregnancy and continued through breastfeeding in a study in Botswana, the overall transmission rate, including transmissions that occurred during breastfeeding, was 1.1%. More than 75% of the HIV infections were detected at birth and had occurred before delivery. Transmission during 6 months of breastfeeding when ART drugs were given was 0.28%. Eliminating breastfeeding will not eliminate HIV transmission. Eliminating breastfeeding will, however, increase infant mortality.

BREASTFEEDING RIGHTS AND WRONGS

It was not clear from the arguments presented by Saloojee et al. what the basis was for the charge that withdrawal of free formula was unconstitutional. It may be the denial of the ‘opportunity to have an HIV-uninfected child’ that will result if women are denied access to formula despite meeting AFFAS criteria. This rhetoric is seductive but not based on fact. Formula will not guarantee that an HIV-positive woman has an uninfected child. Without ART drugs, transmission will occur during pregnancy or delivery in about a quarter of women. With adequate ART drugs given during pregnancy and then stopped, transmission rates would be around 2%, assuming complete abstinence from breastfeeding. A woman may have a ‘right’ (in the broadest sense of the word) to purchase harmful commodities if she so chooses – just as she has a ‘right’ to smoke during pregnancy if she so chooses. However, to claim that a woman has a constitutional right to be given harmful products by the health services simply because they prevent HIV transmission, is wrong. Moreover, Saloojee et al. fail to mention children’s rights, also protected in the South African constitution and detailed in the Convention of the Rights of the Children. This is more than just avoidance of HIV infection.

Health policies should not be decided upon by popularity contest. National health authorities should solicit opinions on policies so that they are sensitive to communities’ needs, but the policies themselves need to be based on biological and public health principles and evidence. Involvement of the HIV-infected and -affected community is central. Children, who can be both infected and affected by HIV, need special lobby groups to attend to their interests. The majority of HIV-positive women care about HIV transmission to their infants and their overall health, well-being and survival. The answer of the health service to an HIV-positive woman’s question about how best to feed her infant should not be a blunt ‘your choice’.

A WAY FORWARD

It is time to put aside polarising debates and conflicts, and come together to address the fundamental public health challenges facing South Africa. Programmes to support breastfeeding need to be strengthened. This includes addressing the education of healthcare workers so that correct information is conveyed to parents, as well as activism to challenge labour and other policies that deny the rights of breastfeeding women. HIV can be treated with ARVs, and those receiving ARVs have a very low risk of transmitting HIV to their child or sexual partners. We should synergise to ensure that all people living with HIV have access to effective care and treatment. Strengthening these ARV programmes can greatly improve maternal and child health in South Africa.

REFERENCES


Whereas individual-level parameters may influence which individuals in a given population acquire infection, it is population-level parameters that affect the prevalence of infection.1 Sexually transmitted infections (STIs) are transmitted via sexual networks, and differences in the structure of these networks constitute the key population-level parameter that determines differences in HIV prevalence.1 The differences in HIV prevalence between South Africa's racial/ethnic groups (19.9%, 3.2% and 0.5% for 15 - 49-year-old blacks, coloureds and whites respectively2 are as big as those between the highest- and lowest-prevalence countries in the world. These large racial/ethnic differences are not related to individual level risk factors such as lifetime number of sexual partners, but are more likely determined by different sexual network structures.3 In African networks, sexual partnerships are more likely to be arranged concurrently, and this increases the interconnectedness of the sexual network in a non-linear fashion.3 Evidence from numerous sources and disciplines shows that these high concurrency rates are a key factor in driving high HIV transmission rates in southern and eastern Africa.4

Two main categories of factors have been advanced as being important in the promotion of these high concurrency rates: cultural and socio-economic factors. Socio-economic factors are unlikely to be the predominant determinants since neither HIV nor concurrency rates are determined by socio-economic factors. Rather, high concurrency rates are maintained by a culturally sanctioned tolerance of concurrency. Why then do socio-economic explanations trump cultural ones in the South African HIV aetiological literature? In this article, we explore how three factors (a belief in monogamy as a universal norm, HIV's emergence in a time of the construction of non-racialism, and a simplified understanding of HIV epidemiology) have intersected to produce this bias and therefore continue to hinder the country's HIV prevention efforts.

EXPLAINING THE UNDER-APPRECIATION OF CULTURAL FACTORS

HIV'S EMERGENCE IN A TIME OF THE RESONANCE AND CONSTRUCTION OF NON-RACIALISM

The first factor relates to the post-apartheid context of the emergence of HIV. Notions of white racial and cultural superiority were central pillars of the apartheid ideology. An uncritical use of race as an analytical variable and on occasion frankly racist views would characterise much South African medical and public health enquiry during the apartheid period. HIV then emerged into prominence during the difficult period while South Africa was attempting to build a new dispensation based on non-racialism. Given this backdrop and the fact that HIV was sexually transmitted, deeply stigmatised and then found to disproportionately affect
black South Africans, it is not difficult to see why many of the investigating experts downplayed the racial differentials in HIV spread and biased their assessments of aetiology towards socioeconomic factors. To have suggested that culturally backed norms were important in HIV spread might well have been construed as racist. An example of the ongoing reluctance to use race or ethnicity as an analytical variable in regard to HIV in South Africa, is the 2008 Human Sciences Research Council HIV Survey. Despite it being South Africa’s only nationally representative HIV-serolinked survey, it does not mention racial differentials in HIV rates anywhere except in one small table in the appendix.8

**MONOGAMY AS A UNIVERSAL NORM**

The second factor derives from the unacknowledged post-Christian ethical foundation of much of the South African HIV epidemiology. One dimension of this is the subtle way that monogamy (either lifetime or serial) is assumed to be normative for all humans. Little consideration is given to the wealth of anthropological and historical evidence as to the normative nature of polygamy in stratified societies across place and time,9 and more pertinently, the fact that polygamy is still far more widely acceptable in sub-Saharan Africa than elsewhere in the world.10 The spread of Christianity in South Africa led to the suppression of polygamy. The historical record is clear that this did not lead to a reduction in the total number of concurrent partners, but only to the non-main partners being kept secret.11 Having main and more or less secret-extra partners is still widely practised and tolerated in the region. Authors who have provided evidence that these high concurrency rates lead to high-risk sexual networks in the region have, however, been portrayed as racist and ‘crypto-racist’.12 If these authors label as racist the argument that monogamy is less prevalent in parts of Africa, then it necessarily follows that these authors regard monogamy as more ethical. Even if this belief in mononormativity exists at a fairly subliminal level, then the cultural explanation for generalised HIV epidemics in Africa may clash with one’s principles of non-racism – one is stating that Africans are more likely to engage in unethical behaviour. Given that mononormatism is protected by its unacknowledged status, this clash should lead to the triumph of the commitment to non-racialism. The theory of cognitive dissonance predicts that given this scenario the mind should then actively search for other theories, such as socio-economic and biological ones, to explain the higher HIV prevalences in Africa (see Fig. 1).

**POORLY DEVELOPED CONCEPTUAL FRAMEWORK FOR HIV SPREAD**

High-risk networks characterised by high concurrency rates are now recognized to be key to the generation of generalised HIV epidemics. Evaluating the strength of these network level effects requires network-level analyses. One of the most dramatic limitations of much of the aetiological literature on HIV epidemiology in South Africa, is the absence of network levels of analysis. A recent example is a study that compared individual level sexual behaviours between South African and United States youth surveys.13 Based on little difference between these parameters in the two countries the authors conclude that differences in sexual behaviour are unlikely to explain South Africa’s generalised HIV epidemic. They ignore network level factors in their analysis and the literature which shows that network level factors are able to explain the magnitude and patterning of South Africa’s epidemic.3 The conclusions of the paper and the accompanying editorial14 are that HIV prevention efforts need to shift away from focussing on sexual behaviour and the norms which underpin these, and instead campaign for conditional cash transfers and a range of biological measures of proven efficacy for HIV prevention.

**TECHNICAL INTERVENTIONS AS THE NEW PANACEA FOR HIV PREVENTION**

In the absence of a national consensus ever having been attained that a culturally sanctioned norm is driving HIV-spread in the

---

**Fig. 1. The cognitive processes involved in evaluating two competing theories for why HIV has spread so extensively in some racial groups in South Africa.**

As illustrated here, the lack of evidence to support the socio-economic thesis should lead to its dismissal (red arrow), while the validity of the evidence to support the cultural hypothesis should serve to strengthen it as an explanatory cognition (yellow arrow). In the setting of the strong ideologies of class-determinism and monogamy-as-a-universal norm and the anchor cognition of wanting to present oneself as antiracist, however, the cultural thesis generates considerable cognitive dissonance (one is implying that Africans are more likely to engage in unethical behaviour), and the theory is therefore rejected (green arrows). Likewise, if one is sufficiently committed to class as the explanation for differing HIV rates, then the dissonance produced by the lack of evidence to back one’s ideologically determined theory up can be reduced by the selective interpretation of evidence to back it up (orange arrows).
country, the majority of contemporary papers on HIV prevention in South Africa continue to focus on socio-economic and technical inventions. The currently favoured interventions include vaginal microbicides, test-and-treat, increased condom usage or STI vaccines. 13–15 Some argue against behaviour change campaigns owing to their futility, 14 while others argue that further research on this topic is immoral. 16 One prominent paper that does mention dealing with high concurrency rates (albeit as one of a long list of factors) goes on to state that conditions created by apartheid were responsible for the genesis and maintenance of high concurrency rates.15 The authors then claim that South Africa’s HIV Strategic Plan ‘is comprehensive’ and ‘highlights that South Africa is not deficient in policy’ (p. 926). Unfortunately, this national plan does not mention the urgency of dealing with concurrency. In fact there is still little more than a few small ad hoc programmes in South Africa to effect the mass social mobilisation necessary to lead to norm and behaviour change in this regard.

CONCLUSION

The key to Uganda’s success in rapidly bringing down HIV rates was the way Uganda fairly rapidly recognised the importance of encouraging ‘zero grazing’ or reducing extra partners. 17 Unfortunately, HIV is still viewed by too many in South Africa as a disease of poverty and inequality. Where concurrency is acknowledged to be important, it is too often regarded as being driven by socio-economic factors. The net effect has been that insufficient focus and research has been directed at the normative cultural factors that sustain the high concurrency rates in South Africa. As a result, there has not been the same pressure brought to bear on effecting the necessary changes in tolerance of extra partners in South Africa as has been the case in Uganda.

REFERENCES

5. Kenyon C, Boule M, Badri M, Axelmen V. I don't use a condom (with my regular partner) because I know that I'm faithful, but with everyone else I do: The cultural and socioeconomic determinants of sexual partner concurrency in young South Africans. Journal of Social Aspects of HIV/AIDS 2010;7:35-43.
DEBATE

WHY IT’S TIME TO SAY GOODBYE TO STAVUDINE ... EVERYWHERE

Isabelle Andrieux-Meyer, MD
Polly Clayden
Simon Collins
Nathan Geffen
Eric Goemaere, MD, DTMH, PhD
Mark Harrington
Sharonann Lynch
Tido von Schoen-Angerer, MD, MSc
Tracy Swan

1 Médecins Sans Frontières Access Campaign
2 HIV i-Base
3 Treatment Action Campaign
4 Médecins Sans Frontières, South Africa
5 Treatment Action Group

Editor’s note: The previous issue of the SAJHIV (December 2011) carried an Opinion piece by Innes, Cotton and Venter regarding the potential value of low-dose of stavudine (20 mg twice a day). They suggested that reduced dosing of stavudine may lead to levels of viral suppression comparable with those achieved with stavudine 30 mg bd but with a lower risk of toxicity and side-effects, and at a fraction of the cost of tenofovir. The Opinion was related to a larger proposal, led by Venter, to conduct a head-to-head trial comparing low-dose stavudine with tenofovir (both in a regimen including lamivudine and efavirenz) on viral suppression and other treatment outcomes over 24 months. There has been considerable debate regarding the advantages and disadvantages of low-dose stavudine, and in turn the value of any such trial. Here the debate continues with a commentary by Isabelle Andrieux-Meyer et al. and a rebuttal by Venter and colleagues.

We read with interest an opinion piece by Innes et al. in the previous issue of the journal, regarding the potential value of low-dose stavudine (20 mg twice daily). The article focused on stavudine use in paediatrics (where there are fewer approved antiretrovirals compared with adults, although there will be greater choice in the near future, as the US Food and Drug Administration (FDA) has recently approved tenofovir for 2 - 12-year-olds, and other regulatory agencies are expected to follow suit). In the article, the authors used the situation with children to argue for a proposal, led by Venter, to conduct a head-to-head non-inferiority study in adults comparing low-dose stavudine with tenofovir (both in a regimen including lamivudine and efavirenz) with a 48-week virological endpoint and other treatment outcomes over 96 weeks. We have serious concerns about this proposed trial, for the following reasons:

1. Stavudine is more toxic than tenofovir, and for this reason it is an inferior treatment option. The proposed trial aims to establish virological non-inferiority, which is a moot point, given the severe adverse events associated with stavudine. Considerable evidence supports the use of tenofovir over stavudine; regulatory bodies and the World Health Organization (WHO) have turned away from the drug. In 2004, stavudine was removed from the list of preferred first-line antiretroviral drugs recommended by the US Department of Health and Human Services (DHHS). Starting in 2006, the WHO recommended that countries start moving away from stavudine, and in 2009 recommended that the drug be phased out in first-line antiretroviral therapy (ART) programmes. Earlier this year, the European Medicines Agency (EMA) revised the indication for stavudine, noting that “… the use of the medicine should be severely restricted in both adults and children ... Prescribers are reminded of the severe side effects seen with Zerit [stavudine] and should only use the medicine when other appropriate treatments are not available. Patients being treated with Zerit should be assessed frequently and switched to appropriate alternatives as soon as possible.”

Médecins Sans Frontières (MSF)/Doctors Without Borders have provided further compelling evidence of stavudine’s toxicity in an operational setting. In a Lesotho cohort, the authors found that “… for patients on stavudine, the risk of a toxicity-driven regimen switch was almost six times higher than tenofovir.” The high incidence of adverse events among patients on stavudine-containing first-line regimens has also been documented in a larger prospective study in South Africa. In that study 30% of patients had to switch from stavudine-based to non-stavudine-based regimens within 3 years.

For good reasons, tenofovir has become the gold standard for today’s first-line antiretroviral therapy. Its introduction in developing countries is an important step towards bringing treatment in poor countries in line with rich ones. As the WHO and all countries are phasing out stavudine, this study will send a confusing message, and it may slow down this transition while countries wait for the results.

*A randomised, double-blind study to demonstrate non-inferiority of stavudine (20 mg BID) compared with tenofovir (300 mg QD) co-administered with lamivudine and efavirenz in antiretroviral-naïve patients over 96 weeks. If funded and approved, the trial is anticipated to start in 2012.
There is no prospect that stavudine 20 mg is a better option than tenofovir. The stavudine parallel track programme, in which over 10 000 patients were randomised to receive 40 (30) mg or 20 (15) mg between October 1992 and February 1994, showed a higher incidence of neuropathy in the high-dose arm (21%). Nonetheless, the incidence of neuropathy observed in the lower-dose arm was also unacceptably high (15%).

Of particular importance in low- to middle-income countries – where tuberculosis (TB) is prevalent among HIV-positive people, who are also receiving stavudine-containing regimens – is a South African study looked at the risk of stavudine substitution for toxicities in 7 066 patients receiving ongoing TB treatment at ART initiation; concurrent initiation of TB treatment and ART and incident TB treatment after ART initiation. The study found people receiving ongoing and concurrent TB treatment to be at increased risk of toxicity leading to stavudine substitution, irrespective of stavudine dose (30 and 40 mg). For ongoing TB treatment, adjusted hazard ratio (aHR) was 3.18 (95% confidence interval (CI) 1.82 – 5.56) in the first 2 months of ART; for concurrent TB treatment, aHR was 6.60 (95% CI 3.03 - 14.37) in the first 2 months of ART.

The stavudine 20 mg study is not being proposed in any developed country. Instead it is planned to include only middle- and lower-income developing countries. Patients enrolling in this trial risk being randomised to receive treatment that may be less effective and is more toxic than the current standard of care. There is therefore no good reason why a properly informed patient should want to enrol in this study.

2. The poor tolerability of stavudine limits therapeutic durability. A person has the best chance of successful treatment with their first-line regimen, making it critical that the medicines are as tolerable as possible. A tolerable first-line regimen enhances therapeutic durability by helping people adhere to treatment, and delays their need to switch to more costly second-line regimens, which are complicated for patients, for health workers and from an operational standpoint.

3. Stavudine’s side-effects cut into stavudine’s savings on cost. A study published by MSF shows that inpatient care and essential drug costs were higher for people on stavudine than for those on tenofovir in a cohort in rural Lesotho. According to MSF’s cost-effectiveness study of switching from stavudine or zidovudine- to tenofovir-based first-line regimens in Lesotho, the tenofovir-containing regimen generated higher life years and quality-adjusted life years than zidovudine or stavudine-based treatment. As the costs of tenofovir and especially efavirenz drop, the cost benefit to patients and to health systems will become clearer. Since the study was completed, the global best price of efavirenz – which partly drives tenofovir costs – has almost halved (US$97 per patient year in 2009 to $52 today).

4. Stavudine can compromise second-line options. When someone does fail their first-line regimen, the longer they remain on stavudine – which is likely in a context with limited access to viral load monitoring – the more their second-line options are compromised. Unlike stavudine, tenofovir does not confer thymidine analogue mutations (TAMs); people taking tenofovir can stay on a failing regimen much longer without compromising efficacy of zidovudine and therefore second-line therapy.

5. Stavudine’s long-term toxicity question will not be answered by this trial. The proposed 20 mg stavudine dose might be acceptable in a short-term 48- or even 96-week virological endpoint study (although Bristol-Myers Squibb studied and rejected 20 mg bd). But, because mitochondrial toxicity is both dose and time dependent, many of stavudine’s most serious side-effects (such as peripheral neuropathy and lipoatrophy) would not necessarily emerge until after such a study was completed. This study does not include monitoring of surrogate markers for mitochondrial toxicity, so it cannot shed light on the incidence of this serious adverse event.

Recently published longer-term Cambodian data on rates of severe stavudine-associated toxicity show 7% of people to have neuropathy within the first year; by year 3 the cumulative incidence was 16.6%, and it rose to 19.0% at year 6. The cumulative incidence of lipoatrophy was 56% by year 3 and 72% by year 6. Stavudine use significantly increased the risk for lactic acidosis among people on concurrent TB treatment; the aHR was 8.6 (95% CI 2.7 - 27.5).

The investigators have agreed that this important question about longer-term toxicity will not be answered in the trial, raising the serious issue that the trial will not be able to answer the primary policy question which drives it – whether long-term 20 mg stavudine twice daily is as good as once-daily tenofovir in first-line ART regimens for use in public health programmes in resource-limited settings.

6. Stavudine must be taken twice a day, compared with tenofovir’s once-daily dosing. A twice-daily dosing regimen (as with stavudine 20 mg) does not have the simplicity of a once-daily fixed-dose combination (as with tenofovir). People are more likely to adhere to simpler regimens and therefore are more likely to have better treatment outcomes, as well as stave off resistance that requires more complex and expensive second-line regimens.

7. A tenofovir-based regimen is recommended for HIV/hepatitis B (HBV) coinfection, because stavudine has no activity against HBV and resistance to lamivudine is inevitable. While HIV/HBV co-infection is an exclusion criterion for this trial, it may encourage persistent use of a suboptimal regimen for HIV/HBV co-infected people. Screening for HBV is not routinely performed before initiation of ART in most resource-limited settings, yet HBV is endemic. For example, in South Africa an estimated 5% of HIV-positive people are HBV co-infected (Dr Mark Sondgerup, personal communication). Giving a stavudine/lamivudine-based regimen to HIV/HBV co-infected people will create lamivudine-resistant HBV in this population (90% at 4 years). Continuing lamivudine in the context of HBV drug resistance may lead to hepatitis flares; these flares can cause serious liver damage, and are potentially life-threatening. Researchers are also concerned about the transmission of drug-resistant HBV that may not be preventable by currently available HBV vaccines, a potential public health catastrophe.

8. Stavudine-related cost savings may become irrelevant by the trial’s end. The rationale for this trial is to lower treatment costs, as stavudine is currently cheaper than its alternatives. However, the price of alternatives, notably tenofovir, has come down dramatically in the last several years, and is expected to decrease further as demand increases. According to MSF’s annual ARV pricing report, tenofovir is now cheaper than zidovudine, with...
the price of single-drug tenofovir having decreased by 52% from 2008 to 2011, and the price of the triple fixed-dose combination of tenofovir, lamivudine and efavirenz having decreased by 53% to US$173 per person per year over that same time period. The price of the double FDC TDF/3TC co-packaged with EFV is $143 per person per year. Furthermore, the Clinton Health Access Initiative (CHAI) is currently working on the in the reformulation of tenofovir, with the goal of increasing bioavailability, hence reducing the required active pharmaceutical ingredient and in turn the cost.

Because the stavudine 20 mg 96-week efficacy trial is expected to be completed at the earliest by 2014 - 2015, and would need to be followed by a larger, longer (perhaps 5-year) field effectiveness trial to determine longer-term tolerability, the drug may not be available for use at the new dose until possibly even 2020. It is therefore likely to take 9 years from now for there to be enough evidence that 20 mg stavudine is safe and non-inferior to tenofovir, and could be used to replace tenofovir in first-line regimens.

If current price trends continue, it is likely the anticipated cost savings associated with stavudine could be overtaken by expected further price reductions for tenofovir and other components of the first-line regimen, by the time stavudine 20 mg would be ready for use. It is worth noting that a three-drug one-pill-once-a-day regimen containing efavirenz and tenofovir is now priced at roughly half what stavudine-based Triomune cost when it was first introduced a decade ago.

Further, even greater potential savings could be achieved if the tenofovir prodrug GS 7340, now in phase II by Gilead Sciences, is approved at a low milligram dose. Results will be available within a similar time frame to those from the 96-week stavudine 20 mg trial. A recent announcement by Gilead of an agreement with Tibotec to develop a fixed-dose combination of darunavir, emtricitabine, GS 7340 and cobicistat with ‘less than one tenth of the amount of the 300 mg of tenofovir disoproxil fumarate contained in Viread and Truvada’ suggests that this is feasible.11 Chimerix Inc. also has a promising tenofovir pro-drug in development, CMX-157.

Other drugs in late-stage development such as the integrase inhibitor dolutegravir (50 mg once daily) also offer potential savings on manufacturing and could end up being cheaper than stavudine 20 mg by the time it would become available.

9. Stavudine has low acceptability in the community. Finally, and most importantly, the continued use of stavudine and the proposed trial has raised opposition from people directly affected by its continued use. As an example, the Malawi Network of People Living with HIV/AIDS (MANET+) recently held a press briefing, as the slow pace for phasing out current use of this drug in their country concerns them. Despite the funding crisis, the Malawi government has a priority for this to be completed by June 2012.12

It is unclear why the Bill and Melinda Gates Foundation—who are in discussion with the investigators about funding the proposal—consider this study to be a priority. It seems an aberration in an otherwise carefully considered strategy for supporting research into the optimisation of ART for resource-limited settings. This includes the ENCORE 1 study of low-dose efavirenz, the reformulation of tenofovir to increase its bioavailability (working with CHAI), and the development of innovative potentially long-acting formulations.

For the reasons outlined above, research and the resources it requires, as well as activist pressure, should focus on increasing access to safer cost-saving alternatives to stavudine, not on seeking a comeback for a drug virtually abandoned in wealthy countries.

Please address all correspondence to: sharonann.lynnch@msf.org polly.clayden@i-base.org.uk markgh@gmail.com

REFERENCES
The debate around relooking at stavudine dosing, both in terms of the adult low-dose stavudine study and more broadly, is welcome. The study being proposed to evaluate low-dose stavudine v. tenofovir is a fairly standard placebo-controlled non-inferiority study. The study design is not controversial; however, the choice of study drug has attracted critical attention. We have previously discussed the issue of stavudine use, cost and access, and the significant implications of stavudine in developing countries, in detail in a recent article.1 We continue to believe that low-dose stavudine clinical trials, in both adults and children, are a priority for developing countries. These studies are being proposed simply because tenofovir is very expensive, and the only available cheaper alternative is stavudine. With recession, uncertainty about donor commitments, the compromise of several treatment projects by lack of money, and the plain fact that many developing countries are unable to take ownership of the antiretroviral programme within their budgets, this and similar studies are urgent and necessary for our region to continue to expand access to antiretroviral therapy.

Stavudine was chosen because it is very cheap, easily co-formulated, very well tolerated initially, and requires no laboratory toxicity monitoring in routine clinical practice. Tenofovir was chosen as the comparator for the trial as it is currently the ‘gold standard’ in many guidelines, and the alternatives (zidovudine (AZT) and abacavir) are now more expensive than tenofovir. Abacavir would be regarded as the equivalent gold standard for paediatric care.

As clinicians involved with antiretroviral (ARV) programmes for the last few years, and having used stavudine in large numbers of both adults and children in the past decade, we identify with the instinctive discomfit of many critics. However, we feel that there is a very strong case for studying stavudine further, and many of the arguments against further trials are not consistent with how we study other antiretrovirals, or indeed medication efficacy in general. Tenofovir certainly has benefits over stavudine: it is dosed daily, and has anti-hepatitis B activity; counter-arguments could be that hepatitis B is easily screened for, and that renal toxicity increasingly recognised with tenofovir has potentially serious consequences in developing countries, where monitoring and access to renal care, dialysis and transplantation are almost non-existent.2 These arguments, however, are likely to be irrelevant when the cost of medication is considered. Broadly, other more substantial arguments fall into the following categories.

**STAVUDINE TOXICITY**

This is the most serious disagreement. Critics maintain that stavudine is too toxic to be used, that it is not possible to monitor toxicity safely, and that the trial duration is only 2 years, hence limiting the usefulness of the data.

Toxicity concerns are exclusively based on data around high dosages of stavudine, largely the (now historic) adult 40 mg twice a day (bd) dosage, for which there are extensive toxicity data, including from our own centres. More limited data suggest that 30 mg bd is significantly safer, at least in the medium term, but still carries significant toxicity. We do not contest this – stavudine causes often irreversible and stigmatising lipo-atrophy, as well as peripheral neuropathy and other mitochondrial toxicity in adults. However, all drugs are toxic in sufficient dosages. AZT was originally dosed at far higher levels than currently; toxicity forced the testing of lower doses that were far better tolerated with equal virological suppression levels, and AZT became the standard of first-line care for almost a decade. However, AZT still has some toxicity at this lower dose, and even lower doses are being tried.3-9

The original dose-finding studies of stavudine (d4T) were a complex affair, and the originator company did not complete what we consider to have been the natural next phase of study, largely as it was clear that the significant investment on another large clinical trial was unlikely to be recouped, as well as a probable internal assessment that the drug had a significant public relations problem related to lipo-atrophy in developed countries, where profits are made. There is some evidence that dosing at 20 mg bd is safer and gives equivalent virological efficacy, but these data have not been tested in a rigorous manner.5

This study plans to repeat those done for drugs such as AZT and many others – optimising the dosage of stavudine so as to minimise toxicity, while maintaining virological efficacy. It also responds to observer calls for more research on stavudine using high-quality trials.5,10

Critics have maintained that stavudine toxicity is impossible to monitor safely. Safety and toxicity monitoring in the proposed study has been extensively examined by many experts, and we are confident that we can pursue the study safely. Ethics committees are currently examining the proposed toxicity monitoring schedules.
The doctors involved in the study have plenty of experience in the early recognition of stavudine toxicity.

We agree with critics that the study will not give us evidence on long-term toxicity. Our study is currently funded to run for 2 years, like most registration ARV studies. For many countries, providing 2 years of safe therapy at reduced cost is of significant importance. However, we have extensive lipoatrophy pre-clinical toxicity monitoring built into the study; if no additional toxicity in the stavudine arm using DEXA scans is demonstrated, it is plausible that the drug could be used for longer, and we therefore intend motivating to extend the study.

Finally, it is contended that this study would never be run in a developed country. We see absolutely no scientific and ethical reasons why not. However, the urgency remains in our region. Rationing has already begun in many programmes, and universal access, with the attendant benefits of HIV prevention suggested by the HPTN 052 trial, is unlikely to be more than a public health dream unless we lower the cost of safe treatment. Even developed countries have had to make rationing decisions, cf. the recent decision to restrict access to first-line tenofovir in the UK; rationing was made easier by having good data to base these decisions on.

TENOFOVIR COST

It is asserted that the price of tenofovir is still dropping and will approach that of stavudine. We have consulted extensively over this with the Clinton Health Access Initiative (CHAI) as well as generic companies, and it seems that price equivalence will not be possible, simply because the daily milligram dosage (a daily 300 mg v. the proposed 40 mg) is so different, as raw chemicals drive the generic manufacturing costs. Furthermore, the incremental drops in the price of tenofovir (the latest announced by the CHAI in January 2012) are unlikely to be as significant as previous ones, as manufacturing efficiencies have largely been realised.

NEW DRUG AVAILABILITY

The next assertion is that new drugs are in the offing and will be available by the time this study has results, making the results irrelevant. We believe that confidence in a new drug that will cost-effectively and timely replace tenofovir is a huge act of faith. Many new medications are indeed being tested, and a small number may show efficacy when this study is completed. However, most drugs fail, even in phase 3 studies, so even this is uncertain. Also, the drugs may not be tested with backbones conventionally used in our context (e.g. raltegravir may be used instead of efavirenz), which may limit agreement on whether we can use the drug safely with available backbone drugs.

Furthermore, the registration process, local regulatory approval, negotiations with generic manufacturers and acceptance into national guidelines mean that it takes many years to go from clinical trial success to broad availability. Tenofovir took over 5 years for registration in South Africa. Prolonged registration is the rule rather than the exception in many developing countries. We may need several more years for adequate costing, price reductions and agreement on priorities for access to this treatment.

We believe that it is responsible to study alternatives to tenofovir and other expensive first-line medications.

OTHER ISSUES

In paediatric care, abacavir is the current preferred first-line drug owing to concerns about tenofovir toxicity and also to to preserve AZT for a second-line NRTI backbone. Abacavir is very expensive. Stavudine remains the most widely used ARV for HIV-infected children in sub-Saharan Africa. Apart from cost, abacavir has other real problems. Although the incidence of abacavir hypersensitivity reactions is probably low, there is no confirmatory test widely available in the public sector. The HLA-B5701 gene test is simply not available outside large tertiary centres, and the gene appears to be a largely Caucasian one anyway. Nonspecific fever and rash are common in childhood, especially during immune reconstitution, and many children are likely to receive the label ‘possible previous abacavir hypersensitivity reaction’, which eliminates the option of ever re-introducing abacavir in their regimen. Once abacavir is eliminated, the remaining options in sub-Saharan Africa are zidovudine or stavudine. The danger of zidovudine-related bone marrow suppression (a common problem with varying degrees of severity) is significant and requires some laboratory monitoring. Stavudine offers almost toxicity-free short-term efficacy. For unknown reasons, thymidine-related peripheral neuropathy has not been documented in pre-pubertal children, and symptomatic lactic acidosis appears remarkably less common than in adults. The incidence of lipo-atrophy is concerning but is strongly dose-related. We think that its frequency and severity will be significantly reduced with the use of low-dose stavudine. In addition, stigmatising lipo-atrophy is avoidable if reasonable awareness is maintained and appropriate drug switches made before lipo-atrophy becomes obvious.11,12 An unrecognised benefit of using children’s regimens similar to adults’ is that it makes children less susceptible to supply-line problems, a huge problem in developing countries. Getting the dose of stavudine right in children is as compelling as it is in adults, and these studies should be prioritised by funders.

In the end, we desperately need alternatives to tenofovir for adults and to abacavir in children in poorer countries. The only current alternative is conventionally dosed stavudine, as AZT and abacavir are more expensive than tenofovir. A minister of health or donor faced with the decision to treat two people with a moderately toxic drug or one with a relatively safe regimen, with the other person definitely dying of AIDS, faces very little choice. Making stavudine safer is an urgent public health issue. We think that doing it safely, efficiently and ethically is possible and should be everyone’s priority.

REFERENCES

Antiretroviral treatment (ART) is a powerful tool for reducing both AIDS mortality and HIV transmission. The monitoring of access to ART is therefore critical to the evaluation of the impact of HIV treatment and prevention programmes. Previous monitoring exercises have shown that, since the announcement of a comprehensive care, management and treatment programme by the South African Department of Health in late 2003, access to ART in South Africa has increased dramatically. These assessments suggested that South Africa was on track to meet the targets laid out in the 2007 - 2011 National Strategic Plan (NSP) for HIV/AIDS and Sexually Transmitted Infections, which aimed to achieve new ART enrolment numbers equal to 80% of the number of newly eligible individuals in each year, by 2011. However, there has not as yet been any formal assessment of whether this target has been met.

The monitoring of access to ART in South Africa is challenging for several reasons. The interpretation of public sector statistics is complicated by changes in reporting practices in late 2009, with most provinces switching from reporting numbers of patients cumulatively started on ART to numbers of patients currently on ART. Statistics from disease management programmes and programmes run by non-governmental organizations (NGOs) have not been routinely collected and reported. In addition, there is generally a lack of information on the age and sex of patients. This is particularly problematic in view of concerns that ART initiation rates may be lower among men than women.

The estimation of ART coverage is also hampered by uncertainty regarding the ‘treatment need’, the denominator in the coverage calculation. Mathematical models have been used to estimate numbers of HIV-positive individuals with CD4 counts below different thresholds, but there is substantial uncertainty surrounding the rates of CD4 decline that are assumed in these models, and there is also growing recognition that these rates of CD4 decline may differ between populations. There is also concern that cross-sectional measures of ART coverage may fail to give a sense of recent programme performance, which is better reflected in the ratio of the number of patients starting ART in a year to the number of individuals becoming eligible for ART in the same year. The latter measure has the advantage of being consistent with the way in which the South African NSP targets are expressed, and is also less sensitive to model assumptions about rates of CD4 decline and ART eligibility criteria.

The objective of this paper is to assess recent changes in access to ART in South Africa, and to evaluate the extent to which the 2007 - 2011 NSP treatment targets have been met. This study also aims to improve on previous work by including more recent programme statistics, by using locally relevant CD4 data in the estimation of the treatment need, by including 95% confidence intervals (CIs) in coverage estimates, and by estimating coverage separately for men, women and children.
METHODS

ART PROGRAMME STATISTICS

Public sector ART programme statistics to the end of June 2011 were obtained from the South African Department of Health, and were adjusted to achieve consistency of definition (cumulative/current), using a previously described formula,\(^4\) for each province. Unpublished data on the sex ratio of adult patients enrolled in public ART programmes in four provinces, collected up to March 2009, were used to estimate the sex ratio of adults starting ART in the public sector.

Private sector data and data from NGOs were obtained through surveys conducted every two years, since mid-2006.\(^5\) Linear interpolation and extrapolation was used to estimate numbers for programmes with missing data and for years in which no survey was conducted. Estimates of the proportion of private sector patients who were men, women and children were obtained from submissions by medical schemes to the Risk Equalization Fund up to March 2008, and the geographical distribution of private sector patients was estimated from early private sector statistics.\(^6\)

Detailed data collected from NGO programmes in the 2008 survey were used to determine the profile of NGO patients by age, sex and province.

MATHMATICAL MODEL

To estimate the numbers of adults needing ART, a mathematical model was developed to simulate the growth of the South African population over time, the incidence of HIV and the decline in CD4 counts in HIV-positive adults. The model stratifies the population by age and sex, and projects the change in population in one-year intervals, starting in the middle of 1985. Assumptions regarding the age- and sex-specific population profile, non-HIV mortality, fertility, migration and HIV incidence are based on the ASSA2008 AIDS and Demographic model.\(^7\)

Once infected, individuals are assumed to progress through a four-stage model of CD4 decline, in the absence of ART (Fig. 1). Individuals are assumed to experience AIDS mortality in the CD4 200 - 349/µl category at a fraction \(κ\) of the AIDS mortality rate in the CD4 200 - 500/µl category, if untreated. Up to mid-2009, adults of sex \(g\) are assumed to start ART only once their CD4 count has dropped below 200/µl, at a rate of \(r_g(t)\) per annum in year \(t\). Between mid-2009 and mid-2011, the model also allows individuals to start ART in the CD4 200 - 349 category if they develop tuberculosis or become pregnant, following the change in South African ART guidelines in early 2010.\(^8\) The \(r_g(t)\) rates in each year are calculated from the ART programme statistics (further detail is provided in the online appendix).

The posterior estimates of the model parameters are summarised in Table 1, and posterior estimates of numbers of patients receiving ART are summarised in Table 2. Over the period mid-2004 to mid-2011, the total number of patients receiving ART in South Africa increased from 47 500 (95% CI 42 900 – 51 800) to 1.79 million (95% CI 1.65 – 1.93 million). Of the latter, 85% were receiving ART through the public health sector, 11% were receiving ART through disease management programmes in the private sector, and the remaining 4% were receiving ART through community treatment programmes run by NGOs. The majority (61%) of patients were women aged 15 or older, men accounted for 31% of patients, and children under the age of 15 comprised the remaining 8% of patients. KwaZulu-Natal and Gauteng were the two provinces with the largest numbers of patients, together accounting for 56% of all patients receiving ART.

Changes over time in numbers of treated and untreated adults in different CD4 stages are shown in Fig. 2. As at mid-2011, untreated HIV-positive adults included 58 000 (95% CI 13 000 – 147 000) individuals who had stopped ART, 385 000 (95% CI 247 000 – 634 000) ART-naive adults with CD4 <200/µl, 1.06 million (95% CI 0.88 – 1.29 million) with CD4 counts of 200 – 349/µl, 0.74 million (95% CI 0.61 – 0.91 million) with CD4 counts of 350 – 500/µl, and 0.94 million (95% CI 0.77 – 1.16 million) with CD4 counts >500/µl. The estimates of annual numbers of new paediatric HIV infections were obtained from a separate model of paediatric HIV in South Africa.\(^14\) Since paediatric ART guidelines recommend ART initiation in all HIV-infected children aged <12 months, regardless of their immunological or clinical status,\(^17\) the annual number of new paediatric HIV infections is used to approximate the annual number of children newly eligible for ART (the denominator in the ART enrolment ratio).

CALIBRATION AND UNCERTAINTY ANALYSIS

The parameters determining the rates of CD4 decline, HIV-related mortality and ART discontinuation are estimated by fitting the model to HIV prevalence data from the 2005 and 2008 Human Sciences Research Council (HSRC) household surveys,\(^9\)\(^10\)\(^9\) and to CD4 data from HIV-positive adults in three South African surveys,\(^10\)\(^22\)\(^23\) using a Bayesian melding procedure.\(^11\)\(^24\) A detailed explanation is provided in the online appendix. Briefly, prior distributions are specified to represent uncertainty regarding the parameters of interest, including the range of plausible values for the average time to starting ART after becoming eligible (1/r_g(t)). Prior distributions are also specified to represent uncertainty regarding the accuracy of the reported ART programme statistics in each year. This uncertainty and the uncertainty regarding ART attrition rates affect the model ART enrolment inputs. A likelihood function is specified to represent how well the model fits the CD4 data and HIV prevalence data, for a given set of parameter values. The posterior distribution, representing the parameter combinations from the prior distributions that have the highest likelihood values, is then simulated by Sampling Importance Resampling.\(^12\)

RESULTS

The posterior estimates of the model parameters are summarised in Table 1, and posterior estimates of numbers of patients receiving ART are summarised in Table 2. Over the period mid-2004 to mid-2011, the total number of patients receiving ART in South Africa increased from 47 500 (95% CI 42 900 – 51 800) to 1.79 million (95% CI 1.65 – 1.93 million). Of the latter, 85% were receiving ART through the public health sector, 11% were receiving ART through disease management programmes in the private sector, and the remaining 4% were receiving ART through community treatment programmes run by NGOs. The majority (61%) of patients were women aged 15 or older, men accounted for 31% of patients, and children under the age of 15 comprised the remaining 8% of patients. KwaZulu-Natal and Gauteng were the two provinces with the largest numbers of patients, together accounting for 56% of all patients receiving ART.

Changes over time in numbers of treated and untreated adults in different CD4 stages are shown in Fig. 2. As at mid-2011, untreated HIV-positive adults included 58 000 (95% CI 13 000 – 147 000) individuals who had stopped ART, 385 000 (95% CI 247 000 – 634 000) ART-naive adults with CD4 <200/µl, 1.06 million (95% CI 0.88 – 1.29 million) with CD4 counts of 200 – 349/µl, 0.74 million (95% CI 0.61 – 0.91 million) with CD4 counts of 350 – 500/µl, and 0.94 million (95% CI 0.77 – 1.16 million) with CD4 counts >500/µl. The
total unmet need in the middle of 2011 (ART-naïve adults with CD4 <350/μl plus all adults who had stopped ART) was 1.50 million (95% CI 1.24 - 1.84 million), which is 32% lower than the total unmet need four years previously. Estimates of adult ART coverage and ART enrolment ratios are shown in Fig. 3. Using previous CD4 thresholds for defining ART eligibility (CD4 <200/μl), the fraction of adults eligible to receive ART who were actually on ART increased from 5.1% (95% CI 4.2 - 6.1%) in the middle of 2004 to 79% (95% CI 70 - 85%) by the middle of 2011. However, using the new CD4 thresholds for defining ART eligibility (CD4 <350/μl), adult ART coverage by the middle of 2011 was 52% (95% CI 46 - 57%).

As noted previously, ART enrolment ratios are similar when using different CD4 thresholds to define ART eligibility. For example,

<table>
<thead>
<tr>
<th>TABLE 1. POSTERIOR ESTIMATES OF MODEL PARAMETERS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symbol</td>
</tr>
<tr>
<td>------------------------------------------------</td>
</tr>
<tr>
<td>Parameters for untreated adults</td>
</tr>
<tr>
<td>Annual rate of progression from CD4 &gt;500 to 350 - 500</td>
</tr>
<tr>
<td>Annual rate of progression from CD4 350 - 500 to 200 - 349</td>
</tr>
<tr>
<td>Annual rate of progression from CD4 200 - 349 to &lt;200</td>
</tr>
<tr>
<td>Annual rate of HIV mortality if CD4 &lt;200</td>
</tr>
<tr>
<td>Ratio of HIV mortality at CD4 200 - 349 to HIV mortality at CD4 &lt;200</td>
</tr>
<tr>
<td>Parameters for treated adults</td>
</tr>
<tr>
<td>Probability of permanent loss to care in first 6 months after ART start</td>
</tr>
<tr>
<td>Annual probability of permanent loss to care after first 6 months of ART</td>
</tr>
<tr>
<td>Proportion of permanent loss to care that is due to death</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TABLE 2. NUMBERS OF PATIENTS RECEIVING ART IN SOUTH AFRICA</th>
</tr>
</thead>
<tbody>
<tr>
<td>-------------------------------------------------------------</td>
</tr>
<tr>
<td>-------------------------------------------------------------</td>
</tr>
<tr>
<td>Currently on ART*</td>
</tr>
<tr>
<td>By sex/age</td>
</tr>
<tr>
<td>Men</td>
</tr>
<tr>
<td>Women</td>
</tr>
<tr>
<td>Children (&lt;15)</td>
</tr>
<tr>
<td>By provider</td>
</tr>
<tr>
<td>Public sector</td>
</tr>
<tr>
<td>Private sector</td>
</tr>
<tr>
<td>NGO programmes</td>
</tr>
<tr>
<td>By province</td>
</tr>
<tr>
<td>Eastern Cape</td>
</tr>
<tr>
<td>Free State</td>
</tr>
<tr>
<td>Gauteng</td>
</tr>
<tr>
<td>KwaZulu-Natal</td>
</tr>
<tr>
<td>Limpopo</td>
</tr>
<tr>
<td>Mpumalanga</td>
</tr>
<tr>
<td>Northern Cape</td>
</tr>
<tr>
<td>North West</td>
</tr>
<tr>
<td>Western Cape</td>
</tr>
<tr>
<td>Started ART last year †</td>
</tr>
<tr>
<td>Men</td>
</tr>
<tr>
<td>Women</td>
</tr>
<tr>
<td>Children (&lt;15)</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

All numbers are rounded to the nearest 1000 (except in the case of 2004 and 2005 totals, which are rounded to the nearest 100). Due to rounding, some rows may not sum to the total. All estimates are posterior averages (95% confidence intervals not shown).

* Totals reflect numbers at the middle of each year.
† Totals reflect ART enrolment over the 12 months up to the middle of the year.
over the period from mid-2010 to mid-2011, the ratio of the number of adults starting ART to the number of adults whose CD4 counts fell below the CD4 threshold was 1.64 (95% CI 1.11 - 2.10) when the CD4 threshold was 200, and 1.56 (95% CI 1.08 - 1.97) when the CD4 threshold was 350. Both ratios are roughly double the target of 80% set in the 2007 - 2011 NSP, and indicate substantial progress in removing the ‘backlog’ of unmet need that accumulated in previous years.

Estimates of ART access are presented separately for men, women and children in Fig. 4. Using the CD4 threshold of 350/μl as the criterion for ART eligibility, the fraction of ART-eligible women who were receiving ART by the middle of 2011 (60%, 95% CI 53 - 65%) was significantly higher than the fraction of ART-eligible men who were on treatment (41%, 95% CI 36 - 46%). A similar difference in magnitude is seen in the ART enrolment ratio over the period mid-2010 to mid-2011: using the same ART eligibility criterion of CD4 <350/μl, the enrolment ratio was 1.96 (95% CI 1.33 - 2.51) in women and 1.23 (95% CI 0.83 - 1.58) in men. Over the same period, the ratio of the number of children starting ART to the number of new infections in children was 1.13 (95% CI 0.74 - 1.48). In most previous years, this ratio was below both the male ART enrolment ratio and the female ART enrolment ratio.

**DISCUSSION**

South Africa has made impressive progress in the rollout of ART since the start of the public sector ART programme in 2004. The number of patients who started ART in 2010/2011 was well in excess of the number of individuals who became eligible to receive ART over the same period, exceeding the targets set in the 2007 - 2011 NSP. The unmet need for ART was also reduced by 32% between 2007 and 2011. According to the ART initiation criteria that were in place at the time, adult treatment coverage by mid-2011 was close to 80%.

However, there appear to be substantial differences between men, women and children in the rate of ART initiation. The low rate of ART initiation in men relative to women may be a reflection of gender differences in health-seeking behaviour and perceptions that men who seek care are ‘weak’. Alternatively, the high rate of ART initiation in women may be due to higher rates of HIV diagnosis through antenatal screening. The relatively low rates of ART initiation in children are probably attributable to the lower rates of HIV testing in children and the greater complexity of paediatric ART relative to adult ART. However, it is difficult to compare adult and paediatric measures of ART access meaningfully because the course of HIV infection is so different in children, with many HIV-infected infants dying in the first few months of life before there is an opportunity for testing.

This analysis extends previous work by including assessment of uncertainty and by incorporating several new data sources. The 95% CIs that have been estimated reflect uncertainty regarding rates of CD4 decline, rates of mortality and rates of ART retention, and also reflect uncertainty regarding the accuracy of reported ART programme statistics. However, the CIs do not reflect the uncertainty regarding the HIV incidence rates that have been estimated from the ASSA2008 model, and this may lead to some exaggeration of precision. CIs around the ART enrolment ratios are considerably wider in 2009/10 and 2010/11 than in previous years, owing to the change in the way that the Department of Health has reported public sector ART programme statistics.

Various attempts were made to validate the reported ART programme statistics using data from external sources, with limited success. Lamivudine sales figures from Aspen Pharmacare, which until recently supplied 80% of lamivudine in the public sector, were used to obtain crude estimates of numbers of public sector patients on treatment in each quarter. These estimates were
The new NSP for the 2012 - 2016 period proposes several measures to strengthen the monitoring and evaluation of South Africa’s ART programme, including the introduction of a single patient identifier in the health sector and a single registry at the primary care level. It is hoped that these measures will lead to greater precision in the estimation of ART coverage in future, as well as a deeper understanding of the factors determining access to care and retention in care.

Fig. 4. Age and sex differences in ART access. Bars represent posterior means and error bars represent 95% confidence intervals. Dashed line in panel (b) represents 2007 - 2011 National Strategic Plan Target.

Many challenges exist, both in achieving future ART rollout targets and in monitoring future progress towards meeting these targets. The new NSP for the 2012 - 2016 period proposes targets that are much more ambitious than those in the previous NSP: the ART enrolment target in 2016 is 80% of the new ART need in that year plus 80% of the unmet need from previous years. High levels of HIV testing and counselling, as well as expansion of capacity to deliver ART, will be required to meet these targets. The

not significantly different from the model estimates in Table 2 up to the end of 2008, and from October 2009 to March 2010, but were substantially lower than the model estimates from January to September of 2009. Numbers of viral load tests performed by the National Health Laboratory Service for public sector clinics were also used to obtain theoretical estimates of numbers of patients receiving ART, on the assumption that patients went for viral load testing twice per annum on average. The resulting estimates were slightly higher than the corresponding model estimates up to 2008, but were 18% lower than the model estimates in 2009. Finally, the model estimate of the fraction of the 15 - 49-year-old population on ART in the middle of 2008 was compared with the corresponding proportion estimated in the 2008 HSRC national household survey,23 based on testing for the presence of antiretroviral drugs in blood samples: the model estimate of 1.8% (95% CI 1.6 - 2.0%) was found to be significantly lower than that measured in the survey (3.0%). External data sources therefore do not provide a clear and consistent assessment of the plausibility of the model estimates derived from reported ART programme statistics.

Although attempts were made to produce estimates of ART coverage for each province, it was not possible to produce plausible results for two provinces (Gauteng and Western Cape) because the estimated numbers of patients starting ART in recent years exceeded the estimated numbers of patients eligible to receive ART, in both of these provinces. This could possibly be due to individuals with advanced HIV migrating to urban areas because of the perceived superiority of health services in the major urban centres of Gauteng and Western Cape. The model assumes migration to be independent of HIV status, and may therefore underestimate the number of HIV-infected ART-eligible individuals who migrate into these provinces. Alternatively, the problems experienced in producing plausible results for Gauteng and Western Cape may be due to assumed HIV incidence rates in these provinces being too low, or reported numbers of ART patients in these provinces being exaggerated.

Many challenges exist, both in achieving future ART rollout targets and in monitoring future progress towards meeting these targets. The new NSP for the 2012 - 2016 period proposes targets that are much more ambitious than those in the previous NSP: the ART enrolment target in 2016 is 80% of the new ART need in that year plus 80% of the unmet need from previous years. High levels of HIV testing and counselling, as well as expansion of capacity to deliver ART, will be required to meet these targets. The
actuarialsociety.org.za (accessed 5 April 2011).


Appendix

This appendix provides further detail regarding the method used to model adult ART initiation. It also includes a detailed explanation of the Bayesian melding procedure: the prior distributions and the data sources on which they are based, the method used to define the likelihood function and the method used to simulate the posterior distribution.

1. Method used to model adult ART initiation

The model inputs are the numbers of adults of sex g starting ART in year t, \(S_g(t)\). These are used to calculate \(r_g(t)\), the rate of ART initiation in adults of sex g, who are ART-naïve and with CD4 counts <200/μl, during year t. Prior to the change in ART guidelines in 2009, the rate \(r_g(t)\) can be calculated in terms of \(S_g(t)\), if we have the model estimates of the numbers of individuals who are ART-eligible at the start of year t, \(E_g(t)\), as well as the numbers who become eligible over the course of year t, \(P_g(t)\), and the rate of mortality in untreated ART-eligible adults, \(\lambda_4\). If it is assumed that the individuals who progress to CD4 <200/μl do so uniformly over the course of the year, it can be shown that

\[
S_g(t) = E_g(t) \frac{r_g(t)}{r_g(t) + \lambda_4} \left[1 - \exp\left(-r_g(t) - \lambda_4\right)\right] + P_g(t) \frac{r_g(t)}{r_g(t) + \lambda_4} \left[1 - \frac{1}{r_g(t) + \lambda_4} \left[1 - \exp\left(-r_g(t) - \lambda_4\right)\right]\right].
\] (1)

This formula is used to calculate the rate of ART initiation, \(r_g(t)\), using Newton’s method. Following the change in ART guidelines in 2009, a modified version of the above equation is used, to exclude individuals with CD4 counts of 200-349/μl who start ART when they have TB or are pregnant. If \(B_g(x,t)\) is the model estimate of the number of adults aged \(x\), of sex g, who have CD4 counts of 200-349/μl at the start of year t, then the left hand side of the above equation changes to

\[
S_g(t) - \Omega(t) \sum_x B_g(x,t)(f_g(x,t) + u),
\]

where \(\Omega(t)\) is the fraction of HIV-positive TB patients and pregnant women who start ART in year t, if their CD4 count is below 350/μl, \(f_g(x,t)\) is the fertility rate in year t at age \(x\), and \(u\) is the TB incidence rate in HIV-positive individuals with CD4 counts of 200-349/μl. Equation (1) is therefore modified so that the left-hand side includes only those individuals who start ART with a CD4 count of <200/μl. A limitation of this formula is that it considers only the number of adults with CD4 200-349/μl at the start of the year, and does not consider movements into and out of the CD4 200-349/μl state during the course of the year. However, since the movements into and out of the CD4 200-349/μl state offset one another to some extent, the change in \(B_g(x,t)\) over a single year is likely to be modest, and the approximation is therefore reasonable.

The parameters in the above equation are estimated from various sources. The fertility rates by age and year are obtained from the ASSA2008 AIDS and Demographic model, and have
been set to zero in the case of males. The annual TB incidence rate \( u \), in adults with CD4 200-349/μl, is set to 0.043, based on data from HIV patients in two provinces, Gauteng and Mpumlanga.\(^2\)\(^3\) The proportion of TB patients and pregnant women who are assumed to start ART, if their CD4 count is in the CD4 200-349 range, is set to 30% in 2009/10 (since the change in guideline was only announced at the end of 2009 and formal guidelines were only published in early 2010) and to 75% in 2010/11. The latter assumption may be optimistic, as South African studies have generally found that the proportion of ART-eligible pregnant women who initiate ART before delivery is typically 30-60%\(^4\)\(^7\) and only one South African study has found the proportion to be as high as 75%\(^8\).

The same assumptions regarding fertility rates and TB incidence rates are used in allowing for ART initiation in the CD4 200-349/μl category.

2. Prior distributions

Prior distributions are specified for various model parameters to reflect the extent of uncertainty that exists a priori, before calibrating the model to CD4 data and HIV prevalence data. The sections that follow provide the justification for the prior distributions that have been chosen for each parameter. In general, gamma distributions have been chosen to represent uncertainty regarding parameters that can take on any positive value, beta distributions have been chosen to represent uncertainty regarding parameters that can take on any value between 0 and 1, and normal distributions have been chosen to represent uncertainty regarding parameters that can take on any value. The distributions have therefore been chosen according to the range of values that they support, since the gamma, beta and normal distributions are defined on the ranges \([0, \infty)\), \([0, 1]\) and \((\infty, \infty)\) respectively.

(a) Rates of CD4 decline in the absence of ART

The model that is used to describe progression to death in HIV-infected adults, in the absence of ART, is illustrated in Figure 1. All infected adults are assumed to have initial CD4 counts above 500 (based on the CD4 distributions observed in HIV-negative South Africans\(^9\)\(^10\)), which decline over the course of HIV infection. The risk of AIDS mortality is assumed to begin after the CD4 count drops below 350.

![Figure 1: Multi-state model of decline in CD4 count in HIV-infected adults](image-url)
In setting the prior distributions on the rates of CD4 decline (\(\lambda_1\), \(\lambda_2\), and \(\lambda_3\)), we review estimates from other studies that have attempted to estimate rates of CD4 decline in untreated HIV-positive individuals, prior to the availability of ART. Estimates obtained from three studies are summarized in Table 1. All of these studies are based on cohorts in high income countries with HIV-1 subtype B epidemics, and the rates of CD4 decline estimated in these studies might therefore not be appropriate to South Africa. Estimates of \(\lambda_1\) lie between 0.38 and 0.60 (average of 0.44), but these values are likely to be over-estimates of the rate of transition out of the CD4 \(>500\) state. This is because in all three studies the CD4 \(>500\) category was actually separated into three states (CD4 \(>900\), CD4 700-900 and CD4 500-699), and the transition intensities shown in Table 1 are the rates of transition from the CD4 500-699 state to the CD4 350-499 state. (Calculating the rate of transition out of the \(>500\) category would mean increasing the denominator without changing the numerator.) Because of the likely over-statement, we set the prior distribution for the \(\lambda_1\) parameter to be gamma with a mean of 0.3 and a standard deviation of 0.06 (so that the 97.5 percentile of the distribution is 0.43, close to the average of the studies in Table 1).

Estimates of the \(\lambda_2\) parameter lie between 0.30 and 0.51 (average value of 0.38). We choose a gamma distribution with a mean of 0.40 and a standard deviation of 0.08 to represent our prior uncertainty regarding this parameter. The 2.5 and 97.5 percentiles of this distribution are 0.26 and 0.58 respectively, an interval that includes all of the empirical estimates in Table 1. Finally, estimates of \(\lambda_3\) lie between 0.26 and 0.49. However, the estimate of Longini et al \(^{11}\) is likely to be an over-estimate because the model assumes that individuals can only develop AIDS-defining illnesses when their CD4 count is <200, so that individuals who develop AIDS with CD4 \(>200\) are artificially “pushed through” into the CD4 <200 category. If the estimate of Longini et al is excluded, the average estimate of the \(\lambda_3\) parameter is 0.36. To represent the prior uncertainty regarding this parameter we therefore choose a gamma distribution with a mean of 0.35 and a standard deviation of 0.07. The 2.5 and 97.5 percentiles of this distribution are 0.23 and 0.50 respectively.

Table 1: Estimated annual rates of CD4 decline in the absence of ART

<table>
<thead>
<tr>
<th>Study</th>
<th>Population</th>
<th>Rate of CD4 decline from</th>
<th>Rate of CD4 decline from</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>&gt;500 to 350-500 ((\lambda_1))</td>
<td>350-500 to 200-349 ((\lambda_2))</td>
</tr>
<tr>
<td>Satten &amp; Longini (^*)</td>
<td>Men in San Francisco, 1984-92</td>
<td>0.42</td>
<td>0.37</td>
</tr>
<tr>
<td>Hendriks et al (^*)</td>
<td>Men in Amsterdam, 1985-90</td>
<td>0.38</td>
<td>0.34</td>
</tr>
<tr>
<td></td>
<td>Men in Amsterdam, 1990-97</td>
<td>0.39</td>
<td>0.30</td>
</tr>
<tr>
<td></td>
<td>Men in Vancouver, 1985-90</td>
<td>0.40</td>
<td>0.37</td>
</tr>
<tr>
<td></td>
<td>Men in Vancouver, 1990-97</td>
<td>0.45</td>
<td>0.36</td>
</tr>
<tr>
<td>Longini et al (^*)</td>
<td>USA army personnel, 1985-90</td>
<td>0.60</td>
<td>0.51</td>
</tr>
<tr>
<td>Average</td>
<td></td>
<td>0.44</td>
<td>0.38</td>
</tr>
</tbody>
</table>

\(^*\) Model allows for reversible CD4 declines, and the transition intensities shown here are the ‘net’ rates of decline after multiplying the reported transition intensity by the probability of not returning to the initial CD4 state.
(b) Rates of HIV mortality in the absence of ART

In setting the prior distributions to represent the uncertainty regarding the mortality rates in different CD4 categories, we review African studies of mortality in HIV-positive individuals not receiving ART. Table 2 summarizes the evidence from various African studies. In one study,\textsuperscript{14} mortality rates were reported both for patients receiving cotrimoxazole (CTX) and patients not receiving CTX, but only the former is included in Table 2, as the provision of CTX to HIV-positive patients has been the standard of care in South Africa for many years. In seven studies that estimated the annual mortality rate in patients with CD4 below 200/μl, who were not receiving ART, estimates ranged from 0.14 to 0.39 (average value of 0.26). To represent the prior uncertainty regarding the $\lambda_4$ parameter we therefore choose a gamma prior with a mean of 0.25 and a standard deviation of 0.07 (this distribution has 2.5 and 97.5 percentiles of 0.13 and 0.40 respectively).

Estimates of the ratio of mortality in the CD4 200-349/μl category to that in the <200/μl category ($\theta$) range between 0.12 and 0.30 (average value of 0.22). However, these are likely to be over-estimates for two reasons. Firstly, because CD4 testing is not frequent, the CD4 count at the time of death is likely to be lower than the last measured CD4 count; this means that the rate of mortality in the CD4 200-349/μl category is likely to be over-estimated if it is assumed that the CD4 count at death is the same as the same as the last measured CD4 count. Secondly, CD4 measurements are highly variable, with recorded CD4 values fluctuating considerably around a notional ‘true’ value, even in the course of a single day.\textsuperscript{15-17} Since we are modelling only the ‘true’ value and not the fluctuations in recorded values, and since the mortality risk is likely to depend on the cumulative incidence of opportunistic infections over several days or weeks, the strength of association between the mortality risk and the true CD4 is likely to be understated by the observed association between the mortality risk and the measured CD4. A stronger association would imply a lower value of $\theta$. We therefore set the prior distribution for $\theta$ to be beta with a mean of 0.15 and a standard deviation of 0.05 (the 2.5 and 97.5 percentiles of this distribution are 0.07 and 0.26 respectively).

Table 2: Estimated rates of mortality in the absence of ART

<table>
<thead>
<tr>
<th>Study</th>
<th>Population</th>
<th>Frequency of CD4 testing</th>
<th>Rate of mortality at CD4 200-349 ($\theta\lambda_4$)</th>
<th>Rate of mortality at CD4 &lt;200 (λ_4)</th>
<th>Ratio of 200-349 to &lt;200 mortality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Badri et al\textsuperscript{18}</td>
<td>Cape Town</td>
<td>At least 6-monthly</td>
<td>0.08</td>
<td>0.27</td>
<td>0.30</td>
</tr>
<tr>
<td>Hargrove et al\textsuperscript{19}</td>
<td>Zimbabwean women, followed postpartum</td>
<td>Only at baseline</td>
<td>0.0293*</td>
<td>0.1442</td>
<td>0.20</td>
</tr>
<tr>
<td>Coutsooudis et al\textsuperscript{10}</td>
<td>Women in KwaZulu-Natal, followed postpartum</td>
<td>Only at baseline</td>
<td>-</td>
<td>0.1714</td>
<td>-</td>
</tr>
<tr>
<td>Mermin et al\textsuperscript{14}</td>
<td>Ugandan patients, receiving CTX</td>
<td>-</td>
<td>-</td>
<td>0.392</td>
<td>-</td>
</tr>
<tr>
<td>Van Oosterhout et al\textsuperscript{20}</td>
<td>Blantyre, Malawi</td>
<td>4-monthly</td>
<td>0.04</td>
<td>0.34</td>
<td>0.12</td>
</tr>
<tr>
<td>Seyler et al\textsuperscript{21}</td>
<td>Abidjan, Côte d’Ivoire (patients receiving CTX)</td>
<td>6-monthly</td>
<td>0.077</td>
<td>0.316</td>
<td>0.24</td>
</tr>
<tr>
<td>eART-linc\textsuperscript{22}</td>
<td>Patients from Uganda and Côte d’Ivoire</td>
<td>-</td>
<td>-</td>
<td>0.192</td>
<td>-</td>
</tr>
<tr>
<td>Average</td>
<td></td>
<td></td>
<td>0.06</td>
<td>0.26</td>
<td>0.22</td>
</tr>
</tbody>
</table>

* Mortality rate in women with CD4 200-399.
(c) Rate of attrition after ART initiation

The probability of attrition in the first 6 months is calculated as \( \kappa_0 = k_0 R \), and the annual probability of attrition after the first 6 months is calculated as \( \kappa_1 = k_1 R \). In these equations, \( k_0 \) and \( k_1 \) represent the ‘base rates’, and \( R \) is the ratio of the true attrition rates to the base rates. The base parameters \( k_0 \) and \( k_1 \) have been set at 0.105 and 0.065 respectively, based on data from the Western Cape public sector programme. A gamma prior distribution is specified for parameter \( R \), to represent the uncertainty regarding the true attrition rates. This prior distribution has a mean of 1, so that the prior means on parameters \( \kappa_0 \) and \( \kappa_1 \) are the same as the values assumed previously. In order to set the variance of the prior distribution, it is necessary to examine the variability in attrition rates between settings. Table 3 compares cumulative rates of attrition by 12 months after starting ART in various South African programmes. Rates of attrition by 12 months reach as high as 25% and 28% in the studies of Ford et al and Rosen et al respectively, almost double the rates of attrition in the Western Cape as a whole. However, it is possible that the rates of attrition observed in the Western Cape may overstate the true rates of attrition, since almost half of all Western Cape patients not retained in care by 12 months are considered ‘lost to follow-up’, and may therefore still be receiving ART elsewhere. Setting the standard deviation for the gamma prior on \( R \) to 0.4 yields 2.5 and 97.5 percentiles of 0.38 to 1.92 respectively, which adequately reflects the extent of uncertainty around the ratio of the true attrition rate to that recorded in the Western Cape province.

Table 3: Cumulative attrition by 12 months after starting ART in South African patients

<table>
<thead>
<tr>
<th>Study</th>
<th>Location</th>
<th>n</th>
<th>Attrition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boule et al 23</td>
<td>Whole of Western Cape</td>
<td>12587</td>
<td>14.7%</td>
</tr>
<tr>
<td>Bekker et al 25</td>
<td>Gugulethu, Western Cape</td>
<td>1139</td>
<td>9%</td>
</tr>
<tr>
<td>Coetzee et al 26</td>
<td>Khayelitsha, Western Cape</td>
<td>287</td>
<td>13.7%</td>
</tr>
<tr>
<td>Fatti et al 27</td>
<td>Sites in KwaZulu-Natal, Eastern Cape, Western Cape, Mpumalanga</td>
<td>29203</td>
<td>17.9%</td>
</tr>
<tr>
<td>Ford et al 28</td>
<td>Lusikisiki, Eastern Cape</td>
<td>1025</td>
<td>24.9%</td>
</tr>
<tr>
<td>Mutevedzi et al 29</td>
<td>Umkhanyakude, KwaZulu-Natal</td>
<td>5719</td>
<td>16.0%</td>
</tr>
<tr>
<td>Rosen et al 30</td>
<td>Sites in Gauteng, Eastern Cape, Mpumalanga</td>
<td>400</td>
<td>28.0%</td>
</tr>
<tr>
<td>Vella et al 31</td>
<td>Sites in KwaZulu-Natal</td>
<td>2835</td>
<td>19%</td>
</tr>
</tbody>
</table>

It is important to note that in our model, the attrition rates \( \kappa_0 \) and \( \kappa_1 \) are defined as the rates at which patients permanently stop ART, either due to death or due to the patient not collecting their medication. Our model does not allow for increases in rates of attrition over time, although various South African studies have noted marked trends towards greater loss to follow-up over time. The reasons for not allowing for a trend towards increased attrition over time are:

- It has been acknowledged that some of the apparent increase in loss to follow-up may be due to increases in administrative error, i.e. because of the increasingly large volumes of patients, it is difficult to keep accurate records of when patients visit the clinics.
- Some of the apparent increase in loss to follow-up is also due to increasing levels of decentralization in the provision of ART, with patients not informing the clinic at
which they started ART of their move to a nearer clinic. These patients who remain on ART despite being lost to follow-up are not included in our definition of attrition.

- Some of the apparent increase in loss to follow-up is also due to patients temporarily stopping ART; patients who started ART in the early years of the ART programme and resumed ART after interrupting therapy would not be considered lost to follow-up now, but patients who recently interrupted ART and have not yet resumed ART would be considered lost to follow-up in ART programme reporting. These patients who temporarily interrupt ART are not included in our definition of attrition.

- To the extent that there is a real increase in the rate at which patients stop taking their medication, it is likely to be partially offset by reductions in mortality rates. These reductions in mortality rates are likely to be the result of trends towards increasing baseline CD4 counts in patients starting ART over time.\(^{34}\) The net effect of rising treatment discontinuation and declining mortality may be relatively little change in overall rates of attrition over time.

(d) Proportion of ART attrition that is attributable to death

Although rates of loss to follow-up are often reported to be higher than rates of mortality in patients receiving ART, a recent review has shown that a substantial proportion of patients ‘lost to follow-up’ are actually dead, and many are in fact still receiving ART in other treatment centres\(^{35}\) (the latter are not included in our definition of ‘attrition’). In this review, the average proportion of patients lost to follow-up after ART initiation who were found to have died when subsequently traced was 46\% in African studies (range 27-87\%). Unpublished data from the IeDEA Southern Africa Collaboration suggest that after correcting for under-reporting of deaths, mortality rates after ART initiation are approximately 75\% of the average ART attrition rates assumed in section (c). The 75\% proportion may be an underestimate of the true mortality fraction, since the IeDEA Collaboration consists mostly of relatively well-resourced ART programmes from urban areas, which are not likely to be representative of the ART programme at a national level, and which may therefore have relatively low mortality. However, 75\% could also be an overestimate of the true mortality fraction, since the model parameter \(\nu\) is defined to exclude individuals who died after they stopped taking ART, and the mortality rate estimated from the IeDEA data does not separate out ‘deaths on ART’ from ‘deaths off ART’. The prior distribution assigned to the \(\nu\) parameter is therefore a beta distribution with a mean of 0.75 and standard deviation of 0.10 (the 2.5 and 97.5 percentiles of this distribution are 0.53 and 0.92 respectively).

(e) Bias in reported ART programme statistics

It is quite likely that there is some degree of bias in reported ART programme statistics, though it is not immediately obvious whether the bias would be towards overstatement or understatement of patient totals. This section discusses the types of reporting error that are likely to emerge and the extent to which they may bias the reported patient totals. We consider separately two types of reported ART programme statistics: reported numbers of patients cumulatively enrolled and reported numbers of patients currently receiving ART.

*Reported numbers of patients cumulatively enrolled*

From the cumulative patient totals, it is possible to calculate \(\overline{S}(t)\), the reported number of patients starting ART in year \(t\), by subtracting the cumulative total at time \(t\) from the
cumulative total at time $t + 1$. The true number of patients starting ART between time $t$ and time $t + 1$, $S(t)$, is assumed to be some multiple of the reported number. Mathematically, $S(t) = \tilde{S}(t)(1 + b_0Z(t))$, where $b_0Z(t)$ is an adjustment factor to allow for bias in the reported data. The $b_0Z(t)$ parameters are unknown, and are therefore included in the Bayesian analysis. Prior distributions are specified for the $b_0Z(t)$ factor in each year, and each prior is assumed to be a normal distribution with a mean of 0 (i.e. assuming that on average the reported totals are correct) and standard deviation of $b_0$, so that $Z(t)$ is a standard normal variate. The standard deviation, $b_0$, is set by noting that there are two potential sources of bias in reporting cumulative numbers of patients enrolled on ART.

The first source of bias is double-counting. This may occur because patients who move from one clinic to another get counted as having started ART twice (once at the first clinic and once at the second clinic). In a worst-case scenario, if all patients moving between clinics are double-counted, the ratio of the number of patients incorrectly counted as starting ART in year $t$ to the true number of patients starting ART in year $t$ is

$$0.5(N(t) + N(t + 1))\alpha\chi \over S(t),$$

where $N(t)$ is the total number of patients on ART at time $t$, $\alpha$ is the annual rate at which patients are lost to follow-up and $\chi$ is the proportion of patients lost to follow-up who move to a different clinic. (Note that $\alpha\chi$ does not include the patients who are known to transfer out, as these patients would typically receive a referral letter informing their new treatment provider of their previous receipt of ART, so that they would not be double-counted.)

Estimates from previous analyses\(^{24}\) suggest that the ratio $0.5(N(t) + N(t + 1)) \over S(t)$ is typically between 1 and 2. Annual rates of loss to follow-up in South African ART programmes are typically between 5% and 15%, though varying according to the maturity of the programme.\(^ {35}\) Two separate studies in Johannesburg have found the proportion of patients lost to follow-up who initiate ART elsewhere to be 11.4% and 25.4%\(^ {36, 37}\). If it is assumed that $\chi$ is unlikely to exceed 0.3, then the ratio of the over-count to the true number of patients starting ART is unlikely to exceed $2 \times 0.15 \times 0.3 = 0.075$.

The second potential source of bias is late reporting of patients who have started ART. If the number of patients cumulatively enrolled on ART is correctly reported at the start of the year, but there is an average delay of one month in the reporting of new patients, by the end of the year, one might expect that the estimated number of new patients $\tilde{S}(t)$ to under-state the true number by one twelfth (0.083), if it is assumed that patients were enrolled uniformly over the course of the year. Conversely, if the cumulative number on ART at the end of the year is correctly reported, but the number at the start of the year is subject to late reporting, the estimated number of new patients $\tilde{S}(t)$ will over-state the true number. However, because numbers of patients enrolled are generally increasing over time, one would generally expect the bias due to late reporting to be a downward bias rather than an upward bias.

The net effect of any over-statement due to double-counting and any under-statement due to late reporting is likely to be relatively small, and we consider it unlikely that the factor $b_0Z(t)$
would be either less than -0.12 or more than 0.12. The standard deviation $b_0$ parameter is therefore set to 0.06, so that the prior 95% confidence interval on $b_0 Z(t)$ is from -0.12 to 0.12.

**Reported numbers of patients currently receiving ART**

Some provinces have recently switched to reporting the numbers of patients currently on ART rather than the numbers of patients cumulatively enrolled on ART – this has also been the standard reporting practice in the private and NGO sectors for several years. In this situation, the true number of patients currently on ART (or temporarily interrupted) is assumed to be some multiple of the reported number of patients on ART at time $t$, $\bar{N}(t)$. Mathematically, $N(t) = \bar{N}(t)(1 + b_1 Z(t))$, where $b_1 Z(t)$ is an adjustment factor to allow for bias in the reported data. As before, prior distributions are specified for the $b_1 Z(t)$ factor in each year, and each prior is assumed to be a normal distribution with a mean of 0 and standard deviation $b_1$.

As in the estimation of $S(t)$, late reporting could lead to under-estimation of $N(t)$. If there is an average delay of 1 month in the reporting of patients who have started ART, then the extent of the under-estimation would be roughly $(1/12) \times 0.5(S(t-1) + S(t))/N(t)$ at time $t$. Estimates from previous analyses\(^{24}\) suggest that the ratio $0.5(S(t-1) + S(t))/N(t)$ is typically around 0.8, so that a 1-month reporting delay would correspond to a bias of roughly $0.8/12 = 6.7\%$.

An important source of bias in the context of cross-sectional data relates to the inclusion and exclusion of individuals who have stopped ART. Patients who have temporarily interrupted ART are included in our modelled number of patients on ART, $N(t)$, but might not be included in the reported number of patients on ART, $\bar{N}(t)$. However, patients who have permanently discontinued ART are excluded from $N(t)$ but might be included in $\bar{N}(t)$. In order to determine the net extent of the bias due to the inclusion/exclusion of treatment interrupters, it is necessary to make the following assumptions:

- Firstly, it is assumed that patients temporarily interrupt ART at a monthly rate of $\beta$. This is distinct from permanent ART interruption. Kranzer et al\(^{38}\) found that in Masiphumelele, the annual rate at which patients discontinue ART is 12.8 per 100 person years, and that 33% of these individuals ultimately resume ART. This suggests a value for $\beta$ of $0.128 \times 0.33 / 12 = 0.0035$.

- Secondly, it is assumed that patients permanently discontinue ART, or transfer to ART at another centre without informing their original treatment centre (so-called ‘silent transfer’), at a monthly rate of $\gamma$. Based on the same data summarized in the previous point, this parameter might be estimated as $0.128 \times (1 – 0.33) / 12 = 0.0071$.

- Thirdly, it is assumed that when reporting numbers of patients currently on ART, programmes count patients as being on ART if they are not more than $w$ months late for a scheduled visit. In the Western Cape, patients are included in the total number of patients on ART if they are not more than 3 months late for their most recent scheduled visit, i.e. $w = 3$.\(^{39}\) Other provinces have been instructed to use the same definition in reporting numbers of patients currently on ART, but it is not clear to what extent this recommendation is being followed (Meg Osler, personal communication). In other ART programmes in developing countries, definitions have varied between 1 and 6 months late for a scheduled appointment.\(^{40}\)
• Fourthly, it is assumed that patients who resume ART after interrupting therapy do so after an average of 7 months, based on the median time to ART resumption observed in the community of Masiphumelele. In a review of other studies that have estimated duration of treatment interruptions, the median delay was found to be 150 days.

Based on these assumptions, patients who permanently discontinue ART are included in \( \bar{N}(t) \) but excluded from \( N(t) \) for an average of \( w \) months, and patients who temporarily discontinue ART are included in \( N(t) \) but excluded from \( \bar{N}(t) \) for an average of \( (7 - w) \) months. (By including in the definition of \( \gamma \) those individuals who have silently transferred, we are also assuming that individuals who silently transfer are double-counted in \( \bar{N}(t) \) for an average of \( w \) months.) The percentage by which \( \bar{N}(t) \) exceeds \( N(t) \) is approximately

\[
\gamma \cdot w - \beta(7-w),
\]

which is 0.7% when \( w = 3 \) months. For values of \( w \) between 0 and 6 months, the percentage would be between -2.5% and 3.9%.

A third potential source of bias is replacement of current totals with cumulative totals in sites that lack the capacity to update their records of numbers of patients currently on ART. This may occur even when the overall policy in the province is to report numbers currently on ART. There is anecdotal evidence of this occurring in the Western Cape, where the overall policy has always been to report numbers of patients currently on ART, but since 2007 lack of technical support has led to certain clinics calculating the numbers of patients ‘currently’ on ART by simply adding the number of patients starting ART in the current period to the number reported as having been on ART in the previous period, without making any allowance for patients who have stopped ART (Meg Osler, personal communication). This would lead to some over-estimation of the numbers of patients currently on ART, but it is difficult to quantify the likely extent of this over-estimation.

In summary, there are three sources of bias to be considered when estimating \( N(t) \) from \( \bar{N}(t) \): bias due to late reporting (bias as low as -6.7% if there is a one-month average delay), bias due to inclusion/exclusion of ART interrupters and double counting of transfers (bias between -2.5% and 4%) and bias due to reporting of cumulative totals in some clinics (bias >0). We assume that the net bias, adding all three sources of error together, would be unlikely to be more than 8% (absolute) of the true total, and have therefore set the standard deviation \( b_1 \) to be 0.04, so that the 2.5 and 97.5 percentiles of \( b_1 \cdot Z(t) \) are -8% and 8% respectively.

**Limitations**

It is important to note that the \( Z(t) \) factor is applied to all data sources in year \( t \), both numbers newly enrolled and current totals (although with different standard deviations). It could be argued that it would be more realistic to allow different values of \( Z(t) \) for different data sources reported in the same year, since different provinces and different providers are likely to have different reporting systems. By using the same value of \( Z(t) \) for all data sources in year \( t \), we are being deliberately conservative, allowing for more uncertainty in the accuracy of the reported ART statistics than would actually exist if the biases in the totals reported by different provinces/providers were independent of one another. However, it is debatable
whether biases are truly independent, and assuming independence could lead to some underestimation of the extent of the uncertainty in the ART programme statistics.

It is also important to note that this approach does not make assumptions about correlation between $Z(t)$ and $Z(t+1)$. Positive correlation might be expected when dealing with reported numbers of patients currently on ART, since the same sources of bias are likely to persist from one year to the next. However, negative correlation might be expected when dealing with reported numbers of patients cumulatively enrolled. For example, if the cumulative patient total is exaggerated at the start of year $t$, but the cumulative totals reported at the start of years $t−1$ and $t+1$ are correct, then the number of patients starting ART in year $t−1$ will be overstated, but the number of patients starting ART in year $t$ will be understated. Because of the uncertainty about whether correlation is positive or negative, zero correlation has been assumed.

(f) Priors on model outputs

Although there is considerable uncertainty regarding the HIV parameters referred to in the previous sections, it is clear that any combination of model parameters that leads to an estimate of ART need smaller than the estimated number of patients starting ART in South Africa cannot be plausible. We therefore specify priors on the model outputs to prevent these implausible parameter combinations from being included in the posterior sample. The model output of interest is $r_g(t)$, the annual rate at which ART is initiated in year $t$, in ART-eligible individuals of sex $g$. The inverse of this rate is the average time from becoming eligible for ART to starting ART (in years). Individuals receiving pre-ART care should receive CD4 testing every 6 months, so that individuals in pre-ART care would be determined to be ART-eligible on average 3 months after their CD4 count drops below the ART-eligible threshold. If individuals are newly-diagnosed and determined to be ART-eligible, there is also likely to be a delay before they start ART. Ingle et al\textsuperscript{42} found that in the Free State ART programme, the median time from enrolment into the ART programme to starting ART was 95 days, though this median dropped from 122 days in 2004 to 78 days in 2007. Some of this delay was due to a 3-week drug readiness training programme that all patients had to complete, and some of the delay was due to patients being referred to clinics where they could start ART. In a Cape Town study,\textsuperscript{43} the median time between enrolment into the ART programme and ART initiation was 34 days, although this did not include the time taken to assess ART eligibility. In a Durban study,\textsuperscript{44} the average time from when the CD4 test was performed (to assess eligibility) to when ART training began was 3.6 months. In another Cape Town study,\textsuperscript{45} 69.2% of patients who were diagnosed HIV-positive and had CD4 counts $<200/\mu l$ started ART within 2 months of HIV diagnosis. Using point-of-care CD4 testing, Faal et al\textsuperscript{46} found that 65% of patients in Johannesburg who were newly diagnosed and ART-eligible started ART within 3 months. However, because point-of-care CD4 testing is not generally used in South Africa, and delays in the collection of CD4 testing are usually significant, this is probably an over-estimate of the proportion of patients who would normally start ART within 3 months of diagnosis. Based on the reviewed evidence, it is considered unlikely that the average time from when individuals become ART-eligible to when they start ART would be less than 2 months (corresponding to $r_g(t) = 6$). The prior distribution for $r_g(t)$ is therefore set to be uniform on the interval $[0, 6]$, for all values of $t$ and $g$. This means that any parameter combination that leads to a modelled rate of ART initiation $> 6$ per annum (for males or females in any year) will automatically be rejected.
3. Likelihood function

The likelihood function measures how well the model fits the data, for a given set of model input parameters. In this analysis, two sources of data are included in the definition of the likelihood: CD4 data from surveys of HIV-positive adults, and age-specific HIV prevalence data from household surveys. A likelihood function is defined separately for each data source, and the total likelihood is the product of likelihood functions calculated for each data source.

(a) Likelihood function for CD4 data

Three different South African studies have measured the proportions of HIV-infected adults in different CD4 categories. The results of these three studies are summarized in Table 4. All three studies were conducted when access to antiretroviral treatment in South Africa was fairly limited, and the proportions can therefore be assumed to be representative of untreated individuals.

<table>
<thead>
<tr>
<th>Study</th>
<th>Year of survey</th>
<th>Population sampled</th>
<th># HIV+ adults</th>
<th>% with CD4 of</th>
</tr>
</thead>
<tbody>
<tr>
<td>Auvert et al</td>
<td>2002</td>
<td>Households in Orange Farm</td>
<td>196</td>
<td>46.0</td>
</tr>
<tr>
<td>Rehle and Shisana</td>
<td>2004</td>
<td>Teachers</td>
<td>444</td>
<td>27.9</td>
</tr>
<tr>
<td>Connelly et al</td>
<td>2005</td>
<td>Health workers</td>
<td>74</td>
<td>35.1</td>
</tr>
</tbody>
</table>

Table 4: Survey estimates of proportions of HIV-infected adults in different CD4 stages

Suppose that if a set of parameters, represented by vector $\psi$, is entered into the model, the model estimates that the proportions of ART-naïve individuals in stage $i$ of HIV infection, in year $t_j$, is $\pi_i(t_j, \psi)$ (there are four possible stages of infection corresponding to the four untreated CD4 stages in Figure 1). Further suppose that the observed numbers of infected individuals in the $j$th study, conducted in year $t_j$, who are in stage $i$ is $n_{ij}$. It can then be assumed that the $n_{ij}$ terms are multinomially distributed, so that the likelihood function in respect of the $j$th study is equal to

$$
\left( \begin{array}{c} n_j \\ n_{1j} n_{2j} n_{3j} n_{4j} \end{array} \right) \pi_1(t_j, \psi)^{n_{1j}} \pi_2(t_j, \psi)^{n_{2j}} \pi_3(t_j, \psi)^{n_{3j}} \pi_4(t_j, \psi)^{n_{4j}},
$$

where $n_j$ is the total number of HIV-positive adults in the $j$th study. This is a fixed effects model, i.e. it is assumed that the CD4 distribution is the same for all South African populations that we might choose to sample. The assumption of a fixed effects framework is probably unrealistic, since some populations may be genetically different from others, some populations may be experiencing more advanced epidemics than others, etc. To account for variation in proportions between sub-populations, we define $\rho_j$ as the true proportion of infected individuals in stage $i$, in the $j$th sub-population, and assume that the $\rho_j$ terms are Dirichlet-distributed, i.e.
\[ p(\mathbf{p}_j \mid \pi(t_j, \psi), \varphi) = \Gamma(\varphi) \prod_{i=1}^{4} \left( \rho_{ij}^{\varphi \pi_i(1, \psi) - 1} \right) / \Gamma(\varphi \pi_i(t_j, \psi)), \]

where \( \mathbf{p}_j \) and \( \pi(t_j, \psi) \) represent the vectors of \( \rho_{ij} \) and \( \pi_i(t_j, \psi) \) values respectively. Note that from the properties of the Dirichlet distribution,

\[ E[\rho_{ij}] = \pi_i(t_j, \psi) \]

\[ \text{Var}[\rho_{ij}] = \frac{\pi_i(t_j)(1 - \pi_i(t_j, \psi))}{\varphi + 1} \]

so that the \( \varphi \) variable controls the variance of the random effects, and hence the variability in CD4 proportions across studies. For the purpose of this analysis, the \( \varphi \) parameter is fixed at 188.5, the maximum likelihood estimate obtained in a previous analysis of the same CD4 data.\(^{49}\)

The likelihood function in respect of study \( j \) is then

\[ p(\mathbf{n}_j \mid \pi(t_j, \psi), \varphi) = \int_{\mathbf{p}_j} p(\mathbf{n}_j \mid \mathbf{p}_j) p(\mathbf{p}_j \mid \pi(t_j, \psi), \varphi) d\mathbf{p}_j \]

where \( \mathbf{n}_j \) represents the vector of \( n_{ij} \) values. This likelihood can be more fully expressed as

\[ \int_{\rho_1} \int_{\rho_2} \int_{\rho_3} \int_{\rho_4} \Gamma(\varphi) \left( \frac{n_j}{n_{1j} n_{2j} n_{3j} n_{4j}} \right) \prod_{i=1}^{4} \frac{\rho_{ij}^{n_{ij} + \varphi \pi_i(t_j, \psi) - 1}}{\Gamma(\varphi \pi_i(t_j, \psi))} d\rho_1 d\rho_2 d\rho_3 d\rho_4. \]

Note that after factoring out the terms that are independent of \( \rho_{ij} \) in the above equation, the integral is itself of a Dirichlet form, and therefore integrates to 1 with the multiplication of an appropriate constant term. Hence

\[ p(\mathbf{n}_j \mid \pi(t_j, \psi), \varphi) = \frac{\Gamma(\varphi) \left( \frac{n_j}{n_{1j} n_{2j} n_{3j} n_{4j}} \right) \prod_{i=1}^{4} \frac{\Gamma(n_{ij} + \varphi \pi_i(t_j, \psi))}{\Gamma(\varphi \pi_i(t_j, \psi))}}{\Gamma(n_j + \varphi)}. \]

For the purpose of simulating the posterior distribution, it is sufficient to calculate the log of the likelihood function and to exclude those terms that are independent of the parameters that we are trying to estimate (i.e. terms independent of \( \psi \)). The total log likelihood is obtained by summing the values of the log likelihood for each individual study:

\[ l_c(\mathbf{n} \mid \psi) = \sum_{j=1}^{3} \sum_{i=1}^{4} \ln(\Gamma(n_{ij} + \varphi \pi_i(t_j, \psi))) - \ln(\Gamma(\varphi \pi_i(t_j, \psi))). \]

where \( \mathbf{n} \) is the matrix of \( n_{ij} \) values.

(b) Likelihood function for HIV prevalence data

This component of the likelihood function is defined with reference to the HIV prevalence levels measured, by age and sex, in the national household surveys conducted by the Human
Sciences Research Council (HSRC) in 2005\textsuperscript{50} and 2008.\textsuperscript{51} Suppose that $H_{g,x,t}(\psi)$ is the model estimate of HIV prevalence in individuals of sex $g$, aged $x$ to $x+4$ in year $t$, where the vector $\psi$ represents the values of the model input parameters. The corresponding prevalence of HIV actually measured in the HSRC survey is represented by $y_{g,x,t}$. It is assumed that if $\psi$ is the true set of parameter values, then the difference between the logit-transformed model estimate and the logit-transformed observed prevalence is normally distributed with zero mean, i.e.

$$
\log\left(\frac{y_{g,x,t}}{1-y_{g,x,t}}\right) = \log\left(\frac{H_{g,x,t}(\psi)}{1-H_{g,x,t}(\psi)}\right) + \epsilon_{g,x,t},
$$

where $\epsilon_{g,x,t} \sim N(0, \sigma_{g,x,t}^2)$, representing the random sampling error in the survey. The logit transformations ensure that the error terms are closer to normality and that the model error terms are roughly independent of the level of HIV prevalence. The $\sigma_{g,x,t}^2$ values are estimated from the 95\% confidence intervals that have been published for the various survey estimates. The likelihood in respect of a single measurement is then calculated as

$$
p(y_{g,x,t} \mid \psi) = \left(2\pi\sigma_{g,x,t}^2\right)^{-0.5} \exp\left[-\frac{\left(\logit(y_{g,x,t}) - \logit(H_{g,x,t}(\psi))\right)^2}{2\sigma_{g,x,t}^2}\right].
$$

As before, it is sufficient to calculate the log of the likelihood and to exclude those terms that are independent of $\psi$. The total log likelihood is obtained by summing the values of the log likelihood for each prevalence measurement:

$$
l_H(y \mid \psi) = -\sum_{g} \sum_{x} \sum_{t} \frac{\left(\logit(y_{g,x,t}) - \logit(H_{g,x,t}(\psi))\right)^2}{2\sigma_{g,x,t}^2}
$$

where $y$ represents the matrix of $y_{g,x,t}$ values. In this analysis, only HIV prevalence measurements in the 15-49 age range are used in calculating the likelihood, so that the summation in relation to $x$ is for age groups 15-19, 20-24, …, 45-49.

Finally, the log likelihood in respect of the CD4 data, $l_C(n \mid \psi)$, is added to the log likelihood in respect of the HIV prevalence data, $l_H(y \mid \psi)$, to obtain the log of the total likelihood.

4. Posterior simulation

The posterior distribution represents the synthesis of the prior distributions assigned to the different parameters and the likelihood values calculated for each parameter combination. The likelihood function is simulated numerically, using Sampling Importance Resampling (SIR).\textsuperscript{52, 53} This is a three-step procedure. In the first step of the algorithm, 10 000 parameter combinations are randomly sampled from the prior distributions described in sections 2(a)-(e) (suppose that $\psi_i$ represents the $i^{th}$ parameter combination that is sampled). In the second step of the algorithm, the log likelihood is calculated for each parameter combination ($l_C(n \mid \psi_i) +$...
and an importance weight is assigned to each parameter combination based on this log likelihood. For the $i^{th}$ parameter combination, this weight is calculated as

$$W_i = \begin{cases} \exp \left( l_c(n | \psi_i) + l_H(y | \psi_i) - \sum_{j \in A} \left[ l_c(n | \psi_j) + l_H(y | \psi_j) \right] \right), & i \in A \\ 0, & i \notin A \end{cases}$$

where $A$ is the set of model parameter combinations that yield estimates of $r_g(t) < 6$ for all $g$ and $t$. Setting $W_i = 0$ if $i \notin A$ means applying the prior distribution described in section 2(f) to the model outputs. In this respect our approach differs from the standard Bayesian approach. Bayesian melding generalizes the standard Bayesian approach by allowing for prior distributions to be assigned to model outputs, and not only to model inputs.\textsuperscript{54, 55}

In the final step of the SIR algorithm, a ‘resample’ of 1 000 parameter combinations is drawn from the original sample, using the $\{W_i\}$ values as sample weights (sampling is with replacement). This sample constitutes a random sample from the posterior distribution. The model is then run for each parameter combination in this posterior sample in order to obtain an approximation to the distribution of model outputs (this distribution is summarized in terms of its mean and 95% confidence interval, for each model output).

In this analysis, the posterior sample of 1 000 parameter combinations included 410 unique parameter combinations, and the most frequently sampled parameter combination accounted for only 13 parameter combinations (1.3% of the total). This indicates that the posterior sample is not unduly influenced by single parameter combinations that have unusually high likelihood values.

Table 5 compares the prior and posterior distributions for each of the model parameters. Posterior estimates suggest a more rapid rate of progression to CD4 <350/μl than assumed \textit{a priori}, based on rates of CD4 decline in developed countries. However, rates of mortality after CD4 drops below 350/μl are marginally lower than those that have been assumed \textit{a priori}, based on data in other African countries. The posterior mean for the $R$ parameter is substantially lower than the prior mean (0.74 compared to 1.00), which suggests that the true rate of attrition in South Africa is lower than has been observed in the Western Cape. This is probably because our definition of attrition excludes individuals who temporarily interrupt ART or transfer between ART services without informing the clinic at which they originally received ART (both groups are likely to comprise a substantial proportion of patients considered ‘lost to follow-up’ in the Western Cape ART programme). The relatively low rate of attrition estimated in this analysis explains why the estimates of total patient numbers (Table 2 of the main text) are slightly higher than those estimated previously.\textsuperscript{24}
Table 5: Comparison of prior and posterior distributions

<table>
<thead>
<tr>
<th>Parameter (symbol)</th>
<th>Prior distribution (mean, 95% CI)</th>
<th>Posterior distribution (mean, 95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameters for untreated adults</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Annual rate of progression from CD4 &gt;500 to 350-500 λ₁</td>
<td>0.30 (0.19-0.43)</td>
<td>0.34 (0.28-0.39)</td>
</tr>
<tr>
<td>Annual rate of progression from CD4 350-500 to 200-349 λ₂</td>
<td>0.40 (0.26-0.58)</td>
<td>0.48 (0.40-0.58)</td>
</tr>
<tr>
<td>Annual rate of progression from CD4 200-349 to &lt;200 λ₃</td>
<td>0.35 (0.23-0.50)</td>
<td>0.32 (0.25-0.39)</td>
</tr>
<tr>
<td>Annual rate of HIV mortality if CD4 &lt;200 λ₄</td>
<td>0.25 (0.13-0.40)</td>
<td>0.21 (0.16-0.27)</td>
</tr>
<tr>
<td>Ratio of HIV mortality at CD4 200-349 to HIV mortality at CD4 &lt;200 θ</td>
<td>0.15 (0.07-0.26)</td>
<td>0.13 (0.05-0.24)</td>
</tr>
<tr>
<td>Parameters for treated adults</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relative rate of attrition (as multiple of WC estimate) R</td>
<td>1.00 (0.38-1.92)</td>
<td>0.74 (0.27-1.34)</td>
</tr>
<tr>
<td>Proportion of permanent loss to care that is due to death ν</td>
<td>0.75 (0.53-0.92)</td>
<td>0.74 (0.53-0.92)</td>
</tr>
<tr>
<td>Bias factors Z(t)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year 2001</td>
<td>0 (-1.96 to 1.96)</td>
<td>0.02 (-2.04 to 1.79)</td>
</tr>
<tr>
<td>Year 2002</td>
<td>0 (-1.96 to 1.96)</td>
<td>0.09 (-1.81 to 2.16)</td>
</tr>
<tr>
<td>Year 2003</td>
<td>0 (-1.96 to 1.96)</td>
<td>0.00 (-1.90 to 1.94)</td>
</tr>
<tr>
<td>Year 2004</td>
<td>0 (-1.96 to 1.96)</td>
<td>-0.07 (-2.32 to 1.99)</td>
</tr>
<tr>
<td>Year 2005</td>
<td>0 (-1.96 to 1.96)</td>
<td>0.06 (-2.08 to 1.99)</td>
</tr>
<tr>
<td>Year 2006</td>
<td>0 (-1.96 to 1.96)</td>
<td>0.11 (-1.90 to 2.11)</td>
</tr>
<tr>
<td>Year 2007</td>
<td>0 (-1.96 to 1.96)</td>
<td>0.05 (-2.03 to 2.22)</td>
</tr>
<tr>
<td>Year 2008</td>
<td>0 (-1.96 to 1.96)</td>
<td>0.07 (-2.15 to 1.97)</td>
</tr>
<tr>
<td>Year 2009</td>
<td>0 (-1.96 to 1.96)</td>
<td>-0.01 (-1.85 to 2.11)</td>
</tr>
<tr>
<td>Year 2010</td>
<td>0 (-1.96 to 1.96)</td>
<td>-0.12 (-1.96 to 1.76)</td>
</tr>
<tr>
<td>Year 2011</td>
<td>0 (-1.96 to 1.96)</td>
<td>-0.39 (-2.41 to 1.44)</td>
</tr>
</tbody>
</table>

WC = Western Cape

In general, average posterior estimates of bias in reported ART statistics are not substantially different from zero. However, in 2011 there appears to be a more substantial upward bias in the reported ART programme statistics (the average values of $b_0 Z(2011)$ and $b_1 Z(2011)$ are -0.023 and -0.016 respectively, suggesting true patient numbers are roughly 2% lower than those reported, on average). This may be because high values of $Z(2011)$ are likely to lead to implausibly high rates of ART initiation in the most recent year ($r_d(t) > 6$), and such parameter values would be excluded from the posterior sample.
References


22. Eligibility for ART in lower income countries collaboration. Duration from seroconversion to eligibility for antiretroviral therapy and from ART eligibility to death in adult HIV-infected patients from low and middle-income countries: collaborative analysis of prospective studies. *Sex Transm Infect*. 2008; 84 (Suppl 1):i31-6.


SAFETY, STRENGTH AND SIMPLICITY OF Efavirenz in Pregnancy

Prinitha Pillay, BSc Hons, MB BCh, Dip Infect Dis
Vivian Black, BSc, MB BCh, Dip Infect Dis
Wits Reproductive Health and HIV Institute (WRHI), University of the Witwatersrand, Johannesburg

The WHO recommends starting lifelong ART for all pregnant women with a CD4 count at or below 350 cells/mm³, which recognises the important component of ‘when to start’ and the role that timing of initiation plays in reducing mortality and disease progression. The data on ‘what to start’ are conflicting, and options for resource-limited settings are limited. The choice of an ART regimen for pregnant women is complicated by the need to take into account the health and safety of both the mother and baby. Particularly contentious is whether to use a nevirapine- (NVP) or efavirenz- (EFV) based regimen. This review presents the latest evidence on the safety and efficacy of EFV and NVP in pregnancy and offers recommendations for improving maternal and child health outcomes and avoid mother-to-child transmission as South Africa moves toward turning back the tide on its HIV epidemic.

Estimates for South Africa for 2010 were that approximately 5.6 million people were HIV-infected,¹ accounting for the largest number of cases in a single country.² According to the latest South African National Antenatal Survey (2010), 30.2% of pregnant women in South Africa were HIV-positive,³ maternal mortality was 6 times higher among HIV-positive women, and more than half of all maternal deaths were attributable to HIV.⁴ About 40 000 children in South Africa are infected with HIV each year, with HIV/ AIDS a major contributor to infant mortality in South Africa.⁵ But amidst the bad news has been some good: more than 1.56 million people in South Africa are now receiving ART, and the introduction of more robust and better-tolerated antiretrovirals (ARVs) such as tenofovir disoproxil fumarate (TDF) for first-line therapy is narrowing the gap between recommended treatment protocols in rich and poor countries. In addition, exciting new knowledge and evidence about the concept of ‘treatment as prevention’ (TasP) has emerged, showing not only the therapeutic but also the potential preventive benefits of ART. Prevention of mother-to-child transmission (PMTCT) as TasP is not new — but it currently lags behind other programme goals and ART scale-up efforts.⁶ Earlier initiation of treatment for pregnant women provides extra benefits in PMTCT. While efavirenz (EFV) has been recommended in the WHO guidelines for initiation of eligible women after the first trimester; its use in pregnant women has been restricted in the South African Clinical PMTCT guidelines, where all pregnant women are initiated on a nevirapine (NVP)-based regimen.⁷ In consequence, as South Africa seeks ways in which new knowledge can be integrated into existing programmes that could have measurable effects on mortality and morbidity,⁸ this review presents the latest evidence of safety and efficacy of EFV in pregnancy.

Is There Really an Option for Women?

To date, limited and complicated PMTCT and treatment options exist for women infected with HIV. The latest WHO PMTCT guidelines offer lifelong ART for those with CD4<350 cells/mm³ and allow resource-limited settings two options for those with CD4 >350 cells/mm³: A or B.⁹ Option B allows all women triple therapy for the duration of pregnancy until the cessation of breast feeding for those with CD4≥350 cells/mm³. The view that option B is superior to option A is emerging, for several reasons:¹⁰

- its simplicity for women and programmes, as option A is especially complicated and requires many regimen changes;⁶
- option B allows more women to have sustained exposure to HAART. For those who may fall pregnant during breastfeeding, HAART allows women to survive longer,¹¹ which is important for survival of their children.
- option B may have an added preventative benefit for pregnant women’s partners in discordant relationships¹²
- the unknown risk of non-nucleoside reverse transcriptase inhibitor (NNRTI) resistance in the mother or infant, despite prophylaxis;¹³
- safety, effectiveness and feasibility of daily infant NVP beyond 6 months of age; and maternal and infant acceptability of daily infant prophylaxis for a long period as well as acceptability in programme settings is largely unknown.
- option B may be better for women, with a growing consensus demonstrating that there are individual benefits for the mother as well as for public health.¹⁰

Option B, although simpler, has some drawback for women who fail pregnant again or become eligible for lifelong ART, as they would need to restart HAART. This essentially translates into treatment interruption. Some countries, such as Malawi, have elaborated on option B. Malawi is now implementing what is termed ‘Option B+’, which is lifelong ART for all pregnant women, irrespective of CD4 cell count, from 14 weeks’ gestation. To achieve this, Malawi has included EFV as part of a fixed-dose once-daily formulation for treatment of pregnant women. This decision was justified on the basis that the limited potential risk of birth defects owing to efavirenz is far outweighed by the increased public health benefit, coverage, and reduced overall mortality of initiating mothers on HAART.¹⁴

Is Efavirenz Safe to Use in Pregnancy?

Efavirenz’s FDA rating was changed from category C to category D in 2005, based on data from animal studies and retrospective case reports of neural tube defects.¹⁵ Evidence of teratogenicity linked...
the use of EFV in pregnancy has been limited since then, and current evidence suggests that the risk is lower than previously thought.16,17

Current WHO guidelines recommend avoiding EFV in the first trimester only, but also note that overall rates of birth defects in infants exposed to EFV, NVP and TDF are similar to those in the general population.9 It is evident that the risk of birth defects on exposure to any of the widely used antiretroviral agents shows a similar risk (NVP 2.5%, EFV 2.7% and AZT 3.3%) (Table 2). In addition, the risks are similar for first, second and third trimester exposures (Table 3). In review of the data till 31 January 2011, among the prospective Antiretroviral Pregnancy Registry (APR) reports, the prevalence of birth defects per 100 live births among women with a first trimester exposure to any of the antiretroviral therapies included in the APR is 2.9% (95% confidence interval (CI) 2.5 - 3.4) i.e. 164 outcomes with defects of 5 555 live births.14 The prevalence of defects is not significantly different from the prevalence of defects among women with an initial exposure during the second and/or third trimester of 2.7% (prevalence ratio 1.08, 95% CI 0.88 - 1.32)/205 birth defects in 7 483 live births.10 The APR result for EFV exposure in the first trimester is 2.7% (95% CI 1.6 - 4.3), and 2.9% (95% CI 0.3 - 10) for second- and third-trimester exposure to EFV. The most recent updated meta-analysis as at July 2011 (which reviews the APR and other prospective cohorts) showed a pooled prevalence of 2% (95% CI 0.82 - 3.18) and relative risk of birth defects in EFV-containing ART regimens to non-EFV-based ART as 0.85 (95% CI 0.61 - 1.20).17 This confirms no increased risk of overall birth defects among women receiving first-trimester efavirenz. Comparatively, the risks in the general population are also quite similar (Table 3): in the USA, the prevalence of birth defects in the general population is approximately 3% of live births; and in South Africa the prevalence is estimated at 5.3%.19

However, concerns have been raised, owing to retrospective reports of myelo-meningocelese received after the FDA category change. The risk of neural tube abnormalities exists before it closes by 28 days. The prevalence of neural tube defects (NTD) globally is 0.1 - 0.4%, while in South Africa it is estimated at 0.23 - 0.36%.19

### TABLE 2. COMPARATIVE RATES OF BIRTH DEFECTS FOR WIDELY USED ARVS IN THE FIRST TRIMESTER

<table>
<thead>
<tr>
<th>First trimester exposure ARV</th>
<th>Defects/live births</th>
<th>Prevalence (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indinavir</td>
<td>6/285</td>
<td>2.1% (0.8% - 4.5%)</td>
</tr>
<tr>
<td>Lopinavir</td>
<td>16/738</td>
<td>2.2% (1.2% - 3.5%)</td>
</tr>
<tr>
<td>Atazanavir sulfate</td>
<td>12/502</td>
<td>2.4% (1.2% - 4.1%)</td>
</tr>
<tr>
<td>Stavudine</td>
<td>19/797</td>
<td>2.4% (1.4% - 3.7%)</td>
</tr>
<tr>
<td>Ritonavir</td>
<td>33/1401</td>
<td>2.4% (1.6% - 3.3%)</td>
</tr>
<tr>
<td>Tenofovir</td>
<td>26/1092</td>
<td>2.4% (1.6% - 3.5%)</td>
</tr>
<tr>
<td>Nevirapine</td>
<td>25/987</td>
<td>2.5% (1.6% - 3.7%)</td>
</tr>
<tr>
<td>Emtricitabine</td>
<td>17/641</td>
<td>2.7% (1.5% - 4.2%)</td>
</tr>
<tr>
<td>Efavirenz</td>
<td>17/623</td>
<td>2.7% (1.6% - 4.3%)</td>
</tr>
<tr>
<td>Abacavir</td>
<td>22/744</td>
<td>3.0% (1.9% - 4.5%)</td>
</tr>
<tr>
<td>Lamivudine</td>
<td>118/3864</td>
<td>3.1% (2.5% - 3.7%)</td>
</tr>
<tr>
<td>Zidovudine</td>
<td>118/3620</td>
<td>3.3% (2.7% - 3.9%)</td>
</tr>
<tr>
<td>Nelfinavir</td>
<td>46/1193</td>
<td>3.9% (2.8% - 5.1%)</td>
</tr>
<tr>
<td>Didanosine</td>
<td>19/406</td>
<td>4.7% (2.8% - 7.2%)</td>
</tr>
</tbody>
</table>

Source: Antiretroviral Pregnancy Register (APR) Interim report 2011."
The recent 2011 meta-analysis shows the incidence of neural tube defects (until July 2011) to be low, at 0.07 (95% CI 0.002 - 0.39). Given the low baseline prevalence of neural tube defects, many more first-trimester efavirenz exposures would be required to quantify the risk. Potentially, it would take a long time for a South African (or another country’s) registry to accumulate sufficient data to allow firmer conclusions to be drawn. In addition, a major problem of retrospective reports is the reporting bias. Retrospective reports can be biased toward the reporting of more unusual and severe cases, and are less likely to be representative of the general population experience. Therefore, the calculation of prevalence from these reports is often inappropriate and needs to be interpreted with caution.

To summarise: current data on efavirenz use in pregnancy shows little and poorly supported evidence of risk to the fetus, with a non-significant relative risk of only 0.85 (95% CI 0.61 - 1.20) with EFV, compared with non-EFV-based exposure in the first trimester. There is no significant increase in risk of NTDs with EFV exposure.

Importantly, as for any ARV drug, it is not possible to conclusively say that EFV is safe, and drug companies and regulatory bodies are therefore unlikely to change the EFV rating out of fear of litigation. Noteworthy is the difference between category X and category D (Table 4); and the latter allows policy decision-makers, clinicians and patients alike to weigh up the evidence and allow judgment in their best interests. The FDA is currently proposing to update its approach to labeling.

**CONSEQUENCES ON COMPREHENSIVE SEXUAL REPRODUCTIVE HEALTH**

Another potentially harmful consequence of the EFV category D rating is reported in data on termination of pregnancy (TOP) for women exposed to efavirenz-containing and non-efavirenz-containing regimens. These reveal a RR of 2.81 (95% CI 0.94 - 8.36) for efavirenz-exposed women. These TOPs are not informed by prenatal screening and could mean that women on EFV are almost 3 times more likely to have a potentially distressing and unnecessary TOP based on the potential risk of teratogenicity and not the actual presence of a birth defect. This has far-reaching harmful consequences for the woman and for clinicians who could be inadvertently ill-advising patients on the basis of poorly supported evidence of risk.

Recent studies in Johannesburg show that issues around providers and information transferred to patients about efavirenz risk in pregnancy are often misunderstood. In one study, 40.7% of 851 women declared that the healthcare provider had not discussed pregnancy options with them. A small proportion (6.4%) said a provider had told them not to have more children, and 36% were unsure whether their provider had approved of them having children. Furthermore, women on both EFV and NVP had similar

---

**TABLE 3. PREVALENCE OF BIRTH DEFECTS**

<table>
<thead>
<tr>
<th>General US pop</th>
<th>General South African pop</th>
<th>1st trimester exposure to any ARV</th>
<th>2nd/3rd trimester exposure to any ARV</th>
<th>1st trimester exposure to EFV</th>
<th>2nd/3rd trimester exposure to EFV</th>
<th>1st trimester exposure to EFV</th>
</tr>
</thead>
<tbody>
<tr>
<td>3%</td>
<td>5.3%</td>
<td>2.9%</td>
<td>2.7%</td>
<td>2.7%</td>
<td>2.9%</td>
<td>2.0%</td>
</tr>
<tr>
<td>95% CI:</td>
<td>(2.5 - 3.4)</td>
<td>(0.88 - 1.32)</td>
<td>(1.6 - 4.3)</td>
<td>(0.3 - 10.0)</td>
<td>(0.82 - 3.18)</td>
<td></td>
</tr>
<tr>
<td>Numbers:</td>
<td>164/555</td>
<td>205/7483</td>
<td>17/643</td>
<td>270</td>
<td>39/1457</td>
<td></td>
</tr>
</tbody>
</table>

Relative risk 1st trimester EFV to non-EFV ART was 0.85 (0.61 - 1.20)

---

**TABLE 4. FDA CATEGORIES OF RISK**

<table>
<thead>
<tr>
<th>Category</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Controlled studies show no risk</td>
</tr>
<tr>
<td></td>
<td>Adequate, well-controlled studies in pregnant women have failed to demonstrate a risk to the fetus in any trimester of pregnancy.</td>
</tr>
<tr>
<td>B</td>
<td>No evidence of risk in humans</td>
</tr>
<tr>
<td></td>
<td>Adequate, well-controlled studies in pregnant women have not shown increased risk of fetal abnormalities despite adverse findings in animals or in the absence of adequate human studies, animal studies show no fetal risk. The chance of fetal harm is remote, but remains a possibility.</td>
</tr>
<tr>
<td>C</td>
<td>Risk cannot be ruled out</td>
</tr>
<tr>
<td></td>
<td>Adequate, well-controlled human studies are lacking, and animal studies have shown a risk to the fetus, or are lacking as well. There is a chance of foetal harm if the drug is administered during pregnancy, but the potential benefits may outweigh the potential risk.</td>
</tr>
<tr>
<td>D</td>
<td>Positive evidence of risk</td>
</tr>
<tr>
<td></td>
<td>Studies in humans, or investigational or post-marketing data, have demonstrated foetal risk. Nevertheless, potential benefits from the use of the drug may outweigh the potential risk. For example, the drug may be acceptable if needed in a life-threatening situation or serious disease for which safer drugs cannot be used or are ineffective.</td>
</tr>
<tr>
<td>X</td>
<td>Contra-indicated in pregnancy</td>
</tr>
<tr>
<td></td>
<td>Studies in animals or humans, or investigational or post-marketing reports, have demonstrated positive evidence of fetal abnormalities or risk which clearly outweighs any possible benefit to the patient.</td>
</tr>
</tbody>
</table>

Source: FDA

---

The recent 2011 meta-analysis shows the incidence of neural tube defects (until July 2011) to be low, at 0.07 (95% CI 0.002 - 0.39). Given the low baseline prevalence of neural tube defects, many more first-trimester efavirenz exposures would be required to quantify the risk. Potentially, it would take a long time for a South African (or another country’s) registry to accumulate sufficient data to allow firmer conclusions to be drawn. In addition, a major problem of retrospective reports is the reporting bias. Retrospective reports can be biased toward the reporting of more unusual and severe cases, and are less likely to be representative of the general population experience. Therefore, the calculation of prevalence from these reports is often inappropriate and needs to be interpreted with caution.

To summarise: current data on efavirenz use in pregnancy shows little and poorly supported evidence of risk to the fetus, with a non-significant relative risk of only 0.85 (95% CI 0.61 - 1.20) with EFV, compared with non-EFV-based exposure in the first trimester. There is no significant increase in risk of NTDs with EFV exposure.
pregnancy intentions – either trying to conceive or planning to do so. Pettifor and Rees found in 2005 that roughly 33% of women planned their pregnancies. Complexity of personal reproductive health issues for women and their relationship with healthcare providers must be acknowledged.

WHAT DO WE KNOW ABOUT THE ALTERNATIVE – NEVIRAPINE?

Current WHO guidelines affirm the role of ARVs for pregnant women, and recommend the use of ARVs in differing combinations, depending on CD4 cell count, in all pregnant HIV-infected women. Consequently, according to current South African guidelines, many more women will be initiated on NVP-based regimens. Today, NVP is the recommended alternative to EFV in women of childbearing age.

The 2NN study (the largest randomised controlled trial (RCT), with more than 1 200 patients) found no difference in efficacy between NVP and EFV, and a systematic review of 7 RCTS also found no difference at 48 weeks. The authors recognise, however, that 48 weeks of follow-up is shorter than other cohort studies, which shows that the difference between EFV and NVP grows larger over time. When the Parkland cohort study data were censored at week 48 (using the endpoint in 2NN), there were no significant differences in time to virological failure (EFV = 38.9 weeks v. NVP = 37.2 weeks, p = 0.20); however, when the patient cohort data were not censored at 48 weeks, significant differences were seen between EFV and NVP at 192 weeks (p <0.001). EFV was specifically found to provide a significantly longer time to virological failure than NVP (EFV = 132 weeks v. NVP = 94.1 weeks, p = 0.027). Additionally, in the 2NN study, fewer patients taking EFV than those taking NVP experienced treatment failure (37.8% v. 47.3%).

These results underscore the need to observe patients for longer periods of time to determine the extended durability of antiretroviral regimens. Since clinical trials are often difficult and expensive to maintain, observational cohort analyses may be an alternative for examining long-term durability. Many observational cohorts show that EFV is superior, with an increased risk of virological failure on NVP-based ART regimens. In June 2011, at the IAS conference, a meta-analysis comparing TDF-containing regimens raised concerns that TDF/3TC/NVP might have decreased virological efficacy compared with the EFV-containing TDF regimens. Therefore, we should be concerned about initiating women or switching them to a NVP-based regimen that might not necessarily be superior because of our poorly supported evidence of teratogenicity.

To date, there is conflicting evidence of severe adverse events (rash and hepatotoxicity) in pregnant women who have higher CD4 cell counts, initiating HAART with a NVP-containing regimen. In 2004, Boehringer-Ingelheim, manufacturers of NVP (Viramune) performed a retrospective analysis of hepatitis events and found no consistent CD4 cell-count cut-off that could be identified in women, that was associated with an increased risk of liver enzyme elevations. The analysis also demonstrated no significant differences in the rate of serious hepatic events among ARV regimens, including between the non-nucleoside reverse transcriptase inhibitors NVP and EFV. Further scrutiny of this analysis revealed that patients with symptomatic events were not included in the subset analysis. It also revealed the risk of rash-associated hepatic adverse events was 3 times higher in women than in men. A rash-associated hepatic event was also associated with a higher CD4 cell count, with women with pre-treatment CD4 count >250 cells/mm³ having a higher risk of hepatotoxicity than women with CD4 <250 cells/mm³. Following these results, the company changed the Summary of Product Characteristics to include a caution that women with higher CD4 cell counts are at increased risk of hepatic toxicity.

Previously, it was not recommended to initiate women on NVP if their CD4 cell count was above 250 cells/mm³. Data are now emerging from both high-income and resource-limited settings, suggesting that it is safe for patients who have experienced good increases in their CD4 cell counts on another ARV regimen to switch to NVP (provided they have an undetectable viral load), even when their CD4 count is above the level recommended for initiating treatment. In 2009, Ouyang and colleagues showed that NVP is not uniquely associated with hepatotoxicity in pregnancy but rather that pregnancy itself may be an independent risk factor. The same study also showed that NVP is not associated with hepatotoxicity at higher CD4 cell counts. Chu et al. found in 2010 no association of CD4 cell count and hepatotoxicity; however, the median CD4 cell count in their cohort was low (112 cells/mm³) and, with resource-limited settings still pervaded by patients presenting late and initiated at low CD4 cell counts, this study highlights one of the possible reasons for the lack of observed difference between high-income and resource-limited settings.

Indeed, a Cambodian cohort study in a resource limited setting found that higher CD4 cell counts at the time of NVP substitution from EFV increased the risk of subsequent NVP toxicity, and that ART-experienced Cambodians appear to have a risk of NVP toxicity comparable with that of ART-naïve patients, despite higher CD4 counts. The analysis from the large randomised clinical trial, the 2NN study, demonstrated that the rate of skin rash and hepatic events was higher in patients with CD4 counts >200 cells/ml, and also that women with CD4 counts >200 cells/mm³ had a significantly greater risk of developing a rash than men. The most recent data from Uganda presented at the IAS conference in June 2011 have documented 3 cases of Stevens-Johnson syndrome in stable experienced HAART patients when switched to NVP. Overall, the meta-analysis of 7 randomised controlled trials (RCTs) show that EFV had a lower incidence of adverse events (AEs) and fewer discontinuations than NVP. Fewer patients taking EFV discontinued therapy because of any AE or HIV event than patients taking the other treatment regimens. Two deaths were directly associated with NVP use (one from toxic hepatitis and the other from Stevens-Johnson syndrome); no deaths were associated with EFV. Overall, EFV was associated with a more favorable tolerability profile than NVP, with less grade 3 or 4 clinical AEs, fewer discontinuations for AEs, and numerically less treatment changes with EFV than with NVP.

There therefore seems to be insufficient evidence to recommend that it is safe to switch NVP for EFV, in particular in settings such as South Africa with higher co-infection rates of TB i.e. women who are switched to and from.
IS EFAVIRENZ AFFORDABLE AND COST-EFFECTIVE?

The prohibitively high cost of EFV had prevented its widespread use in the early part of the decade, and the price evolution is demonstrative (Fig. 1). The Medicins Sans Frontieres (MSF) report Untangling the Web reveals that the cost of EFV has been driven down from the originator price of $347 in December 2002 to a WHO-prequalified generic price of $52 in July 2011 (per patient per year).45 Despite cost, perhaps more important is a recent study looking to quantify the benefit (life expectancy gains) and risk, that shows that the use of non-efavirenz-based initial ART in HIV-infected women of childbearing age may reduce life expectancy gains from ART.46 The mean life expectancy for women who would start ART at a CD4<250 on NVP-based HAART was 25.49, compared with 27.08 for EFV-based ART, with a resultant 1.6-year life expectancy gain on EFV compared with NVP.46 In addition, survival of women who received an EFV-based ART regimen was 0.89 years greater than all non-EFV-based regimens.46 Policymakers do indeed need to take into account cost and cost-effectiveness, but the benefit to women and their families favours EFV-based ART when reduced survival and potential life-threatening severe adverse events on NVP are quite stark. Today, the fixed-dose combination of tenofovir, lamivudine and efavirenz in a once-a-day pill is likely to have positive spill-over effects for those women who need to take treatment every day for the rest of their lives, without jeopardising their own health and further resistance through poor adherence.

CONCLUSIONS AND RECOMMENDATIONS

The above describes the most recent evidence available; could we expect more robust evidence? Randomised controlled trials (RCTs) (gold standard) are not feasible, and it would take a very long time to truly assess and compare outcomes. Modeling exercises can help to inform some potential future outlook for those questions not answered by available evidence today or when RCTs are not feasible. Ouattara et al.'s (2012) latest projections found that starting ART with EFV, which has a lower adverse events on NVP are quite stark. Today, the fixed-dose combination of tenofovir, lamivudine and efavirenz in a once-a-day pill is likely to have positive spill-over effects for those women who need to take treatment every day for the rest of their lives, without jeopardising their own health and further resistance through poor adherence.

Therefore, the risk-benefit question for women is: Does the risk of birth defects (knowing that we have low and poorly supported risk to the fetus and enough data to say we don’t have a tenfold increase in risk of NTDs) after the organogenesis period on EFV outweigh the risks of life-threatening toxicity, regimen changes and a potential risk of failure when switching women from EFV to NVP? Particularly as South Africa has moved to earlier initiation of HAART at CD4<350 cells/mm², many more women will be picked up early at antenatal clinic with the risk of severe adverse events being potentially higher in women with higher CD4 counts if switched to NVP.

It is important to bear in mind that most studies are confounded by HIV disease stage, smoking, co-morbidities and other medication. Generally, an HIV-infected population is possibly at increased risk of adverse outcomes of pregnancy unrelated to teratology, and in South Africa there is an extra burden of fetal alcohol syndrome. ‘Fetal alcohol spectrum disorder is the most common birth defect in South Africa, by far more common than Down syndrome and neural-tube defects combined;’ according to Professor Denis Viljoen of the Foundation for Alcohol Related Research (FARR).48

Based on the evidence, there are several policy recommendations that the South African government should consider at this critical juncture while heading towards the ‘getting to zero’ goal.

- Firstly, it should allow for already on HAART who fall pregnant to continue on EFV-based HAART instead of switching to NVP. Most pregnancies are not detected until at least one month after conception; switching to NVP after this point may not protect against birth defects, and needs to be balanced against the risk of serious adverse events caused by switching to NVP.
- Secondly, it could allow only women who are on ART and who want to conceive to switch from EFV to NVP before falling pregnant.
- All women of child-bearing age should be encouraged to plan their pregnancies and be tested before conception.
- The South African government should consider moving to embrace Option B as preferred PMTCT, and to initiate all women in need of HAART them on the superior combination of TDF/3TC/EFV from 14 weeks’ gestation. This has an added benefit of simplification for nurse-initiated ART as it is consistent with adult preferred first-line treatment, and has the potential to simplify the supply chain, thereby preventing potential stock-outs.
- Consider pilot projects that could ascertain the benefits and risks for individuals and at the population level, as well as programmatic implications for putting all pregnant women on HAART (Option B+).
- Regulatory bodies and the government should fast-track the registration of the fixed-dose once-daily formulation of TDF/3TC/EFV for all patients.
- Lastly, increased pharmaco-vigilance and a South Africa-wide prospective Antiretroviral Pregnancy Registry are needed. With the number of women exposed to EFV in the first trimester, however, it would take a very long time for a South African registry to accumulate enough data to allow firmer conclusions to be drawn; therefore, this should not be done at the expense of women in need of treatment now.

This paper has argued that, although we could never claim any ARV to be completely safe, weak associations in some studies are far outweighed by the benefits of HAART in pregnancy. The consideration to use EFV in the first trimester of pregnancy in resource-limited settings such as South Africa needs to move beyond concerns of poorly supported evidence to recognising new evidence of survival gains, efficacy, toxicity, direct medical and programmatic costs (including costs of simplification and scaling.
up coverage) – as well as indirect costs e.g. unnecessary and distressing termination of pregnancies. This allows policymakers an opportunity to narrow the evidence accumulated to date and focus on pursuing an effective strategy based on evidence and balancing risks and benefit of best prevention and treatment options for women and their families.

REFERENCES


CASE STUDY

CHRONIC GENITAL ULCER DISEASE WITH SUBSEQUENT DEVELOPMENT OF METHICILLIN-RESISTANT STAPHYLOCOCCUS AUREUS (MRSA) URETHRITIS AND BACTERAEMIA IN AN HIV-SEROPOSITIVE PERSON – A CASE OBSERVATION

Christine Katusiime, MB ChB, PGDPPM, MIPH
Andrew Kambugu, MB ChB, MMed
Infectious Diseases Institute, College of Health Sciences, Makerere University, Kampala, Uganda

HIV-seropositive persons are at increased risk of methicillin-resistant Staphylococcus aureus (MRSA). Genital ulcerative disease and sexually transmitted infection with subsequent MRSA infection in HIV-seropositive persons have been documented only once. We report a case of a 44-year-old man who presented to the Infectious Diseases Institute, Kampala, Uganda, with chronic genital ulcer disease and who subsequently developed MRSA urethritis and bacteraemia. This case also demonstrates that persistent genital ulcer disease in HIV-seropositive persons may be as a result of concurrent MRSA infection.

HIV-seronegative persons are less likely to become infected with Staphylococcus aureus and associated S. aureus septicaemia than their HIV-seropositive counterparts. Studies have demonstrated an increased risk of MRSA infection among HIV-seropositive persons of up to 18 times. MRSA infection therefore presents a public health concern. MRSA infection has also been found to be associated with sexually transmitted infections (STIs) and genital ulcer disease (GUD) in HIV-seropositive persons. GUD with associated MRSA infection has only been documented once. The case we present in this report is of chronic genital ulceration complicated by MRSA urethritis and bacteraemia. This case of chronic GUD with subsequent MRSA infection in an HIV-seropositive host also highlights the importance of ruling out concurrent MRSA infection in chronic GUD in HIV-seropositive persons.

CASE PRESENTATION

A 44-year-old HIV-seropositive man with World Health Organization (WHO) stage IV disease, who had been on highly active anti-retroviral therapy (HAART) of tenofovir, lamivudine and efavirenz, and co-trimoxazole prophylaxis since March 2008, presented with an 8-month history of genital ulceration and a 4-month history of urethral pus discharge. He initially noticed a small papule on the glans of his penis which increased in size and ulcerated. He subsequently developed a yellow non-odorous urethral discharge. On review of symptoms, he denied associated fevers and chills, or trauma to the site. He had no significant previous medical history and denied cigarette smoking, alcohol consumption and intravenous drug usage. He was heterosexual, denied unsafe sexual practices, and had no history of previous STIs, antibiotic use, and recurrent STIs or GUD.

Physical examination revealed a body temperature of 36.1°C, pulse rate 86 beats per minute, blood pressure 120/70 mmHg, and respiratory rate 12 breaths per minute. Genital examination revealed an ulcer involving the glans of the penis with a yellow urethral discharge. Precordial, ophthalmic, gastrointestinal and respiratory examinations were unremarkable. Neurologically, there were no cranial nerve deficits, reflexes were symmetrical and normal, and distal sensation was intact.

Laboratory data revealed a white blood cell (WBC) count of 5 000/mm³ and an elevated erythrocyte sedimentation rate (ESR) of 80 mm/hr. His urine revealed 500 leukocytes/ul, trace protein, 66 white blood cells/high-power field (hpf) and 1 cast/hpf. The rest of the routine laboratory tests including renal function tests, liver function tests, random blood sugar (RBS), Treponema pallidum haemagglutination assay (TPHA), serum cryptococcal antigen (CRAG) titres, hepatitis B surface antigen (HBsAg) and brucella agglutination assay were unremarkable. CD4+ cell counts were 944 cells/mm³ (30%), and plasma HIV RNA levels were undetectable. Herpes simplex virus-2 (HSV-2) serology could not be done because of the high costs involved.

A penile wedge biopsy taken for histopathological examination showed features suggestive of a chronic penile ulcer. The urethral pus swabs and blood cultures grew MRSA sensitive to gentamicin, ciprofloxacin and vancomycin but resistant to oxacillin, tetracycline, penicillin and erythromycin. Urine cultures depicted no growth. A clinical diagnosis of chronic ulcerative genital herpes was made, following consultation with an STD specialist. The patient was then admitted and commenced on a course of oral acyclovir 400 mg twice daily for 6 months and intravenous vancomycin for 2 weeks for chronic ulcerative genital herpes and MRSA urethritis and bacteraemia respectively. Intravenous ceftriaxone was also administered to treat the urinary tract infection (UTI). Continual adherence to co-trimoxazole, acyclovir and HAART was continually emphasised.
The course of treatment was successful; examination at 1 month and 2 months follow-up revealed no urethral discharge and completely healed penile ulcerations.

**DISCUSSION**

HIV-infected persons have a higher risk of MRSA infection than the general population. The reason why our patient was predisposed to MRSA infection could have been a result of the chronic genital ulceration.

Prior studies have implicated lack of co-trimoxazole prophylaxis, intravenous drug usage, low CD4 T-cell count, high HIV viral load and hospitalisation as risk factors for MRSA colonisation in HIV-seropositive patients. These factors, however, were not noted in our patient. Ramsetty and colleagues demonstrated that HIV-seropositive patients with CD4 T-cell counts <200 cells/mm³ were at significant risk of MRSA infections. Our patient was the exception to this finding, as his CD4 T-cell count was 944 cells/mm³.

A ramification of this case is the importance of considering concurrent MRSA infection in HIV-seropositive patients with chronic GUD. Other factors that make this case unique are the development of MRSA urethritis and bacteraemia despite the patient’s high CD4 T-cell counts and good virological control.

**CONCLUSION**

Our patient developed MRSA infection following chronic GUD that was not effectively managed. Clinicians need to maintain vigilance in the management of chronic GUD in HIV-seropositive persons, as MRSA co-infection may become an increasing complication in the future.

**REFERENCES**

Two CPD points are awarded for the correct completion and submission of the questions below.

CPD questionnaires must be completed online via www.cpdjournals.co.za.
After submission you can check the answers and print your certificate.
Questions may be answered up to 6 months after publication of each issue.

This programme is available free of charge to members of the HIV Clinicians Society and SAMA only.

Regarding chronic genital ulcer disease in HIV-infected individuals
1. True (A) or false (B) — click on the correct answer:
   Chronic genital ulcer disease in HIV-infected individuals is only observed at low CD4 cell counts and resolves after aetiological treatment with the use of antiretroviral therapy.
2. True (A) or false (B) — click on the correct answer:
   MRSA superinfection is a common complication of chronic genital ulcer disease in HIV-infected men.

Regarding the use of efavirenz during pregnancy
3. Click on the correct answer:
   The current labelling of efavirenz in pregnancy is:
   A: Controlled studies show no risk
   B: No evidence of risk in humans
   C: Risk in pregnancy can not be ruled out
   D: Evidence of risk in pregnancy
   X: Contra-indicated in pregnancy
4. True (A) or false (B) — click on the correct answer:
   There is substantial evidence from clinical research that EFV is associated with birth defects in humans.
5. True (A) or false (B) — click on the correct answer:
   Data from the Antiretroviral Pregnancy Register demonstrate that efavirenz is associated with a higher rate of birth defects than both AZT and 3TC.
6. True (A) or false (B) — click on the correct answer:
   Data show that among NNRTIs, efavirenz is associated with equal or greater levels of long-term viral suppression than nevirapine.
7. True (A) or false (B) — click on the correct answer:
   Women with high CD4 cell counts using nevirapine are at an increased risk of hepatotoxicities, and therefore nevirapine may not be an ideal choice of NNRTI for a strategy for universal ART use in pregnancy (e.g. “Test and Treat”).
8. True (A) or false (B) — click on the correct answer:
   Nevirapine is more expensive than efavirenz.

Regarding the coverage of antiretroviral therapy across South Africa
9. True (A) or false (B) — click on the correct answer:
   Calculating the number of patients on ART in South Africa is a simple addition because of consistently high quality health statistics reported uniformly by all ART services.
10. True (A) or false (B) — click on the correct answer:
    By the middle of 2011, more than 1.5 million South Africans had been started on ART.
11. True (A) or false (B) — click on the correct answer:
    The proportion of ART-eligible men who are receiving ART is higher than the proportion of eligible women receiving ART.
12. True (A) or false (B) — click on the correct answer:
    The numbers of individuals starting ART nationwide fell far short of the targets set by the 2007-2011 National Strategic Plan.

Regarding socioeconomic v. cultural explanations for the spread of HIV in South Africa
13. True (A) or false (B) — click on the correct answer:
    Concurrent partnerships are likely to play a role in understanding the spread of HIV infection at a population level.

Regarding the impact of breastfeeding by HIV-infected mothers on the health of exposed infants
14. True (A) or false (B) — click on the correct answer:
    Antiretroviral drug interventions such as nevirapine prophylaxis for infants, and/or maternal use of multi-drug ARV regimens, can be used during breastfeeding to reduce the risk of transmission.
15. At a population level, the benefits of formula feeding (related to prevention of HIV infection) are likely to outweigh the risks (in other sources of morbidity and mortality) only in countries where the infant mortality rate is less than 10 per 1 000.

Regarding the use of stavudine
16. True (A) or false (B) — click on the correct answer:
    Both stavudine and tenofovir can lead to thymidine analogue mutations (TAMs).
17. Although lipo-atrophy occurs with stavudine, it is highly reversible after stopping stavudine use.
18. Neuropathy is a complication of stavudine use that may be more likely in patients receiving TB co-treatment.
19. Most toxicities associated with stavudine use take place in the first year on therapy.
20. Most toxicities associated with stavudine use appear to be dose-dependent.