HIV, antiretroviral therapy, and tuberculosis: outcomes in Johannesburg, South Africa

Daniel Westreich

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Epidemiology in the School of Public Health.

Chapel Hill
2008

Approved by:

Annelies Van Rie

Joseph Eron

Jay Kaufman

Michele Jonsson Funk

Patrick MacPhail

ABSTRACT

DANIEL WESTREICH: HIV, antiretroviral therapy, and tuberculosis: outcomes in Johannesburg, South Africa
(Under the direction of Annelies Van Rie)

The rollout of highly active antiretroviral therapy (HAART) in sub-Saharan Africa poses many challenges. In one of the largest HAART clinics in Africa, we described overall outcomes of HAART, the impact of treated pulmonary tuberculosis (PTB) at the time of HAART initiation on all-cause mortality, and the impact of prevalent and incident tuberculosis treatment on the risk of the substitution of the antiretroviral drug stavudine.

In these analyses, robust methods including inverse probability weighted marginal structural models were used where appropriate to control for confounding and bias due to losses to follow-up and competing risks. This analytic approach made these results more robust compared with more “traditional” analyses.

Relatively few patients died after HAART initiation. Risk of death during follow-up increased in individuals who at baseline were older or had low CD4 cell counts, low body mass index, low hemoglobin, or advanced disease stage.

The crude hazard ratio (HR) for the effect of treated PTB at time of HAART initiation on mortality was 1.71 (95% confidence interval [CI] 1.31-2.23) and the adjusted HR was 1.06 (95% CI 0.75-1.49), a difference due to the fact that individuals with PTB at time of HAART initiation had more severe

immunodepression. Additionally, individuals who started HAART within 30 days of initiation of treatment for PTB were not at higher risk of death than other individuals, suggesting that early HAART initiation can be safely recommended.

In analysis of the effect of treatment for pulmonary or extrapulmonary tuberculosis (TB) on risk of stavudine substitution, we found that if TB treatment is ongoing at the time of initiation of HAART, there was between a two- and sevenfold increase in risk of stavudine substitution during the early months of HAART. However, incident TB treatment after HAART initiation, did not raise the risk of stavudine substitution. These results were robust to sensitivity analysis, and suggest that the use of stavudine might be reconsidered in patients receiving treatment for TB.

Overall, these results suggest that early initiation of HAART is warranted in patients with TB, but that clinicians need to monitor these patients closely for indications for stavudine substitution.

DEDICATION

To my loving and ever-supportive wife, Katie Davis Westreich.

ACKNOWLEDGEMENTS

In addition to my wife, to whom this document is dedicated, I must thank several people, without whom this dissertation would never have come to pass. First, my advisor, Annelies Van Rie, who put an enormous amount of work and effort into this document, and into my graduate career more generally. Thank you. The remainder of my committee also provided outstanding support for me both on this document and throughout my graduate career. Thanks to Dr. Michele Jonsson Funk, Dr. Jay S. Kaufman, and Dr. Joseph Eron at UNC-Chapel Hill, and Professor Patrick MacPhail in Johannesburg at the Themba Lethu Clinic.

I thank Nancy Colvin and Carmen Woody for their general support in the Department of Epidemiology.

I thank Dr. Ian Sanne for trusting me with these data. I thank Babatyi Malope-Kgokong, Shashi Nagar, Mhairi Maskew, Prudence Ive, Dennis Rubel, Francesca Conradie, Ina Dejongh, Francois Venter, AJ Edge, and Michele Dean for their support at Themba Lethu and in Johannesburg throughout this process. I thank all unnamed clinical and general staff of Themba Lethu for their trust in these analyses, and for their unyielding efforts on behalf of the patients; and I thank the patients of Themba Lethu Clinic for the trust they have placed in receiving treatment at the clinic.

I thank Dr. Charles Van Der Horst, for assisting my efforts to connect with Dr. Sanne in Johannesburg. I thank Dr. Stephen R. Cole for his methodological support in preparing these analyses.

The clinical activities of the Helen Joseph Hospital are supported by the National and Gauteng Department of Health and the United States President's Emergency Plan for AIDS Relief (PEPFAR) in a grant by USAID to Right to Care and the Institution (674-A-00-08-00007-00). The research activities of this publication are supported by the National Institute for Health in a grant from the NIAID, DAIDS division (CIPRA IU19 AI53217-01, PEPFAR protocol #3 U19 AI 053217-04SI-R2C01). I also received support from an unrestricted educational training grant from the UNC-GlaxoSmithKline Center for Excellence in Pharmacoepidemiology and Public Health, UNC School of Public Health. I thank the Center, GlaxoSmithKline, and Dr. Suzanne West.

Last, I thank my family on both coasts for their support and encouragement throughout this process.

TABLE OF CONTENTS

LIST OF TABLES

Table

LIST OF FIGURES

Figure

KEY ABBREVIATIONS

HIV	Human immunodeficiency virus type 1
HAART	Highly active antiretroviral therapy
ARV (drug)	Antiretroviral
ART	Antiretroviral therapy
TLC	Themba Lethu Clinic
TLCC	Themba Lethu Clinical Cohort
CD4 (cell)	CD4-positive T-lymphocyte
BMI	Body mass index (measured kg/m^2)
TB	Tuberculosis
PTB	Pulmonary tuberculosis
d4T	Stavudine
AZT	Zidovudine
EFV	Efavirenz
NVP	Nevirapine
LPVr	Lopinavir-ritanovir (Kaletra ®)
NNRTI	Non-nucleoside reverse transcriptase inhibitor
NRTI	Nucleoside reverse transcriptase inhibitor
PI	Protease inhibitor

CHAPTER 1

CRITICAL REVIEW OF THE LITERATURE

GENERAL BACKGROUND

Global and sub-Saharan African epidemiology of HIV. In 2007, there were an estimated 33.2 million individuals living with human immunodeficiency virus (HIV) worldwide (UNAIDS/WHO 2007), 2.5 million of whom were newly infected (UNAIDS/WHO 2007); since the disease was first described in 1981, over 20 million individuals have died as a result of HIV infection (WHO 2006). It is by now common knowledge that sub-Saharan Africa bears a disproportionate share of the world's burden of current and new HIV infections: 1.7 million of the 2.5 million new infections in 2007 occurred in sub-Saharan Africa, and the average adult prevalence in sub-Saharan Africa is 5.0%, an order of magnitude greater than most other regions in the world, and perhaps five times the prevalence in the Caribbean, the region with the next-highest prevalence (UNAIDS/WHO 2007). Sub-Saharan Africa likewise bears a disproportionate burden of AIDS deaths, with an estimated 1.6 of 2.1 million deaths due to AIDS in 2007 (UNAIDS/WHO 2007).

Epidemiology of HIV in South Africa. The South African National AIDS Council in the Department of Health estimates that the adult (15-49) prevalence of HIV was 18.8% in 2006, and that there were approximately 5.5 million cases of HIV

prevalent in the country (Kapp 2007; South African National AIDS Council 2007) though other sources have offered lower prevalence estimates, including 22% of the adult population and 12.9% (UNAIDS 2006), and between 10.2% and 16.2% (Garcia-Calleja, Gouws, and Ghys 2006) of the general population, depending on location of the survey.

While there is evidence that the epidemic is decreasing in intensity in parts of the eastern and even southern African regions (UNAIDS/WHO 2006), the evidence in South Africa is more mixed: in 2006 the prevalence of HIV among pregnant women presenting in ante-natal care clinics was 29%, only 1% lower than the highest prevalence ever recorded in that population, in 2005, and 7% higher than the estimate from 1998 (UNAIDS/WHO 2007). And while HIV prevalence has decreased in prevalence in pregnant women ages 15-24, this population – especially at ages 20 and 21 – appear to be at highest risk of HIV acquisition (Pettifor et al. 2005). While there is some risk that population prevalence is overestimated due to over-reliance on antenatal clinic data, it is nevertheless clear that prevalence of HIV in South Africa is extremely high by global standards, and that this situation is not likely to change substantially in the near term.

HIV: BIOLOGY AND TREATMENT

Pathobiology of HIV. Human immunodeficiency virus (HIV) is a single-stranded RNA virus (Tronchet and Seman 2003). HIV is a retrovirus, and replicates by using reverse-transcriptase to convert RNA into DNA, which is integrated into the

DNA of infected host cells. HIV mutates rapidly, largely due to error-prone copying strategies (Stebbing and Moyle 2003), and to the production of a large number of copies per day, on the order of 10^9 (Ho et al. 1995). This has resulted in the evolution of many distinct clades, or sub-species, of HIV, which have slightly different clinical implications (Stebbing and Moyle 2003). HIV is classified into two major families, HIV-1 and HIV-2; the M, or Main clade of the former is responsible for the large majority of cases of HIV in the world, and is abbreviated as HIV in this paper (Stebbing and Moyle 2003). The South African epidemic is chiefly HIV-1 C.

HIV uses the viral glycoprotein gp120 to target cells of the human immune system that present the CD4 co-receptor, including CD4+ T-cells, dendritic cells, and macrophages (Janeway et al. 2005). Long term HIV infection results in acquired immune deficiency syndrome, or AIDS, the consequence of a long-term depletion of immune system cells, especially CD4-cells. When a patient's CD4 cell counts fall below 200, "opportunistic" infections with bacteria and other pathogens become common, eventually leading to death (Janeway et al. 2005). CD4 count and viral load monitoring are widely considered the key markers of HIV disease progression.

Opportunistic and co-infections common to sub-Saharan Africa. There are many opportunistic infections and co-infections that impose significant burdens of morbidity and mortality; among the most important are Hepatitis B (Hoffmann and Thio 2007), Hepatitis C (Rockstroh and Spengler 2004), and HHV8 (strongly associated with Kaposi's Sarcoma) (Malope et al. 2007). Perhaps the two most important co-infections to consider in sub-Saharan Africa are tuberculosis (TB) (WHO 2007) and malaria (WHO 2007).

In South Africa, and Johannesburg in particular, TB is clearly the more important of these two co-infections. TB infects an estimated two billion people worldwide and killed an estimated 1.6 million people in 2005 (WHO 2007). The majority of those deaths were in southeast Asia and Africa, particularly South Africa, Zimbabwe, Tanzania, and Kenya (WHO 2007, 2007). In Africa generally, TB mortality in 2005 was 77 per 100000 individuals, more than 3 times the global rate, more than twice the rate in southeast Asia (31/100000), and 10 times the European rate. TB can be treated with a standard six-nine month course of pharmacotherapy, typically including at least two drugs such as isoniazid, rifampicin, pyrazinamide, and ethambutol. Many of these drugs have significant side effects, including hepatitis, rash, anemia, nausea, and peripheral neuropathy (Subbaraman et al. 2007) – we defer a complete discussion of these issues for the literature review for Specific Aim 3.

Therapeutic strategies for HIV infection, including HAART. A consequence of the innate variability of HIV is that an effective HIV vaccine has so far proved elusive (Ho and Huang 2002). Treatment for (though not cure of) HIV therefore relies on therapeutic drugs, known collectively as antiretroviral drugs. Antiretroviral drugs currently comprise six distinct classes and over 20 individual agents. The first class of ARV agents developed were the nucleoside analogues, or *nucleoside reverse transcriptase inhibitors* (NRTIs), which inhibit reverse transcriptase by chain-termination (Portsmouth, Stebbing, and Gazzard 2003). The two additional classes of ARVs in routine use are *protease inhibitors* (PIs), which act by inhibiting the cleaving of HIV particles from host cells (Wynn et al. 2004), and

non-nucleoside reverse transcriptase inhibitors (NNRTIs), which bind directly to – and thus inhibit – reverse transcriptase. Additional classes of approved antiretroviral drugs include *entry* and *fusion inhibitors* (Portsmouth, Stebbing, and Gazzard 2003; Zapor et al. 2004) and *integrase* inhibitors (Evering and Markowitz 2008), while other classes (e.g., maturation inhibitors) are in development.

The first of these drugs to be used clinically was zidovudine (AZT). Early trials showed a significant virologic, immunologic, and clinical impact on HIV+ individuals of both AZT monotherapy and NRTI dual therapy (Collier et al. 1993; Fischl et al. 1990; Fischl et al. 1990). Subsequently, the introduction of new drugs and drug classes allowed for the advent of the era of highly active antiretroviral therapy (HAART), usually involving at least three distinct ARVs, usually in at least two classes (Wynn et al. 2004). It was hoped three drugs would result both in a greater viral suppression as well as an increase in the length of that suppression, the latter by delaying the emergence of HIV resistance mutations.

Indeed, HAART has proven to be an enormous success (Chen, Hoy, and Lewin 2007; Detels et al. 1998). While HAART is not able to eradicate HIV because of persistent infection in resting CD4 T-cells (Lehrman et al. 2005) and possibly additional reservoirs, methodologically definitive studies have demonstrated that HAART results in marked reductions in time to AIDS or death (Cole et al. 2003; Detels et al. 1998), as well as providing a longer duration of viral suppression compared to dual or monotherapy (Cole et al. 2003), albeit often with increased toxicities and side-effects of the medications (Wynn et al. 2004; Zapor et al. 2004),

and increased CD4 count reconstitution compared with dual therapy and monotherapy (Chen, Hoy, and Lewin 2007; Detels et al. 1998; Hammer et al. 1997).

Even with multiple drugs and drug classes, adherence is a key determinant of patient success on HAART, because risk of therapy failure (generally defined as a confirmed viral load of > 400 copies/ml, six months after the start of therapy or after a suppression of virus under that level (Republic of South Africa Department of Health 2004; United States of America Department of Health and Human Services 2005)) is strongly associated with imperfect adherence (Cote and Godin 2005; Paterson et al. 2000; Republic of South Africa Department of Health 2004). Although resistance mutations will eventually emerge in the course of any long-term antiretroviral therapy regimen, the rate of the emergence of those mutations is related directly to the evolutionary cost of those mutations, where such cost is measured in terms of replicative capacity, and also to the rate of replication of HIV given the current drug regimen. In consequence, imperfect adherence to a drug regimen will allow HIV to replicate at a higher rate, which in turn will lead to a more rapid turnover of virus, and thus a faster evolution of resistance to the current drug regimen. Monitoring of adherence – and indeed, counseling to improve that adherence – is thus a key component of good HIV clinical care (Cote and Godin 2005; Paterson et al. 2000).

An additional set of concerns with HAART is that co-treatment with antibiotics for TB – common in many rollout settings including throughout South Africa – can complicate and affect both treatment regimens significantly. First, TB drugs, especially rifampicin, may interfere with the pharmacoavailability of PIs, NNRTIs

(including an approximately 25% reduction in concentrations of EFV), and some limited reports of the NRTI AZT (Chideya 2007; McIlleron et al. 2007; Patel et al. 2007), through rifampicin induction of cytochrome P-450 enzymes; while this can be avoided by use of an alternative formulation of rifampicin (rifabutin), this formulation is currently too expensive for use in rollout settings (McIlleron et al. 2007). Second, HIV infected patients appear to maintain lower concentrations of oral TB drugs including isoniazid, rifampicin, pyrazinamide, and ethambutol (Chideya 2007; McIlleron et al. 2007), especially in individuals with lower CD4 counts. Third, a review of relevant studies suggest that major TB drug toxicities, particularly hepatotoxicity, are more common among HIV-positive than HIV-negative patients, and among HIV positive patients more common among those receiving HAART, though results are not conclusive nor easily generalized to the developing world (McIlleron et al. 2007). TB-associated immune reconstitution disease (IRIS) may accompany rapid recovery of immune function with HAART in TB co-infected individuals, and may pose a serious threat to patient health and ability to maintain adherence to HAART and/or TB drugs (McIlleron et al. 2007). Last, some reports suggest that use of isoniazid, known to cause peripheral neuropathy on its own, may lead to greater incidence of severe peripheral neuropathy when used in combination with HAART, and especially the NRTI stavudine (Forna et al. 2007), a subject germane to aim 3 of this dissertation.

Rolling out HAART worldwide. The extraordinary burden of HIV in sub-Saharan Africa contrasts sharply with the resources available to cope with the HIV epidemic in this region. Despite early (and not evidence-based) fears that

adherence would be difficult to maintain in the developing world, trials and field experiences have demonstrated excellent adherence in resource-poor settings, especially where drugs are provided for free (Coetzee et al. 2004; Orrell 2005; Orrell et al. 2003; Oyugi et al. 2004), and have demonstrated convincingly that HAART can effect a substantial improvement of outcomes in the developing world (Coetzee et al. 2004).

There are serious challenges in providing HAART in the developing world, however. Some cohorts have had difficulty maintaining adherence over time (Byakika-Tusiime et al. 2005; Laurent et al. 2002), although some of this poor adherence may be attributable to the need to pay for drugs (Byakika-Tusiime et al. 2005). In addition, while viral load and CD4 cell count monitoring are commonplace in the United States and Western Europe and considered essential to clinical management in those settings, these tests are often financially or logistically beyond the means of many clinicians in many sub-Saharan African or other developed world settings (Cheruiyot Bii 2007; Colebunders et al. 2006). Likewise, toxicity of antiretroviral medications is a significant problem in the developed world, but lack of alternative medications may make toxicity a more serious problem in resource poor settings (Boulle et al. 2007). In many developing world settings, simple malnutrition and poverty are serious problems in their own right (Cheruiyot Bii 2007; Muula 2004; Ncayiyana 2004), compounding HIV-related diseases as well as potentially making toxicities more severe.

Despite these difficulties, the global community has begun to move towards providing HAART to all those in need in both sub-Saharan Africa and worldwide

(United Nations Special Assembly 2001), with the goal of universal access to HAART by 2010 (WHO 2007, 2006), an extension of "3x5 Initiative", which planned to put 3 million individuals in resource-poor settings on HAART by 2005 (WHO 2005). An estimated 2 of the 7 million people in need of HAART in low- and middle-income countries in 2006 received treatment, including 1.3 of the 4.6 million in need in sub-Saharan Africa(WHO 2007).

Rapid expansion of access to HAART has been enabled by international initiative such as the Global Fund (The Global Fund 2007) and the United States of America's President's Emergency Plan For AIDS Relief (PEPFAR) (USAID 2005) have committed and invested heavily to providing HAART to millions of individuals worldwide. PEPFAR, for instance, has committed US$15 billion to efforts to provide HAART to 2 million individuals in 15 target resource-poor countries, including South Africa (USAID 2005). Collectively, these efforts to expand HAART to all those in need are referred to interchangeably as a "rollout" and a "scale-up".

Rolling out HAART in South Africa. South Africa has one of the largest unmet needs for HAART in the world. An estimated 325,000 of the approximately 1 million people in need of HAART in South Africa were receiving HAART by December of 2006 (WHO 2007). South Africa is perhaps unique in sub-Saharan Africa in the relative robustness of its public health and general infrastructure, making it an ideal staging ground for many early HAART programs. Many of the technical and logistical difficulties of HAART provision in sub-Saharan Africa are less acute in South Africa due to greater healthcare infrastructure and resources, though of course substantial challenges remain, especially in providing HAART to

individuals in the lower-income “townships” and rural South Africa (South African National AIDS Council 2007).

While the South African response to the HIV epidemic has been complicated and slowed by political controversy (AVERT.ORG 2007; Smith and Novella 2007), nonetheless by 2004 the South African government had begun providing antiretroviral drugs free of charge. The “government era” of HAART provision began in the spring of 2004 in Gauteng Province (AVERT.ORG 2007, 2007), when the government dedicated US $62 million to the antiretroviral program rollout in 2004-2005 (UNAIDS 2006) and establishing guidelines for the national HAART rollout program (Republic of South Africa Department of Health 2004).

By the end of 2005, an estimated 200,000 individuals were receiving HAART in South Africa (Nattrass 2006), though many more remain in need. The government of South Africa has continued to show a commitment to HIV and AIDS palliative and preventive care over the past few years, issuing a new, extremely ambitious five year plan to combat and treat HIV and AIDS (Kapp 2007), which hopes to “appropriate packages of treatment, care and support to 80% of HIV positive people and their families by 2011” (South African National AIDS Council 2007). Indeed, the effort to initiate large numbers of patients on HAART is only part of the larger effort to fight HIV and AIDS in South Africa (WHO 2007); other aspects of this effort concentrate on prevention of new infections through education, poverty reduction, and addressing gender based violence; addressing mother-to-child transmission of HIV; encouraging research into new prevention and treatment

technologies; reducing HIV-related stigma; and supporting the human rights of those with HIV and AIDS (South African National AIDS Council 2007).

HAART provisioning programs in South Africa provide two full HAART regimens, one an NNRTI-based regimen, one a PI-based regimen, both with a double nucleoside analogue backbone; but, except where circumstances permit, no additional medications. First-line HAART typically comprises stavudine (d4T), lamivudine (3TC),and either efavirenz (EFV) or nevirapine (NVP), and second-line HAART comprises zidovudine (AZT), didanosine (ddI), and lopinavir-ritanovir (Kaletra®) (Republic of South Africa Department of Health 2004). In addition, some pregnant women receive lopinavir-ritanovir in place of nevirapine.

HAART rollout and scale-up in resource-poor settings: goals and challenges. The overall goal of HAART is to optimize the treatment benefit for patients, maximizing survival as well as quality of life. However, large scale rollouts of HAART in developing world settings – and in sub-Saharan Africa in particular, the region with the greatest need of HAART – pose a unique set of challenges due to several factors. Among these are an extremely high rate of new patients initiating HAART and high throughput per clinic more generally, not enough doctors to treat these high numbers of individuals, geographically diverse settings in which HAART must be provided, limited options for alternative HAART regimens, and possibly limited available of some laboratory tests (Harries et al. 2007; Muula 2004; Ncayiyana 2004; Van Damme, Kober, and Laga 2006; Carpenter 2006; McCarthy, O'Brien, and Rodriguez 2006).

Specific concerns, both public health and medically-oriented, have arisen from these general challenges, finding voice in numerous editorials and research reports. Among the most prominent public health fears are whether a rollout might lead to the emergence of greater amounts of HIV drug resistance (Blower et al. 2005; Ncayiyana 2004), and related concerns about public health infrastructure and the ability to ensure an uninterrupted drug supply to rural settings (Harries et al. 2007; Muula 2004). Others have voiced the concern that a HAART rollout might exacerbate existing socio-economically-based health inequalities (Egger et al. 2005). Underlying all these concerns, however, lies the more fundamental question: can we maintain a high quality of patient care in a rollout situation, despite the aforementioned obstacles and challenges?

The WHO has proposed a public-health approach to solving these problems (Gilks et al. 2006), emphasizing standardized treatment regimens, simplified clinical decision making, emphasizing the "four Ss" ("start; substitute for toxicity; switch treatment after failure; and stop, moving to end-of-life care"), simplified management of the threat of the emergence of HIV drug resistance, and decentralized provision of care, including shifting of many clinical responsibilities from doctors to mid-level health care workers including nurses and physicians assistants. In practice many of these guidelines are being followed, and problems solved, but many questions remain unresolved within the context of this approach. A report from WHO investigators notes that "Gaps in knowledge have limited the standardisation [sic] of some treatment approaches; others have not yet been adequately assessed." Among the key questions that remain unanswered are "How is failure best

defined?...[How can we] construct clear and simple protocols for switching therapy[?]" (Gilks et al. 2006)

Other unresolved challenges are hybrids of substantive and methodological issues. The nature of a rollout is such that there are rarely enough personnel or resources to follow-up all individuals who fail to return for a scheduled visit (though attempts are made); thus, there is likely to be significant loss to follow-up in most scale-up situations, a situation which may worsen with time, as total enrollment increases and resources are stretched ever thinner. At the same time, cohorts which track down a portion of those who do not return to clinic, typically find that between ½ and 2/3 of those patients on whom vital status can be ascertained are dead. For instance, in one study in Botswana, of 68 patients categorized as lost in passive follow-up, 40 were categorized as dead in active follow-up (Bisson et al. 2007); the Themba Lethu Clinical Cohort in Johannesburg, South Africa has found similar results. Likewise, in a conference abstract, Stafford et al. reported that Zambian patients who were lost to follow-up resembled deceased patients in measured covariates and time-to-event analysis (Stafford et al. 2007), and suggest that the majority of patients lost to follow-up ought to be included in mortality estimates.

Thus, using complete case analysis in a situation of substantial loss to follow up – characteristic of rollouts in general and our rollout in Johannesburg particularly – can result in biased estimates of mortality (Anglaret et al. 2004) or of risk factors for mortality. One possible solution to this problem is suggested by Martin et al. of the IEDEA Consortium in a conference abstract (Martin et al. 2007): a validation

sub-study, in which investigators try very hard to find the true vital status of a random sample of lost to follow-up patients; because the validated patients are a random sample, their vital status can be regarded as an unbiased estimate of the vital status of all those lost to follow-up.

There are two key problems with this approach: first, many people who become lost to follow-up – especially in longer term historical cohorts such as the TLCC – will be difficult to track down, and continuing stigma attached to HIV diagnosis may make cause of death or even vital status itself difficult to ascertain with confidence. Second, estimate from this technique will remain unbiased only if those who cannot be tracked down despite these efforts (22% in Martin et al.) are *themselves* an unbiased sample, which will be difficult to verify in the small sample size mandated by this technique. While this may well improve our estimates of raw mortality, it may introduce substantial bias in causal estimates for outcomes if reasons for unverifiable loss to follow-up correlate with reasons for the outcome of interest. In addition, while validation sub-study approaches may work well for unambiguous outcomes such as vital status, their use of such an approach in a study of an outcome such as virologic or immunologic success or drug toxicity is likely to introduce substantial misclassification of the outcome. We propose that inverse probability weighting and multiple imputation to solve the problem of informative censoring more robustly.

A lesser methodological issue – addressed by some reports but not others – is that of repeated outcomes measures. Most reports on rollout outcomes concentrate on six-month increase in CD4 count; it is clear that twelve-month

increase is also an important measure of treatment success. Because CD4 measures are strongly correlated within individuals, appropriate models must be used to account for inter-individual correlation. In addition, there is good evidence to support the idea that CD4 count rises more rapidly in the first six months of therapy than in the next 12-18 months; and then plateaus around 24 months after HAART initiation. Thus, modeling CD4 count over time requires both random effects models to account for correlations within individuals, as well as static change points to account for changing slopes of CD4 count over time.

Last, many rollout reports have taken a somewhat narrow, medical view of public health problems. For instance, it is obviously important to characterize the reasons for drug substitution, especially those reasons which relate to toxicity. Enumerating the particular risk factors for a particular drug toxicity – say, lactic acidosis associated with stavudine – allows clinicians to monitor obese women (at particular risk of the condition) more closely. But because lactic acidosis is rare, the positive predictive value of obesity and female gender for the condition is likely to be quite low; thus, we will not necessarily achieve better outcomes at the population level by removing stavudine from the first line regimens of all obese women. Another question, perhaps better for the whole population, is "who is likely to require a stavudine substitution within 18 months of initiation on the drug?" This question, answered well, may lead us more quickly to an ability to consider customization of initial HAART regimen in those at high risk of substitution.

Thus, key data analysis challenges of rolling out HAART on the massive scale required to meet the stated goals of the WHO and other organizations include

evaluating overall effectiveness of HAART in settings with continuing stigma of HIV infection, significant loss to follow-up, and a lack of resources to track down patients who do not return to clinic; lack of options, resources, and time to enable the customization of HAART regimens to individuals; the related problems of drug toxicity compounded by poor nutritional status of HIV-infected individuals; basic problems of maintaining quality of care while expanding treatment access as rapidly as possible; and possible local constraints on laboratory test availability.

REPORTS OF HAART ROLLOUTS

Approach to literature review. The main search to find evaluations of HAART rollouts in the developing world was performed in the NLM PubMed database with the search term [(“scale-up” or “scale up” or “roll out” or “rollout” or “rollout” or "scaling up" or "scaling-up") AND (“antiretroviral therapy” or “antiretroviral therapy” or “ART” or "HAART”) AND ("HIV" or "AIDS")]. This search returned 120 articles, which were reviewed by hand for inclusion. Other searches for other topic areas were performed similarly, with search terms such as “drug substitution antiretroviral toxicity” (18 results), “stavudine substitution antiretroviral toxicity” (5 results), and “stavudine switching antiretroviral toxicity” (9 results). Additional references were found by examining relevant references from articles already found, as well through use of the PubMed “related articles” search option. Finally, abstracts from recent relevant conferences were searched and reviewed, including 14th Conference on Retroviruses and Opportunistic Infections (CROI 2007) and the 4th

International AIDS Society Meeting on HIV Pathogenesis and Treatment (IAS 2007), using similar search strategies.

The main literature review in each of the following sections includes only reports from sub-Saharan Africa, which is the focus of our investigations and of the bulk of scale-up efforts. In addition, in the main literature review we wish to address only HAART rollouts or efforts to scale-up HAART access. Rollouts are characterized chiefly by very high volume of new enrollments of patients – while there is no absolute rate of new patient enrollment and initiation that meets this definition, a review of the literature suggests that defining a rollout as a cohort which enrolls approximately 50 new patients per month for 9 or more months is an acceptable lower limit for defining a rollout.

After the main literature review is completed, selected additional studies are reported, including conference abstracts and reports from non-sub-Saharan rollouts, where they provide appropriate and valuable context.

Reports of HAART rollouts: methodological challenges and substantive gaps. Articles relevant to evaluating outcomes of HAART in sub-Saharan Africa in a scale-up context are reviewed below. It is worth noting that in most of the studies reviewed below, the baseline characteristics of those initiating HAART are similar: more than half and as many as 70% of those initiating HAART are women; many individuals have a body mass index (BMI) < 20, CD4 counts less than 50 cells/mm3, and/or are in WHO stage III or IV HIV disease. The vast majority of patients in these studies initiated HAART with d4T or AZT, 3TC, and EFV or NVP.

In most settings, as well, the overall conclusions are the same; a summary of the major findings of outcomes from key studies is shown in Tables 1.1, 1.2, and 1.3. Rates of viral suppression are typically 75%-85% at 6 months or earlier among those still in therapy and with viral loads available for analysis; median CD4 gain by month 6 (among those remaining in therapy with CD4 counts available) is approximately 100-110 cells. More qualitatively, nearly all studies found similar associations among baseline risk factors and key outcomes. For instance, for the outcome of death, baseline characteristics of low CD4 count, low BMI, WHO stage IV (and to a lesser extent III) and – in some studies – male gender – was found to be associated with increased hazard or risk of death consistently in nearly every study that examined such associations.

There is substantially more variation in mortality and loss to follow-up (as shown in Table 1.1), with values ranging from 0%(Severe et al. 2005) to 19%(Ferradini et al. 2006) depending on setting and rollout design (though nearly all studies saw the majority of deaths occur in the first three to six months of follow-up). The issue of loss-to-follow-up – which was often found to be associated with key characteristics predictive of death in these studies including low CD4 count and low BMI at baseline – is a key issue for a large scale rollout of HAART. Furthermore, it seems likely that if differential and informative loss to follow up were accounted for, six and twelve month CD4 gains as well as other outcomes (such as weight gain) might also change. Accordingly, we will review key methodological aspects of published literature evaluating rollouts of HAART (or equivalent, high-capacity

HAART provisioning situations) in the developing world, to see where bias may be introduced into estimates of key outcome parameters.

Outcomes in rollouts: peer reviewed reports. The following will review only peer-reviewed reports of rollouts of HAART in the sub-Saharan Africa; conference abstracts and additional peer-reviewed studies from non-sub-Saharan Africa locations are reviewed below.

Stringer and colleagues (Stringer et al. 2006) initiated 16198 individuals onto HAART in 18 clinical centers in the Lusaka district of Zambia over approximately 18 months, with funding from PEPFAR, the Global Fund, and other sources, a mean enrollment rate of 50 patients/month/site initiated. All drugs and clinical care were provided for free. There was significant loss to follow up in this cohort, with over 3400 of the 16198 individuals recorded as more than 30 days late for a pharmacy appointment. Enrollment characteristics were typical of a rollout setting, and 11% of patients had active tuberculosis at baseline; outcomes are described in Table 1.1 above. Median follow-up time was less than six months.

Stringer et al. used Cox proportional hazards models, and "accounted for imperfect ascertainment of death increasing over time by including a term to adjust for calendar time at the initiation of therapy", noting that follow-up teams may have less opportunity to trace individuals who go missing if they disappeared later in the 18 month assessment period. They used mixed (random effects) models with individual random intercepts to assess CD4 gain over time. They performed sensitivity analysis indicating that that "some predictors of failure or death also

predicted loss to follow-up", concluding that "loss to follow-up likely represents a heterogeneous mixture of individuals, including some who have died".

The main limitation of this work was the authors' approach to analyzing individuals lost to follow-up was somewhat naïve; given that 21% of patients became lost during follow-up, there is an enormous potential for bias in estimates of mortality and relative risks of various covariates for outcomes. As noted previously, this problem can be addressed with inverse probability weighting (see Methods section), both of which allow data analysts to estimate the outcomes without bias in the presence of informative censoring given an assumption of no unmeasured predictors of both outcome and missingness. While they concluded that "some" of those lost to follow-up had died, Cox proportional hazards models assume that those who are missing died in the same proportions of those who died; that is, that perhaps 7% of lost individuals had died. Our experiences in the Themba Lethu Clinical Cohort suggest strongly that the proportion of lost individuals who have died is as high as 50%, suggesting that – sensitivity analysis aside – their estimates of death are significant underestimates.

Other shortcomings of the work included relatively short median follow-up time (207 days among those who initiated ART); lack of control for correlation within clinical sites with either random effects or GEE models; lack of control for duration of active TB treatment at baseline; and that some categories used in their analyses do not appear to coincide well with internationally accepted standards of (for instance, BMI < 16 is not a standard category). However, in most respects aside from dealing

with informative censoring, the authors were admirably methodologically responsible.

Libamba and colleagues (Libamba et al. 2005) report on 13183 patients (12527 of them adults) in Malawi who were initiated on antiretroviral therapy at 24 health facilities over approximately 10 months, a mean enrollment rate of slightly more than 50 patients/month/site. They note that approximately as many people who were found to be dead were found to be missing or lost to follow-up (in both cases, 8% at one year). They reported secondary outcomes (including quality of life measures) among those alive and still on therapy, excepting those for whom there was poor or incomplete information, or where clinics were too busy to record the information. They reported no results from any kind of multivariable analysis.

The lack of multivariate analysis is obviously the largest methodological deficit in this study. Such an analysis would need to control for correlations within site of treatment, as well as for site-specific effects and characteristics – for instance, the authors noted that patients had to pay for medications only in some clinics. In addition, their reports of secondary outcomes could be systematically flawed if patients at busier clinics are less often reported and have worse outcomes.

Fairall et al. (Fairall et al. 2008) and colleagues presented a sophisticated analysis of a population of 14267 individuals of whom 3619 received HAART during follow-up, initiating about 48 patients per site per month on HAART. Fairall et al. used marginal structural models (see Methods) to examine the effect of HAART and (separately) cotrimoxazole therapy on HR for death in their cohort, finding that both markedly reduced HR of death, while use of HAART increased weight and CD4 from

baseline, and was protective for incidence of tuberculosis. While this study reports some high level outcomes such as CD4 count only in passing, it is a rigorous and thorough study of survival and other outcomes in the context of HAART in the developing world.

In a model including both individuals who initiated HAART and those who did not, Fairall et al. found a reduced hazard of mortality among individuals who had TB at baseline in a model which controlled for a number of other important factors (HR = 0.71, 95% CI 0.56-0.90). This is a counterintuitive finding, and the authors suggest that, despite the fact that it was reported as applying to all individuals (Table 1.1), this reduced hazard of mortality associated with baseline TB "was only true for patients who did not receive HAART." They also suggested that the finding might be due to selection bias due to left truncation, in that the sickest patients with tuberculosis may have died before entering care.

Wools-Kaloustian (Wools-Kaloustian et al. 2006) and colleagues reported on 2059 individuals initiated on HAART between November 2001 and February 2005 in Kenya, a mean of 52 patients/month enrolled. They reported 5.4% mortality and 24.5% loss to follow-up; other outcomes are summarized in Table 1.1. They also assessed weight gain directly as an outcome. The authors used Kaplan-Meier methods and log-rank tests, and Cox models, for survival outcomes; they used random effects models to model change in square-root of CD4 count with a change-point mixed model (apparently implemented using a constant or linear b-spline) with within-subject random intercepts and slopes. They used similar models for

longitudinal models of change in weight. The authors excluded all pregnancies from their analysis, and median follow-up time was 40 weeks.

The methods used for change in CD4 count and weight were appropriate, accounting for the correlation of CD4 counts within individual, while allowing slope of CD4 increase to change with time. This represents directly the way that CD4 changes are typically large in the first three to six months of HAART, but rate of change declines thereafter, eventually flattening out around 2 years post-HAART initiation (Chu et al. 2005). Another strength of this paper is that it assessed weight gain directly as a clinical outcome. This is important because in many rollout settings, CD4 monitoring may not be available; as such, weight may be an important means of assessing success on HAART.

However, this study suffered from large loss to follow-up over the study period in comparison to total recorded mortality (24.5% vs. 5.4%); this strongly suggests that without imputations of outcomes for those lost to follow-up, conclusions drawn from survival analysis of these data are likely to be biased. Additionally, they note that self-pay patients in the cohort did not have all laboratory testing done at all time-points. A large multi-center study (Braitstein et al. 2006) saw a clear association between the need to pay for drugs and worse outcomes of HAART; if those in the Kenya study who had to pay for medications also had worse outcomes, this would result in a bias of laboratory outcomes (CD4 gain, viral suppression) that would make the program look as if it had performed better than it was indeed performing, a limitation acknowledged by the authors.

Ferradini and others (Ferradini et al. 2006) published on 1308 patients in a scale-up in rural Malawi, enrolled over a period of several years (rate unclear, but apparently accelerating). They found 19% mortality at a median of 36 weeks of follow-up, 7% loss to follow up, and median CD4 gains of 138 and 176 at six and twelve months. Their analysis was intention-to-treat, and they analyzed both known deaths as an outcome and the combined outcome of dead or lost to follow-up. Otherwise their methods were straightforward, making use of Cox proportional hazards models, logistic regression, and similar methods. Virologic outcomes were assessed among those still in therapy at 6 months. Last, they used the outcome of virologic success to validate their adherence measures.

There were a number of methodological shortcomings of this work. Their Cox models to find important predictors of early mortality (within six months) did not account for informative loss to follow-up – since early losses to follow up is exactly those most likely to be related to death, their results are likely to be biased in this analysis. Similarly, if early virologic failure is related to death, then assessing virologic outcomes only among those still in therapy will provide an overestimate of true virologic outcomes. They failed to report or control for tuberculosis in any analyses. Last, they apparently performed no modeling of predictors of CD4 gain, an important outcome.

Hawkins et al. (Hawkins et al. 2007) reported on 1286 patients enrolled over 24 months from 2004 to 2006, a mean enrollment rate of 54 patients/month. They observed an extremely low risk of death (1.1% in median 350 days follow-up), but 435 (34%) patients were lost to follow-up. They saw mean gains in CD4 counts of

169 and 208 cells/mm3 at 6 and 12 months, respectively. They used Cox models, including switching therapy as a time-varying covariate. The authors did not attempt to model LTFU. HAART was not provided for free in this cohort.

There are two serious methodological issues with this analysis. First, the extremely high loss-to-follow-up rate, due in part to passive patient reporting systems which they note are characteristic of low-resource settings, undoubtedly explains the very low observed risk of death; in addition, this loss to follow-up rate likely contributes to higher than typical mean CD4 gain at six and twelve months by selecting for those with better immunologic response to HAART. Second, using traditional Cox models to account for time-varying therapy – which is confounded in a time-varying way by CD4 count, a confounder which is also affected by prior exposures – will result in biased estimates of effect if proper methods (e.g., inverse probability of treatment weighted marginal structural Cox proportional hazards models) are not utilized (Hernan, Brumback, and Robins 2000; Robins, Hernán, and Brumback 2000).

Lawn and colleagues have published a series of papers describing a rollout cohort in Cape Town, South Africa, including the burden and incidence of tuberculosis and effect of tuberculosis on HAART outcomes (Lawn, Badri, and Wood 2005; Lawn et al. 2006), CD4 gain among those who initiated HAART by baseline CD4 counts (Lawn et al. 2006), early mortality (Lawn et al. 2005), and reasons for both mortality and "nondeath losses" (Lawn et al. 2006). Separately, **Bekker et al.** (Bekker et al. 2006) reported outcomes on many of the same individuals.

In a paper examining early mortality (Lawn et al. 2005) Lawn et al. reports that 66% of first-year deaths in this cohort occur within the first 90 days of HAART initiation. They found that major predictors of death included WHO clinical stage 3 or 4 disease and a baseline CD4 count < 50. In contrast to other studies, they did not find any significant effects of age or sex on risk of death. They used (what appears to be) generalized estimating equations-based maximum likelihood regression accounting for inter-individual correlation with an exchangeable working correlation structure robust standard errors to model associations between mortality rate and treatment status *as treated*.

The authors did not account for informative loss to follow-up in this study, possibly because loss to follow-up was quite small (4%). The as-treated nature of this analysis makes it unique among papers reviewed here; all other reports have taken an intent-to-treat approach. While this alternative approach is admirable in that it may give us a better estimate of the true effect of antiretroviral therapy on mortality, an intent-to-treat analysis may well give us a better idea of the outcome of individuals initiated on HAART in a rollout setting, taking into account all management available to them. In addition – as with Hawkins et al., above – because the relationship of treatment status to mortality may be confounded by CD4 count in a time-varying manner, and CD4 counts themselves may by affected by treatment, then these results may be biased when the stratification techniques typical of standard maximum likelihood regression are used rather than inverse probability of treatment weighting techniques used with marginal structural models

(Cole et al. 2003; Hernan, Brumback, and Robins 2000; Robins, Hernán, and Brumback 2000).

Lawn and colleagues also analyzed CD4 count increases in pre-specified time-spans (0-16 weeks, 16-32 weeks, and 32-48 weeks), thus accounting for changing slopes over time (Lawn et al. 2006). In addition, they examined a dichotomous "immunologic non-response" CD4 cutoffs of 50 cells/mm3 gained by 48 weeks, as well as dichotomous cutoffs of CD4 count ≥ 200 cells/mm3 and 500 cells/mm3 by 48 weeks. They used linear regression to examine rate of CD4 increase within time period, and logistic regression to examine factors predicting success reaching the dichotomous targets. One potential drawback of this methodology is that it fails to account for intra-individual correlation, which might be addressed using a random effects or GEE model.

In addition the authors do not account explicitly for the possibility of differential CD4 recovery rates among those lost to follow-up in this cohort analysis. However, they do comment on the potential bias that might result from such informative censoring, arguing that resulting bias is likely to be small because (among other things) the majority of deaths that occurred in this cohort happened earlier (and thus are unlikely to be related to immune reconstitution) and also that, although "WHO clinical stage of disease is the strongest predictor of death in this cohort…this variable was not associated with CD4 cell responses." If those who died early were profoundly immunocompromised, and such immune status affected their CD4 recovery rates, then estimates of CD4 recovery will be biased. However, as the primary concern with individuals who do not receive follow-up treatment may

be their overall survival, the as-treated assessment of CD4 count among only those who remained in the cohort may be useful in its own right for assessing (for instance) whether CD4 gain is adequate among those who remain in care. For these purposes, this modeling approach is likely adequate. However, if we were to decide that the CD4 gain in the entire cohort was important, then modeling the combined risk of death and CD4 reconstitution is a situation which might be better served by a zero-inflated (or hurdle) linear model, where the data analyst first models the dichotomous outcome of survival, and then subsequently models the increase in CD4 cell count among the survivors.

Separately, logistic regression may have been a poor choice for predicting virologic success, because odd ratios derived from logistic regression are known to overestimate the (more-interpretable) risk ratios when events are common, and the dichotomous events being modeled here are fairly common (for instance, 49% of patient with baseline CD4 count < 50 failed to achieve a CD4 count of 200 cells/mm3 at 48 weeks).

Last, **Bekker and colleagues** (Bekker et al. 2006) reported on 1139 individuals initiated on HAART over three years in Guguletu South Africa, with an initiation rate of 58 patients/month in the last year of enrollment. Their reporting of outcomes were quite limited, noting only that virologic suppression was assessed in those who were not dead or lost to follow up. This cohort experienced relatively little loss to follow-up: only 3% of patients were lost, compared to 7% who were confirmed dead. In addition, they compared outcomes across all three years of enrollment using various non-modeling statistical techniques. Compared to the first

two years of enrollment, during the third year fewer patients presented in WHO Stage IV and patients presented with higher median CD4 count, suggesting that patients were presenting earlier in disease course.

Due to the paucity of methods and more-advanced outcomes presented, it is difficult to comment further on this report.

Outcomes in rollouts: reports from conference posters and abstracts. Very little information is available on methods used in any of these reports, due to their being derived from brief conference reports, from the 4th International AIDS Society Conference on HIV Pathogenesis, Treatment and Prevention and the 2007 Conference on Retroviruses and Opportunistic Infections.

El-Sadr and colleagues (El-Sadr et al. 2007) reported at the 4th International AIDS Society Conference on HIV Pathogenesis, Treatment and Prevention (2007) on 171259 patients enrolled at 151 sites throughout seven African countries (approximately 1100 individuals per site); unfortunately, because these results are only available in abstract form, a close evaluation of the methods used here are not available. In addition, the authors reported significant variation in results by country.

Marlink et al. (Marlink et al. 2007) reported at CROI 2007 on 134120 people who initiated HAART in Côte d'Ivoire, South Africa, Tanzania, and Zambia (apparently including those who Stringer (Stringer et al. 2006) reported on below). After 28 months of follow-up, 93312 individuals remained in care, with 5% recorded deaths, 7% transferred or lost to follow-up, and 5% unknown. There was little additional information given by Marlink et al. on methods or definitions.

Meloni et al. (Meloni et al. 2007) reported in a poster at CROI 2007 on 9070 people who initiated HAART in six different sites Nigeria. After 6 months of follow-up, median CD4 had increased from baseline 128 cells/mm3 to 252 cells/mm3, and between 80 and 90% of patients were still in care (depending on site). After 12 months, median CD4 was 286 cells/mm3, and between approximately 73 and 92% of individuals were still in care. The authors reported significant by-site heterogeneity of results.

Mosoko et al. (Mosoko et al. 2007) reported on 2920 individuals in Cameroon, with fairly low rates of "alive and in care" outcomes at follow-up (see Table 1.2). They found that age < 30, male gender, large distance (>150km) from clinic, and self-pay status for ARV drugs were associated with increased hazard of the combined dropout or death outcome. However, HAART was not provided for free in this cohort.

Outcomes in rollouts: other peer reviewed reports (non-rollouts)

The following reports are not from sub-Saharan African rollouts (though some are from sub-Saharan Africa, and some might arguably considered rollouts), but are helpful in assessing the wider context for this work.

Calmy and colleagues from Médecins Sans Frontieres(MSF) (Calmy et al. 2006) report outcomes of fixed-dose combination HAART (comprising d4T-3TC-NVP) in 6861 patients at 21 MSF sites in Africa (~75%) and Asia (~25%). They saw comparable proportions of deaths and individuals lost to follow up in a median of 4.1 months of observations. In the sub-cohort with at least 12 months of exposure, there was substantially more loss to follow up (from 4.8% overall to 9.3%). The

authors caution against over-interpretation of their data due to short (median 18 weeks) follow-up times overall.

Calmy et al. used sensitivity analysis to ensure that their Kaplan-Meier estimates of risk factors were well-bounded. They first used death as the outcome; and secondarily used death or lost to follow-up as an outcome, thus bounding the effect of loss to follow up on causal estimates of risk factors. This approach has the advantage of giving us a worst-case estimate; but it does not help the reader assess what values are more likely within the range of best case to worst case scenarios. It is unclear how much influence losses may have on accounting for those lost to follow-up.

The ART-LINC collaboration (Antiretroviral Therapy in Lower Income Countries) is an international collaboration of treatment cohorts (Dabis et al. 2005). The goal of ART-LINC is

> to define the prognosis of HIV-1 infected patients treated with HAART in resource-limited settings; (ii) to compare the experience between different settings, delivery modes and types of monitoring; and (iii) to compare prognosis in resource-limited settings with that observed in industrialized nations. Importantly, both individual and programme level characteristics are of interest in this context." (Dabis et al. 2005)

They found an increased hazard of death in the first six months of HAART comparing low to high income countries (adjusted HR 4.3, 95% CI 1.6-11.8), and that in low-income countries the adjusted hazard ratio associated with free treatment compared to not-free treatment was 0.23 (95% CI 0.08-0.61).

None of the developing world sites appeared to be rollout sites, which may limit generalizability to HAART rollout in sub-Saharan Africa. The authors used random-effects Weibull regression to model mortality with random effects accounting

for different treatment programs, with bootstrapped confidence intervals; this is a flexible proportional hazards model appropriate for highly stratified data (with strong heterogeneity between programs) (Carlin and Hodges 1999).

In addition, the authors used multiple imputation techniques to fill in missing values of WHO disease stage, allowing them to use much more data than otherwise would have been available in a complete case analysis. Last, the extremely strong, significant hazard ratio for death associated with free treatment compared with treatment-for-fee (0.23, 95% CI 0.08-0.61) in low income countries suggests that studies which do not either report or control for need to pay for HAART in rollouts will be difficult to interpret.

Koenig, Léandre, and Farmer (Koenig, Leandre, and Farmer 2004) report on scaling-up directly observed HAART in rural Haiti, in which approximately 1050 individuals were enrolled on HAART in the first year of scale-up. Koenig et al. saw almost no mortality or loss to follow-up among the first 100 of the 1050 individuals receiving directly observed therapy (DOT)-HAART, and living in the area served by community health workers (also called *accompagnateurs*). While these results are very impressive in light of substantially higher mortality seen in other scale-up efforts, several aspects of these data may limit their generalizability in both data and implementation practices. First, while there was no mortality in this group, the background of malnutrition and co-infections may differ from Haiti to other developing world locations, and in particular may differ from conditions in Africa, where the scale-up is most important in terms of sheer numbers of individuals in need. Second, Koenig et al. report that individuals who lived outside of the area

served by *accompagnateurs* were not eligible for DOT-HAART, but that 29 in 100 evaluated patients in this group died in the first year. This opens the question of whether a less personnel-intensive scale-up program – one that did not use *accompagnateurs* and DOT-HAART – might have reached more people and minimized deaths overall, or whether the *accompagnateur*-focused effort is useful in both ensuring good outcomes among those who start HAART and limiting onward transmission or the transmission of drug resistant HIV. We conclude that at present the circumstances of the Haitian rollout are not easily generalizable to African rollouts.

Severe and colleagues (Severe et al. 2005) reported on initiation and outcomes of HAART in 1004 individuals in Haiti beginning in March 2003 over 14 months, a rate of approximately 71 individuals per month (Severe et al. 2005). Individuals were initiated on therapy when their CD4 counts were below 200 cells/mm3, or if they had an AIDS defining illness, and were treated with AZT-3TC-EFV, and d4T was sometimes substituted for AZT. All clinical care, treatment, and antiretroviral drugs were provided free of charge. CD4 tests were performed regularly, but viral load measurements were not available clinically. Only 8% of adult and adolescent patients (hereafter, adult patients) had tuberculosis in this study, limiting this study's generalizability to sub-Saharan Africa. Analysis was intent-to-treat and used straightforward regression methods such as Cox proportional hazards modeling for multivariable survival analysis.

Severe et al. reported an overall one-year survival of 87% using a Kaplan-Meier estimator. Approximately 7% of individuals were lost to follow-up by analysis;

however, the authors report that individuals lost to follow-up most often because they returned to their villages from the urban center where the clinic was located, and did not differ from those not lost to follow up by any baseline characteristics including age, sex, CD4 T-cell count, body weight, or stage of HIV infection, suggesting that censoring these individuals will not seriously bias the results of this study.

However, patients who return to their villages might be better classified as "transferred out" rather than "lost", in that we indeed have information on them. It is possible that individuals truly lost to follow up – a subset of those identified as "lost" in this study – have different outcomes, and that the mixing of the populations of truly lost and transferred out patients causes Severe to mistakenly conclude that those who are lost are more like those who are not lost than is truly the case. However, the overall percentage of those lost to follow up is relatively low, and so such an effect is unlikely to substantially change mortality estimates.

Etard et al. (Etard et al. 2006) reports on 7-years of follow-up (median 199 weeks of follow-up) in 404 patients. This study is too small to be properly considered a rollout, and thus the methodological challenges of analysis are somewhat different. Their follow-up was quite good (7 individuals lost to follow-up in the first year). It is worth noting though, that Etard et al. took the same "sensitivity analysis" approach to loss to follow-up, assuming that the patients were dead to bound possible mortality estimates.

One additional study worth noting here (not included in Table 1.3) is by **Van Cutsem et al.** (Van Cutsem et al. 2007), reviewing five years of follow-up in a public-

sector antiretroviral clinic in Khayelitsha, Cape Town, South Africa. They enrolled 2565 patients (70% female) over the five years, or approximately 40 new patients per month (compared to 200 per month in Themba Lethu). Six month mortality declined from 12.4% in 2001 to 5.4% in 2005; some of that may be attributable to loss to follow-up, which increased from none to 3% in that same time period – very low losses by rollout standards. At 54 months, they estimated 78% (95% CI 74-82) of patients remained in care. They do not report tuberculosis co-infection rates or rates of switching, and do not describe their methods in depth in this brief conference report.

Summary of literature review on outcomes of rollouts and other large HIV cohorts. Many reports have indicated that rollouts can achieve significant success in improving laboratory markers and clinical outcomes of HIV disease status, even while rolling out HAART quickly and on a massive scale. Despite widely varying estimates of mortality among studies, there is a qualitative consensus among these studies that low CD4 count, low BMI, poor adherence to HAART, and HAART-for-pay all contribute to early mortality. Likewise, there is generally qualitative and quantitative agreement about median CD4 cells gained among individuals initiating HAART within six and twelve months and about percentage of individuals who can be expected to suppress virus within six months of initiation.

Yet taken together, these studies suffer from serious flaws: they lack extensive documentation of outcomes after six months; they are frequently statistically and methodologically naïve, especially with regards to dealing with informative loss to follow-up, a major challenge in a rollout setting, and thus give

inaccurate pictures of mortality, other key outcomes, and of estimates of predictors of those outcomes; many reports ignore tuberculosis co-infection (see below); and while most of these studies achieve qualitative consensus as mentioned, there are serious disagreements among them as to the magnitude of effect measures relative to each other, due to differing methodologies, poor follow-up, and other issues.

Our research on the Themba Lethu Clinical Cohort will remedy many of these shortcomings. We have documentation of outcomes greater than 6 months, in some individuals as much as 3 years of follow-up. Large sample size will allow us to control for many confounders of these outcomes simultaneously while maintaining good power. And will deal with loss-to-follow-up in a methodologically more-sophisticated way that will help us estimate mortality and relative influences on mortality both more accurately and more precisely.

TUBERCULOSIS: BIOLOGY AND TREATMENT IN THE CONTEXT OF HIV

Mycobacterium tuberculosis. Tuberculosis (TB) is the infection of the lungs (pulmonary TB) or other organs (miliary TB) with the aerobic bacterium *Mycobacterium tuberculosis*; here we will focus exclusively on pulmonary TB, hereafter TB. TB is typically transmitted by direct inhalation of aerosolized droplets produced by an infectious individual, often through coughing. Bacteria which reach the lungs of an exposed individual commence primary infection, which is contained – though not eliminated – by the immune system, particularly by alveolar macrophages, in about 95% of cases; the remainder develop active disease. In the

majority of the remaining 95% of individuals, the mycobacteria remain contained by the immune system for the rest of the infected individual's life, without incident. However, in about 5% of total individuals infected, TB will reactivate later in life, often at a time of immune compromise (such as older age) (Weinberger 2004; WHO 2007). In addition, TB can disseminate through the blood or lymph, and a small number of people develop TB in organs other than the lungs or miliary (disseminated) TB (Weinberger 2004), but these diseases are quite rare among non-immunocompromised individuals. However, in HIV-infected individuals, the probability of developing active TB of any sort is much higher, due to the attendant immunosuppression (Weinberger 2004; WHO 2007), with some estimates of annual incidence going as high as 15% per year (Meya and McAdam 2007).

Clinically, TB presents with symptoms including coughing (and coughing up blood), weight-loss (hence the classical nickname, consumption), night-sweats, and general fatigue. Diagnosis of TB is somewhat difficult; the purified protein derivative skin test (PPD) cannot distinguish active disease from previous exposure (Weinberger 2004). Chest x-rays are more useful, but can still be likewise non-specific, failing to distinguish inactive from active disease. Acid-fast staining of M. tuberculosis can also diagnose TB, but is non-specific to other Mycobacteria, and can be insensitive when the number of bacteria in the lungs is small. Smear-positive diagnosis is useful for identifying the most-infectious cases of TB, but will not be able to find every clinically significant case. Most sensitive, and indeed, definitive is to culture the TB bacterium itself, from sputum samples coughed up by a potentially infected individual, a process which requires appropriate laboratory infrastructure

and in addition takes six weeks to return results and can be difficult in very sick individuals (Weinberger 2004).

Treating tuberculosis. Drug susceptible TB is generally treated with a six to nine month course of antibiotics. In South Africa, for existence, this course involves two months of rifampicin, isoniazid, pyrazinamide, and ethambutol (WHO 2007), followed by four months of rifampicin and isoniazid alone; in other settings in sub-Saharan Africa, the drug course may include only rifampicin and isoniazid. TB pharmacotherapy poses significant challenges when co-administered with HAART, for reasons of pill burden, adherence, and overlapping drug toxicities (Dean et al. 2002; Subbaraman et al. 2007) (relevant issues of TB/HIV drug toxicites will be dealt with in literature review for Specific Aim 3, below). In addition, especially in resource challenged settings such as South Africa, more and more cases of TB are resistant to these first-line TB drugs. Though the emergence of drug resistant TB (especially multiply-drug resistant, or MDR-, TB) is the immediate and direct consequence of the onward transmission of TB in individuals who did not complete their (arduous) course of medication, such emergence is perhaps more properly seen as the consequence and signal of a breakdown in public health systems and infrastructure (Van Rie and Enarson 2006). The recent emergence of XDR-TB – resistant to fluoroquinolone and another, injectable second-line drug – makes treatment still more difficult (Van Rie and Enarson 2006).

Tuberculosis and immune reconstitution inflammatory syndrome. One potentially important complication of TB in patients initiating HAART is immune reconstitution inflammatory syndrome (TB-IRIS), in which tuberculosis paradoxically

worsens or is “uncovered” soon after initiating HAART due to an immune “dysregulation” and inflammatory response (McIlleron et al. 2007; Murdoch et al. 2008; Murdoch et al. 2007). While some investigators have reported incidence of IRIS (due to TB or other opportunistic infections) as high as 33% of patients in some cohorts and settings (Murdoch et al. 2007), a well-done prospective study in Johannesburg (and similar to the patient population of Themba Lethu Clinic in several significant ways) found a substantially lower incidence, of 10.4% total IRIS and 4% TB-IRIS by six months of follow-up (Murdoch et al. 2008).

Incidence of TB-IRIS may be related to length of time from initiation of TB therapy to initiation of HAART (Murdoch et al. 2007; Navas et al. 2002). It is partly for this reason that the SA National ARV Treatment Guidelines recommend waiting at least two weeks between initiation of TB and HIV therapies even in very sick patients (Republic of South Africa Department of Health 2004). The only other strong risk factor for development of IRIS is of lower CD4 count at baseline and at IRIS diagnosis (Murdoch et al. 2008).

While there are fears that IRIS may result in significant morbidity and mortality, preliminary reports of studies suggest that IRIS is more of a clinical nuisance than a life-threatening condition (Lawn 2008). Indeed, the prospective clinical study referenced earlier noted that only two deaths were attributable to IRIS out of 22 cases, only 1 died (from Cryptococcal meningitis); given that individuals with IRIS had substantially lower CD4 counts than those without IRIS (median 79 vs. 115), it is unlikely that IRIS would fall out of a multivariate model for mortality as an

independent risk factor. It seems unlikely, therefore, that IRIS is an important risk factor for mortality.

Clinical management of tuberculosis in South Africa. There are four main drugs used in treatment of TB in South Africa, isoniazid, rifampicin, pyrazinamide, and ethambutol. Isoniazid is most effective against metabolically active TB bacilli; rifampicin can kill semidormant bacilli; pyrazinamide kills bacilli inside cells such as macrophages; and ethambutol is bacteriostatic, and interferes with growth of the tuberculosis cell wall (WHO 2004). Standard treatment for tuberculosis in SA is a six month course of chemotherapy, beginning with two months of isoniazid, rifampicin, pyrazinamide, and ethambutol, and then continuing with four months of isoniazid and rifampicin (WHO 2004). If the patient has relapsed, defaulted, or failed treatment, streptomycin (which is typically injected into the muscle, which makes it painful and difficult to administer to severely wasted TB/HIV patients) is added to the combined drug regimen, which is repeated for a longer course (WHO 2004).

In South Africa, presenting for TB treatment is often an HIV patient's entry into the health care system; many patients therefore present for initiation of antiretroviral therapy having already initiated TB therapy. This is in line the South African National ARV Treatment Guidelines, which state that a patient should delay two months (or a minimum of two weeks if the patient is extremely sick) between initiation of TB therapy and HAART, and warns further that patients may be at substantially higher risk several side-effects shared between TB and HIV therapies, that the high pill burden of the dual pharmacotherapy may cause problems with adherence to both regimens, and that immune reconstitution may be a problem

(Republic of South Africa Department of Health 2004). However, as noted previously, TB-IRIS may be more of a clinical nuisance than serious complication of therapy.

TB and HIV in the developing world. Despite the fact that it is largely treatable, TB is the most common cause of death among HIV-infected individuals in the developing world (Meya and McAdam 2007). Accordingly, a major component of the international effort to stop the spread of TB focuses co-infection issues (WHO 2007). The burden of pulmonary tuberculosis (TB) among HIV-positive individuals is extremely high in South Africa, rollout settings in general, and indeed worldwide.

Despite the widely appreciated importance of TB in HIV-infected individuals and the relatively high prevalence of TB in areas which have seen HAART rollout in the last few years, some articles (e.g., (Bekker et al. 2006; Ferradini et al. 2006)) do not report TB prevalence at all, and fewer still have reported comprehensively on the relative contribution of prevalent TB to risk of death. For instance Lawn(Lawn et al. 2006) and Moore (Moore et al. 2007) both report only a crude estimate of the impact of tuberculosis on death, which, given the strong association reported in many studies between tuberculosis and low CD4 counts and low BMI, is not a useful measure. Relatively few studies of rollouts have reported hazard or risk ratios for mortality (or other outcomes) on HAART given active TB at HAART initiation. Further, to our knowledge none have accounted explicitly for length of ongoing TB treatment at time of HAART initiation (an important modifier of the TB-death relationship, because most death from TB occurs early in TB treatment) and none have accounted explicitly for informative loss to follow-up in such as assessment; in

situations where TB or factors associated with TB are predictive of loss to follow-up, analysis of risk factors for death that fails to explicitly account for loss to follow up may provide biased estimates of the impact of tuberculosis on death in individuals receiving HAART.

TUBERCULOSIS AND MORTALITY IN THE CONTEXT OF TREATMENT FOR HIV

The relatively few studies of rollouts that do report the impact of tuberculosis on mortality, controlled for important confounders, offer up somewhat mixed results. Studies that report at least TB prevalence among those initiating HAART are summarized in Table 1.4. These studies, and a few additional studies, are discussed below.

The report from **Stringer et al.** (Stringer et al. 2006) was notable in that, unlike many other sources, the authors found no effect of prevalent TB on risk of death. They attribute this to two factors; first, that many patients have been on TB treatment for several weeks before HAART initiation, suggesting a selection bias for fewer TB-related deaths after the start of HAART. Second, they note that incident TB after HAART initiation is not well measured in this cohort.

Reports from **Libamba** (Libamba et al. 2006), **Wools-Kaloustian** (Wools-Kaloustian et al. 2006), and **Hawkins** (Hawkins et al. 2007) all report from different regions of sub-Saharan Africa, and as such report widely variable TB prevalence at baseline, though all report more than 10% prevalence. Given the high prevalence of

pulmonary TB in these cohorts and the well-recognized risk of death associated with the disease, none of these reports comment on either crude or relative hazard of death associated with prevalent or incident TB infection.

Lawn et al. (Lawn et al. 2006) explored characteristics associated with tuberculosis (or concurrent TB treatment) at baseline, concluding “Although prevalent and incident TB were associated with greater than two-fold increased mortality risk, they did not compromise immunological and virological outcomes among survivors at 48 weeks” (Lawn et al. 2006). Frustratingly, the two-fold increased mortality risk they reported was completely unadjusted for baseline covariates such as body mass index, CD4 count, and hemoglobin, to say nothing of length of TB therapy prior to HAART initiation. An earlier report from the same authors (Lawn et al. 2005) reported that tuberculosis was a “major attributed [cause] of death”, and reported results of a multivariate analysis of risks for early mortality in overlapping patients, but did not include tuberculosis in that analysis, reporting only on WHO stage (mortality rate ratio compared to stage I or II: 3.44, 95% CI 0.80-14.85 for stage III, 5.93, 95% CI 1.36-25.89 for stage IV).

The former paper by Lawn et al. (Lawn et al. 2006) reports on the impact of TB on CD4 and viral load outcomes in this cohort, in a complete case analysis, where the outcomes of those who died were apparently ignored (though this was not made explicit). Multivariate risk analysis in this paper was performed using Poisson rate regression by follow-up interval, using GEE and robust standard errors to adjust properly for inter-individual correlation of CD4 cell counts and viral load measurements. While this analysis is unbiased insofar as the outcome is well-

defined, a full consideration of the impact of tuberculosis on outcomes may wish to also consider the impact of tuberculosis-related deaths on outcomes such as CD4 count. This might imply a zero-inflated or “hurdle” model which incorporates a dichotomous survival model followed by a dichotomous outcome model. Additionally, there is the potential for bias to be introduced here in that CD4 cell counts may both confound the tuberculosis-CD4 exposure-outcome relationship in a time-dependent fashion, and in addition CD4 may be changed by the exposure of tuberculosis, and thus standard adjustment methods for assessing outcomes may return biased results (Cole et al. 2005; Hernan, Brumback, and Robins 2000; Robins, Hernán, and Brumback 2000).

The **ART-LINC study** (Braitstein et al. 2006) was only able to control for the presence of an onsite TB clinic in the HAART clinic; however, they found a strong correlation (HR = 3.76, 95% CI 1.01-14.02) between the presence of such an onsite clinic and risk of death. While it is tempting to claim that this ecological association is evidence of a relationship between TB infection at baseline and risk of death, such a claim requires us to both assume that the presence of such a clinic correlates with substantially higher prevalence of TB in those sites, and also that the TB-infected individuals are in fact dying earlier. There is little evidence to suggest that either of these claims is true; indeed, we might equally well assume, on this evidence, that the easy availability of TB treatment accelerates time to death.

Severe (Severe et al. 2005) reported an 8% prevalence of TB at HAART initiation in Haiti, and that 16% of deaths during follow-up were attributable to TB, but that there was no difference in risk of death between those with TB (who all received

TB treatment) and those without TB; however, this appeared to be a crude analysis and was not commented upon further. Likewise, **Etard** (Etard et al. 2006) reported a baseline TB prevalence of 11.6% in Senegal, and that TB was the leading cause of death among their patients, but reported no relative effect of TB on risk of death.

Zachariah et al. (Zachariah et al. 2006) studied 1507 individuals initiating HAART in Thyolo district, Malawi over two years, finding in multivariate logistic regression that TB was associated with a slightly reduced risk of mortality (OR = 0.7, 95% CI 0.4-1.6). This study appears to be well-done, and the result, while not statistically significant, is counterintuitive.

López-Gatell et al. (Lopez-Gatell et al. 2007) studied the effect of incident TB on HIV-related death in the United States, finding a hazard ratio of 4 for women who had incident TB during follow-up in the WIHS Cohort from 1995-2002, a hazard ratio that was not affected by whether those women were on HAART or not (4.3 among HAART); this estimate was highly imprecise because of a very small number of women developing TB while on HAART (around 2% in all of follow-up). In addition, the WIHS Cohort is based in the United States, and so the relevance of these results to the developing world is highly questionable.

Moore et al. (Moore et al. 2007), reporting from Tororo, Uganda, found a prevalence of pulmonary TB of 7.2% among 1044 individuals initiating HAART, and an incidence of 3.9 cases/100 person-years during a median of 1.4 years of follow-up. 37 of 75 TB-prevalent individuals were already receiving TB treatment (of an unknown duration) at HAART initiation; of the remaining half, 23 began HAART within 2 weeks starting TB treatment. Individuals with TB accounted for 35% of

deaths within 18 months of HAART initiation, and the crude hazard ratio for TB for the outcome of death was 4.7. The authors estimate that HAART reduced TB-associated mortality by 52%. However, the authors did not estimate the adjusted hazard of death associated with either incident or prevalent TB in this cohort; as prevalent TB is strongly associated with BMI < 18 (adjusted odds ratio 4.95, 95% CI 2.95-8.31), a risk factor for death, a multivariate analysis would be very useful here.

Several additional conference presentations focused on the issue of TB co-infection and HAART outcomes. One study from India (Patel et al. 2007), presented at IAS 2007, noted that individuals with TB at initiation of HAART, who received EFV-based HAART along with rifampicin-containing TB medications had no clinically meaningful differences in treatment outcomes compared with those who were not receiving TB medications. However, they did experience higher rates of hepatitis, and did not monitor viral loads; as they report that TB may impact the pharmacokinetics of EFV in the body – reducing available EFV – viral load outcomes may be significant in this context. In addition, they did not report on mortality or loss to follow-up; this, along with the fact that we should expect outcomes to be substantially different in India and sub-Saharan Africa, may limit the applicability of these results to our cohort.

A report from a study performed from 1997-2001 in Botswana (Chideya 2007) observed that HIV positive individuals with CD4 counts below 200 (not receiving HAART) were more likely to have sub-therapeutic concentrations of rifampicin in the blood during TB therapy, though TB treatment failure itself did not appear to be associated with CD4 count among HIV-positive individuals. However, low

pyrazinamide drug concentrations were associated with poor treatment outcomes and risk of death during follow-up, controlling for CD4 counts and HIV status. While this study was done in a setting where there was no HAART provided, it does strongly reinforce the idea that TB/HIV co-infection may put patients at higher risk of death from each.

Summary of literature review on tuberculosis and mortality. In rollout settings, despite numerous reports of high prevalence of active, pulmonary tuberculosis at initiation of HAART throughout sub-Saharan Africa, and widespread acknowledgement of the large contribution of TB to HIV-related deaths in these settings, no reports from rollouts have rigorously assessed the relative contribution of prevalent TB at initiation of HAART to measures of mortality. The few studies that assessed this measure at all both failed to account for losses to follow-up, nor did they account for length of TB treatment prior to initiation of HAART. Finally, none of the studies appeared to assess interactions among strong confounders of the TB-death relationship for interaction or effect measure modification, especially low BMI and low CD4 counts.

The impact of TB on mortality after initiation of HAART is clearly important and essentially not yet evaluated to any rigorous degree. Going forward, we will estimate the marginal relative hazard of death for the exposure of prevalent tuberculosis at baseline, accounting for length of TB treatment prior to HAART initiation. In sub-analyses we will assess the potential for interactions with low BMI and low CD4 count.

Antiretroviral drug toxicities and single-drug substitutions within first-line HAART. First, we will review background and literature on drug toxicities and drug substitution within first-line HAART in general and with regards to stavudine in particular, with special attention to reports of toxicity in rollout cohorts in sub-Saharan Africa. Then we will review the state of knowledge with regards to how current medication for tuberculosis (particularly with isoniazid) affects drug toxicities in general and stavudine toxicities and substitutions in particular.

As suggested previously, drug toxicities may play a significant role in the success of a HAART regimen. If a patient cannot tolerate her drugs, her adherence to the drug regimen may wane, leading to faster accumulation of resistance mutations (due to different half-lives of different antiretroviral drugs) and eventually to virologic, immunologic, or clinical failure of the drug regimen (Saag 2003). For example, the management of stavudine toxicity frequently involves brief periods in which patients receive no HAART at all, followed by a re-start of HAART with a new drug – most often zidovudine (Boulle et al. 2007; Zachariah et al. 2005) – in place of stavudine, after which symptoms usually resolve and recurrences of toxicity are rare (Bolhaar and Karstaedt 2007; Llibre et al. 2006; Lonergan, Barber, and Mathews 2003; Lonergan et al. 2004; Marceau et al. 2003; McComsey et al. 2005) and underlying metabolic pathologies appear to improve (McComsey et al. 2005). Such inadvertent toxicity mediated drug holidays may put individuals at higher risk of death, as in one study that saw increased mortality during treatment interruption due

to the development of opportunistic infections among patients with low CD4 counts (Bolhaar and Karstaedt 2007); as well, these treatment interruptions appear highly likely to compromise virologic efficacy of HAART. One study showed an HR of 22.5 for developing an NNRTI resistance mutation that could compromise HAART after drug holidays (Parienti et al. 2004). In addition, drug toxicities may themselves cause significant morbidity or mortality, e.g. (Currier 2007).

Problems of drug toxicities and concomitant management of those toxicities are of particular public health concern in rollout settings for several reasons. First, some of the drugs that compose first-line HAART in rollout settings – especially the NRTI stavudine – are associated with greater rates of toxicities than comparable NRTIs common to first-line HAART in the developed world (including AZT and TDF) (Bolhaar and Karstaedt 2007; Geddes et al. 2006). Several reports have reported higher-than-expected rates of stavudine-related toxicities in developing world settings (Bolhaar and Karstaedt 2007; Geddes et al. 2006; Wester et al. 2007). There are typically fewer drug options in rollouts than in other settings; with fewer drug options, toxicities (especially serial toxicities) may be more difficult to manage clinically. Last, greater demands on limited clinical personnel and resources intrinsic to the rollout setting, and less favorable clinician-to-patient ratios, may make clinical management of both toxicities and drug substitution significantly more difficult than in other settings, leading both to worse outcomes for individual patients, and worse outcomes for rollout cohorts as a whole, as clinical staff are stretched even further.

For these reasons and others, a thorough assessment of risk factors for drug substitution within first-line HAART – particularly, substitution of stavudine, the most

frequently substituted drug in TLCC – may help lead to more-customized therapy regimens at the population level and thus save effort, clinic visits, and (hopefully) ensure longer durability of first- and second-line HAART, as well as lower costs, better care provisioning, and reduced drug-related morbidity. This has been suggested by others; for instance, Currier suggests that it may be better to start women, and especially obese women, on TDF or AZT instead of d4T when those drugs are available (Bolhaar and Karstaedt 2007); the recommendation for TDF receives preliminary further support from a recent conference report on long-term safety of TDF in Central and South America (Madruga et al. 2007), which indicates little to no lipoatrophy after six years of TDF. Both of these drugs have been shown to produce less direct biological toxicity than d4T (Birkus, Hitchcock, and Cihlar 2002). However, Currier appears to consider only the risk of symptomatic hyperlactatemia and lactic acidosis. There are a number of other important stavudine-related toxicities, and decisions to customize initial HAART regimen by demographic characteristics might do better to consider other toxicities and reasons for substitution (or dose reduction) of stavudine.

In particular, peripheral neuropathy and lipodystrophy are both common and serious stavudine toxicities (Boulle et al. 2007), and both can be an indication for stavudine substitution. Despite this, risk factors for these two toxicities have not been well characterized, in part because they do not always present in a way serious enough to warrant drug substitution. Peripheral neuropathy, in particular, is of specific concern in developing world settings because it is patients being treated concurrently for tuberculosis with the common anti-TB drug isoniazid (INH), which

itself can cause peripheral neuropathy (Breen, Lipman, and Johnson 2000; Moyle and Sadler 1998). Nonetheless, few studies have rigorously examined the contribution of active tuberculosis pharmacotherapy (including isoniazid) on stavudine-related toxicities, and none have rigorously examined the impact of both prevalent and incident tuberculosis on time-dependent patterns of stavudine substitution.

It is worth noting that other drugs key to first line HAART are substituted for reasons of toxicity (most notably NNRTIs for hepatotoxicity). However, the majority of drug substitutions on first line HAART in TLCC (as well as in comparable cohorts) have been substitutions for stavudine (most often, with zidovudine); thus, we will concentrate on that substitution in this investigation.

Here we will review reports from the literature that explore briefly the biology and epidemiology of major stavudine toxicities that are cause for substitution in rollout settings; and then we will review recent reports of risk factors for drug substitution, drug dose changes, and toxicity-related drug regimen discontinuation with a focus on tuberculosis as a main risk factor. Finally, we will review additional literature on the impact of tuberculosis on stavudine toxicity and substitution.

Stavudine toxicities of concern in the developing world. Many of the most important drug toxicities on first-line HAART in the developing world are associated with the use of NRTIs in general (Wester et al. 2007), and with the use of stavudine (d4T) in particular (Subbaraman et al. 2007). These toxicities are generally regarded as the result of NRTI-mediated inhibition of mitochondrial polymerase-gamma (White 2001), which results in malfunction of the process by

which mitochondria produce adenosine triphosphate (ATP), which is essential to cell function (Montaner et al. 2003; Pinti, Salomoni, and Cossarizza 2006; Wester et al. 2007; Wohl et al. 2006). The most important of these toxicities are symptomatic hyperlactatemia (SHLA) and lactic acidosis (LA); lipoatrophy and lipohypertrophy; dyslipidemia; and peripheral neuropathy (PN) (Colebunders et al. 2005; Subbaraman et al. 2007).

Symptomatic hyperlactatemia and *lactic acidosis* occur when this mitochondrial toxicity results in anaerobic glycolysis (Wohl et al. 2006), of which lactate is a byproduct. LA has a substantial mortality rate (as high as 57% in one review of the literature (Falco et al. 2002)), and SHLA and LA both result in significant morbidity, with non-specific but potentially significant symptoms including fatigue, nausea, vomiting, abdominal pain, weight loss, numbness, and lipodystrophy (see below) (Falco et al. 2002; Marceau et al. 2003; Wester et al. 2007; Wohl et al. 2006). Both conditions are relatively common in developing world settings, reaching 10.6 cases per 1000 patient-years in one South African cohort (Bolhaar and Karstaedt 2007) and 2% of all patients in an RCT in Botswana (Wester et al. 2007). Risk factors for SHLA and LA are very similar, and include female gender, pregnancy, high BMI, DDI use, and duration of therapy (Bolhaar and Karstaedt 2007; Wester et al. 2007; Wohl et al. 2006).

There is some evidence to suggest that – in absence of symptoms – elevated lactate levels are not clinically significantly, and that routine lactate monitoring may be of limited usefulness in surveillance for SHLA and LA (Currier 2007) in part because it is extremely difficult to measure lactate accurately in clinical settings

(Wohl et al. 2004). One study found that when lactates were collected in a standardized fashion, 82 of 83 apparent cases of asymptomatic HLA resolved (Wohl et al. 2004). Along these same lines, two European studies found no demographic characteristics predicting laboratory-diagnosed asymptomatic hyperlactatemia (Boubaker et al. 2001; Moyle et al. 2002), though in both cases the generalizability to developing world rollouts is questionable. Nonetheless, the evidence suggests that in absence of solid evidence of SHLA, elevated lactate levels are not a cause for serious concern.

Lipodystrophy is the re-distribution of body fat around the body, and takes two main forms. *Lipoatrophy* is the loss of subcutaneous fat from one or more of several areas of the body, including the face, torso, or limbs (Wohl et al. 2006), and differs from so-called "AIDS wasting" in that it follows immune recovery rather than immunosuppression, and additionally that it does not typically involve loss of lean body mass, but only accumulated fat (Wohl et al. 2006). In contrast, *lipohypertrophy* is increased fat accumulation in the abdomen, breasts, or the so-called "buffalo hump" (Heath et al. 2003). Lipodystrophy in general is more common among older patients and lower CD4 counts prior to initiation of HAART (Wohl et al. 2006) and with increasing duration of HAART (Heath et al. 2003).

Both forms of lipodystrophy are relatively common in many settings and populations; lipodystrophy affected 10% of patients in a South African rollout by 3 years of follow-up (Boulle et al. 2007) and were a relatively common indication for regimen change in Botswana (Wester et al. 2007). In Rwanda, at least some lipodystrophy was observed in 34% of 571 subjects receiving stavudine, including

70% of those who had received HAART for more than 72 weeks (Mutimura et al. 2007). It is common, as well, to see both lipoatrophy and lipohypertrophy in the same subject, as in (Aurpibul et al. 2007; Mutimura et al. 2007).

Mitochondrial toxicity can result in *dyslipidemia*, the dysregulation of blood lipids that typically manifests as increased triglycerides, though less frequently will manifest as increased low-density cholesterol and/or decreased high-density cholesterol (Currier 2002; Heath et al. 2003). Dyslipidemia is associated with duration of HIV infection, duration of NRTI-based therapy, lipoatrophy, and use of EFV (Parienti et al. 2007; Wohl et al. 2006).

Peripheral neuropathy is a consequence of both HIV disease and HAART although it considered to occur more abruptly and be more painful (White 2001). PN can be distinguished by bilateral, distal pain and numbness, particularly in the lower extremities (Moyle and Sadler 1998), as well as by "paraesthesia, numbness and tingling in a symmetrical pattern in the distal extremities, and signs of diminished sensation to pinprick and vibration, loss of ankle jerks, but preservation of joint position sense and motor function" (Breen, Lipman, and Johnson 2000). Major risk factors for drug-associated peripheral neuropathy include time on therapy, baseline CD4 count, prior alcohol use, malnutrition, history of disease-associated PN, and older age (Moore et al. 2000; Moyle and Sadler 1998; White 2001). It is important to emphasize that because HIV disease can cause PN, it may be difficult to definitively identify PN as disease or HAART related; as well, other factors such infection with CMV, malnutrition and/or vitamin B deficiency, and (as previously noted) use of isoniazid for TB therapy (Colebunders et al. 2005; Moyle and Sadler 1998). About

6% of individuals substituted a drug in first-line HAART for reasons of PN within 2 years of initiating HAART in one large South African cohort (Boulle et al. 2007), and it is frequently cited as a main reason for stavudine discontinuation in other studies as well (e.g. (Amoroso et al. 2007; Hawkins et al. 2007)).

All of these toxicities – common or otherwise – are cause for serious concern: they may cause significant morbidity or mortality, earlier virologic failure, or less time- and cost-efficient treatment of cohorts as a whole. This serious concern has led to ongoing and sometimes acrimonious debate at the WHO and CDC. Stavudine is less expensive than its main alternative zidovudine ($80 difference per patient per year in low income countries (WHO 2007)) , so there is continuing pressure to use it in those countries; at the same time, the high level of toxicities makes an argument against the cost-effectiveness of that solution. Regardless, the lower cost of stavudine makes it very likely that it will remain in use in the developing world for at least the next five and likely ten years, and thus that this research will remain relevant.

Stavudine substitutions and toxicities. Here we review several studies – rollouts, where possible – of rates of stavudine toxicities in sub-Saharan African HAART cohorts. There are relatively few such reports, and so (below) we review additional studies from other locations as well; all data are summarized in Table 1.5. Discussions of TB are deferred to the next section.

Boulle et al. (Boulle et al. 2007) reported on 2679 individuals receiving HAART for a median of 11 months, looking at cumulative probability of substitution of NVP, EFV, AZT, and d4T and risk factors for many of those substitutions. By

three years, Kaplan-Meier estimation indicates that 21% of those who began on a d4T-based regimen would have substituted another drug for d4T. They found that female gender and baseline weight of at least 75kg were strongly associated with substitution of d4T for reasons related to lactic acidosis or symptomatic hyperlactatemia (SHLA) (HR=10.7 95% CI 1.5-79.0 for female gender; HR=36.1, 95% CI 10.6-122.8 for weight>75kg) and lipodystrophy (HR=10.0, 95% CI 1.3-74.5 for female gender; HR=4.1, 95% CI 1.6-10.5 for weight>75kg); they also noted that early weight gain was a risk factor for discontinuation of d4T for reasons of SHLA. Risk of d4T discontinuation for reasons of peripheral neuropathy were not affected by weight or gender in their analysis, but increased slightly with age and ordinal WHO stage.

A chief drawback of this analysis is that Boulle et al. are modeling based on relatively few actual switching events; though their Kaplan-Meier estimate of total switches by 3 years is 21%, they in fact saw only 125 stavudine substitutions in the entire cohort of 1996 individuals at risk in the cohort, a point which is not emphasized in their discussions. This is likely the reason that their hazard ratio confidence intervals (above) are so large.

Methodologically, Boulle et al. use a mix of naïve and more sophisticated techniques. For example, in the former category, their control for weight was undoubtedly biased because they controlled only for weight, and not for weight-for-height or BMI; this is of particular concern in modeling risk of LA/SHLA, because men typically weigh more than women, but women are at higher risk of the toxicity. Lack of control for height, and lack of control for interaction between gender and

weight (or BMI) may produce biased results in this context. After this, however, Boulle et al. report results from multivariate Cox models of risk factors for each major area of stavudine toxicity separately, the results of which are properly interpretable as hazard ratios and which are controlled for a number of important baseline factors.

Boulle et al. do not account for losses to follow-up in this analysis, but as only 5% of those who entered into care at their site in South Africa were lost, their complete case approach to the data is unlikely to result in substantial bias. In addition, they note that "associations with each toxicity-mediated substitution could therefore potentially be biased by censored observations (related to stopping stavudine for other toxicities)", and "explored estimating the associations jointly across all three toxicity-related reasons for stopping stavudine" in a competing risks framework, and said that "the results from this exercise were consistent with those presented here", though they did not give further detail.

Despite these methodological and substantive drawbacks, however, Boulle et al. present comprehensive, useful data which addresses a number of important points and elucidates risk factors for stavudine substitution. Though not a final word on the subject, this report helps move us forward in our own investigation.

A conference report from the same site by **Osler et al.** (Osler et al. 2007) noted a referred rate of approximately 10 cases/1000 person-years of severe SHLA, presenting at a median of 10 months post-HAART initiation and the vast majority presenting within 18 months, with an acute mortality rate of 15%. 86% of cases were receiving stavudine; 95% were female compared to approximately 66% of the overall cohort. In univariate analysis, co-morbidity with peripheral neuropathy and/or

lipodystrophy were both common. Case-control analysis indicated that those with SHLA were more likely to have higher baseline weight (with referent weight < 60 kg: OR 5, 95% CI 1.6-15.4, for 60kg ≤ weight < 75kg and OR 19.3 95% CI 4.8 – 77.1 for weight ≥ 75kg), female gender (OR 44.2, 95% CI 6.4 – 303.8), decreasing CD4 nadir (OR 1.1, 95% CI 1.1-1.25 per 10 cells) and increasing WHO stage (OR 1.9, 95% CI 0.9-3.7 per stage).

These results suggest overlap in risk factors for several important stavudine related toxicities. However, the imprecision of the results combined with methodological shortcomings of this report (no description of control sampling procedures; claims that the study was a "matched" case-control study without any description or indication of what they matched on), makes these estimates difficult to interpret.

Bolhaar and Karstaedt (Bolhaar and Karstaedt 2007), report a high incidence of LA and SHLA among 1735 individuals receiving HAART in the Baragwaneth Hospital in Johannesburg; the great majority of toxicity was among women (63% of the cohort) receiving stavudine. Overall, they saw a rate of 10.6 cases of LA per 1000 patient years of follow-up, and 20.2 cases of SHLA per 1000 patients years. Replacement with zidovudine after a median ~8 weeks treatment interruption resulted in no recurrences of LA or SHLA in a mean follow-up time of 9.6 months among 56 patients. Their findings suggested that high BMI (>30) might well be a risk factor for developing both conditions. However, these findings were largely descriptive, because a small number of total outcomes (23) precluded extensive multivariate modeling.

Hawkins et al. (Hawkins et al. 2007) studied 1286 patients in Kenya, 59.1% of whom were female, on d4T-3TC-NVP, with a median time on HAART of 350 days. These patients experienced a cumulative incidence of toxicity-related drug substitution estimated at 40.6% at 12 months, with 26.5% of their patients actually experiencing such a toxicity during follow-up. Any drug substitution was predicted by older age (≥40), AIDS symptoms at baseline, male gender, and lower baseline CD4 count (≤ 100 cell/mm3); toxicities in particular were predicted only by lower CD4 count and older age. Peripheral neuropathy was common in this cohort, with 20.7% of patients experiencing PN in follow-up; PN accounted for two-thirds of observed toxicities. The authors report that this toxicity was attributable to stavudine, though they do not explain adequately how they attributed PN to stavudine in particular. Interestingly, the authors do not report any symptomatic hyperlactatemia or lactic acidosis in this cohort, a lack-of-finding they do not comment on.

Forna et al. (Forna et al. 2007) studied 1029 adult patients on d4T-3TC-NVP (96%) and d4T-3TC-EFV in rural Uganda, finding a rate of 4.47 toxicities per 100 patient-years (median follow-up time 11.2 months). The vast majority of toxicities found were peripheral neuropathy (322 of 411 total toxicities, having excluded those with evidence of disease-associated PN), which was found to be associated with older age (≥35) and tuberculosis treatment at baseline (severe PN only); they drew attention to this latter association as a matter of serious concern in the developing world. 222 toxicity-related single drug substitutions were made in this cohort, for any reason; 178 patients replaced zidovudine with stavudine (and four of those had an

anemia toxicity). The authors estimated that 24% of those who began on stavudine would substitute for stavudine by 18 months. The authors saw a single case of lactic acidosis, but did not look for symptomatic hyperlactatemia as a potential toxicity.

Wester et al. (Wester et al. 2007) reported rates of 2% of SHLA and 1% of LA during follow-up of 659 adults (69% female) in a randomized controlled trial in Botswana in 2 years of follow-up; in crude analysis, the major predictors of both conditions included female gender, BMI > 25 kg/m2, and older age (>40 years). No multivariate modeling was done in this study. The authors report that individuals who were survived their toxicity were switched onto a completely new drug regimen.

There is only one conference abstract worth reviewing in depth. **Amoroso and colleagues** (Amoroso et al. 2007) reported on 6520 patients treated in PEPFAR programs in Uganda, Kenya, and Zambia, of whom 1164 experienced a toxicity driven drug substitution in a maximum of 20 months of follow-up. The population as a whole was 68% female, while those who experienced a toxicity driven substitution were 67% female. Of 2149 patients treated with stavudine, 24% required a substitution in a median of 141 days; by far the most common reason for these substitutions was peripheral neuropathy (53.0%), followed by lipoatrophy (7.5%), lactic acidosis (2.4%), and pancreatitis (0.7%); 36% did not have a documented reason. Because they do not specify the median follow-up time, it is difficult to extrapolate from these reported switching probabilities to rates if the full cohort were followed up for a year (for example). Nonetheless, these results confirm what we have seen in several other large African cohorts.

Additional conference abstracts (not listed in Table 1.5) report similar findings. A Nigerian study indicated that rate of drug substitution was three times higher for d4T-based regimens compared with AZT- and TDF-based regimens, though they do not specify who, in particular, required those substitutions (Shepherd et al. 2007). A study in Cambodia found that 75% of drug substitutions were for reasons of toxicity, and went as far as to recommend that a programmed switch from d4T at 6 months might be appropriate (Hing et al. 2007). A retrospective review from a Kenyan cohort showed that while only 16% of individuals required drug regimen changes within four years of follow-up (median follow-up not specified), approximately 81% of those switches were due to stavudine toxicity, primarily lipodystrophy, SHLA and LA, and PN (Kocholla et al. 2007). A study in India found that 27% of subjects initiated on HAART experienced a serious adverse event within a median of 48 weeks, most of those attributable to stavudine, primarily SHLA and PN (Abraham et al. 2007). Finally, a London study (Davidson et al. 2007) estimated an "observed toxicity switch rate (OTSR)" for stavudine of 304(95% CI 174 to 495) per 1000 patient years, but also a high OTSR for zidovudine (224(180 to 276) per 1000 p-y).

Several additional studies (not included in Table 1.5) are worth noting here. **Forna** (Forna et al. 2007) reported the effects of substitution of AZT for stavudine in 261 of the 1029 patients noted above and followed up for a median of 17.3 months. The authors found that risk of leukopenia (a key zidovudine toxicity) was higher among those with a history of more than six months of d4T before substitution of AZT (HR = 2.63, 95% CI 1.51-4.60), suggesting that earlier switching to AZT might alleviate risk of some AZT toxicities in a population of people who experienced

stavudine toxicities. This, in turn, provides limited support for replacing d4T with AZT in initial HAART regimen among people most likely to have d4T toxicities.

Kumarasamy et al. (Kumarasamy et al. 2006) studied 1443 individuals initiating HAART in India from 1996-2004, 72% of them on d4T-containing regimens. 183 of those individuals had a drug substitution due to a toxicity in a median of 40 days, and an additional 27 patients discontinued first-line HAART due to a toxicity in a median of 49 days. Only 27 of the total patients (about 1.8%) modified their HAART for reasons of d4T-related toxicities; this is much lower than in other studies (such as Boulle et al., above). This may be in part due to the much higher proportion of men in this cohort (82%, compared to 29% in Boulle et al.).

Dieleman et al., writing for the ATHENA Cohort in the Netherlands (Dieleman et al. 2002), studied the causes of subsequent drug substitution among individuals who substituted a drug on first-line HAART. They did not analyze risk factors for first substitution, noting only that 1573 of 2096 individuals at risk experienced some drug substitution between 1998 and 2002 (though there are varying lengths of follow-up among individuals in the cohort) and that the most frequent indication for substitution was toxicity. Risk factors for toxicity-driven second substitutions of antiretroviral drugs included female gender, toxicity-driven (rather than virologic failure or contraindication-based) first-substitution, and not switching underlying NRTIs while switching a PI during first substitution. Risk of toxicity-driven second substitutions fell with calendar time. They saw no influence of BMI or duration of first HAART on risk of toxicity-driven second substitutions. This cohort largely comprised men with relatively high CD4 counts, and additionally the HAART regimens used were mostly

PI-based on first line, and it is unclear how many used stavudine; thus, it is difficult to generalize these results to a developing world rollout situation. Nonetheless, we view these results as hypothesis generating of the possibility that previous toxicity-driven drug substitution may act as a risk factor for subsequent toxicity-driven drug substitutions; this would be an additional motivation to avoid drug substitutions by customizing baseline HAART.

Summary of findings on stavudine toxicities and substitution. As we noted above, there is ongoing controversy surrounding use of stavudine and support from WHO for moving away from use of the drug. Several reports allude to the need for increased drug regimen options – for instance, Forna et al. (Forna et al. 2007) conclude their 2007 paper by saying:

> Our findings indicate the need to expand drug formularies in resource-limited settings, because the large number of patients experiencing toxicity and requiring drug substitution shows that programs need to have alternate drugs available to switch to if toxicity occurs. In resource-limited settings, toxicity from ART regimens containing stavudine or nevirapine is manageable but more tolerable regimens are needed.

Nonetheless, as we noted above, stavudine is $80 less expensive per patient per year than the next cheapest alternative, zidovudine (WHO 2007), in developing countries. While WHO recognizes the prevalence of stavudine toxicities and as a consequence "now recommends zidovudine as one of the preferred NRTI options to be considered by countries" (WHO 2006), stavudine is not likely to disappear from use in the developing world due to price considerations. As WHO notes, price considerations for these drugs will remain central to the choices made in national programmes." (WHO 2006). Unfortunately, therefore, the proposed work will remain useful for some time to come.

Stavudine toxicity is an important concern in both developed and developing world settings; while a number of papers have explored frequency of and reasons for stavudine substitution within first-line HAART, there are still several outcomes that have not been rigorously examined – among them risk factors for time to all-cause stavudine substitution. All-cause stavudine substitution is of particular interest in rollout settings, in fact, because in rollouts we are often concerned not only with effect of (for instance) peripheral neuropathy on patient quality of life and (therefore) on the patient's drug adherence, but also about the optimal use of already-overworked clinician resources and drug supplies. By the same token, if peripheral neuropathy is not serious enough to justify stavudine substitution, then perhaps the clinical event is of lower interest in terms of drug regimen management. Thus, focusing on all-cause stavudine substitution therefore may be more appropriate in many ways than assessing risk factors for individual drug toxicities in the context of large scale public health rollouts.

In the next section, we will review literature which examines the contribution of tuberculosis pharmacotherapy containing isoniazid to risks of all-cause stavudine substitution and peripheral neuropathy. First, however, we will briefly review the standard treatment plan and pharmacological regimen for TB in South Africa.

Side effects of isoniazid and other TB drugs. Each of these drugs has side-effects, some serious, and many which overlap with side-effects and toxicities of anti-retroviral drugs. While the rest of this review will focus on isoniazid, it is worth noting at this point that there is a growing literature on interactions between rifampicin and nevirapine. Because rifampicin induces the cytochrome p450 enzyme, results in reduced pharmacoavailability of nevirapine (Colebunders et al. 2005; Zachariah et al. 2005). This appears to be a less-serious concern with efavirenz (Manosuthi et al. 2005; Patel et al. 2004), hence the strong recommendation that patients on TB pharmacotherapy be placed on EFV-containing regimens when at all possible (Republic of South Africa Department of Health 2004).

The most common side-effect reported with the common and highly potent anti-TB drug isoniazid is peripheral neuropathy (Lawn, Edwards, and Wood 2007; Subbaraman et al. 2007), also a well-known and common side-effect of stavudine (Lawn, Edwards, and Wood 2007). Despite the fact that these overlapping toxicities have been recognized for 10 years (Lawn, Edwards, and Wood 2007; Moyle and Sadler 1998), there are few published reports of the effect of isoniazid on stavudine toxicity: a Medline search for “tuberculosis stavudine” returns only 17 matches, not all of them relevant, and a search for “isoniazid stavudine” returns only 3. We review

the relevant literature here, beginning with reports reviewed in the previous section, on stavudine toxicities more generally.

Isoniazid for tuberculosis, stavudine-related toxicities, and drug substitution. Here we will re-examine literature already reviewed above for content regarding the particular problem; then we will review other relevant literature.

Forna et al. (Forna et al. 2007) observed a strong association between receiving TB pharmacotherapy (including isoniazid) at baseline and experiencing severe peripheral neuropathy (n=83 events) during follow-up was 2.35 (95% CI 1.18-4.68). They found no association with non-severe peripheral neuropathy, and the association between TB and any severe toxicity was also attenuated at 1.7 (95% CI 0.97-2.98), and in their discussion specifically suggested that this increased risk might be due to co-medication of isoniazid and stavudine.

In contrast to the report by Forna, no other report of toxicities reviewed previously specifically addressed issues of tuberculosis and stavudine toxicity or substitution. **Hawkins et al.** (Hawkins et al. 2007) reported that 28% of the observed drug substitutions in their cohort were due to TB-HAART drug interactions, but imply that these were proactive drug regimen modifications to avoid rifampicin/nevirapine interactions. Although 40.6% of their observed drug substitutions were caused by drug toxicity, and 2/3 of those were caused by peripheral neuropathy, they do not mention TB as an independent risk factor for substitution or peripheral neuropathy (it is not clear whether this is because it was not significant in the model or because they did not check for it). **Kumarasamy et al.** (Kumarasamy et al. 2006) concentrated on the interaction of rifampicin and nevirapine, which in their study was

shown to elevate risks of adverse events as the cause of first-line HAART modification, and made no mention of TB-stavudine drug interactions. **Boulle et al.** (Boulle et al. 2007) mentions the same issue in passing.

Very few other studies shed light on interactions between concurrent use of stavudine and pharmacotherapy for tuberculosis including isoniazid. A very small study of 28 individuals receiving d4T in the UK, **Breen et al.** (Breen, Lipman, and Johnson 2000) found an significant increase in cases of peripheral neuropathy among those who were also receiving isoniazid. **Patel et al.** (Patel et al. 2004) studied 255 patients with and without TB in India, most of them receiving stavudine for HIV, and found no increase in incidence of peripheral neuropathy with combination of stavudine and isoniazid, although their numbers of total cases of peripheral neuropathy were small in both groups (7 among TB, 4 among non-TB). However, they did find that rates of hepatitis were greatly increased in patients receiving anti-TB medications, from 0 cases in the non-TB group to 17 cases in the TB group. It is unclear how to apply these Indian findings to a sub-Saharan African context; the very small numbers of cases of peripheral neuropathy suggest that the results indeed may not be generalizable. **Dean et al.** noted that among English patients receiving isoniazid, those who also received stavudine were at 1.7 times the risk of peripheral neuropathy (p=0.07) (Dean et al. 2002).

Summary of research on tuberculosis and stavudine-related toxicities and stavudine substitution. While there is a substantial body of research examining particular stavudine toxicities and particular reasons for stavudine substitution, and despite the clear importance of tuberculosis to HIV treatment

outcomes in the developing world, there are no reports which rigorously examine the confounding-adjusted population-level impact of ongoing or incident tuberculosis pharmacotherapy on the risk of stavudine substitution during first-line HAART.

Table 1.1. Major outcomes of rollout studies published 2003-2007.

Study	Location	n	Mortality		Lost to follow-up	CD4 gain	
			6 mo	12 mo		6 mo	12 mo
(Stringer et al. 2006)	Zambia	16,198	16 / 100 person-years		21%	155 *	175 *
(Libamba et al. 2006)	Malawi	13,183	n/r	8%	8%	n/r	n/r
(Fairall et al. 2008) ‡	South Africa	3,619	5%	10%	n/r	90	170
(Wools-Kaloustian et al. 2006)	Kenya	2,059	5.4%		24.5%	n/r	160 *
(Ferradini et al. 2006)	Malawi	1,308	19% at median 36 weeks follow-up		7%	138 #	176 #
(Hawkins et al. 2007)	Kenya	1,286	1.1% at median 350 days follow-up		33.8%	169 *	208 *
(Bekker et al. 2006; Lawn, Badri, and Wood 2005; Lawn et al. 2006, 2006; Lawn et al. 2006; Lawn et al. 2005)	South Africa	1,139	4.4%	7%	≤3%	116 #	164 #
			12/100 person-years in 563 person-years at risk				

n indicates number reported as initiating HAART. n/r = not reported. * indicates mean values, # indicates median values. ‡ Mortality and CD4 count gains estimated from Figures in article.

Table 1.2. Major outcomes from conference reports of HAART rollouts.

Study	Location	n	Mortality		Lost to follow-up	CD4 gain	
			6 mo	12 mo		6 mo	12 mo
(El-Sadr et al. 2007)	Multiple	171,259	n/r		n/r	119 #	n/r
(Marlink et al. 2007)	Multiple	134,120	5% total recorded during follow-up		7-12%	101 *	n/r
(Meloni et al. 2007)	Nigeria	9,070	Alive and in care: 80-90% at 6 months; 73-92% at 12 months.			124 #	158 #
(Mosoko et al. 2007)	Cameroon	2,920	Alive and in care: 66% at 6 months; 58% at 12 months; 47% at 24 months			189 #	233 #

n indicates number reported as initiating HAART. * indicates mean values, # indicates median values.

Table 1.3. Major outcomes from additional studies relevant to study of sub-Saharan African HAART rollouts.

Study	Location	n	Mortality		Lost to follow-up	CD4 gain	
			6 mo	12 mo		6 mo	12 mo
(Calmy et al. 2006)	21 sites, global	6,861	14.2 / 100 person-years		n/r	102 #	144 #
(Braitstein et al. 2006)	Africa, Asia, Brazil	4,810	n/r	6.4%	15% (year 1)	106 #	n/r
(Koenig, Leandre, and Farmer 2004)	Haiti	1,050	0%	0%	n/r	n/r	n/r
(Severe et al. 2005)	Haiti	1,004	10%	13%	7.8%	128 #	163 #
(Etard et al. 2006)	Senegal	404	6.3/100 person-years		<1%	104 #	n/r

n indicates number reported as initiating HAART. n/r = not reported. # indicates median values.

Table 1.4. Prevalence and impact on risk of death of pulmonary tuberculosis in selected large cohorts. No study controlled for length of tuberculosis treatment.

Study	n	Baseline TB prevalence	Hazard ratio of TB for death	Notes
(Stringer et al. 2006)	16,198	11%	No significant difference	Post-hoc estimated 80% power to detect HR of 1.26 at α=0.01
(Libamba et al. 2006)	13,183	11%	n/r	Lower than expected prevalence; suggest due to drug interactions and structural issues
(Wools-Kaloustian et al. 2006)	2,059	<27% (see Notes)	n/r	27% of cohort received treatment in median 40 weeks follow-up
(Hawkins et al. 2007)	1,286	34.3%	n/r	11.4% first-year incidence of TB
(Lawn et al. 2006)	944	25%	Crude >2	High early incidence of TB
(Braitstein et al. 2006)	4,810	11 sites of 18 had TB clinic onsite	3.76 (95% CI 1.01-14.02)	Purely ecological analysis; HR is for death in HAART clinics with onsite TB clinic
(Severe et al. 2005)	1,004	8%	No significant difference	16% of deaths attributed to TB
(Etard et al. 2006)	404	11.6%	"Leading cause of death"	Prevalence is from % receiving anti-TB drugs at baseline
(Lopez-Gatell et al. 2007)	1,412	n/r	4.3 (0.9-22.0)	HR is incident TB (2% incidence); study conducted in the United States
(Moore et al. 2007)	1,044	7.2%	4.7 (crude)	Incidence 3.9/100 person-years

This table includes all studies in Tables 1.1, 1.2, and 1.3 which reported at least TB prevalence; in addition, this table includes two studies not included previously. n/r, not reported.

Table 1.5. Incidence of lactic acidosis, peripheral neuropathy, and stavudine substitution in large cohorts from sub-Saharan Africa.

Study	n	Incidence			
		Lactic acidosis or symptomatic hyperlactatemia	Peripheral neuropathy	Lipo-dystrophy	d4T substitution
(Boulle et al. 2007)	2,679	4.7%#	6.2%#	9.0%#	20.8% in 3 years
(Osler et al. 2007)	7,080 p-y in 28 months	LA: 10 / 1000 p-y	n/r	n/r	n/r
(Bolhaar and Karstaedt 2007)	1,735	LA: 10.6 / 1000 p-y SHLA: 20.2 / 1000 p-y	n/r	n/r	n/r
(Hawkins et al. 2007)	1,286*	n/r	20.7%	2.2%	40.6% by 350 days (median follow-up)
(Forna et al. 2007)	1,029	LA: 1 case SHLA: n/r	~3.5/100 p-y	n/r	18% by 1 year
(Wester et al. 2007)	659	LA: 1% in 2 years SHLA: 2% in 2 years	n/r	n/r	n/r
(Amoroso et al. 2007)	2,149	LA: 0.6% SHLA: n/r	13.0%	1.9%	24.6% in median 141 days

P-y is patient-years. # incidence measured only in events which led to drug substitution.

CHAPTER 2

STATEMENT OF THE SPECIFIC AIMS

SPECIFIC AIMS AND JUSTIFICATIONS

Specific aim 1. To assess outcomes among patients who initiated HAART in the Themba Lethu Clinic in Helen Joseph Hospital, Johannesburg, South Africa, between 1 April 2004 and 31 March 2007.

Justification. Rapid rollout of ART is necessary to achieve the largest possible number of patients on ART. Concerns have been raised of the quality of such large programs, with fears of poor adherence and poor outcome. We will assess the outcomes of survival, “alive and still in care”, viral suppression, and CD4 gain among 7587 patients who presented for care in the first 3 year of rollout. We will use maximum-likelihood regression modeling to explore the impact of baseline characteristics including gender, age, CD4 count, tuberculosis status at baseline, and WHO stage on these outcomes.

Specific aim 2. To estimate the effect of currently treated baseline pulmonary tuberculosis on survival after the initiation of HAART among patients in the Themba Lethu Clinic.

Justification. Tuberculosis is a key predictor of mortality among individuals initiating HAART, especially in South Africa, but few studies have been able to

provide good estimates of the hazard ratio associated with the condition in a rollout setting. We will use inverse probability weighted marginal structural discrete time hazard models to estimate the marginal effect of TB on mortality, as well as perform an analysis which estimates the effect of time lag between start of TB pharmacotherapy and HAART initiation. In secondary analysis, we will examine interactions of TB with key confounders for effect measure modification to identify subpopulations of individuals with TB who may be at differential risk of death. We will adjust for both confounding and informative censoring using inverse probability weights to obtain accurate and precise estimates of hazard attributable to TB in this cohort, results which (given the assumption of no uncontrolled confounding) will approximate those that would be observed if we were to perform a randomized trial of 'assigning' TB to individuals at random.

Specific aim 3. To estimate the effect of treatment for tuberculosis on hazard of stavudine substitution within first line HAART among patients in the Themba Lethu Clinic.

Justification. Stavudine is a key component of first line HAART throughout sub-Saharan Africa, but many individuals require a substitution for stavudine due to toxicities within relatively short windows of follow-up; many of those same toxicities are exacerbated by concurrent TB pharmacotherapy. We will use inverse probability weighted history adjusted marginal structural models to estimate the marginal effect of TB on the outcome of all-cause stavudine substitution, controlling for time since initiation of TB pharmacotherapy. We will adjust for both confounding and informative censoring using inverse probability weights to obtain accurate and

precise estimates of hazard attributable to TB in this cohort, results which (given the assumption of no uncontrolled confounding) will approximate those that would be observed if we were to perform a randomized trial of 'assigning' TB to individuals at random.

RATIONALE

The HIV/AIDS pandemic continues to claim lives at an astounding pace in all corners of the developing world; this is especially true in sub-Saharan Africa, where the World Health Organization (WHO) estimates that nearly 2 million individuals died of HIV-related causes in 2006 alone (UNAIDS/WHO 2006). Timely treatment with antiretroviral therapy (ART) has enormous potential to reduce the morbidity and mortality of AIDS. Despite efforts such as the WHO/UNAIDS 3x5 initiative (WHO 2005), many more people need ART than currently receive it; in South Africa, for instance, it is estimated by the WHO that only 20% of the one million individuals in need of ART are receiving the therapy (WHO 2007).

Meeting the UNAIDS goal of universal access to HAART by 2010 (WHO 2006) will require very large scale "public-health style" rollouts of HAART, in which thousands or tens of thousands of individuals are initiated on ART over a very short period of time. Rollouts of this sort have been a major part of expanding access to treatment throughout sub-Saharan Africa, where hundreds of thousands of individuals have initiated HAART in the last several years. Because of the immense scale of such rollouts, monitoring and sound evaluation of HAART outcomes

including survival, and virologic and immunologic status, predictors of those outcomes, and assessment of issues surrounding antiretroviral drug toxicity are both challenging and extremely important to ensure that a high quality of care is delivered.

Pulmonary tuberculosis is a particular concern among people suffering from HIV in South Africa and throughout the developing world (WHO 2004), even in the presence of HAART. The prevalence of TB among people presenting for HAART is as high as 25% in some locations (Lawn et al. 2006); an estimated 50 to 80% of TB patients are HIV positive in southern Africa (South African National AIDS Council 2007). In addition, there are serious concerns about possible interactions between components of first-line HAART and standard TB pharmacotherapy (Subbaraman et al. 2007). Nonetheless, the contribution of TB to both mortality and durability of first-line HAART remain uncharacterized in a methodologically rigorous way. Indeed, many previous reports of HAART rollouts ignore issues of TB entirely.

Thus, despite numerous reports in the literature, serious gaps remain in our knowledge about outcomes in the context of rapid scale-up and delivery of HAART in the developing world, especially as relates to tuberculosis co-infection and co-medication.

Taken together, these three specific aims will describe the impact of TB on the outcomes of the ongoing HAART rollout at the Themba Lethu Clinic, adding to a growing body of results from other rollouts; these results will contribute to the literature both substantively and methodologically, and will provide essential evidence for rational clinical decision making and policy changes. In particular,

proper evaluation of the HR associated with death among individuals with TB in this cohort may contribute to policy around patient care and follow-up immediately after HAART initiation. Similarly, a thorough understanding of the impact of TB on risk of stavudine substitution within first-line HAART may help guide decisions to initiate certain individuals on alternative drugs (such as zidovudine or tenofovir), leading to better outcomes and lower costs and burden of care overall.

HYPOTHESES TO BE TESTED

Specific aim 1. No hypotheses to be tested.

Specific aim 2. A. Is the effect of current treatment for pulmonary tuberculosis on all-cause mortality among individuals initiating HAART different than the null? B. Does the effect of current treatment for pulmonary tuberculosis on all-cause mortality among individuals initiating HAART differ by the time lag between initiation of treatment for tuberculosis and initiation of HAART?

Specific aim 3. Does current treatment for pulmonary or extrapulmonary tuberculosis increase the hazard of substitution of the antiretroviral drug stavudine for any reason?

CHAPTER 3

METHODS

STUDY SETTING AND CLINCAL PROCEDURES

Overview of study setting. The rollout of HAART in South Africa provides an excellent resource to perform the analyses proposed in this dissertation. In particular, this dissertation concentrated on a central site for the South African rollout, the Themba Lethu Clinic in Helen Joseph Hospital, Johannesburg, Gauteng Province, South Africa (Clinical HIV Research Unit 2005). The Themba Lethu ("Our hope") Clinic (TLC) is the single largest clinic providing HAART to HIV-infected adults drugs in South Africa, with over 10000 individuals in care in January of 2008, all of whom were enrolled since the start of the so-called "government era" of HAART availability in April of 2004.

Along with Johannesburg General Hospital and Chris Hani Baragwanath Hospital (Bara), TLC is one of the three major centers for where HAART is provided within Johannesburg. Patients are referred to TLC through voluntary counseling and testing clinics, hospitals, prenatal care facilities, and by self-referral. Unlike Bara, which is located in the township of Soweto, TLC is relatively far from the areas where TLC patients live and work. TLC is a joint project of Right to Care, a South African non-profit organization dedicated to providing HAART (Right To Care 2005),

and the South African government through the CCMT initiative; TLC is funded by the Gauteng Province Department of Health and the President's Emergency Plan for AIDS Relief (PEPFAR), and works closely with the Clinical HIV Research Unit of the University of Witwatersrand.

The effort to initiate large numbers of patients on HAART at TLC is part of the larger effort to fight HIV and AIDS in South Africa, which has one of the largest unmet needs for HAART in the world (WHO 2007). The widespread scale-up of HAART in South Africa is important not only because of the lives it will save, but also because of the political statement it makes in a country with a tangled political history around AIDS (AVERT.ORG 2007) and whose leader has sown considerable doubt about the fact that HIV is the cause of AIDS (Smith and Novella 2007).

Study population and enrollment. Patients were enrolled upon initiation of HAART in TLC between 1 April 2004 and 31 March 2007. In all analyses, patients were excluded if they had a history of HAART before 1 April 2004; if they had missing or incomplete information on initial HAART regimen; or if they initiated a non-HAART ART regimen. There are approximately 7500 patients who meet inclusion criteria for baseline analyses, though there were fewer individuals for some analyses due to missing data or inadequate opportunity for follow-up. Specific aim 3 considered only patients on one of three main HAART regimens (see Antiretroviral therapy regimens, below).

All patients in TLCC receive antiretroviral medications (ARVs) free of charge; until 1 October 2006 patients paid 35 Rand (less than $5 US) per visit to clinic. We accounted for this change in policy in analysis with an indicator variable for date.

Antiretroviral therapy regimens. Following the South African national ARV treatment guidelines, first line HAART regimens in this cohort were most frequently stavudine (d4T), lamivudine (3TC), and either efavirenz (EFV) or nevirapine (NVP) (Regimens 1a and 1b, respectively). Some individuals received Kaletra® (lopinavir boosted with ritanovir, LPVr) in place of NVP or EFV, and zidovudine (AZT) replaced d4T in some cases. Second-line HAART regimens included AZT, didanosine (DDI), and LPVr.

For purposes of these analyses, *drug substitution* was defined as the substitution of one or two drugs within a single drug class (for instance, substituting AZT for D4T), or switching between NVP and EFV. A *drug regimen switch* (or *switch*) was defined as a change between any WHO-approved first line HAART regimen to any WHO-approved second line HAART regimen. Collectively, all changes to drug regimen (including drug substitutions, drug regimen switches, the addition or subtraction of a drug to a drug regimen, or the adjustment of a dose of a drug within an established drug regimen), were referred to as *drug regimen changes*. Drug regimens were assumed to be continued until termination date given in the database. If a drug regimen was discontinued for more than 30 days and then subsequently restarted, it was deemed a treatment interruption (TI).

Other medications, including TB medications. A large number of other medications and treatments are provided to individuals receiving antiretroviral therapy including a wide variety of antibiotics and vitamins. TLC is not recognized as a TB treatment clinic. Patients with TB are referred in to TLC or referred out to TB clinics, thus making perfect ascertainment of adherence to and completion of TB

therapy difficult for most patients. The non-integration of TB and HIV treatment services is widely acknowledged as a serious barrier to optimal treatment in South Africa.

Clinical procedures; collection of laboratory and clinical measurements. Patients attended TLC at initiation (month 0), and at months 4, 10 and every 6 months thereafter. Patients attend additional visits at months 1, 5, 11 and every 6 months thereafter to receive test results, and come to clinic in the intervening months to pick up drugs from the pharmacy, or if they become sick or experience drug side effects. Patients who had not suppressed virus by month 4 were scheduled for monthly visits until viral suppression was achieved; patients who experienced drug toxicities or other health problems were seen as required. In addition to clinical visits, patients also visited the TLC pharmacy monthly or bi-monthly to receive antiretroviral drugs and other prescription medications (including tuberculosis drugs and other antibiotics). A typical clinical visit involved an assessment by a nurse, evaluation by a doctor or nurse-practitioner, a pharmacy visit, additional adherence counseling if the nurse, doctor, or pharmacy deemed it necessary; individual or group wellness counseling if desired, and a blood draw if necessary.

Patient demographics were taken at baseline; patient weight and other basic measurements were taken at every visit to the clinic (though not at pharmacy-only visits). Patient viral load, CD4 count, hemoglobin, liver enzymes, and other lab values were evaluated at every scheduled visit; viral loads were not taken at

baseline. All laboratory assays were performed by the Auckland Park Regional Lab of the South African National Health Laboratory Service at Helen Joseph Hospital.

Adherence. Adherence data was not routinely or systematically collected in TLC. Records of pharmacy visits by the patients, which indicate when they picked up their ARV drugs and their next scheduled visit, have been used by other cohorts to ascertain adherence (**reference**). Using these data, adherence can be measured in two ways: first, as the proportion of intervals between two consecutive visits in which the patient was no more than 2 days late for the latter visit; and second, as the days captured in the pharmacy when the patient had drugs available to them. These adherence values constitute an upper bound on possible adherence, in that patients cannot be adherent on days when they have insufficient drugs; in contrast, patients can be non-adherent on days on which they have sufficient drugs.

These adherence measures are problematic for other reasons. First, they are more likely to be missing among those who become lost to follow-up very early, and second, they give high adherence scores to individuals with fewer opportunities to miss visits. For instance, an individual who shows up to their first visit on time and then dies before the next scheduled visit would have perfect adherence. This may suggest that high adherence causes death, and especially early death.

We did not control for adherence, but instead for successful suppression of viral load by month 6. As virologic suppression of virus is a direct result of adherence to therapy, we assumed that in this way, we have controlled the majority of any confounding due to adherence.

Loss to follow up and patient tracing. Individuals who did not return to TLCC in the 90 days after next scheduled visit were automatically flagged as potentially lost. These patients were called at least three times at telephone numbers provided at entry to the cohort; patients who could not be contacted in this way were visited by the home-based care team (HBCT) at the address they provided. If the HBCT failed to locate the patient, the patient was declared "confirmed" lost to follow-up.

In time to event analysis in which multiple imputation or inverse probability weighting was not used to deal with informative censoring, individuals lost to follow up were censored midway between the last recorded visit and the next scheduled visit; and potential lost to follow up patients were analyzed as confirmed lost to follow-up.

DEFINITIONS

Laboratory and clinical definitions for analysis. In general, we defined *baseline covariates* as those collected closest to time of initiation of HAART, and no more than four months before initiation of HAART. If no value was available previous to the initiation of HAART, we used values collected up to one month post HAART initiation only for those variables not likely to be substantially affected by the first month of HAART. Thus, while a value of BMI or CD4 count from two weeks post HAART initiation might be included, as these values rise only somewhat during

the first month, viral load measurements were never included if taken after the initiation of HAART.

Other variables were defined as follows.

Sex was coded dichotomously.

Age was coded categorically five-year exclusive increments, as <25, 25-30, 30-35, 35-40, 40-45, 45-50, 50-55, and ≥55.

Ethnicity was coded as African or otherwise.

Employment status was coded dichotomously as employed, student, or retired; or otherwise.

Virologic suppression was defined as viral load ≤ 400 copies/ml; patients were considered to have experienced initial virologic success if they achieved virologic suppression at any point within seven months of HAART initiation.

WHO disease stage was coded categorically as 1 or 2; 3; and 4, although we modeled all four categories separately in some analyses. In analyses that include tuberculosis as a main exposure, we controlled only for WHO stage IV, because TB is WHO stage III defining and thus there is substantial co-linearity between the exposure and WHO stage (WHO 2005).

Body mass index (BMI) was coded as a linear predictor or categorically as indicator variables as in the WHO classification as Underweight (<18.5kg/m^2), Normal (18.5-24.99kg/m^2), Overweight (25-29.99kg/m^2), and Obese (>30kg/m^2) (WHO 2008). Where height was missing but weight was present (n=28), we imputed height as the median height by gender. When both height and weight were missing (n=20), we imputed median BMI by gender.

Low hemoglobin was defined as a hemoglobin (Hb) value below 13.0 g/dl (men), 12.0 g/dl (women), or 11.0 g/dl (pregnant women) (WHO 2001). However, Johannesburg is at an elevation of 1753 meters (Wikipedia 2008) (higher than Denver, Colorado), and thus measured hemoglobin counts must be adjusted for altitude (Beutler and Waalen 2006; Cook et al. 2005; Ruiz-Arguelles 2006). Thus, the final cutoffs for low hemoglobin were 12.35, 11.35, and 10.35 g/dl, respectively.

History of antiretroviral therapy (ART) is rare in this cohort, with a total of 231 individuals (about 3% of the cohort) reporting any history of ART at all, including mono-, dual-, or triple-therapy. As such, we coded history of ART as a simple dichotomous indicator.

Initial HAART regimen was classified simply as efavirenz-containing, nevirapine-containing, Kaletra-containing.

CD4 count was included as both a categorical and a linear predictor in most models. Categories were widely accepted clinical categories of <50 cells/mm^3, 50-199 cells/mm^3, 200-349 cells/mm^3, and >350 cells/mm^3 (Badri et al. 2001; Churchyard et al. 2000; Mermin et al. 2004; Mocroft et al. 2003; Schim van der Loeff et al. 2002; van der Sande et al. 2004; Whalen et al. 2000; WHO 2005).

Pulmonary TB status at baseline was coded dichotomously, with the understanding that everyone recorded as having TB in this cohort was also being treated for TB. Thus, the exposure of pulmonary TB should be understood as the exposure of prevalent, treated, pulmonary TB.

Baseline pregnancy status and *baseline complaint of peripheral neuropathy* were both coded dichotomously.

Year of enrollment was coded as a three-category/two-indicator variable, measuring from first day of follow-up (April 1 2004) to April 1 in subsequent years.

DATA QUALITY

Data were entered into the TherapyEdge™ (TherapyEdge Inc., USA) clinical database. This database system is comprehensive, in that it has fields for every laboratory test imaginable in a clinical HIV context, as well as some advanced patient-monitoring display tools. The TherapyEdge system is unfortunately somewhat needlessly complicated for a rollout setting: it is somewhat too comprehensive, as well as not designed intuitively (though the design of data entry forms changed in small ways many times throughout the enrollment period, and in a substantive way at least once). As a result, there have been serious backlogs of data entry over the three years of enrollment. Due to these backlogs, as well as because of very high patient load – and thus high volume of data collected and entered into the database – double entry of data was not an option in this database (although it occurred inadvertently – see below).

One additional complicating factor in this database is that many laboratory values which are key to analyses were not entered directly into the database, but instead were provided to the data management team in the form of flat text files from the Auckland Park Regional Lab of the NHLS, mentioned above, without unique identifiers that could be traced back to our database directly. These values were matched with unique ids within the TherapyEdge database by means of two

independently computer programs written by data management personnel, which matched values between the two files by first initial, last name, date of birth, and gender. Because there were a number of misspellings as well as confusion over names (misspellings of last names are common; and in addition many black South Africans have two first names, and may report their first name differently depending on a number of cultural and interpersonal factors), we could not reject imperfect matches as necessarily incorrect. Thus, each computer program ranked matches by quality based on closeness of match, and included flags for whether the last name matched; to what degree, if any, the last name was mismatched (by both leftmost mismatched character, and the Levenshtein edit distance, the latter with the SAS complev function); whether the date of birth matched; and whether the sex matched. The two ranked lists of definite and possible matches were sent to me. I integrated only those observations ranked highly on both lists into the existing database.

Thus, there are undoubtedly at least some incorrect laboratory values in the database for analysis. However, some of the laboratory values in the NHLS data had already been entered into the TLCC database for lab results most essential to the clinic, especially for CD4 count. After merging the two sets of data by unique ID and month in which CD4 count was performed (because recorded days are often one day apart in the two databases), we compared CD4 counts from the unintended double entry database. Comparison of the 22645 doubly entered CD4 counts (out of 37116 total CD4 counts available for analysis) indicated that 22063 (97.43%) were identical, while the remaining 582 had CD4 counts that were divergent. Among those 582 divergent CD4 pairs, 192 were labs recorded within 1 day of each other;

71 of the 192 CD4 count pairs are within 20 CD4 cells of each other. Of those "mismatches" recorded between 2 and 10 days apart (n=224), 128 pairs are within 20 CD4 cells of each other. Of those where labs were taken more than 10 days apart (n=166), 112 were within 50 CD4 cell counts. Thus of the 582 mismatches, there are 311 plausible values; we therefore judge the prevalence of badly entered or mismatched CD4 counts to be approximately 1.2% comparing those values already in the Therapy Edge database to those that matched those values from outside values. Matching by week of CD4 instead of month indicates a 98.83% perfect matching percentage among CD4 counts; 95 of the 246 mismatched CD4 counts differed by less than 20 CD4 cells, indicating a 99.5% plausible match rate. We therefore believe that data quality is very high in this database.

STATISTICAL METHODS FOR CHAPTER 4

The purpose of Chapter IV was to address Specific aim 1, by characterizing outcomes over three years of a rollout of antiretroviral therapy in the Themba Lethu Clinical Cohort. We have previously described the study site, enrollment, eligibility, treatment, clinical procedures, data capture and quality ascertainment, and covariate definitions. Here we will concentrate on statistical analyses.

Baseline demographics were characterized using standard descriptive statistics. Chi-square tests were used to compare categorical variables and Wilcoxon rank-sum tests for continuous variables by category. Crude rates were estimated from simple Poisson models. Risks were estimated from simple

proportions. Risks and rates of drug toxicities were characterized using simple descriptive statistics, crude rates and rates controlled for sex using Poisson regression. Poisson regression was implemented in all cases using proc genmod in SAS with link=log and dist=Poisson. The Poisson model is formally stated as:

$$\ln(\text{rate of death} \mid X=x) = \beta_0 + \beta_j * x_j + e$$

where outcome is rate of death, e is distributed as Poisson, β_0 is the intercept, β_i is a vector of j modeled parameters, and x_i is a vector of j observed covariates. As is customary in the Poisson model, rate of death will be broken out into an indicator variable for death (yes/no), and the log of person-time follow-up for each individual will be included as the "offset" variable.

We examined four main outcomes in this study: initial virologic success, death, combined death or LFTU, and CD4 gain during follow-up. Analyses were intent-to-treat with regards to HAART, ignoring treatment interruptions and drug substitutions in time to event analyses.

Virologic outcomes. Initial virologic success was evaluated using descriptive statistics including risk of virologic suppression by six months by baseline HAART regimen and (separately) history of ART, evaluated by the chi-square test.

Mortality and loss to follow-up. Risk of death and combined risk of death or LTFU was evaluated using descriptive statistics and Kaplan-Meier curves stratified by baseline CD4 counts and examining twelve month mortality risk stratified by both baseline CD4 count and WHO stage (four separate categories).

We then evaluated multivariate hazard ratios for both outcomes using Cox proportional hazards model. Predictors in these models included gender, age, ethnicity, employment status, and baseline measurements of BMI, hemoglobin, WHO stage, CD4 count, pulmonary tuberculosis (TB) status, history of ART, pregnancy status, initial HAART regimen, and peripheral neuropathy at baseline. We used univariate log-log-survival plots to examine the univariate proportional hazards assumption (PHA) for each variable included in Cox models and examined linear and plausible categorical time interactions with each covariate using chi-square tests at an alpha of 0.10. We used time-stratified Cox models to estimate hazard ratios in the presence of violations of the PHA. The form of the Cox model to be used is:

$$h(t \mid X=x) = h(t \mid X=0) * e^{\beta jxj}$$

for all t, where (as above) β_j is a vector of estimated parameters, and x_j is a vector of observed covariates.

In the time-varying case with a categorical time interaction, this model is extended to

$$h(t \mid X=x) = h(t \mid X=0) * e^{\beta j1*xj1+ \beta j2*xj2(t)}$$

where $x_{j2}(t)$ is a function of x_{j1} such as $x_{j2}(t) = x_{1j} * I(month>3)$.

Where we are modeling the hazard at time t as a function of covariates X and as a fraction of the underlying (unspecified) hazard function h(t | X=0) such that hazard ratio at time t is constant and thus

$$HR(t) = h_x(t)/h_0(t) = e^{\beta x}$$

which in the time-varying case is extended to

$$HR(t) = e^{\beta j1 * xj1 + \beta j2 * xj2(t)}.$$

All the above models for mortality were complete case analyses with regards to missing data. Because baseline CD4 count was missing for approximately 650 individuals at baseline, we performed a sensitivity analysis to establish the impact of complete case analysis on multivariate estimates of mortality for baseline CD4 count. We created 10 datasets (using SAS proc mi) in which missing baseline CD4 values were imputed based on other variables in the model (WHO stage, BMI, sex, age, and hemoglobin count). Then we compared the effect of baseline CD4 count on hazard of death during follow-up in the complete-case and multiple imputation scenarios.

In addition, because we were concerned by the large ratio of recorded LTFU to recorded deaths, we used baseline inverse probability of censoring weights to estimate the effect of loss to follow-up on total crude incidence of death. The

implementation of IPCW is described in the next section (General statistical methods used in chapters IV and V).

CD4 gain on antiretroviral therapy. CD4 gain over time was first assessed using simple dichotomous measures at four months of follow-up, using t-tests to assess the crude effect of baseline CD4 on four-month CD4 gain. Mean and median CD4 count during follow-up for the entire cohort were estimated non-parametrically by month.

Mean CD4 count and CD4 count gain during follow-up was estimated by initial CD4 count category using two separate static change-point generalized estimating equations (GEE) models with an independent correlation matrix. In the first model, we included the main variable of months since HAART initiation, indicator variables for baseline CD4 count, and static change points at months 4, 10, and every 6 months thereafter until 28 months (5 changepoints) to allow the slope of CD4 increase to change in time (Chu et al. 2005). These changepoints were selected because they were the scheduled appointment times for regular follow-up visits; thus data at these time points should have been least affected by selection bias, whereas data at other time points was typically generated as a result of patients who arrived for care for specific reasons and thus may have been less representative of the cohort as a whole.

The second model included all terms from the first model, as well as interaction terms between all baseline CD4 count categories and continuous time as well as all changepoints (total 24 interaction terms), to allow slope to change in time and by initial CD4 count. Last, we performed one GEE analysis which controlled for

sex, age, and virologic success after month 6 among only those patients with a measure of virologic success at that time point.

The need to control for within-patient correlation suggests that either a generalized estimating equations (GEE) model or a general linear mixed model (random effects model) is necessary for this modeling. The former provides a population average parameter, while the latter provides subject-specific estimates; however, when estimating a collapsible linear outcome (CD4 count), these two quantities converge to the same estimate given large sample size. Since we have large sample size in this investigation, we conclude that either model is acceptable. Because we are using GEE models to implement marginal structural models for Chapters IV and V, we will use GEE models here for simple consistency.

The general form of the GEE model is

$$S(\beta) = \Sigma_{i=1 \text{ to } n} (d\mu_i/d\ \beta_i)*V_i^{-1} (y_i - \mu_i) = 0$$

where there are n subjects (i=1 to n) and t time points (j=1 to t), y_i is the vector of outcomes for individual i at t times points, μ_i is the vector of $E(y_i)$ for t time points, V_i^{-1} is the matrix estimating the variance-covariance structure over time, the heart of which is the correlation matrix R(α). We will have thousands of clusters even if we stratify the model completely by initial CD4 count, which will help ensure the robustness property of GEE using R(α); we will initially use the Unstructured correlation matrix specification, and examine estimated pairwise correlations to see if a more efficient correlation matrix is appropriate.

We should note here that the main disadvantage of the GEE model compared with random effects models is that the latter have implicit assumptions of missing at random rather than the implicit assumption of GEE models that missingness is completely at random (MCAR). Missing CD4 counts in this analysis are likely informative in this analysis, because low CD4 counts are a risk factor for death. Thus, because we are modeling only CD4 count by time, the missing at random (MAR) assumption of random effects models would imply that missingness is at random (and in particular, independent of the missing value) given time alone, which is not likely. As such, both models are likely to be biased when considering the unconditional estimate of CD4 gain among all individuals who initiated HAART in this cohort.

However, the overall goal of this analysis is to allow clinicians to understand where a particular, observed patient falls in the spectrum of observed CD4 gain at a particular time point. As such, we are interested in explicitly estimating CD4 gain *contingent on remaining in the cohort*. Under this interpretation, we can ignore missingness (or rather, assume that missingness is MCAR); either model will provide an unbiased estimate of this contingent outcome. Furthermore, this approach is typical of published approaches taken to the problem of estimation of CD4 gain (e.g. (Stringer et al. 2006)), and so – while clearly not ideal – we feel that it is an appropriate compromise position for this investigation.

If we *were* interested in the unconditional estimate of CD4 gain, we could take one of two approaches: first, we could implement a “hurdle model”, in which we first model probabilities of death and/or loss to follow-up, and then model the linear

outcome contingent on the outcome of the first model. Separately, we could use inverse probability weighting (see next section) to allow observations to "stand-in" for missing observations.

INTRODUCTION TO METHODS USED IN CHAPTERS V AND VI

Chapters V and VI both make extensive use of marginal structural models. Here we describe marginal structural models (MSM) briefly, and explain the ideas behind inverse probability weighting, which is used to estimate MSM.

What are marginal structural models? Marginal structural models (MSM) are a necessary methodology for assessing outcomes in longitudinal settings in which confounders are both time-varying and affected at later time points by the exposure (e.g., confounders of interest are on the causal pathway) (Hernán, Brumback, and Robins 2000; Robins, Hernán, and Brumback 2000).

For instance, consider a study of the relationship between receipt of antiretroviral therapy (ART) and death. CD4 count may cause the initiation of ART, and a low CD4 count is a risk factor for death; thus, at time of initiation of ART, CD4 is a confounder. But, as Figure 3.1 below, the receipt (or not) of antiretroviral therapy at time 0 affects subsequent CD4 count (at time 1), which in turn confounds the relationship between antiretroviral therapy (at time 1) and death at some later time point. As a result, CD4 count is on the causal pathway between ART (at Time 0) and Death (at Time 2).

Traditional regression analysis – a so-called 'stratification' approach - will control for CD4 count as it changes over time; but because CD4 is on the causal pathway, the resulting estimate of effect will be biased (Hernán, Brumback, and Robins 2000). Marginal structural models avoid this problem by using standardization rather than stratification, not dissimilar to the approach implicitly taken when we report age-standardized mortality rates.

There are at least two additional advantages of a marginal structural approach compared to more traditional regression analyses. First, given several important assumptions (Cole et al. 2003), the results of a marginal structural model are interpretable as the result we would have seen had we conducted a randomized controlled trial (RCT) of the exposure on the outcome. The assumptions that must hold true for this interpretation to be valid include the assumptions that we have correctly measured all variables, that we have sufficient covariate information to control for confounding, that we have correctly specified all models (Hernán, Brumback, and Robins 2000), and that the positivity condition holds (Hernán, Brumback, and Robins 2002; Hernán and Robins 2006). Positivity states that there is a non-zero probability that those who were exposed could have been unexposed, and that those who were unexposed could have been exposed. This condition intuitively holds true when we actually assign an exposure to individuals at random (as in an RCT) but it is not always true in observational settings. Intuitively, then, if the two groups were so different from each other that they were not comparable, we would have non-positivity. If these conditions hold, however – and many of them are no more heroic than assumptions commonly made in observational studies (Hernán,

Brumback, and Robins 2000), MSM can be used to construct (that is to say, estimate) post-hoc RCTs of exposures which were not – or could never be – randomized, such as tuberculosis infection or pregnancy.

Second, marginal structural models are fit with inverse probability weights (Robins, Hernán, and Brumback 2000), which are a general tool to ameliorate biases to missing information, such as is imparted by selection and information bias, as well as confounding; thus, a MSM approach lends itself easily to simultaneous control of multiple types of systematic error (e.g., (Hernán, Brumback, and Robins 2000)). This is a specific reason that MSM might be useful in a situation in which exposure and/or confounding are only important at baseline.

What are inverse probability weights? Inverse probability weights are a generalized methodology for controlling both confounding and selection bias. We will here concentrate on inverse probability of treatment weights, which control the confounding of a treatment-outcome relationship by (a vector of) covariates Z. We note without further comment here that "exposure" can substitute for "treatment" in all the following explanations.

Properly then, an individual's inverse probability of treatment weight (IPTW) is defined as exactly equal to $1 / P(T=t \mid Z=z)$, that is, the inverse of the probability that the individual received the treatment she actually received, given her covariates. Ideally, the vector of covariates $Z=z$ includes all possible confounders of the relationship between treatment; in reality, the IPTW is calculated based on only observed data (whereas an RCT controls confounding by all factors, observed and unobserved). While IPTW can be estimated non-parametrically (and we will do so

below), in reality they are typically estimated using parametric maximum likelihood regression. The estimation of IPTW is exactly analogous to the estimation of propensity scores, and is therefore a rich area for the application of recursive portioning algorithms (Breiman 2001; D'Agostino 1998; Glynn, Schneeweiss, and Sturmer 2006; McCaffrey, Ridgeway, and Morral 2004). This is not our concern at present, however.

Intuitively (and in some cases, mathematically (Hernán and Robins 2006)) inverse probability weights (IPW) are equivalent to standardization, a common basic epidemiologic technique for controlling confounding. When used in the context of marginal structural models, IPW are used to control confounding and also to control selection bias induced by loss to follow-up. While any attempt here to explain IPW will fall far short of the superlative explanation given by Hernán and Robins (Hernán and Robins 2006), the basics of IPW can be illustrated with a series of simple 2x2 tables.

In Tables 3.1a and 3.1b we show crude stratified 2x2 tables for an exposure-outcome relationship that is confounded by Z (only). Issues of sample size aside, the crude relationship of 3.000 is substantially different than the estimate adjusted for confounding by the Mantel-Haenszel method.

To apply inverse probability of treatment (exposure) weights to these data, we first estimate the conditional probability of observed exposure given value of Z, and the inverse of that quantity. Note that P(E=1 | Z=1) is calculated simply as 300 / (300+200) = 0.600; the results of these calculations are shown in Table 3.1c below.

The far right column of Table 3.1c is the inverse of P(E=e | Z=z); that is, exactly the inverse probability of treatment weight for this analysis.

Then the original populations are multiplied by their respective weights to construct pseudopopulation in which exposure and Z are independent (Table 3.1d). In this simple example, this step is equivalent to standardizing exposure to the distribution of Z in the whole population (Hernán and Robins 2006; Sato and Matsuyama 2003).

Note in Table 3.1d that the conditional probability P(E=e | Z = z) is now equal to 0.5 in both strata of Z; this demonstrates the independence of E and Z in the calculated pseudopopulation. Because E and Z are (now) independent, the crude estimate of the effect of the exposure on the outcome in the pseudopopulation will be unbiased with respect to Z. The last step of this analysis, therefore, is to combine the two populations back into a single population and calculate the crude risk ratio (Table 3.1e).

Note that the IPTW-risk ratio of 2.370 is very close to the MH-risk ratio of 2.375. Thus, we have shown how IPTW can control for confounding. Similar arguments apply to inverse probability of censoring weights.

Basic implementation of time-to-event marginal structural models using IPW. Generally speaking, the implementation of a marginal structural model for time-to-event data is as follows, assuming we are trying to estimate the effect of a Treatment on an Outcome given a set of k covariates Z measured as z. Assuming this, the MSM will attempt to estimate the quantities

P(Outcome | set T=1) and

P(Outcome | set T=0)

where we borrow “set T=1” terminology from Sato and Matsuyama (Sato and Matsuyama 2003) to indicate the situation had everyone in the target population (the whole population) were exposed, compared with everyone in the target population having been unexposed. This terminology implies the positivity assumption holds; that is that it is (at least probabilistically) sensible to talk about setting the exposure status of everyone in the population to one exposure or another. Our IPTWs are therefore attempting to estimate the conditional probability P(T=t | Z=z).

First, the data analyst constructs a discrete time hazard model (DTHM) for the data, such that each person-time-unit of experience is a separate observation. When analyzed using pooled log-binomial regression the DTHM will approximate proportional hazards regression, and with infinite sample size the two models are equivalent as the size of the person-time-unit approaches zero (or rather, the point where each person-time-unit is small enough that each contains only a single event). While DTHM for MSM are often analyzed using logistic regression (e.g. (Hernán, Brumback, and Robins 2000)), the use of logistic regression means that the model will estimate the hazard odds ratio rather than the hazard ratio itself (Muthén and Masyn 2005); the hazard odds ratio, in turn, should approximate the hazard ratio under the same conditions as the odds ratio will estimate the risk ratio: that is, when prevalence of the outcome is low.

Then the data analyst estimates weights for each person-time-unit in the DTHM (a process which is computer intensive and may well be a limiting factor on the size of the time-unit used in the DTHM). Probability of treatment (which is then trivially transformed into weights, see Appendix of (Hernán, Brumback, and Robins 2000)) are typically estimated using simple or polytomous logistic regression to predict the treatment, in a model such as:

$$\text{logit}(\text{treatment}) = \beta_0 + \beta_j * z_j$$

where treatment is (typically) categorical and often binary, logit is the log-odds transformation of risk, and i goes from 1 to k. (Continuous treatments are more difficult to deal with, and we do not address them in this dissertation, so discussion will be deferred.) If treatment is time-varying (as in Figure 3.1) then this model will contain both baseline and time-updated covariates.

If we also wish to estimate the effect in absence of censoring, then the causal effect being estimated is:

P(Outcome | set T=1 & set C=0) and
P(Outcome | set T= 0 & set C=0)

As a result, our weights must control for P(T=t & C=0 | Z=z). Using symbolic logic, we find that this statement is equivalent to P(T=t | Z=z) * P(C=0 | T=t, Z=z). Thus, in estimating the inverse probability of censoring weights, we use the model

$$\ln(\text{uncensored}) = \beta_0 + \beta_1 * T_i + \beta_j * x_j$$

IPTW and IPCW can be combined by simple multiplication; at this point, every observation of person-time has its own weight. Then a weighted GEE is used to implement the crude structural model

$$\ln(\text{pr}(\text{Outcome}_{jk}=1 \mid \text{Outcome}_{j(k-1)}=0) = \beta_0 + \beta_1 * T_{tk}$$

where Outcome_{jk} is the outcome for patient j at time k, and T_{jk} is the treatment (or exposure) for patient j at time k. In the case where the exposure violates the proportional hazards assumption, this generalizes to:

$$\ln(\text{pr}(\text{Outcome}_{jk}=1 \mid \text{Outcome}_{j(k-1)}=0) = \beta_0 + \beta_1 * T_{tk} + \beta_2 * T_{tk} * f_t(k)$$

where $f_t(k)$ is some function of current time k, such as I(k>3 months). These general models were adapted from Cole et al.(Cole et al. 2007).

EXTENSIONS OF MARGINAL STRUCTURAL MODELS

With simple IPTW (and IPCW, but here restricting discussion to IPTW for simplicity; all arguments apply analogously to IPCW), occasionally we will estimate a weight which is extremely large. As Robins says, “These few subjects will contribute

a very large number of copies of themselves to the pseudopopulation and thus will dominate the weighted analysis, with the result that our IPTW estimator will have a large variance and will fail to be approximately normally distributed." (Robins, Hernán, and Brumback 2000). Experience shows that this is especially true when including factors in the IPTW model which are strongly associated with exposure but only weakly associated with outcome.

We can stabilize by multiplying our weights by the unconditional probability of the treatment received (Robins, Hernán, and Brumback 2000), P(T=t), which has the effect of forcing the mean of weight to equal 1 and reducing the variance of the IPTW. However, stabilization by unconditional probability of treatment received may not sufficiently stabilize weights – and this, unfortunately, is where the implementation of IPTW becomes as much an art as a science. In this dissertation, we have chosen to use a traditional regression backwards-elimination strategy (in a non-weighted preliminary analysis) to determine which risk factors for exposure are true confounders of the outcome, and therefore which ought to be included in weight models, but other strategies are possible (and perhaps more desirable). For instance, van der Laan et al. have used recursive partitioning algorithms with multiple potential interaction and polynomial terms to search the space of possible model specifications, a method which makes fewer assumptions than the backwards-elimination method used here; however, their approach uses custom software and is neither well-documented nor independently well-validated.

Typically and more generally, when examining a time-varying exposure, covariates measured at baseline are used in both the stabilization (numerator) and

weighting (denominator) components of IPTW, while time-varying covariates are included only in the denominator. Note that this implies that the limit as time-varying covariates become randomized with respect to the exposure in this model (that is, are not predictive of exposure), stabilized IPTW will converge to 1. In general, a factor which is included in both the numerator and denominator of IPTW will not be fully controlled for in the structural model. Thus, true confounders of the exposure-outcome relationship must be included in the structural model as well, which transforms the structural model into the following:

$$\ln(\text{Outcome}_{ij} \mid x_{i0}) = \beta_0 + \beta_1 * T_{ij} + \beta_j * x_{j0}$$

where x_{ij} is the observed covariates for person i measured at time 0 – that is, baseline covariates (again, formula adapted from Cole et al.(Cole et al. 2007)). Note that in this model – which is common in the MSM literature (Cole et al. 2007; Cole et al. 2005; Cole et al. 2003; Hernán, Brumback, and Robins 2000; Lopez-Gatell et al. 2007) – the effect being estimated is no longer purely marginal, but is instead a marginal effect *conditional on* baseline covariates. This point is often lost in discussions in the literature. Additionally, according to Cole it is acceptable to include non-confounding variables in both the numerator and denominator of IPTWs to improve stabilization (i.e., bring variance of weights down) (Cole 2008).

While most published examples of MSM assume that the proportional hazards assumptions holds, there is no bar to relaxing the PHA by – for instance – including a time-interaction term in the main model; an example can be found in

Cole et al. (Cole et al. 2007). However, if MSM are applied to baseline exposure only, the analyst must include a time-interaction variable in the structural model to ensure a hierarchically well-specified model.

Finally, we note that in relaxing the PHA we face the same potential for selection bias as is faced in RCTs which evaluate outcomes in a subset of individuals selected after randomization. In either case, selection bias occurs when evaluating the effect of baseline covariates on the effect of a baseline exposure on an outcome which is conditional on not having had an earlier outcome. For example, if we are examining an outcome from months 0-2 and then from 3 onwards, the only individuals who will be at risk of the outcome at month 3 are those who did not experience the outcome by month 2. As a result, there is a selection for those at risk of the outcome in months 3 and onward; due to this selection bias, randomization is not guaranteed to hold after baseline (Shepherd et al. 2006).

METHODS AND DEFINITIONS SPECIFIC TO CHAPTER V

This section will concentrate on statistical methods, with references to the previous sections on MSM and IPTC weights. A directed acyclic graph-causal diagram (DAG) for the investigation in Chapter V is shown in Figure 3.2. We examined the effect of prevalent, treated, pulmonary TB at the time of HAART initiation on hazard of mortality during follow-up.

We described differences in baseline exposure category using descriptive statistics including chi-square tests to compare categorical variables, t-tests to compare means, and the Wilcoxon rank-sum test to compare medians.

We assessed the proportional hazards assumption for the exposure variable in univariate and multivariate analysis through visual inspection of survival curves, log-survival, and log-log-survival plots, as well as Cox tests of both continuous and categorical time interactions.

The first main analysis (model 1) examined the marginal effect of any TB at time of initiation of HAART on all-cause mortality during follow-up, adjusting for baseline confounders. The second analysis (model 2) examined the effect of time lag between initiation of TB therapy and initiation of HAART on mortality on HAART, where time lag was categorized in approximate quartiles of duration of TB treatment prior to initiation of HAART: >120 days, 60-120 days, 30-60 days, and <30 days.

Both models were implemented using IPTC weighted-MSM as described above, using log-binomial discrete time hazard model with one observation per person-month and an "independent" correlation matrix (Hernán, Brumback, and Robins 2000). We performed an equivalent analysis using one observation per person-week to ensure that one observation per person-month did not lose important information; person-months were found to be adequate.

In model 1, IPTW were estimated using simple logistic regression and stabilized with the unconditional prevalence of observed exposure (TB). In model 2, IPTW were estimated using polytomous logistic regression and stabilized similarly. We considered the following as predictors for both the weight and censoring models:

gender, ethnicity, employment status, age, history of antiretroviral therapy, history of TB; and baseline measures of pregnancy, peripheral neuropathy, hemoglobin (adjusted for sex, pregnancy status, and altitude), BMI, CD4 cell count, and WHO stage (IV or other), year of enrollment (measuring years starting at baseline), and whether treatment was initiated after October 2006 (when all fees for treatment were eliminated). In addition, time-varying censoring models accounted for time-updated measures of hemoglobin, CD4 count, and BMI, as well as for month of follow-up with both a linear month and cubic spline terms for month of follow up (derived using the SAS macro RCSPLINE(Harrell 2004), with knots at the approximate 5^{th}, 25^{th}, 50^{th}, 75^{th} and 95^{th} percentiles of survival time (Fairall et al. 2008; Hernán, Brumback, and Robins 2000)).

We did *not* control for HAART adherence, successful virologic suppression, or initial HAART regimen (NVP vs. EFV) because these factors either occur during follow-up or may be affected by the exposure (and would therefore be part of the causal mechanism of the exposure on the outcome).

If patients initiated TB therapy outside of the HAART clinic and then died before starting HAART, they would never enter the study (i.e., they would be left-truncated), and thus the mortality attributable to TB could have been underestimated. Consequently, our results must be interpreted conditionally - given that an individual who initiated TB therapy survived until initiating HAART, there was no effect of being on treatment for TB on mortality on HAART. At the same time, however, many (if not most) individuals initiate HAART because of the diagnosis of an opportunistic infection such as pneumonia or Cryptococcal meningitis; deferring

therapy among such individuals can markedly increase risk of death or progression to AIDS (Zolopa et al. Abstract 142 CROI 2008). Thus, it is unlikely that left truncation due to TB would bias these results unless left truncation due to TB acted differently than left truncation due to these other indications. To the extent that left-truncation remains a problem in this analysis, other have proposed inverse probability weighting approaches to left truncation-like problems (e.g. (Cole et al. 2004)); unfortunately, we did not have the data in this cohort to take such an approach and instead dealt with this in sensitivity analysis.

A related problem is that we may have adjusted for intermediates on the causal pathway between TB and mortality, thus biasing the effect estimate towards the null. If TB treatment, initiated 120 days ago, affects key risk factors for death (e.g., CD4 count or BMI) then controlling for baseline risk factors amounts to controlling for a factor on the causal pathway. We hypothesized that if this effect were significant, we would see evidence of increasing bias with increasing lead time; this was examined in the study. Similarly, in sensitivity analysis we limited exposure to recent TB only and to recent and early incident TB only, situations which should not suffer from this bias.

Another limitation of the study resulted from the lack of data on culture- or smear-confirmation of TB in most cases and likewise no information on drug resistance; as a result, there may have been misclassification of the exposure. We investigated the possibility of a sensitivity analysis using only those TB cases on which we had smear results, but the numbers were too small to enable such a sub-analysis. Thus, this is a limitation of the study to be discussed.

Sensitivity and secondary analyses. We performed seven sensitivity analyses. To limit the potential effects of individuals dying between initiation of PTB treatment and initiation of HAART (i.e., left-truncation), we restricted PTB to cases who initiated HAART within 30 days of start of PTB treatment (S1) and separately to cases who initiated PTB treatment 30 days before or after initiation of HAART (S2). To assess the effect of PTB on early mortality, we restricted follow-up to the first four months of HAART (S3). To determine the effect of the exposure among the exposed, we calculated the weighted standardized mortality ratio for PTB (S4).(Sato and Matsuyama 2003) We used a traditional proportional hazards regression approach (with adjustment rather than reweighting) (S5). To ascertain the extent of bias due to LTFU, we examined the main effect estimate without censoring weights (S6). Last, to ensure that the effects of this study were not limited to PTB, we repeated analysis 1 using both pulmonary and extrapulmonary tuberculosis as the exposure category (S7).

Last, we performed a secondary analysis to examine effect measure modification of PTB by key baseline risk factors for mortality in people living with HIV, including severe immunosuppression (CD4≤50 cells/mm^3 or WHO stage IV disease), malnutrition (underweight, BMI<18.5 kg/m^2), and anemia.

METHODS AND DEFINITIONS SPECIFIC TO CHAPTER VI

This specific aim requires several specific definitions, after which we will describe the statistical methods used in this analysis. A DAG for the baseline

analysis is shown in Figure 3.3, and a simplified DAG showing only components of the time-varying analysis are shown in Figure 3.4. Figures 3.2 and 3.3 are similar, largely because (a) the exposures in the two figures are very similar (though not identical), and (b) the outcome of stavudine substitution is related to issues of toxicities and virologic failure, which are in turn closely related to and causal of mortality.

The key differences between the exposure in this study and the previous study is that while in the previous study we examined treated pulmonary TB at time of HAART initiation, here we examine the exposure of TB treatment (for pulmonary and/or extrapulmonary TB) at both time of HAART initiation and during follow-up.

Study site, clinic procedures, data collection. This study was conducted on the Themba Lethu Clinical Cohort, a prospective cohort of adults initiating HAART. The clinic is located in Helen Joseph Hospital, Johannesburg, South Africa, and funded by the South African Government, Gauteng Province SA, PEPFAR, and USAID, and operated as a collaborative project with the Clinical HIV Research Unit of the University of the Witwatersrand and the non-profit Right To Care. Themba Lethu is one of the largest single clinics providing HAART in South Africa and has initiated patients on HAART at a rate of approximately 200 per month since April 2004. After initiating HAART, patients are scheduled for monthly pharmacy visits and clinical visits to Themba Lethu Clinic at 4 months and every 6 months thereafter. Clinical data as well as CD4 counts, hemoglobin, and other laboratory diagnostics are collected at all visits, except for viral loads which are not collected at baseline.

All data were stored in the TherapyEdge™ (TherapyEdge Inc., USA) database, and analyzed in SAS v9.1.3 (SAS Institute, NC, USA).

Therapy for TB is usually initiated outside of Themba Lethu Clinic. The standard course of TB treatment in South Africa is six months: two months of isoniazid, rifampicin, pyramizade, and ethambutol followed by four months of isoniazid and rifampicin (Republic of South Africa Department of Health 2004). The course of therapy is longer if the patient does not respond to treatment, has relapsed, or has drug resistant TB. Vitamin B6 (pyridoxine) is commonly prescribed at time of initiation of TB treatment. Amitriptyline is commonly prescribed for individuals experiencing peripheral neuropathy (Saarto and Wiffen 2007) in the Themba Lethu Clinic.

Definitions. Patients were eligible for this study if they initiated a stavudine-containing HAART regimen at Themba Lethu Clinic between April 1 2004 to March 31 2007 (Republic of South Africa Department of Health 2004).

The exposure of interest for the main analysis was TB treatment, defined as therapy for pulmonary or extrapulmonary TB which included isoniazid. There were three specific exposure categories: ongoing TB treatment at time of HAART initiation; recent initiation of TB treatment; and incident TB treatment during HAART. *Ongoing TB treatment* was defined as TB treatment started at least two weeks before initiation of HAART. A two-week cutoff was chosen based on the South African National ARV Guidelines recommendation to wait at least two weeks between initiation of TB treatment and HAART. *Recent initiation of TB treatment* was defined as initiation of TB treatment 0 to 14 days before initiation of HAART.

Last, *incident TB treatment* was defined as any TB treatment which began after initiation of HAART. A schematic diagram for these exposure categories is shown in Figure 6.1 in Chapter 6.

In analyses of ongoing and recent TB treatment, patients were censored at time of incident TB. In the analysis of incident TB treatment, individuals who had ongoing or recent TB treatment became eligible for incident TB one month after the recorded completion of their previous course of TB treatment, or nine months after the initiation of the previous course of therapy, if no end date was available. When these patients became eligible for the incident TB analysis, they were marked as having a history of TB treatment (just as patients with ongoing or recent TB treatment might have such a history). Individuals who experienced multiple episodes of incident TB were censored at the second episode.

The outcome of *all-cause d4T substitution* was defined as the event of substitution of a single drug (most often zidovudine [AZT]) for d4T while the rest of the regimen remained unchanged. Patients were still at risk during treatment interruptions and after drug substitutions for drugs other than d4T. Drug substitution events in which d4T and other drugs were substituted simultaneously (multi-drug substitutions) were considered competing events and cause for censoring in the main analysis, but were included as outcomes in a sensitivity analysis.

Statistical analyses. There were three main analyses, corresponding to the three exposures defined above. In all analyses, we controlled for confounding and censoring using inverse probability of treatment and censoring (IPTC) weights (Cole et al. 2007; Hernán, Brumback, and Robins 2000; Robins, Hernán, and Brumback

2000). Analyses were implemented as discrete time hazard models, and evaluated using log-binomial generalized estimating equations models, which controlled for within-individual correlation using an "independent" correlation matrix for robust standard errors in the presence of repeated observations (Hernán, Brumback, and Robins 2000).

Stabilized weights for TB treatment were estimated using ordinary logistic regression, and histograms of multivariate treatment weight models were examined to ensure that the positivity assumption (see Appendix) was met (Hernán and Robins 2006). The estimation of stabilized censoring weights was complicated because of the issue of competing risks (Matsuyama and Yamaguchi 2007) of multi-drug substitution, death, and lost to follow-up, all of which may be informative with respect to the outcome of stavudine substitution. In a marginal structural model, the most intuitive solution to this problem is to extend IPTC weights to encompass multiple causes of censoring (Matsuyama and Yamaguchi 2007). Because initial analysis showed little change in estimates and confidence intervals when modeling outcomes of lost to follow-up and death separately, we estimated a single weight for both outcomes and a separate weight for the competing risk of multi-drug substitution (except in sensitivity analysis).

Possible predictors for both the weight and censoring models were gender, ethnicity, employment status, age, history of antiretroviral therapy, previous history of TB treatment; and baseline measures of pregnancy, peripheral neuropathy, hemoglobin (adjusted for sex, pregnancy, and altitude), BMI, CD4 cell count, and WHO stage (IV or other), calendar date, and whether treatment was initiated after

October 2006 (when all consultation fees were eliminated by the Department of Health). In addition, time-varying censoring models accounted for time-updated measures of hemoglobin, CD4 count, BMI, and initial virologic success (measured at month 6), as well as for month of follow-up as both a continuous and categorical variable. Predictors were removed from weight models if they were not found to confound the outcome in traditional proportional hazards regression using backwards elimination (with an alpha of 0.10 for removal). Some of these variables were subsequently added back to both the weight and stabilization models to increase the stability of the final IPTC weights.

We did not account for administrative (right) censoring (AC) in our IPCW. This is in contrast to some reports of MSM in the literature (e.g. (Hernán, Brumback, and Robins 2000)) but in concert with others (e.g. (Cole et al. 2007)). Because all observed person-time is included for analysis in this cohort, to introduce selection bias AC would have to act such that (for instance) a first person-month of observation which took place from March 1 to March 31 2007 (the month before AC) was not exchangeable with a first person-month of observation which took place from April 1 to May 1 2004 (the first possible month of observation), conditional on the observed covariates. For this bias to be significant, then, there would need to be at least one important unmeasured confounder, the effect of which changes substantially in time. We believe that this situation is relatively unlikely, and that to the extent that it is possible, controlling for year of enrollment and for whether or not all clinical treatment was free of charge at time of HAART initiation, in addition to a

large number of demographic and clinical factors, makes it unlikely that AC will introduce significant selection bias into this investigation.

The components of the final treatment model for all three exposures are shown in Table 3.2, and further detail for these models are given in Appendix III. Note that the final model for incident TB included baseline factors in the final structural model, which was not true of the ongoing/concurrent weight models. As such, while the estimates from the ongoing and concurrent models can be interpreted as purely marginal estimates of effect, the estimate from the incident model must be interpreted as conditional on baseline covariates included in the final model.

Proportional hazards assumption in this analysis. Preliminary analysis indicated that the effect of TB treatment on hazard of stavudine substitution decreased markedly with time and thus violated the proportional hazards assumption. We therefore we included a categorical time-interaction term (Cole et al. 2007) to allow the hazard ratio (HR) to change at the beginning of the third month of concomitant TB treatment and HAART. We included an additional time-interaction term for months 3-6, as TB treatment lasts 6 months except in cases of drug resistance or retreatment (Republic of South Africa Department of Health 2004).

Thus, in this framework, the effect estimate for "7+ months" estimates the combined effect of (a) TB therapy from month seven onwards and of (b) the effect of having completed TB therapy (after a standard course of six months).

Secondary and sensitivity analyses. For each of the three main models, we performed several sensitivity analyses. We included multi-drug substitutions (excluding switches to second-line HAART) as outcomes (S1); we replaced the treatment portion of the IPTC weights with simple adjustment for confounding (S2); we truncated IPTC weights to the 99th and 1st percentile to ensure that no individual observations contributed excessively to effect estimates (S3); we estimated the relative effect of the exposure among the exposed (the standardized morbidity ratio) for each group (Sato and Matsuyama 2003) (S4); and we restricted the referent group to those without recorded history of TB (S5) to ensure that lingering peripheral neuropathy from previous drug exposures did not bias the results.

We performed one additional sensitivity analysis (S6) for ongoing TB treatment only, stratifying results by length of TB treatment prior to initiation of HAART in two categories chosen based on treatment guidelines (Republic of South Africa Department of Health 2004): 2 weeks to 2 months after initiation of TB treatment (15-60 days), and after two months (the continuation phase of TB treatment, 61+ days).

3.1a-e. Step-by-step example of how inverse probability of treatment weights work in a 2x2 table

Table 3.1a. Crude example 2x2 table.

Crude	Disease	No disease	Total	Risk	Risk ratio
Exposed	120	280	400	0.300	3.000
Not exposed	40	360	400	0.100	

Table 3.1b. Table 3.1a stratified by confounder Z.

Z=1	Disease	No disease	Total	Risk	Risk ratio
Exposed	110	190	300	0.367	2.444
Not exposed	30	170	200	0.150	
Z=0	**Disease**	**No disease**	**Total**	**Risk**	**Risk ratio**
Exposed	10	90	100	0.100	2.000
Not exposed	10	190	200	0.050	
Mantel-Haenszel estimate					2.375

Table 3.1c. Calculation of conditional probability of exposure.

Z=1	Disease	No disease	Total	P(E=e | Z=1)	1 / P(E=e | Z=1)
Exposed	110	190	300	0.600	1.667
Not exposed	30	170	200	0.400	2.500
Z=0	**Disease**	**No disease**		**P(E=e | Z=0)**	**1 / P(E=e | Z=0)**
Exposed	10	90	100	0.333	3.000
Not exposed	10	190	200	0.667	1.500

Table 3.1d. Application of inverse probability weights from Table 3.3.

Z=1	Disease	No disease	Total	P(E=e | Z=1)
Exposed	183.33	316.67	500	0.500
Not exposed	75.00	425.00	500	0.500
Z=0	**Disease**	**No disease**	**Total**	**P(E=e | Z=0)**
Exposed	30.00	270.00	300	0.500
Not exposed	15.00	285.00	300	0.500

Table 3.1e. Crude risk ratio in combined pseudopopulation.

ALL Z	Disease	No disease	Total	Risk	Risk ratio
Exposed	213.33	586.67	800	0.267	2.370
Not exposed	90.00	710.00	800	0.113	

Table 3.2. Variables included in final IPTW models for three exposures.

Covariate	Ongoing and concurrent		Incident		
	Weight	Stabilization	Weight	Stabilization	Marginal
Female	√		√	√	√
Ethnicity	√	√	√	√	
Employment	√	√	√	√	
Age	√		√	√	√
History of TB	√	√	√	√	
History of ART	√		√	√	√
Pregnant	√		√	√	√
Baseline PN	√		√	√	√
Hemoglobin					
Baseline	√	√	√	√	
Current			√		
BMI					
Baseline	√*	√*	√	√	√
Current			√		
CD4					
Baseline	√	√	√	√	
Current			√		
WHO stage	√	√	√	√	
Year/enrollment	√	√	√	√	√
Free treatment	√	√	√	√	√
Month/follow-up	√		√	√	
Virologic failure	√		√	√	

* indicates that this variable was coded differently (different referent groups) in the two models.

Figure 3.1. Directed acyclic graph for an analysis for which a marginal structural model is indicated.

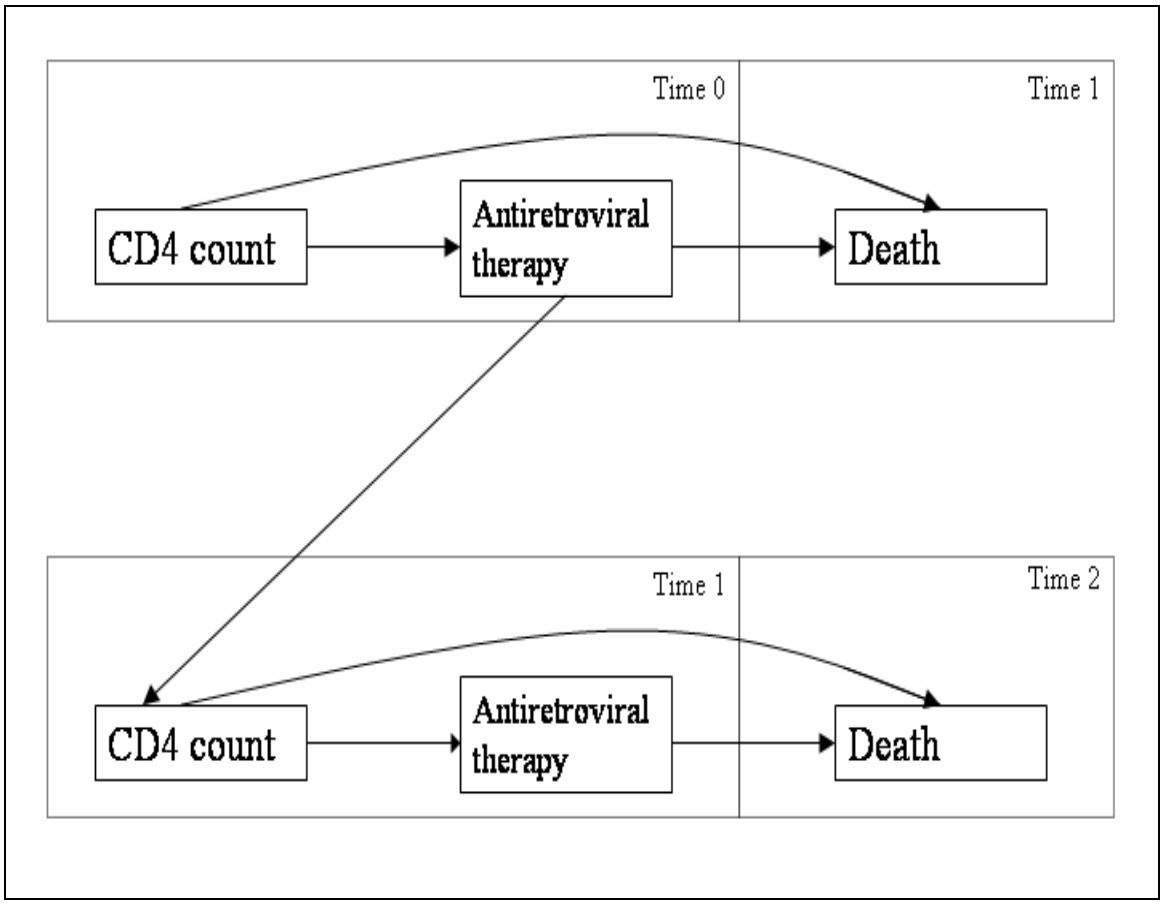


Figure 3.2. Directed acyclic graph of conceptual causal model for the effect of prevalent therapy for pulmonary TB at time of HAART initiation, and subsequent risk of mortality.

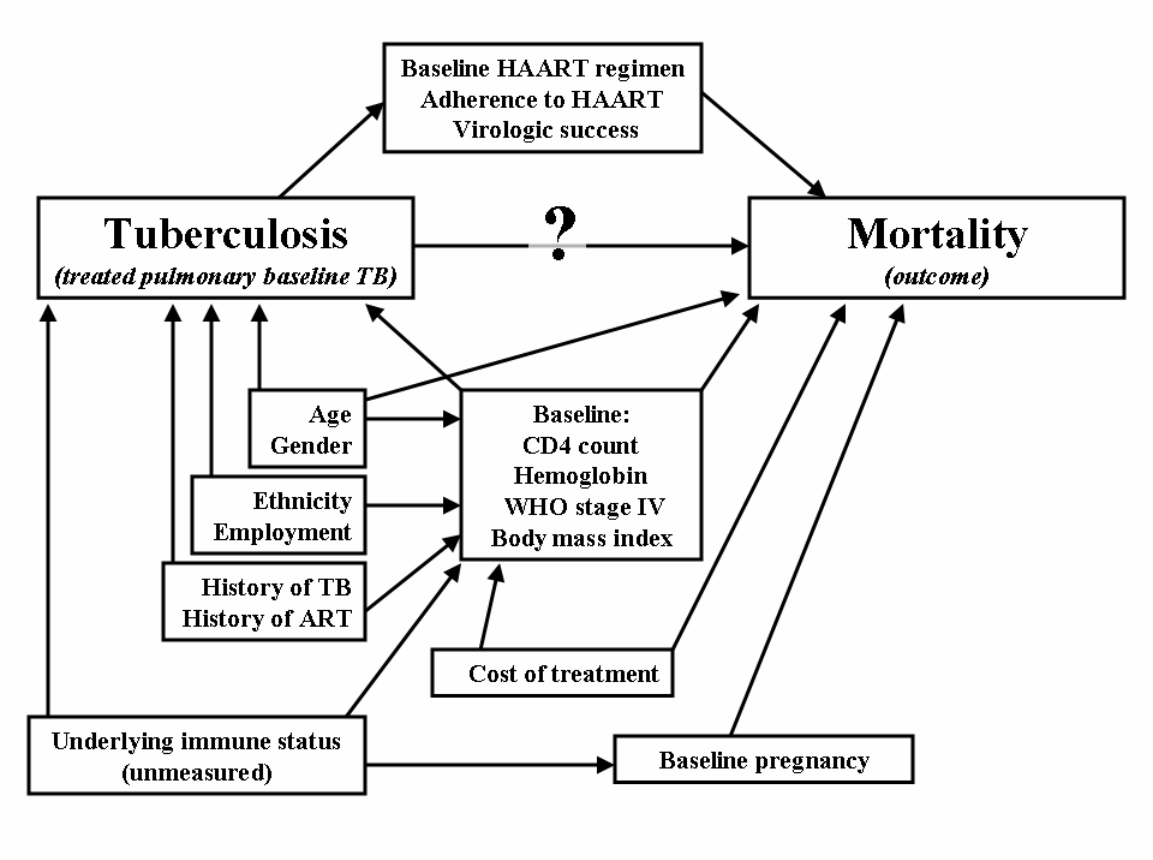


Figure 3.3. Directed acyclic graph of conceptual causal model for the effect of ongoing or concurrent TB therapy and risk of subsequent substitution of stavudine.

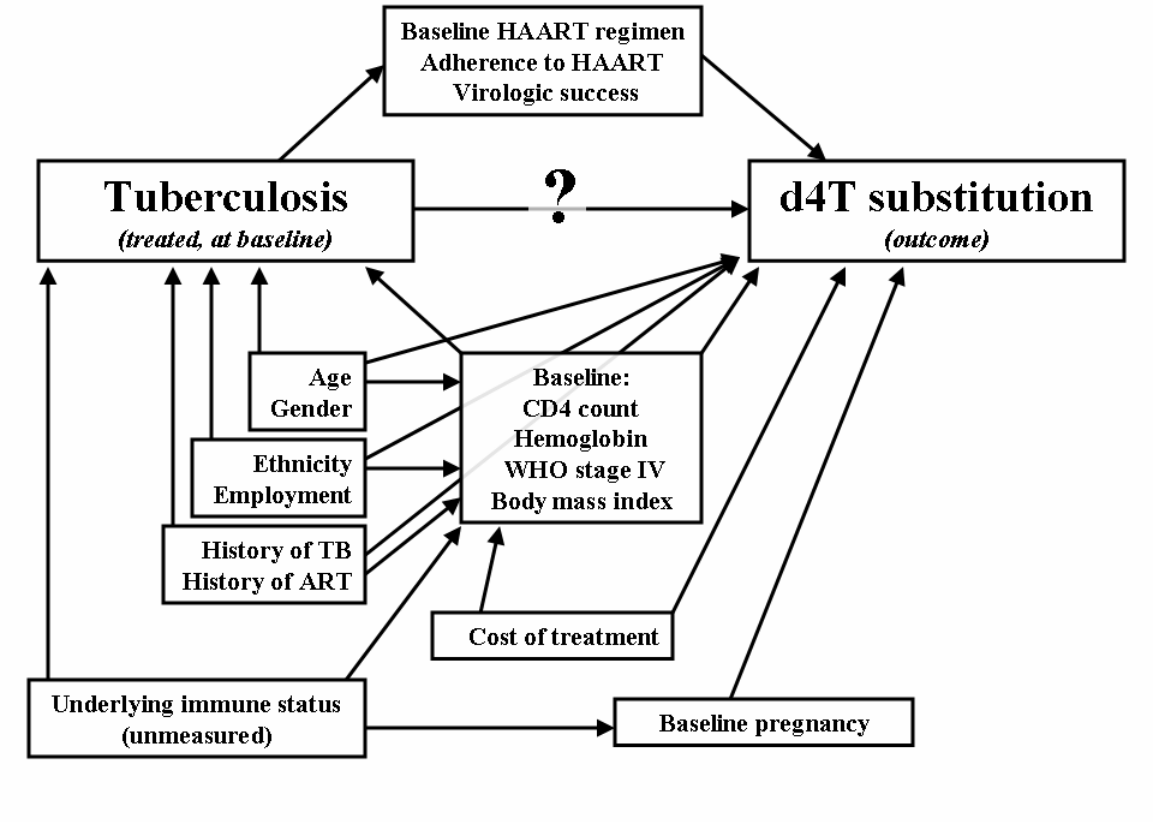
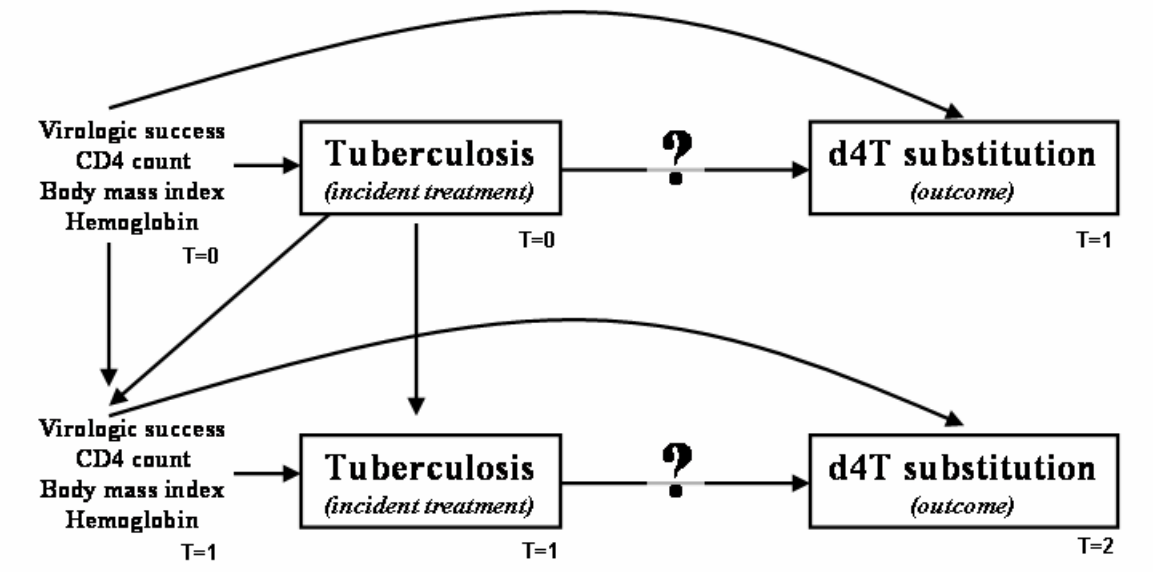


Figure 3.4. Simplified directed acyclic graph of conceptual causal model for the effect of incident TB therapy and risk of subsequent substitution of stavudine. This diagram is simplified in that it includes only time-varying confounders; baseline confounders of the relationship are shown in Figure 3.3.



CHAPTER 4

OUTCOMES OVER THREE YEARS OF ANTIRETROVIRAL THERAPY SCALE-UP IN JOHANNESBURG, SOUTH AFRICA: THE THEMBA LETHU CLINICAL COHORT

INTRODUCTION

Infection with human immunodeficiency virus (HIV) affects over 33 million people globally (UNAIDS/WHO 2007). Although global access to highly active antiretroviral therapy (HAART) has increased dramatically, the majority of those in need remain untreated, especially in sub-Saharan Africa. In response, the World Health Organization (WHO) has set the goal of universal HAART access by 2010 (WHO 2006). Large scale HAART rollouts using a public health approach are essential to meet this goal. Several reports from sub-Saharan Africa have confirmed the success of these rollouts and have provided evidence that HAART can save the lives of countless people living with HIV/AIDS (Bekker et al. 2006; Ferradini et al. 2006; Hawkins et al. 2007; Lawn, Badri, and Wood 2005; Lawn et al. 2006, 2006; Lawn et al. 2006; Lawn et al. 2005; Libamba et al. 2006; Stringer et al. 2006; Wools-Kaloustian et al. 2006).

Despite the success of these programs, much of current need for HAART remains unmet, a situation exemplified by the Republic of South Africa (SA). With

more than five million individuals living with HIV and AIDS, SA has the single largest population of HIV-infected individuals in the world (Kapp 2007; UNAIDS/WHO 2007). In August 2003, the South African Cabinet made a commitment to provide HAART in the public Health sector and an Operational Plan on Comprehensive HIV and AIDS Care, Management and Treatment for SA was published in November 2003. Unfortunately, although South Africa has the largest number of people receiving HAART in the world (South African National AIDS Council 2007), scale-up of HAART has been slower than anticipated, and the so-called treatment gap in South Africa remains in excess of 500000 individuals.

Extremely rapid expansion and rollout of HAART is required to meet the need in SA and throughout sub-Saharan Africa; however, the published literature still lacks comprehensive, rigorous assessments of outcomes of HAART in such situations. To this end, we describe clinic procedures, enrollment characteristics, and basic outcomes of individuals initiating HAART over three years of rollout and treatment in the Themba Lethu ("Our Hope") Clinic in Johannesburg, SA.

METHODS

Study site. The Themba Lethu Clinic (TLC), located in Helen Joseph Hospital in urban Johannesburg, is the single largest clinic providing HAART to HIV-infected adults in Africa, with over 10000 individuals in care by January of 2008. TLC is a joint project of the South Africa Department of Health Provincial HIV and AIDS Comprehensive Care Management & Treatment program, and Right to Care, a

non-profit organization dedicated to providing HAART and receives additional funding from the Gauteng Province Department of Health, United States PEPFAR, the United States Agency for International Development. TLC is supported by researchers and clinicians from the Clinical HIV Research Unit of the University of Witwatersrand. Patients were referred to TLC from voluntary counseling and testing clinics, hospitals, prenatal care facilities, and by self-referral. All patients in TLC received antiretroviral medications free of charge. Patients paid 35 Rand (about $5 US) per clinic visit prior to October 2006, after which all fees were eliminated.

Treatment and clinical care. As per the South African national guidelines, the majority of first-line HAART regimens given in the TLC included stavudine (d4T), lamivudine (3TC), and either efavirenz (EFV) or nevirapine (NVP) (Regimens 1a and 1b, respectively). Some individuals – mostly pregnant women – received Kaletra® (lopinavir boosted with ritanovir, LPVr) in place of EFV or NVP; a small number of individuals initiated zidovudine (AZT) instead of d4T. Second-line HAART typically included AZT, didanosine (DDI), and LPVr.

Patients visited TLC at time of HAART initiation (month 0); regular follow up visits were scheduled at months 4 and every six months thereafter. A typical clinic visit involved initial assessment by a nurse, evaluation by a doctor or nurse-practitioner, pharmacy visit, blood draw, and adherence counseling if necessary; in addition, patients could opt to receive individual or group wellness counseling. Patients also visited the clinic one month after each assessment visit to receive test results and adjust medications. Patients who had not suppressed HIV by month 4 were scheduled for monthly visits until viral suppression was achieved, and patients

who experienced drug toxicities or other health problems were seen as required. In addition to clinical visits, patients visited the pharmacy monthly or bi-monthly to receive prescription medications.

Patient demographics and clinical measurements were recorded at every clinic visit; laboratory measures including viral load, CD4 count, and hemoglobin were collected at baseline (except viral load) and at each regular follow-up visit. All laboratory assays were performed by the South African National Health Laboratory Service (NHLS).

Data capture and quality. Throughout the three years, data from clinical records were entered into the TherapyEdge™ (TherapyEdge Inc., USA) database. Results were regularly quality controlled by a senior nurse or data quality assurance manager. Some laboratory values (CD4 counts, viral loads, and hemoglobin) were integrated into existing data from an NHLS database. The duplicate entry of a subset of laboratory values in both TherapyEdge and the NHLS database served as a de facto validation study of quality of data entry into TherapyEdge.

Definitions. Patients who did not show up for a scheduled appointment (including pharmacy visits) were categorized as *defaulters*; patients more than three months late were categorized as *lost to follow-up* (LTFU). LTFU designations were considered *confirmed LTFU* if the follow-up team was unable to locate the patient after three phone calls and a home visit. LTFU patients were censored midway between the last recorded visit and the next scheduled visit if available, or at last recorded date of receipt of therapy.

Collectively, *drug regimen changes* comprised both individual drug substitutions (e.g., the substitution of AZT for d4T), or drug regimen switches, from first to second line HAART. *Virologic suppression* was defined as viral load ≤ 400 copies/ml, and *initial virologic success* was defined as virologic suppression within six months of HAART initiation. *Body mass index* (BMI) was categorized as underweight (< 18.5 kg/m^2), normal (18.5-24.9 kg/m2), overweight (25-29.9 kg/m2), and obese (≥30 kg/m2) (WHO 2008). *Anemia* was defined as a hemoglobin (Hb) value below 13.0 g/dl (men), 12.0 g/dl (women), or 11.0 g/dl (pregnant women) (WHO 2001). Because Hb increases with elevation (Cook et al. 2005; Ruiz-Arguelles 2006; WHO 2001) and Johannesburg is 1753 meters above sea level (Wikipedia 2008), Hb values were down-adjusted by 0.65 g/dl (WHO 2001). *Baseline covariates* were those collected closest to time of initiation of HAART, and no more than four months before initiation of HAART or one month after initiation.

Statistical analysis. Patients who initiated HAART at TLC between 1 April 2004 and 31 March 2007 were eligible for the analysis. Patients were excluded if they had missing or incomplete information on initial HAART regimen or initiated a non-HAART ART regimen in TLC. Baseline demographics were characterized using standard descriptive statistics. Chi-square tests were used to compare categorical variables and Wilcoxon rank-sum tests for continuous variables by category. Rates were estimated from crude Poisson models; risks were estimated from simple proportions.

We examined four main outcomes in this study: initial virologic success, death, combined death or LFTU, and CD4 gain during follow-up. Analyses were

intent-to-treat with regards to HAART, ignoring treatment interruptions and drug substitutions in time to event analyses.

Initial virologic success was evaluated using descriptive statistics. To examine risk of death and combined risk of death or LTFU, we created Kaplan-Meier curves stratified by baseline CD4 counts and examined mortality risk stratified by both baseline CD4 count and WHO stage. In addition, we used Cox proportional hazards models to estimate hazard ratios for predictors of death and combined death or LFTU; predictors in these models included gender, age, ethnicity, employment status, and baseline measurements of BMI, hemoglobin, WHO stage, CD4 count, pulmonary tuberculosis (TB) status, history of ART, pregnancy status, initial HAART regimen, and peripheral neuropathy at baseline. We used univariate log-log-survival plots to examine the proportional hazards assumption (PHA) for each variable included in Cox models and examined linear and plausible categorical time interactions with each covariate using chi-square tests at an alpha of 0.10. We used time-stratified Cox models to estimate hazard ratios in the presence of violations of the PHA.

CD4 gain over time was assessed using two linear generalized estimating equations (GEE) models with an independent correlation matrix. In the first model, we included the main variable of months since HAART initiation, indicator variables for baseline CD4 count, and static change points at months 4, 10, and every 6 months thereafter until 28 months (5 changepoints) to allow the slope of CD4 increase to change in time (Chu et al. 2005). These changepoints were selected because they were the scheduled appointment times for regular follow-up visits; thus

data at these time points should have been least affected by selection bias, whereas data at other time points was typically generated as a result of patients who arrived for care for specific reasons and thus may have been less representative of the cohort as a whole (see Discussion). The second model included all terms from the first model, as well as interaction terms between all baseline CD4 count categories and continuous time as well as all changepoints (total 24 interaction terms), to allow slope to change in time and by initial CD4 count.

Sensitivity analysis and missing data. In a small number of individuals missing baseline height (n=48) or weight (n=20), median height and weight for gender was assigned for the missing value. Where baseline CD4 counts (N=639) or outcomes were missing, complete case analysis was performed. We tested the validity of complete case analysis for CD4 counts by performing a sub-analysis in which we used multiple imputation (n=10) to fill in values of baseline CD4 count based on other variables in the model (including WHO stage, BMI, sex, age, and hemoglobin count), and compared the effect of baseline CD4 count on hazard of death during follow-up in the complete-case and multiple imputation scenarios. Last, because we were concerned by the large ratio of recorded LTFU to recorded deaths, we used baseline inverse probability of censoring weights to estimate the effect of loss to follow-up on total crude incidence of death.

RESULTS

Enrollment and baseline characteristics. Between 1 April 2004 and 31 March 2007, a total of 7583 or approximately 210 patients per month initiated first-line HAART in TLC. Data from 7519 patients were available for analysis; 40 were excluded due to missing demographic information, and 24 were excluded because of errors in recorded dates of therapy. Baseline characteristics of the 7519 individuals are summarized in Table 4.1. Two-thirds (66.5%) of patients were women; mean age at start of HAART was 35.4 years in women and 38.7 years in men ($p < 0.0001$). More than half (56%) were unemployed. At time of ART initiation, 18.4% had a BMI < 18.5. Median CD4 count at ART initiation was 88 (IQR 31-158); 34.4% of individuals initiating HAART had CD4≤50, 55% had CD4≤100 and 88% had CD4 < 200.

Therapy initiation and durability of first-line HAART. The majority (78.1%) of patients initiated d4T-3TC-EFV; smaller numbers initiated d4T-3TC-NVP (7.8%) and d4T-3TC-LPVr (8.2%), and 5.9% initiated other regimens. Those initiating LPVr regimens were more likely to be pregnant women (unadjusted OR 147.8, 95% CI 114.0-191.4).

During 99008 person-months follow up (mean 13.2 months per patient, median 11 months), 1888 (25.1%) patients experienced any change in their drug regimen, at a median of 273 (IQR 141-431) days of therapy. The crude rate of drug regimen change was 28.5 (95% CI 27.2 – 29.8) per 100 person-years, and women were twice as likely to experience a drug change than men (HR for change 2.08,

95% CI 1.86-2.33). 172 patients (2.3%) patients were switched to second-line HAART, a rate of 2.1 (95% CI 1.8-2.5) drug regimen switches per 100 person-years. Among those who switched to second line therapy, median time on first drug regimen was 336 days (IQR 236-553).

Viral suppression. Among individuals who had viral load measured within six months of HAART initiation (n=5077), 4502 (88.7%) had suppressed virus to < 400 copies/ml. Patients on d4T-3TC-NVP suppressed virus at slightly lower rates than patients on other first-line regimens (85.7% vs. 88.9%, chi-square p=0.05). In addition, a smaller proportion of patients with any history of anti-retroviral therapy previous to initiation of HAART in TLC suppressed virus (no vs. yes, 85.2% vs. 88.8%, chi-square p=0.15).

Survival and loss to follow-up. 7519 individuals were available for survival analysis. 298 (4.0%) patients died during follow-up; 111 (37.3%) within 60 days and (cumulatively) 204 (68.5%) within 180 days of initiating HAART. 1425 patients (19.0%) were lost to follow up, 382 (26.8%) of those within 30 days and (cumulatively) 878 (61.1%) of those within 180 days. The crude rate of death over all of follow-up in this cohort was 3.6 deaths per 100 person-years (95% CI 3.2-4.1) over all of follow-up; in the first three months of follow-up crude mortality rate was 8.6 (7.3-10.2) per 100 person-years; in first six months of follow-up, crude mortality rate was 6.8 per 100 person-years (95% CI 6.0-7.8). Crude mortality rate rose to 7.3 (95% CI 6.3-8.5) per 100 person years in the first six months when we accounted for LTFU using baseline IPMWs. Kaplan-Meier curves for survival and combined outcome of alive and in care, stratified by CD4 count, are shown in Figure 4.1. In

addition, crude twelve month risk of death stratified by WHO stage and CD4 count is shown in Figure 4.2.

In univariate proportional hazards analysis, several baseline characteristics were associated with increased risk of mortality, including male gender, older age, low BMI, low hemoglobin, WHO stage 3 or 4, low CD4 count, prevalent TB, and receiving EFV-based HAART (compared to LPVr- or NVP-based HAART). Pregnancy and history of ART were protective. Estimated univariate hazard ratios for the same risk factors were similar but attenuated (closer to the null) in analysis for the combined outcome of death or LTFU (not alive or not in care); exceptions were pregnancy and use of LPVr, both of which were both associated with increased LTFU.

Hazard ratios from multivariate Cox proportional hazards models for the outcomes of death and death or LTFU are summarized in Table 4.2. Key predictors of death included age > 55 (HR 2.36, 95% CI 1.41-3.94), CD4 count ≤ 50 compared to 50<CD4≤200 (HR 2.54, 95% CI 1.94-3.32), and WHO stage 3 (HR 1.47, 95% CI 1.10-1.97) or 4 (1.96, 95% CI 1.37-2.80), compared to WHO stages 1 and 2. Hazard ratios were similar but generally closer to the null in a model for the combined outcome of death or LTFU; however, pregnancy at baseline was associated with twice the risk of death or LTFU (HR=2.03, 95% CI 1.54-2.68), but not death alone (HR=0.84, 95% CI 0.27-2.59). Hazard ratios for these risk factors on death alone are reported separately for first three months of follow-up and the remainder of follow-up in Table 4.3.

CD4 response. A “four month” measurement of CD4 count was available for 5199 patients after a median of 118 days (IQR 110-137 days, range 61-213 days); median and mean CD4 count at four months were 205 and 226 cells/mm3, respectively, and 23.9% of patients with a CD4 measurement at four months had gained fewer than 50 CD4 cells/mm3 at this time. Four-month CD4 gain differed by baseline CD4 count: individuals with a baseline CD4 count of ≥200 gained a mean of 73 CD4 cells/mm3 while those with CD4<200 gained a mean of 123 CD4 cells/mm3 ($p<0.0001$). Mean and median CD4 count during follow-up for the entire cohort was estimated non-parametrically by month and is shown in Figure 4.3. Mean CD4 count and CD4 count gain during follow-up was estimated by initial CD4 count category using a static change-point generalized estimating equations model; these estimates are given in Tables 4.4a and 4.4b. The predicted mean CD4 counts for the whole population at baseline and at four months from this model are 111 and 228 cells/mm3 respectively, very close to the observed means of 109 and 226 cells/mm3.

Mean gain in CD4 cell count at six months was estimated at 134 (95% CI 130-138) cells/mm3 and 188 (95% CI 183-193) cells/mm3 at twelve months. Patients who suppressed virus within six months experienced substantially greater CD4 count gains at six months compared to those who did not suppress virus. In a model for the entire cohort that that controlled for baseline CD4 count, sex, and age, the estimated mean CD4 gain was 43 (95% CI 34-52) cells/mm3 higher at six months post HAART initiation among those who achieved virologic success compared to those who did not.

Major drug toxicities. 845 individuals experienced incident peripheral neuropathy after a median of 173 days (IQR 88-329) of HAART, corresponding to a rate of 11.6 cases per 100 person-years of follow-up (95% CI 10.8-12.4). Women experienced a slightly lower rate of this toxicity than men (incidence rate ratio 0.80, 95% CI 0.69-0.92).

404 individuals experienced incident lactic acidosis (n=203) or symptomatic hyperlactatemia (n=201), after a median of 378 days (IQR 302-519), corresponding to an overall rate of 5.1 (95% CI 4.6-5.6) per 100 person-years. Women experienced this toxicity at 2.91 times the rate of men (95% CI 2.20-3.85).

367 individuals experienced incident lipodystrophy, after a median of 477 days (IQR 322-614), corresponding to a crude rate of 4.6 (4.1-5.1) cases per 100 person-years. Women experienced lipid disorders at 4.36 (95% CI 3.09-6.16) times the rate of men.

Data quality and sensitivity analysis. We evaluated the laboratory values entered by hand into TherapyEdge to those matched by patient id and date from the NHLS databases; laboratory values matched between the two databases in 98.8% of cases evaluated. In sensitivity analysis, the crude estimates of effect for baseline CD4 category on hazard of death did not change after multiple imputation analysis for missing baseline CD4 (results not shown).

DISCUSSION

Rapid scale-up of HAART is essential to confronting the HIV/AIDS epidemic in South Africa and throughout the world, but will only succeed if that scale-up does not come at the cost of a high level of care. These results demonstrate that a high level of clinical care can consistently be delivered in a rollout setting. This is particularly significant given the very high enrollment rates seen at TLC, which initiated an average of 210 new patients on HAART every month over the period of this study; this rate of enrollment is nearly 4 times higher than the estimated 50-60 patients/month/clinic initiation rate that can be calculated from comparable, published reports (Bekker et al. 2006; Ferradini et al. 2006; Hawkins et al. 2007; Lawn, Badri, and Wood 2005; Lawn et al. 2006, 2006; Lawn et al. 2006; Lawn et al. 2005; Libamba et al. 2006; Stringer et al. 2006; Wools-Kaloustian et al. 2006).

Our results show relatively low mortality, below 5% during three years of follow-up, suggesting that HAART can be extremely effective at reducing mortality in resource poor settings with very high rates of enrollment. However, sensitivity analysis by inverse probability weighting suggests true incidence of death is likely to be moderately higher than our observed rate of death. This is because, while the TLC staff undertook great efforts to track down individuals who became LTFU in this cohort, it is highly likely that at least some deaths were mistakenly classified as LTFU in this analysis. This misclassification is likely to be non-differential with regard to predictors of this outcome; thus, predictors in Tables 4.2 and 4.3 are likely slightly biased towards the null. Use of IPMWs increased the overall estimate of death rate

in the first six months only slightly from 6.8 to 7.3 per 100 person-years, suggesting that the magnitude of bias is likely to be small. However, the IPMWs will only correct bias due to missing data to the extent that there are no unmeasured predictors of missingness, an unverifiable assumption.

As with comparable studies, we found that the majority of both deaths and LTFU occur early (Braitstein et al. 2006; Ferradini et al. 2006; Lawn et al. 2005; Stringer et al. 2006), Extending the results of other studies, we found several variables affected risk of death differently in early and later follow-up, violating the proportional hazards assumption. In particular, we found that while the effect of baseline CD4 count and low BMI did not appear to change over time, the effect of older age and WHO stage were both more strongly associated with risk of death in the first three months of follow-up.

Our estimates of CD4 gain over time are concordant with other estimates from the literature, suggesting that these results are generalizable to other situations and may therefore be of use to clinicians seeking guidance on what CD4 response is "normal" over several years of HAART. However – as with other models of CD4 gain in the literature – these results must be considered conditionally, keeping in mind that they are representative only of those individuals who remained alive and in care and had CD4 counts measured. However, as clinicians typically want to apply such estimates of CD4 count to exactly those patients who remain alive and in care, this limitation should not affect the usefulness of these data.

Despite potential limitations of these data, however, these results are in accord with other key reports of HAART scale-up in the literature, and extend

previous findings in several key ways. This report describes results on the most rapid scale-up of HAART yet reported in the literature, and demonstrates that even with a rate of HAART initiation approximately four times that previously reported, high quality of care can be maintained in resource challenged settings.

Table 4.1. Baseline characteristics of 7,543 individuals enrolled in Themba Lethu Clinical Cohort (TLCC) from 1 April 2004 to 31 March 2007.

Characteristic, lab value, condition	Count (%) by sex		P-value
	Female (n=5015)	Male (N=2528)	
Age (Median and IQR)			
Mean (SD)	35.4 (8.4)	38.7 (8.4)	<0.0001
Median (IQR)	34.0 (29.4-40.2)	37.2 (32.8-43.3)	
Ethnicity			
African or black	4817 (96.1)	2378 (94.1)	<0.0001
Coloured	167 (3.3)	97 (3.8)	
Asian, White, or Missing	31 (0.6)	53 (2.1)	
Employment			
Employed, student, retired	1650 (32.9)	1086 (43.0)	<0.0001
Unemployed	2977 (59.4)	1263 (50.0)	
Unknown or missing	388 (7.7)	179 (7.1)	
BMI			
< 18.5	796 (15.9)	642 (25.4)	<0.0001
18.5 – 24.9	2593 (51.7)	1543 (61.0)	
25.0 – 29.9	1064 (21.2)	280 (11.1)	
≥ 30	550 (11.0)	55 (2.2)	
Missing	12 (0.2)	8 (0.3)	
Hemoglobin*			
Normal	2496 (49.8)	1202 (47.6)	0.1892
Low	2463 (49.1)	1297 (51.3)	
Missing	56 (1.1)	29 (1.2)	
WHO stage			
I	2372 (47.3)	1012 (40.0)	<0.0001
II	631 (12.6)	286 (11.3)	
III	1558 (31.1)	929 (36.8)	
IV	454 (9.1)	301 (11.9)	
CD4 count			
Median (IQR)	*94 (37-164)*	*74 (23-146)*	*<0.0001*
Mean (95% CI)	*114 (111-118)*	*98 (94-102)*	*<0.0001*
<50	1423 (28.4)	953 (37.7)	<0.0001
50 – 99	976 (19.5)	453 (17.9)	
100 – 199	1569 (31.3)	703 (27.8)	
200 – 349	459 (9.2)	169 (6.7)	
≥ 350	135 (2.7)	61 (2.4)	
Missing	453 (9.0)	189 (7.5)	
Pulmonary tuberculosis	700 (14.0)	504 (19.9)	<0.0001
History of ART	146 (2.9)	78 (3.1)	0.3224
Pregnant	574 (11.5)	NA	NA
Peripheral neuropathy (at baseline)	177 (3.5)	143 (5.7)	<0.0001

All figures are expressed as number (% total). Marital status, age, ethnicity, employment status, residency, education, referral status, alcohol, and smoking are by self-report. P-values are by chi-square test of general association across all categories by sex (all variables are categorical). *Lower limit of normal hemoglobin is 12.9 g/dl for men; 11.5 g/dl for non-pregnant women; and 10.5 g/dl for pregnant women.

Table 4.2: Multivariate hazard ratios for baseline risk factors for outcomes of mortality and mortality or LTFU (combined outcome)

Baseline risk factor	Outcome of mortality (only)		Outcome of mortality or LTFU (combined)	
	HR	95% CI	HR	95% CI
Male gender	1.13	0.88-1.45	1.27	1.14-1.42
Age > 55 vs. age ≤ 55	2.31	1.38-3.86	1.24	0.94-1.63
African ethnicity vs. other	0.81	0.48-1.37	0.94	0.75-1.18
Employed vs. other	0.98	0.76-1.28	0.87	0.78-0.97
Body mass index				
BMI < 18.5	1.80	1.39-2.33	1.38	1.22-1.56
BMI >18.5 and BMI < 30	Reference			
BMI ≥ 30	0.87	0.45-1.66	0.92	0.75-1.14
Hemoglobin low vs. normal	1.60	1.21-2.10	1.34	1.20-1.49
WHO stage				
I or II	Reference			
III	1.53	1.15-2.05	1.10	0.98-1.25
IV	2.01	1.41-2.88	1.29	1.09-1.52
CD4 count				
CD4 ≤ 50	Reference			
50 < CD4 ≤ 100	0.46	0.33-0.65	0.72	0.63-0.83
100 < CD4 ≤ 200	0.35	0.25-0.49	0.74	0.65-0.84
CD4 > 200	0.28	0.15-0.55	0.84	0.70-1.01
Pulmonary tuberculosis	0.89	0.66-1.20	0.88	0.77-1.02
History of ART	0.71	0.26-1.98	0.57	0.39-0.82
Pregnancy	0.84	0.27-2.59	2.03	1.54-2.68
Baseline HAART regimen				
Contains efavirenz	Reference			
Contains nevirapine	0.62	0.34-1.14	0.90	0.73-1.10
Contains lopinavir-ritanovir	0.65	0.25-1.70	1.03	0.79-1.35
Peripheral neuropathy	1.09	0.67-1.77	1.00	0.80-1.27

Table 4.3. Hazard ratios (HRs) for baseline risk factors for outcomes of early (≤3 months) and late (>3 months) mortality.

Baseline risk factor	Early mortality (within 3 months)		Late mortality (more than 3 months)	
	HR	95% CI	HR	95% CI
Male gender	1.13	0.79-1.61	1.14	0.80-1.61
Age > 55 vs. age ≤ 55	2.93	1.47-5.84	1.81	0.84-3.92
African ethnicity vs. other	0.77	0.35-1.65	0.84	0.41-1.73
Employed vs. other	1.02	0.70-1.49	0.95	0.66-1.36
Body mass index				
BMI < 18.5	1.76	1.22-2.54	1.84	1.28-2.66
BMI >18.5 and BMI < 30	Reference			
BMI ≥ 30	0.41	0.10-1.70	1.22	0.58-2.57
Hemoglobin low vs. normal	1.61	1.07-2.42	1.60	1.10-2.32
WHO stage (vs. stage I or II)				
I or II	Reference			
III	2.00	1.29-3.10	1.23	0.83-1.83
IV	3.03	1.83-5.01	1.36	0.81-2.31
CD4 count				
CD4 ≤ 50	Reference			
50 < CD4 ≤ 100	0.47	0.29-0.76	0.46	0.29-0.73
100 < CD4 ≤ 200	0.31	0.18-0.51	0.39	0.25-0.60
CD4 > 200	0.12	0.03-0.48	0.43	0.20-0.90
Pulmonary tuberculosis	0.77	0.50-1.18	1.04	0.69-1.58
History of ART	Not estimable due to small sample size		1.21	0.43-3.37
Pregnancy	0.27	0.03-2.32	1.86	0.54-6.42
Baseline HAART regimen				
Contains efavirenz	Reference			
Contains nevirapine	0.41	0.13-1.28	0.76	0.37-1.58
Contains lopinavir-ritanovir	1.41	0.47-4.24	0.25	0.06-1.04
Peripheral neuropathy	1.21	0.63-2.31	0.92	0.45-1.90

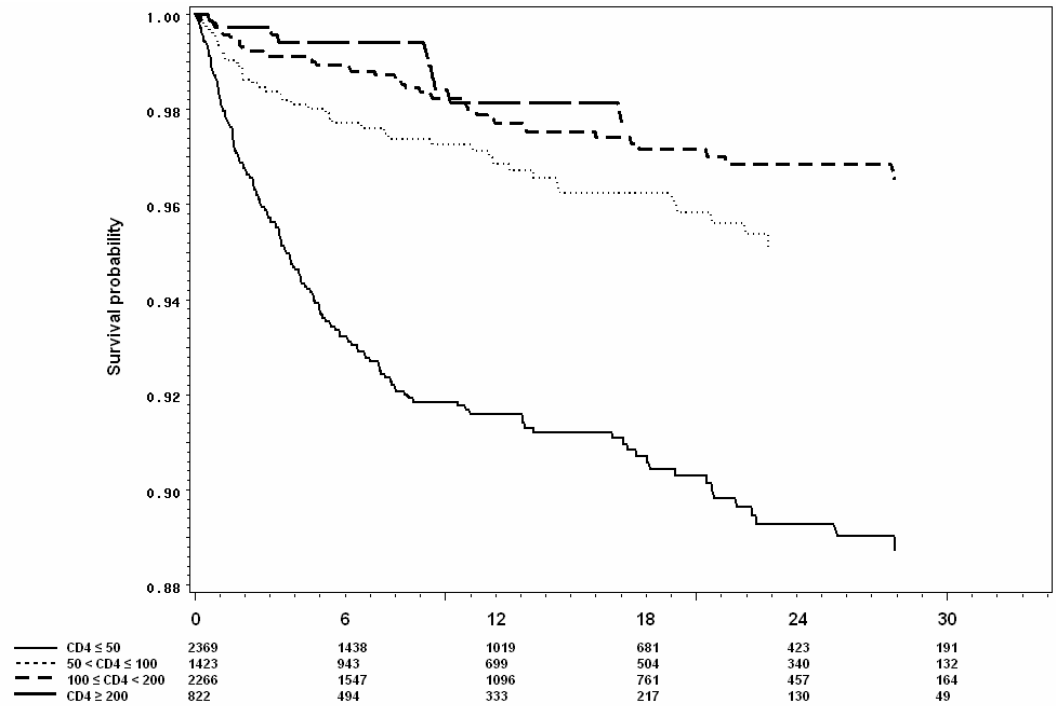
Tables 4.4a and 4.4b. Mean (SD) estimated CD4 count (4a) and gain (4b) in the Themba Lethu Clinical Cohort from two linear generalized estimating equations models.

Baseline CD4 count	N	Months post HAART initiation						
		0	4	6	12	18	24	36
Total population	6880	111 (1)	228 (2)	245 (2)	300 (3)	348 (4)	383 (5)	469 (33)
CD4 ≤ 50	2369	22 (1)	143 (3)	166 (3)	235 (4)	295 (6)	334 (9)	359 (43)
50 < CD4 ≤ 100	1423	76 (1)	205 (4)	221 (3)	267 (5)	304 (7)	335 (9)	482 (63)
100 < CD4 ≤ 200	2266	149 (1)	277 (3)	290 (3)	334 (4)	378 (6)	418 (8)	303 (48)
CD4 > 200	822	317 (6)	367 (8)	393 (8)	451 (10)	485 (12)	524 (16)	782 (116)

Baseline CD4 count	N	Months post HAART initiation						
		0	4	6	12	18	24	36
Total population	6880	N/a	116 (2)	134 (2)	188 (3)	237 (4)	272 (5)	358 (33)
CD4 ≤ 50	2369	N/a	121 (3)	144 (2)	213 (4)	273 (6)	312 (9)	336 (43)
50 < CD4 ≤ 100	1423	N/a	129 (4)	145 (3)	191 (5)	228 (7)	259 (9)	406 (62)
100 < CD4 ≤ 200	2266	N/a	128 (3)	141 (3)	185 (4)	229 (6)	270 (8)	154 (48)
CD4 > 200	822	N/a	49 (8)	76 (7)	133 (9)	167 (12)	205 (16)	464 (115)

Model for total population values (CD4 model 1) included changepoints at 4 months and every 6 months thereafter. Model for by-strata estimates (CD4 model 2) includes all variables in CD4 model 1 and interaction terms for each CD4 category and continuous time, and each CD4 category and each changepoint term (total 34 terms in model).

Figures 4.1a and 4.1b. Kaplan-Meier plots of survival (4.1a) and combined outcome "alive and in care" (4.1b), by strata of CD4 count. For readability, graphs have different X and Y axes.



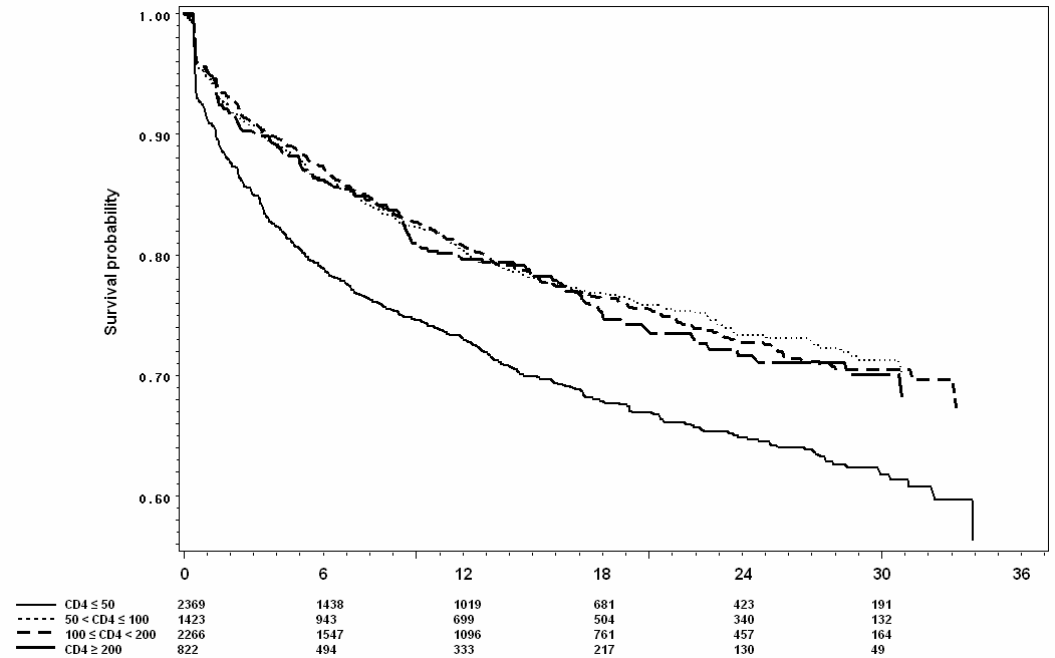


Figure 4.2. Twelve month risk of death by CD4 count category and WHO stage at baseline.

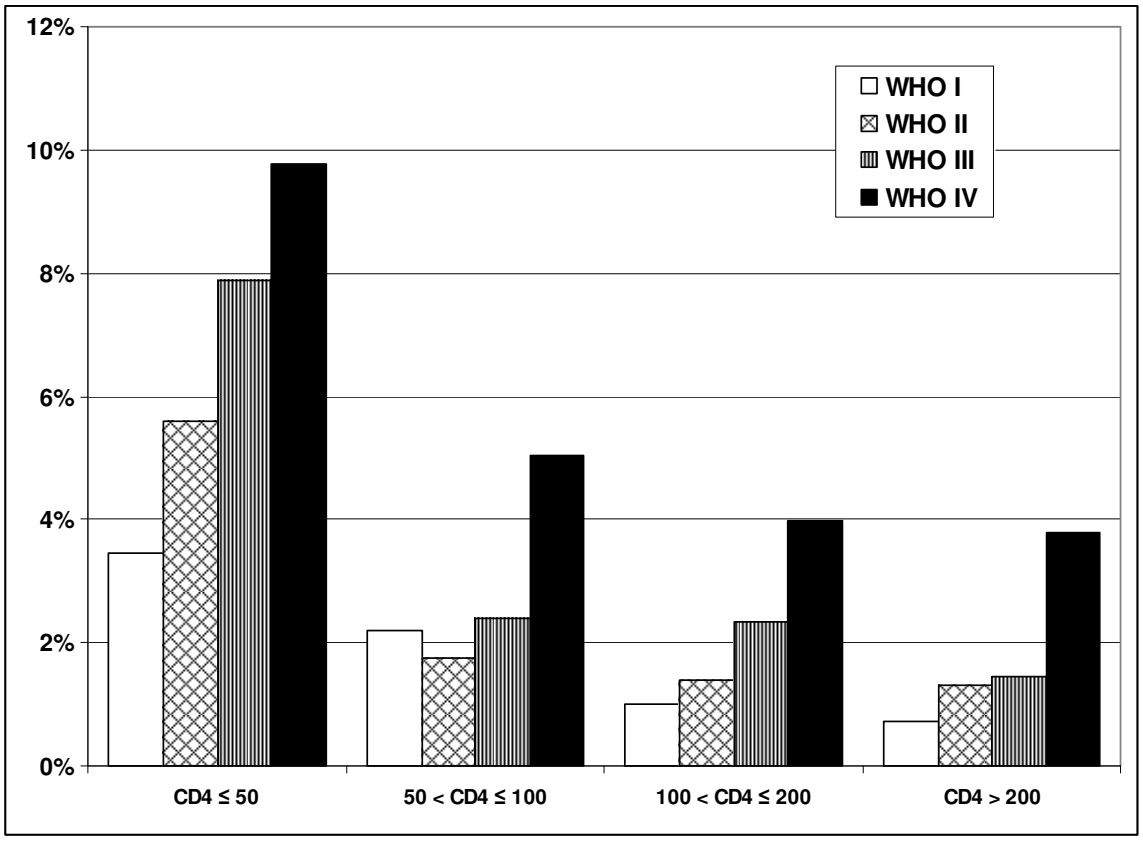
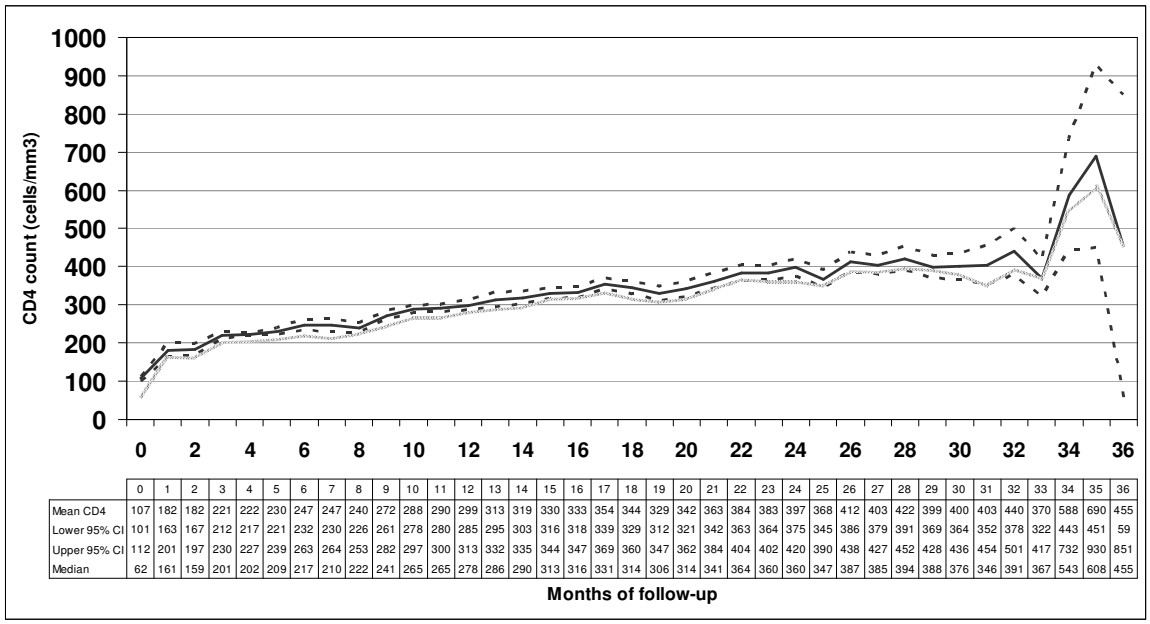


Figure 4.3. Predicted mean CD4 gain from baseline with 5% and 95% confidence bands for all individuals. Values are conditional upon being alive and in care.



	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36
Mean CD4	107	182	182	221	222	230	247	247	240	272	288	290	299	313	319	330	333	354	344	329	342	363	384	383	397	368	412	403	422	399	400	403	440	370	588	690	455
Lower 95% CI	101	163	167	212	217	221	232	230	226	261	278	280	285	295	303	316	318	339	329	312	321	342	363	364	375	345	386	379	391	369	364	352	378	322	443	451	59
Upper 95% CI	112	201	197	230	227	239	263	264	253	282	297	300	313	332	335	344	347	369	360	347	362	384	404	402	420	390	438	427	452	428	436	454	501	417	732	930	851
Median	62	161	159	201	202	209	217	210	222	241	265	265	278	286	290	313	316	331	314	306	314	341	364	360	360	347	387	385	394	388	376	346	391	367	543	608	455

CHAPTER 5

EFFECT OF PULMONARY TUBERCULOSIS ON MORTALITY IN PATIENTS RECEIVING HIGHLY ACTIVE ANTIRETROVIRAL THERAPY

INTRODUCTION

Pulmonary tuberculosis (PTB) is a leading cause of death among HIV-infected individuals in sub-Saharan Africa and throughout the world.(WHO 2004) PTB is of particular concern in South Africa, where an estimated 44% of new adult PTB cases arise among HIV-positive individuals.(WHO 2008) Prior to the introduction of highly active antiretroviral therapy (HAART), PTB in HIV-positive individuals was identified as a leading cause of death even among those receiving PTB treatment.(Badri et al. 2001; Mukadi, Maher, and Harries 2001; Nunn et al. 1992) Access to HAART for individuals with HIV-PTB co-infection reduces both mortality(Dean et al. 2002; Dheda et al. 2004; Manosuthi et al. 2006; Meya and McAdam 2007) and incidence of PTB.(Badri, Wilson, and Wood 2002; Meya and McAdam 2007)

Whether there are differences in prognosis between patients with PTB initiating HAART and patients free of active PTB when initiating HAART remains unresolved. Several studies have identified PTB as a leading cause of death among people receiving HAART.(Etard et al. 2006; Lawn et al. 2005) A study in Uganda

found that among individuals receiving HAART, those who had prevalent or incident PTB died at 4.7 times the rate of those without PTB, although the authors acknowledge that the relationship was likely confounded by body mass index (BMI) and CD4 cell count.(Moore et al. 2007) In contrast, a large study in Zambia reported no difference in mortality rates by PTB status(Stringer et al. 2006), but these results may have been affected by the substantial numbers of patients lost to follow-up (LTFU).

Controversy also remains over the effect of time between initiation of PTB treatment and initiation of HAART on the risk of death. The South African National Comprehensive Care, Management and Treatment Guidelines for HIV and AIDS recommend waiting two weeks to two months to initiate HAART, (Republic of South Africa Department of Health 2004; WHO 2007) because of concerns of drug-drug interactions, pill burden, adherence, and immune reconstitution inflammatory syndrome (IRIS) (McIlleron et al. 2007) However, the WHO suggests separately that while it may be safer to defer the initiation of HAART, concurrent initiation of HAART and PTB treatment may be necessary in patients at high risk of death.(WHO 2004)

In this study of the Themba Lethu Clinical Cohort, a large (>7500 patients) Department of Health HAART treatment site in Johannesburg, South Africa, we used state of the art methods (inverse probability of exposure and censoring weighted marginal structural models) to control for both confounding and LTFU to comprehensively assess the impact of PTB and the treatment for PTB at the time of HAART initiation on subsequent all-cause mortality, both overall and by length of

delay between initiation of PTB treatment and initiation of HAART. In addition, we assessed the impact of the combined presence of PTB and other risk factors for mortality on survival during HAART.

METHODS

Study site, clinic procedures, data collection. The Themba Lethu Clinical Cohort is a prospective clinical cohort of adults initiating HAART at the Helen Joseph Hospital in Johannesburg, South Africa. The program is funded by the South African Government, Gauteng Province, PEPFAR, and USAID. Themba Lethu has initiated over 10,000 patients on HAART since April 2004 and is the largest single clinic providing HAART in South Africa, and one of the largest antiretroviral clinics in the world. HAART-eligible patients attend educational and adherence sessions, and are assessed by a physician prior to initiating treatment. After HAART initiation, patients are scheduled for monthly pharmacy visits and clinical visits at month 4 and every 6 months thereafter and whenever needed clinically. Clinical and laboratory (including CD4 count, viral load and hemoglobin) data are collected at all scheduled visits, except for viral load which is not collected at baseline. All data were stored in the TherapyEdge™ (TherapyEdge Inc., USA) database, and analyzed in SAS v9.1.3 (SAS Institute, NC, USA).

South African national policy for the treatment of PTB follows the DOTS strategy. Patients with PTB typically receive treatment at primary health care clinics (outside of Themba Lethu); data on treatment adherence was therefore not available

for the majority of patients. Date of initiation of PTB treatment is by patient self-report.

Study eligibility and definitions. Patients were eligible for this study if they initiated HAART at Themba Lethu Clinic between April 1 2004 and March 31 2007. The main exposure was prevalent, treated PTB (hereafter, PTB) at the time of HAART initiation; the main outcome for analysis was all-cause mortality. Individuals more than three months late for a scheduled clinic appointment and who could not be located by the follow-up team after three phone calls and a visit to the last known address were classified as LTFU. These patients were censored midway between the last recorded visit and the next scheduled or at the last recorded date of receipt of therapy.

Statistical analyses. Descriptive analyses were performed using chi-square tests to compare categorical variables, t-tests to compare means, and Wilcoxon rank sum test to compare medians. We assessed the proportional hazards assumption for PTB in univariate and multivariate analysis through visual inspection of survival and log-log-survival curves, as well as Cox tests of continuous and categorical time interactions.

Analysis 1 examined the effect of PTB at time of initiation of HAART on all-cause mortality during follow-up. Analysis 2 examined the effect of time lag between initiation of PTB treatment and initiation of HAART on subsequent mortality, where time lag was categorized by approximate quartile of duration of PTB treatment prior to initiation of HAART: >120 days, 61-120 days, 31-60 days, and ≤30 days.

In both analyses we controlled for confounding and LTFU using inverse probability of exposure and censoring weights.(Cole et al. 2007; Hernán, Brumback, and Robins 2000; Robins, Hernán, and Brumback 2000) Analyses were implemented as discrete time hazard models with one observation per person-month and time-updated covariates. Models were evaluated using log-binomial generalized estimating equations, controlling for within-individual correlation using an "independent" correlation matrix for robust standard errors in the presence of repeated observations. Log-binomial regression was preferred to logistic regression because the former directly estimates the hazard ratio (HR) rather than the hazard odds ratio.(Muthén and Masyn 2005)

In both analyses, stabilized weights were estimated separately for exposure and censoring using ordinary (Analysis 1) or polytomous logistic regression (Analysis 2). We considered the following factors as predictors for both the weight and censoring models: gender, ethnicity, employment status, age, history of antiretroviral therapy, history of PTB, baseline measures of pregnancy, peripheral neuropathy, hemoglobin (adjusted for sex, pregnancy status, and altitude), BMI, CD4 cell count, and WHO stage (IV or other), year of enrollment (measuring years starting at baseline), and whether treatment was initiated before or after October 2006 (when consultation fees of ranging between $2-$5 per consultation were eliminated by the Department of Health). In addition, the time-varying censoring model accounted for time-updated measures of hemoglobin, CD4 count, and BMI, as well as for month of follow-up with both a linear month and cubic spline terms(Harrell 2004) for month of follow up with knots at the approximate 5th, 25th,

50th, 75th and 95th percentiles of survival time.(Fairall et al. 2008; Hernán, Brumback, and Robins 2000).

We did *not* control for HAART adherence, successful virologic suppression, or initial HAART regimen (NVP vs. EFV) because these factors either occur during follow-up or may be affected by the exposure and therefore would be part of the causal mechanism of the exposure on the outcome. Full descriptions of the components and validation of the stabilized weight models for both exposure and censoring are given in the Appendix.

Sensitivity and secondary analyses. We performed seven sensitivity analyses. To limit the potential effects of individuals dying between initiation of PTB treatment and initiation of HAART (i.e., left-truncation), we restricted PTB to cases who initiated HAART within 30 days of start of PTB treatment (S1) and separately to cases who initiated PTB treatment 30 days before or after initiation of HAART (S2). To assess the effect of PTB on early mortality, we restricted follow-up to the first four months of HAART (S3). To determine the effect of the exposure among the exposed, we calculated the weighted standardized mortality ratio for PTB (S4).(Sato and Matsuyama 2003) We used a traditional proportional hazards regression approach (with adjustment rather than reweighting) (S5). To ascertain the extent of bias due to LTFU, we examined the main effect estimate without censoring weights (S6). Last, to ensure that the effects of this study were not limited to PTB, we repeated analysis 1 using both pulmonary and extrapulmonary tuberculosis as the exposure category (S7).

Last, we performed a secondary analysis to examine effect measure modification of PTB by key baseline risk factors for mortality in people living with HIV, including severe immunosuppression (CD4≤50 cells/mm^3 or WHO stage IV disease), malnutrition (underweight, BMI<18.5 kg/m^2), and anemia.

Ethics approval. This study was approved by the human subjects review boards of both the University of the Witwatersrand and the University of North Carolina at Chapel Hill.

Role of the funding sources. The funding sources had no involvement in the design of the study, in the collection, analysis, or interpretation of the data, in the writing of this report, or in the decision to submit this report for publication.

RESULTS

There were 7,512 individuals included in this study, who contributed a total of 106,795 person-months to this analysis. Of these patients, 1,197 (15.9%) were being treated for PTB at the time of HAART initiation. Individuals who had PTB at baseline were more likely to be male, unemployed, and underweight, and had more severe immunosuppression (lower CD4 counts, more WHO stage IV) than those without PTB at baseline (Table 5.1a). The risk of death during follow-up was higher in individuals receiving PTB treatment at time of HAART initiation than in individuals not receiving PTB treatment at HAART initiation (6.2% vs. 3.6%, chi-square $p<0.0001$). Similar proportions of LTFU were observed in the two populations (19.5% and 18.8% respectively, $p=0.6150$).

Of the 1,197 individuals with PTB at baseline, 254 (21.2%) had been receiving PTB treatment for more than 120 days, 320 (26.7%) for 61-120 days, 289 (24.1%) for 31 to 60 days, and 334 (27.9%) for 30 days or less. Key differences in baseline characteristics by amount of lead-time are summarized in Table 5.1b. Patients with shorter lead time had more advanced disease, including lower CD4 cell counts, hemoglobin, and BMI.

As there was no evidence of violation of the proportional hazards assumption (Figure 5.1a), we summarized the effect of PTB at time of HAART initiation on mortality as a single estimate. In the overall crude (unweighted and unadjusted) analysis (Table 5.2), individuals with PTB at baseline were more likely to die during follow-up than those without PTB (crude HR for death 1.71; 95% CI 1.31-2.23). The crude HRs by lead time ranged from 1.34 for those with 60-120 days of lead-time, to 2.02 for those with <30 days of lead-time, but all 95% CIs were relatively wide and there were no discernible trends by lead-time. Crude survival curves comparing individuals with PTB to those without PTB are shown in Figure 5.1a. HRs in Figure 5.1 differ slightly from those presented in Table 5.2 because the figure required the use of weighted Cox proportional hazards models.(Cole and Hernán 2004)

In models which adjusted for confounding and LTFU using inverse probability of exposure and censoring weights, the HR comparing risk of death in those with PTB to risk of death in those without PTB at baseline was 1.06 (95% CI 0.75-1.49). The results were again similar when stratified by lead-time category, with adjusted HRs ranging from 0.89 to 1.27 and confidence intervals including the null (Table 5.2). Semi-parametric survival curves adjusted for confounding (but not censoring)

with inverse probability of exposure weights(Cole and Hernán 2004) are shown in Figure 5.1b.

Results of sensitivity and secondary analyses. Analyses S1-S7 yielded estimates of effect similar to the results of analysis 1, and none yielded an adjusted estimate of effect different from the null (Figure 5.2).

In the secondary analysis in which interaction terms were included in the model, we found several subpopulations in which PTB appeared to increase risk of death. First, this analysis showed that the HR for PTB on mortality among individuals who had CD4>50, BMI≥18.5, normal hemoglobin, and were not in WHO stage IV was 1.70 (95% CI 0.83-3.49), suggesting a raised hazard due to PTB among these otherwise relatively "healthy" individuals. Second, the estimates of the HR for PTB on mortality for individuals who had both PTB and a BMI<18.5, and individuals who had both PTB and CD4≤50 differed when interaction terms were considered. The HR for PTB and BMI<18.5 increased from 2.06 (95% 1.37-3.11) to 4.43 (95% CI 1.91-10.27) when interactions were added to the model; likewise, the HR for PTB and CD4≤50 increased from 2.91 (95% CI 1.95-4.37) to 4.01 (95% CI: 2.09-7.72). These results suggest a synergistic effect of PTB and other risk factors for death.

DISCUSSION

In a large urban clinic providing HAART in South Africa, we have shown that treatment for PTB at the time of HAART initiation does not put patients at increased

risk of death during follow-up. Additionally, we have shown that among those eligible for HAART, the period of time between initiation of PTB treatment and initiation of HAART does not change the relationship between prevalent PTB and risk of death on HAART. Our approach to the problem of LTFU makes it highly unlikely that these results can be explained by differential, informative LTFU among PTB patients. Instead, the increased risk seen in the unadjusted models is due to the fact that PTB patients are more likely than others to have severe immunosuppression (low CD4 counts, WHO stage IV illness), low BMI, and anemia.

In addition, we have shown that the overall lack of effect of PTB on mortality persists when we restrict our model to those who started PTB within only 30 days of HAART initiation, and also when we restrict follow-up to the first four months of HAART, a period in which the majority of deaths occur in our clinic. Taken with the main analysis, these results suggest strongly that starting HAART soon after initiation of PTB treatment will not put patients at higher risk of death.

While it is tempting to conclude from these results that length of time between initiation of PTB treatment and initiation of HAART does not affect risk of death, our findings must be interpreted cautiously. Patients in South Africa most often initiate PTB treatment outside of HAART clinics, and may die before starting HAART (in analytic terms, they would be left-truncated). The mortality attributable to PTB could therefore have been underestimated. Consequently, our results must be interpreted conditionally - *given* that an individual who initiated PTB treatment survived until initiating HAART, there was no effect of current PTB treatment on mortality.

However, the size of this potential left-truncation bias is likely to be small. Results from a recent randomized controlled trial of time between treatment for opportunistic infections (chiefly *Pneumocystis carinii* pneumonia, Cryptococcal meningitis, and pneumonia; the trial excluded all tuberculosis) showed that deferring HAART among such co-infected individuals can significantly increase risk of death or progression to AIDS (Zolopa et al. 2008). As is suggested by universally low CD4 counts and high prevalence of WHO stage III and IV disease in this cohort, many patients in our cohort initiate HAART with an acute opportunistic infection. Unless left truncation due to PTB acted differently than left truncation due to other opportunistic infections, then, it is unlikely that left truncation due to PTB would substantially bias these results. Additionally, if left-truncation created substantial bias, we would expect individuals with shorter lead-time to be at higher overall risk of death from PTB. We saw no trends in hazard of death with changing lead-time, nor did we observe a significant effect of PTB on mortality among the patients with the shortest lead-time (S1), again suggesting that the effect of left-truncation bias in this study is likely to be modest.

Another limitation of the study resulted from the lack of culture- or smear-confirmation of PTB in most cases, reflecting a high rates both of diagnoses made on clinical suspicion and of patient referral with previous PTB diagnosis. As a result, there may have been misclassification of the PTB status, which would likely lead to a bias towards the null. However, even if smear microscopy were available, the problem of misclassification would remain due to the low sensitivity of smear microscopy among HIV-positive individuals.(Tsiouris et al. 2006) PTB drug-

resistance status was also not available to this study; the inclusion of cases of drug-resistant PTB among individuals being treated for PTB would likely raise the effect of PTB on death and bias the results away from the null.

One additional limitation of this work is that we did not consider diagnoses of extrapulmonary TB in the main analysis. Although a finding of extrapulmonary TB is sufficient for a diagnosis of WHO stage IV disease(WHO 2005), there is a wide range in the severity of extrapulmonary TB from lymph node TB to TB meningitis. As such, a summary estimate for the effect of extrapulmonary TB on mortality would likely be misleading. Nonetheless, the sensitivity analysis which included both pulmonary and extrapulmonary TB (S7) showed no independent effect on risk of death in the population as a whole. In addition, we did not control for adherence to either HAART or treatment for PTB in these analyses.

Key risk factors for death after HAART initiation include advanced disease with severe immunosuppression, poor nutritional status and anemia(Ferradini et al. 2006; Lawn et al. 2005; Stringer et al. 2006; Wools-Kaloustian et al. 2006), conditions which are more prevalent among PTB patients than non-PTB patients in our cohort. The results of the secondary (hypothesis-generating) analysis suggested that patients with both PTB and one of these key risk factors (low BMI, low CD4 count), may be at increased risk of death in the setting of current treatment for PTB and HAART therapy. Future investigators should rigorously test these hypotheses.

Although a substantial number of people in this cohort became LTFU, rates of LTFU did not differ between those with or without PTB and did not appear to be

informative with regard to the PTB-mortality relationship; the effect of PTB on mortality was very similar with or without control for LTFU (see Figure 5.1b and analysis S6). This should increase our confidence in other reports about the PTB-mortality relationship, e.g., Stringer et al.(Stringer et al. 2006), which suffered from high rates of LTFU but did not model the censoring process explicitly.

In conclusion, the robust analysis of this large South African cohort strongly indicates that the effect of treatment for PTB on mortality following initiation of HAART may be of limited or no clinical importance and that initiation of HAART soon after start of PTB treatment does not increase risk of death. These results complement several other pieces of emerging evidence, including the randomized trial results (mentioned above) arguing for a earlier initiation of HAART in patients with non-TB opportunistic infections(Zolopa et al. 2008), reports suggesting that PTB-associated IRIS is not as dangerous as previously thought(Lawn et al. 2008; McIlleron et al. 2007; Murdoch et al. 2008), and evidence that virologic response can be achieved in the presence of co-administration of rifampicin with efavirenz.(Manosuthi et al. 2008; Manosuthi et al. 2005) While only the results of an ongoing randomized controlled trial (ACTG 5221) will provide a definitive answer to the important question of optimal time of initiation of HAART in PTB patients, results of this study, taken in context with these other recent findings, argue that the presence of PTB treatment should not delay initiation of HAART.

Table 5.1a. Characteristics of 7,512 individuals initiating HAART in Themba Lethu Clinical Cohort from 1 April 2004 to 31 March 2007 by pulmonary tuberculosis (PTB) treatment status at baseline.

	Prevalent PTB (n = 1197)	**No PTB** (n = 6315)	**p-value**
Female gender	695 (58.1)	4300 (68.1)	<0.0001
Age (Median, IQR)	35 (30-40)	35 (31-42)	0.0134 ‡
African ethnicity	1144 (95.6)	6022 (95.4)	0.7483
Employment			
Employed, student, retired	358 (29.9)	2372 (37.6)	<0.0001
Not employed or unknown	760 (63.5)	3460 (54.8)	
Missing	79 (6.6)	483 (7.7)	
BMI			
<18.5	393 (32.8)	1041(16.5)	<0.0001
18.5-24.9	663 (55.4)	3473 (55.0)	
25.0-29.9	110 (9.2)	1229 (19.5)	
≥30	31 (2.6)	572 (9.1)	
Hemoglobin			
Normal	330 (27.6)	3353 (53.1)	<0.0001
Low	864 (72.2)	2881 (45.6)	
Missing	3 (0.3)	81 (1.3)	
CD4 count (cells/mm^3)			
Median (IQR)	58 (22-116)	94 (34-165)	<0.0001 ‡
Mean (95% CI)	78 (74-83)	115 (112-118)	<0.0001 †
≤50	501 (41.9)	1865 (29.5)	<0.0001
51-100	271 (22.6)	1152 (18.2)	
101-200	284 (23.7)	1981 (31.4)	
201-350	44 (3.7)	581 (9.2)	
>350	12 (1.0)	183 (2.9)	
Missing	85 (7.1)	553 (8.8)	
WHO stage IV at baseline	207 (17.3)	548 (8.7)	<0.0001
History of ART (any)	11 (0.9)	211 (3.3)	<0.0001
Pregnant	20 (1.7)	552 (8.7)	<0.0001
Peripheral neuropathy	89 (7.4)	231 (3.7)	<0.0001
History of PTB	60 (5.0)	796 (12.6)	<0.0001

Figures are expressed as number (% total), except where noted. P-values are 2-sided by chi-square test, † t-test or ‡ Wilcoxon rank sum test. After adjustment for altitude, lower limit of normal hemoglobin is 12.35 g/dl for men; 11.35 g/dl for non-pregnant women; and 10.35 g/dl for pregnant women.

Table 5.1b. Selected characteristics of 1,197 individuals receiving treatment for pulmonary tuberculosis (PTB) at time of initiation of HAART by length of time between initiation of PTB treatment and initiation of HAART.

Characteristic at initiation of HAART	Time between initiation of PTB treatment and HAART				p-value
	0-30 days (n=334)	31-60 days (n=289)	61-120 days (n=320)	> 120 days (n=254)	
BMI < 18.5	108 (32.3)	117 (40.5)	103 (32.2)	65 (25.6)	0.0032
Low hemoglobin	259 (77.5)	234 (81.0)	232 (72.5)	139 (54.7)	<0.0001
CD4 count					
Median (IQR)	43 (15-88)	52 (21-104)	82 (36-133)	68 (27-138)	<0.0001†
Mean (95% CI)	60 (53-67)	69 (62-77)	95 (86-105)	92 (80-104)	<0.0001†
CD4<50	176 (52.7)	134 (46.4)	101 (31.6)	90 (35.4)	<0.0001
WHO stage IV	63 (18.9)	49 (17.0)	60 (18.8)	35 (13.8)	0.3525
Peripheral neuropathy	20 (6.0)	20 (6.9)	30 (9.4)	19 (7.5)	0.4108
History of PTB	18 (5.4)	18 (6.2)	15 (4.7)	9 (3,5)	0.5281

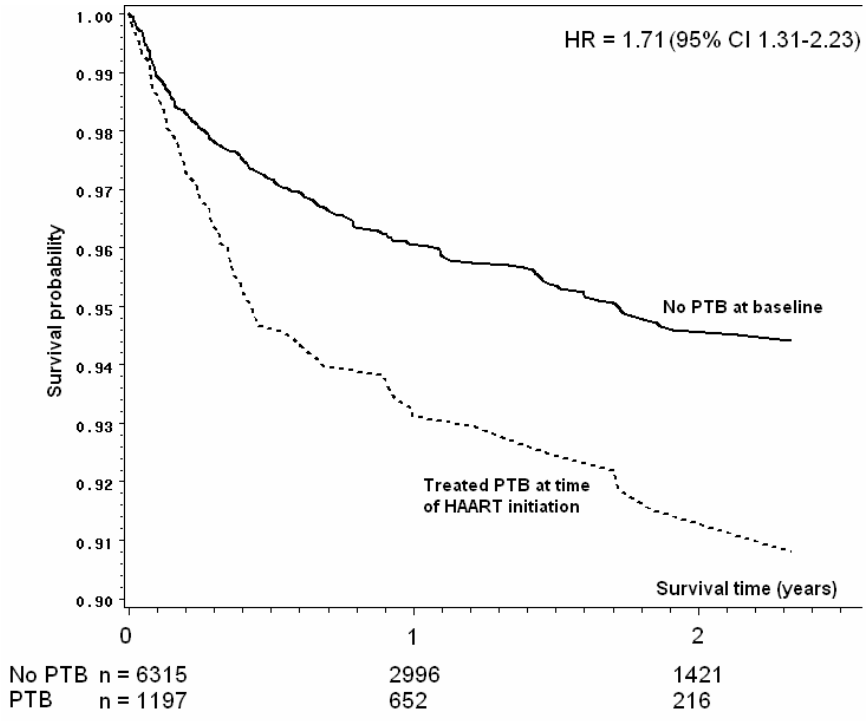
Figures are n(%) unless otherwise noted. P-values are 2-sided by chi-square test or † one-way ANOVA.

Table 5.2. Estimates of hazard ratios from marginal structural models of the main analyses of the effect on all-cause mortality of current treatment for pulmonary tuberculosis (PTB) at time of HAART initiation.

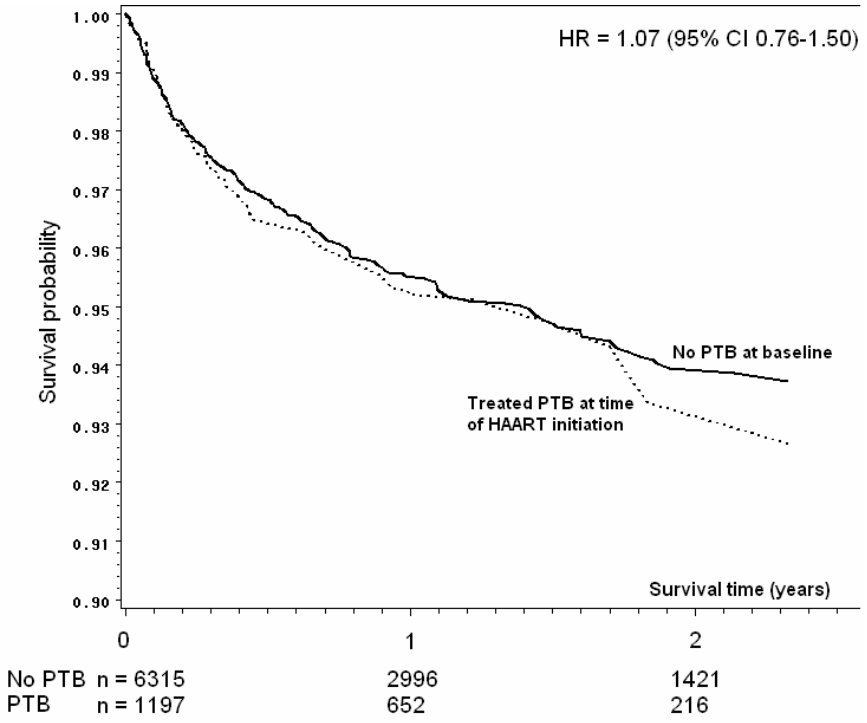
	Main exposure	Crude (unweighted) estimates		Adjusted for confounding and censoring	
		HR	95% CI	HR	95% CI
Model 1	All PTB	1.71	1.31-2.23	1.06	0.75-1.49
Model 2	> 120 days	1.89	1.17-3.06	1.26	0.65-2.43
	61-120 days	1.34	0.80-2.23	0.89	0.48-1.63
	31-60 days	1.62	0.97-2.72	1.07	0.58-1.96
	≤ 30 days	2.02	1.31-3.11	1.27	0.78-2.09

Model 1 estimates the effect of prevalent, treated PTB on outcome of mortality. Model 2 estimates the effect of the same exposure by length delay between initiation of PTB treatment and initiation of HAART. Both models adjust for both confounding and LTFU using inverse probability of exposure and censoring weights; components and specification of weight models are described in the Appendix.

Figure 5.1. Survival curves by status of current treatment for pulmonary tuberculosis (PTB) at baseline, unadjusted (5.1a), and reweighted for confounding only (5.1b).

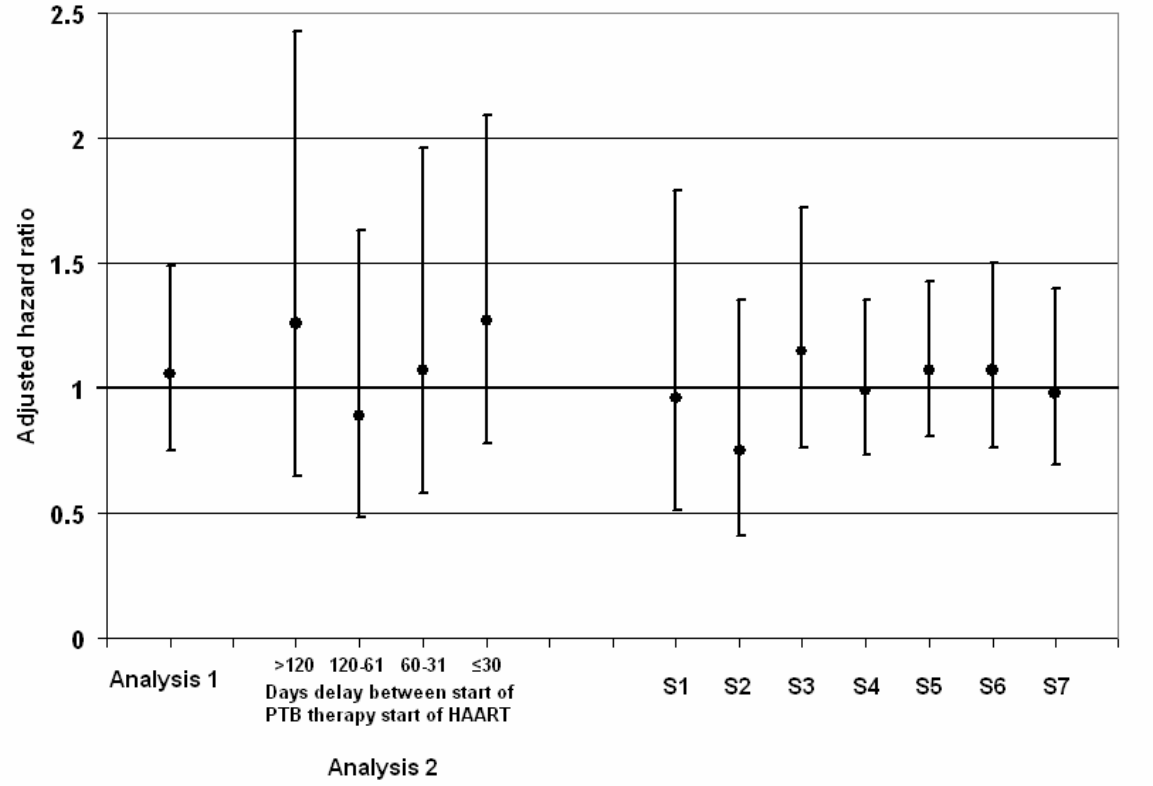


1a



1b

Figure 5.2. Estimates of adjusted hazard ratios from main analyses 1 and 2 and sensitivity analysis S1-S7.



S1: Pulmonary tuberculosis (PTB) treatment started within 30 days before HAART. S2: PTB treatment started within 30 days before or after HAART. S3: restricted follow-up to first four months of HAART. S4: standardized mortality ratio (effect of the exposure among the exposed). S5: Traditional Cox proportional hazards model. S6: Exposure weights only; no censoring weights. S7: Pulmonary and extrapulmonary TB together.

CHAPTER 6

TUBERCULOSIS TREATMENT AND RISK OF STAVUDINE SUBSTITUTION IN FIRST LINE ANTIRETROVIRAL THERAPY

INTRODUCTION

The rollout of highly active antiretroviral therapy (HAART) to individuals living with human immunodeficiency virus (HIV) has expanded rapidly in the last five years. In sub-Saharan Africa alone, rapid scale-up of HAART expanded access from 100,000 in 2003 to 1.3 million people or about 28% of those in need by the end of 2006 (WHO 2007).

An important challenge faced by rollout programs is the management of adverse drug reactions in the face of limited resources (both people and time) and limited choice of antiretroviral drugs (Bolhaar and Karstaedt 2007; Boulle et al. 2007; Ferradini et al. 2006; Stringer et al. 2006). In particular, the inclusion of the nucleoside reverse transcriptase inhibitor stavudine as a component of standard first-line HAART throughout sub-Saharan Africa remains controversial because of relatively high rates of adverse reactions. While inexpensive and highly potent, stavudine causes significant toxicity and has been implicated as a chief cause of symptomatic hyperlactatemia, lactic acidosis, lipid disorders, lipodystrophy, and peripheral neuropathy in various populations of HIV-positive individuals (Boulle et al.

2007; Colebunders et al. 2005; Currier 2002; Currier 2007; Currier and Havlir 2005; Subbaraman et al. 2007). In one study from South Africa, over 20% of patients required substitution of stavudine because of toxicity by 36 months, most often for reasons of lipodystrophy or peripheral neuropathy (Boulle et al. 2007).

These toxicities may have significant implications for the success of HAART. They cause substantial morbidity (Boulle et al. 2007; Falco et al. 2002; Subbaraman et al. 2007), can impact adherence to HAART (Bolhaar and Karstaedt 2007; Boulle et al. 2007), and may lead to treatment interruptions and resulting in increased risk of death (Bolhaar and Karstaedt 2007) and virological failure (Parienti et al. 2004). While the risk factors for hyperlactatemia, the most clinically severe stavudine toxicity, have been well-characterized, the risk factors for all-cause stavudine substitution itself remain largely unexamined (Boulle et al. 2007; Forna et al. 2007; Hawkins et al. 2007).

Despite the high prevalence of tuberculosis (TB) among individuals receiving HAART in sub-Saharan Africa and the widely recognized importance of TB in the HIV epidemic (WHO 2004), the impact of both HAART and TB treatment on the incidence of stavudine substitution remains almost entirely unexplored. A search in PubMed for “tuberculosis and stavudine” performed in May 2008 yielded only 17 citations, of which only two addressed issues of toxicity (PubMed). Two small studies from England (Breen, Lipman, and Johnson 2000; Dean et al. 2002) observed higher than expected rates of adverse events with stavudine and TB treatment. No studies have comprehensively addressed these issues in the developing world. This gap in the literature is particularly concerning because

isoniazid, which is given for six months as part of standard TB treatment (Republic of South Africa Department of Health 2004; Subbaraman et al. 2007), has been implicated in peripheral neuropathy (Moyle and Sadler 1998; Subbaraman et al. 2007), one of the most commonly reported stavudine toxicities (Amoroso et al. 2007; Boulle et al. 2007; Hawkins et al. 2007; Moyle and Sadler 1998).

Our objective was to estimate the effect of TB treatment on incidence of all-cause stavudine substitution in a large HAART rollout clinic in Johannesburg, South Africa.

METHODS

Study site, clinic procedures, data collection. This study was conducted on the Themba Lethu Clinical Cohort, a prospective cohort of adults initiating HAART. The clinic is located in Helen Joseph Hospital, Johannesburg, South Africa, and funded by the South African Government, Gauteng Province SA, PEPFAR, and USAID, and operated as a collaborative project with the Clinical HIV Research Unit of the University of the Witwatersrand and the non-profit Right To Care. Themba Lethu is one of the largest single clinics providing HAART in South Africa and has initiated patients on HAART at a rate of approximately 200 per month since April 2004. After initiating HAART, patients are scheduled for monthly pharmacy visits and clinical visits to Themba Lethu Clinic at 4 months and every 6 months thereafter. Clinical data as well as CD4 counts, hemoglobin, and other laboratory diagnostics are collected at all visits, except for viral loads which are not collected at baseline.

All data were stored in the TherapyEdge™ (TherapyEdge Inc., USA) database, and analyzed in SAS v9.1.3 (SAS Institute, NC, USA).

Therapy for TB is usually initiated outside of Themba Lethu Clinic. The standard course of TB treatment in South Africa is six months: two months of isoniazid, rifampicin, pyramizade, and ethambutol followed by four months of isoniazid and rifampicin (Republic of South Africa Department of Health 2004). The course of therapy is longer if the patient does not respond to treatment, has relapsed, or has drug resistant TB. Vitamin B6 (pyridoxine) is commonly prescribed at time of initiation of TB treatment. Amitriptyline is commonly prescribed for individuals experiencing peripheral neuropathy (Saarto and Wiffen 2007) in the Themba Lethu Clinic.

Study eligibility and definitions. Patients were eligible for this study if they initiated a stavudine-containing HAART regimen at Themba Lethu Clinic between April 1 2004 to March 31 2007 (Republic of South Africa Department of Health 2004).

The exposure of interest for the main analysis was TB treatment, defined as therapy for pulmonary or extrapulmonary TB which included isoniazid. There were three specific exposure categories: ongoing TB treatment at time of HAART initiation; recent initiation of TB treatment; and incident TB treatment during HAART. *Ongoing TB treatment* was defined as TB treatment started at least two weeks before initiation of HAART. A two-week cutoff was chosen based on the South African National ARV Guidelines recommendation to wait at least two weeks between initiation of TB treatment and HAART. *Recent initiation of TB treatment*

was defined as initiation of TB treatment 0 to 14 days before initiation of HAART. Last, *incident TB treatment* was defined as any TB treatment which began after initiation of HAART. A schematic diagram for these exposure categories is shown in Figure 6.1.

In analyses of ongoing and recent TB treatment, patients were censored at time of incident TB. In the analysis of incident TB treatment, individuals who had ongoing or recent TB treatment became eligible for incident TB one month after the recorded completion of their previous course of TB treatment, or nine months after the initiation of the previous course of therapy, if no end date was available. When these patients became eligible for the incident TB analysis, they were marked as having a history of TB treatment (just as patients with ongoing or recent TB treatment might have such a history). Individuals who experienced multiple episodes of incident TB were censored at the second episode.

The outcome of *all-cause d4T substitution* was defined as the event of substitution of a single drug (most often zidovudine [AZT]) for d4T while the rest of the regimen remained unchanged. Patients were still at risk during treatment interruptions and after drug substitutions for drugs other than d4T. Drug substitution events in which d4T and other drugs were substituted simultaneously (multi-drug substitutions) were considered competing events and cause for censoring in the main analysis, but were included as outcomes in a sensitivity analysis.

Statistical analyses. There were three main analyses, corresponding to the three exposures defined above. In all analyses, we controlled for confounding and censoring using inverse probability of treatment and censoring (IPTC) weights (Cole

et al. 2007; Hernán, Brumback, and Robins 2000; Robins, Hernán, and Brumback 2000). Analyses were implemented as discrete time hazard models, and evaluated using log-binomial generalized estimating equations models, which controlled for within-individual correlation using an "independent" correlation matrix for robust standard errors in the presence of repeated observations (Hernán, Brumback, and Robins 2000).

Stabilized weights for TB treatment were estimated using ordinary logistic regression, and histograms of multivariate treatment weight models were examined to ensure that the positivity assumption (see Appendix) was met (Hernán and Robins 2006). The estimation of stabilized censoring weights was complicated because of the issue of competing risks (Matsuyama and Yamaguchi 2007) of multi-drug substitution, death, and lost to follow-up, all of which may be informative with respect to the outcome of stavudine substitution. In a marginal structural model, the most intuitive solution to this problem is to extend IPTC weights to encompass multiple causes of censoring (Matsuyama and Yamaguchi 2007). Because initial analysis showed little change in estimates and confidence intervals when modeling outcomes of lost to follow-up and death separately, we estimated a single weight for both outcomes and a separate weight for the competing risk of multi-drug substitution (except in sensitivity analysis).

Possible predictors for both the weight and censoring models were gender, ethnicity, employment status, age, history of antiretroviral therapy, previous history of TB treatment; and baseline measures of pregnancy, peripheral neuropathy, hemoglobin (adjusted for sex, pregnancy, and altitude), BMI, CD4 cell count, and

WHO stage (IV or other), calendar date, and whether treatment was initiated after October 2006 (when all consultation fees were eliminated by the Department of Health). In addition, time-varying censoring models accounted for time-updated measures of hemoglobin, CD4 count, BMI, and initial virologic success (measured at month 6), as well as for month of follow-up as both a continuous and categorical variable. Predictors were removed from weight models if they were not found to confound the outcome in traditional proportional hazards regression using backwards elimination (with an alpha of 0.10 for removal). Some of these variables were subsequently added back to both the weight and stabilization models to increase the stability of the final IPTC weights.

Preliminary analysis indicated that the effect of TB treatment on hazard of stavudine substitution decreased markedly with time and thus violated the proportional hazards assumption. We therefore we included a categorical time-interaction term (Cole et al. 2007) to allow the hazard ratio (HR) to change at the beginning of the third month of concomitant TB treatment and HAART. We included an additional time-interaction term for months 3-6, as TB treatment lasts 6 months except in cases of drug resistance or retreatment (Republic of South Africa Department of Health 2004).

Secondary and sensitivity analyses. For each of the three main models, we performed several sensitivity analyses. We included multi-drug substitutions (excluding switches to second-line HAART) as outcomes (S1); we replaced the treatment portion of the IPTC weights with simple adjustment for confounding (S2); we truncated IPTC weights to the 99th and 1st percentile to ensure that no individual

observations contributed excessively to effect estimates (S3); we estimated the relative effect of the exposure among the exposed (the standardized morbidity ratio) for each group (Sato and Matsuyama 2003) (S4); and we restricted the referent group to those without recorded history of TB (S5) to ensure that lingering peripheral neuropathy from previous drug exposures did not bias the results.

We performed one additional sensitivity analysis (S6) for ongoing TB treatment only, stratifying results by length of TB treatment prior to initiation of HAART in two categories chosen based on treatment guidelines (Republic of South Africa Department of Health 2004): 2 weeks to 2 months after initiation of TB treatment (15-60 days), and after two months (the continuation phase of TB treatment, 61+ days).

RESULTS

A total of 7067 individuals contributed 88916 person-months of follow-up to this analysis, while 445 patients were excluded because they were not receiving one of three standard first-line HAART regimens. More than half of the excluded patients (n=253, 57%) were receiving AZT-3TC-EFV, which is clinically indicated for patients suffering peripheral neuropathy at time of HAART initiation. However, only 65 of these 253 patients were being treated for TB at baseline, and these patients were only slightly more likely to have TB than other patients (risk ratio [RR] 1.25 95% CI 1.01-1.54).

Ongoing therapy for TB at time of HAART initiation was present in 1272 individuals (18.0%), of which 964 (75.7%) had pulmonary TB. TB treatment was started recently with respect to HAART in 202 individuals (2.9%), of which 164 (81.2%) had pulmonary TB. Characteristics of those with and without TB at baseline are summarized in Table 6.1. Compared to those who were not receiving therapy for TB at initiation of HAART, individuals who were being treated for TB at baseline or who initiated TB recently with respect to HAART were less likely to have a history of TB, and were more likely to be male and have advanced disease (lower CD4 counts, BMIs, hemoglobin, and/or more WHO stage IV).

There were 404 first episodes of incident TB during follow-up, at a median of 99 (intraquartile range [IQR] 31-273) days of follow-up and a mean of 179 days. There were 15 second episodes of incident TB recorded. The rate of first episode of incident TB was 5.5 (95% CI 5.0-6.0) per 100 person-years of follow-up. Of these 404 first episodes, 371 occurred in patients still at risk of stavudine substitution.

There were 842 recorded single-stavudine substitutions. There were an additional 375 multi-drug substitutions which included stavudine; of these, 172 were drug regimen switches to second-line HAART (AZT-DDI-LPVr), and most of the remainder were switches to LPVr-EFV. Median time to single-stavudine substitution was 347 (IQR 175-535) days, and to multi-drug substitution was 357 (IQR 254-488) days. Crude rate of single-stavudine substitution was 12.4 (95% CI 11.6-13.3) per 100 person-years. Kaplan-Meier curves for single-stavudine substitutions stratified by TB treatment status at baseline are shown in Figure 6.2.

Of the 842 stavudine substitutions, 714 were attributed clinically to either peripheral neuropathy (n=362, 43% of 842), lipodystrophy (n=205, 24%), lactic acidosis or symptomatic hyperlactatemia (n=168, 20%), or two of these three factors (n=21, 2%). Those who switched for peripheral neuropathy were more likely to have TB either at baseline or during follow-up (RR=1.53, 95% CI 1.33-1.75); those who switched for lactic acidosis or lipodystrophy were less likely to have TB (RR=0.58, 95% CI 0.48-0.71).

A total of 260 individuals (3.7%) died during follow-up and 1252 (17.7%) became lost to follow-up.

Results from the main analysis are shown in Table 6.2. Compared to those who were not being treated for TB, patients with ongoing TB treatment had a crude hazard ratio (HR) for stavudine substitution of 2.78 (95% CI 1.61-4.79) for the first two months of follow-up, 2.24 (95% CI 1.61-3.12) for months 3-6, and 1.10 (95% CI 0.88-1.37) thereafter. The adjusted (weighted) HR estimates, corrected for confounding, loss to follow-up, death, and the competing risk of multi-drug substitution were similar but slightly closer to the null.

Those with recent TB treatment had crude estimates of HR of 5.05 (95% CI 2.14-11.93) for the first two months, 2.33 (95% CI 1.14-4.79) for months 3-6, and 0.95 (95% CI 0.56-1.63) thereafter. In the adjusted analysis the estimates were 6.54 (95% CI 2.58-16.57), 1.39 (95% CI 0.63-3.06), and 1.03 (95% CI 0.55-1.93).

For patients with incident TB treatment, the crude estimates of HR were 1.26 (95% CI 0.72-2.22), 1.15 (95% CI 0.64-2.08), and 1.66 (95% CI 0.67-4.07) for months 0-2, 3-6, and 7+ respectively, with similar estimates in the adjusted analysis.

Sensitivity analysis results. Results from the sensitivity analysis confirmed the main results both in magnitude and overall trend (effect in the first two months, but a smaller or null effect thereafter) (Appendix II, Figures A2.1- A2.3), and suggest that the effects of ongoing TB treatment do not depend strongly on length of TB treatment prior to initiation of HAART (Table 6.3).

DISCUSSION

Concomitant administration of TB treatment and HAART is common in sub-Saharan Africa, especially at time of initiation of HAART. In our cohort, the prevalence of TB treatment at time of HAART initiation was 20.9%. Our results show that the presence of TB treatment at the time of HAART initiation puts patients at higher risk of requiring a substitution for stavudine, especially if HAART is initiated soon (≤14 days) after the start of TB treatment. Initiation of TB treatment in patients who were already using stavudine did not appear to raise the risk of stavudine substitution.

The effect of TB on d4T substitution is strongest in the first two months of treatment when the two drug regimens are initiated close together in time, with a 2-month adjusted HR estimated at 6.54. After two months of HAART, the effect was smaller, probably due to depletion of susceptible patients. With recent initiation, the short term risk of stavudine substitution is very high, and thus patients at risk of stavudine substitution by TB co-medication are removed from the population sooner, explaining the large drop in effect in this population. After completion of TB

treatment (≥ 7 months), there was no longer an effect of TB treatment on the risk of stavudine substitution.

In light of the strong effect we observed of prevalent TB treatment (ongoing or recent) on risk of stavudine substitution, we were surprised to see little to no effect of incident TB treatment on the outcome. There are several possible explanations for this observation. First, because these patients have been taking stavudine prior to initiation of TB treatment, it is possible that these patients have grown accustomed to stavudine, and are thus at lower risk from TB treatment. Another possibility is that the effect of TB treatment on risk of stavudine substitution is mediated by HIV viral load; control of HIV replication may thereby reduce the risk of d4T in conjunction with TB treatment. Last, if patients in whom TB treatment would have most increased risk were also at high baseline risk of stavudine substitution, these patients might have experienced substitution before developing incident TB, and were thus not included in our study due to left-truncation; in this last case, depletion of susceptible patients could have caused our analysis of incident TB treatment to underestimate the true effect of that exposure on the outcome.

Analysis S5 provides additional evidence that these results may have underestimated the true impact of TB treatment on risk of stavudine substitution. Because peripheral neuropathy does not disappear the day TB treatment ends, it is possible that in the main analysis, individuals with a somewhat recent history of TB treatment may be at higher risk of stavudine substitution due to residual peripheral neuropathy. When we restrict the reference group to individuals without history of TB, the point estimate for the effect of TB treatment on risk of stavudine substitution

in the first two months of co-treatment increases slightly for all exposures (ongoing, recent, incident) (Appendix II, Tables A2.1-A2.3).

The largest number of single-stavudine toxicities were caused by peripheral neuropathy and individuals receiving TB treatment were more likely than others to switch for this reason. This is consistent with our original hypothesis that peripheral neuropathy is a key pathway for the interaction of TB drugs and HAART, because peripheral neuropathy is caused by both the isoniazid (Moyle and Sadler 1998; Subbaraman et al. 2007) and stavudine (Amoroso et al. 2007; Boulle et al. 2007; Breen, Lipman, and Johnson 2000; Dean et al. 2002; Hawkins et al. 2007; Moyle and Sadler 1998). This result is especially striking given that the great majority of TB patients in South Africa are prescribed vitamin B6 (pyridoxine) at time of initiation of TB treatment for the prevention of peripheral neuropathy; in addition, amitriptyline is frequently prescribed to manage incident peripheral neuropathy. It is likely that the effect of TB on the risk of stavudine substitution would be even higher if these two drugs were unavailable. Conversely rates of peripheral neuropathy may be further reduced with additional micronutrient supplementation (Villamor et al. 2008).

There were several limitations of this study. First, while focusing on stavudine substitution as an outcome ensures that we have captured the most severe toxicities, we did not examine the impact of TB treatment on risk of (for instance) peripheral neuropathy specifically. It seems likely that treatment for TB resulted in substantial low level peripheral neuropathy, which may remain undiagnosed, unreported, or resolves without stavudine substitution. Thus, the impact of TB

treatment on risk of peripheral neuropathy (or other toxicities) may be different (and in particular, higher) than the effect of TB treatment on stavudine substitution.

Second, we estimated the relative hazard rather than absolute risk differences. Because the absolute risk of stavudine substitution is relatively low early on in this study, even a large hazard ratio may translate into a somewhat small absolute risk. For instance, only 220 stavudine substitutions took place during the first six months following initiation of HAART; the absolute risk in those with any TB treatment at baseline (ongoing or recent) was 5.6%, and in those without was 2.4% (risk difference 3.2%, 95% CI 1.9-4.4%). Thus, in the first six months after HAART initiation, there is one stavudine substitution for every 18 patients with TB treatment, and one for every 42 patients without. However, because there are large number of individuals who initiate HAART while on TB treatment, we believe these results still have considerable relevance to public health.

Last, while we strove to control adequately for confounding in this study, it is likely that uncontrolled confounding remains; this is especially true for alcohol and other drugs which may be implicated in peripheral neuropathy. Similarly, while we attempted to account for several competing risks – including multi-drug substitution, death, and lost to follow-up – complete control of these competing risks requires an assumption that competing risks can be completely accounted for by observed variables, a somewhat heroic assumption analogous to the assumption of no uncontrolled confounding (Hernán, Brumback, and Robins 2000; Matsuyama and Yamaguchi 2007).

Despite these limitations, these results show that co-medication of TB at time of initiation of HAART is an important risk factor for stavudine substitution in the Themba Lethu Clinical Cohort. The issue of co-medication with TB drugs and with stavudine is a vital issue in the rollout of HAART in sub-Saharan Africa, and deserves more attention than it has so far received. While clearly the results of a single observational study should never serve as the sole basis for changes in policy, the magnitude and consistency of these results suggest strongly that we may wish to reconsider the use of stavudine in first-line HAART among patients with ongoing or recent initiation of TB treatment.

Table 6.1. Characteristics of 7,067 individuals initiating HAART in Themba Lethu Clinical Cohort from 1 April 2004 to 31 March 2007 by tuberculosis treatment status at baseline. Figures expressed as n (%) unless noted.

	No TB treatment (n=5593)	Ongoing TB treatment (n = 1272)	p-value	Recent TB treatment (n = 202)	p-value
History of TB	855 (15.3)	64 (5.0)	<0.0001	12 (5.9)	0.0003
Female gender	3861 (69.0)	749 (58.9)	<0.0001	114 (56.4)	0.0002
Age > 40	1656 (29.6)	347 (27.3)	0.0991	60 (29.7)	0.9769
Employed	2107 (37.7)	396 (31.1)	<0.0001	72 (35.6)	0.5587
BMI					
< 18.5	901 (16.1)	412 (32.4)	<0.0001	60 (29.7)	<0.0001
18.5 - 24.9	3085 (55.2)	707 (55.6)		111 (55.0)	
25.0 - 29.9	1100 (19.7)	124 (9.8)		24 (11.9)	
≥ 30.0	507 (9.1)	29 (2.3)		7 (3.5)	
Hemoglobin					
Normal	3014 (53.9)	356 (28.0)	<0.0001	51 (25.3)	<0.0001
Low	2508 (44.8)	913 (71.8)		151 (74.8)	
Missing	71 (1.2)	3 (0.3)		0 (0)	
CD4 count					
Median (IQR)	94 (35-164)	59 (23-116)	<0.0001‡	48 (13-106)	<0.0001‡
Mean (95% CI)	111 (108-113)	80 (75-84)	<0.0001†	70 (59-81)	<0.0001†
< 50	1647 (29.5)	525 (41.3)	<0.0001	102 (50.5)	<0.0001
50 - 99	1029 (18.4)	303 (23.8)		38 (18.8)	
100 - 199	1824 (32.6)	306 (24.1)		43 (21.3)	
200 - 349	506 (9.1)	46 (3.6)		7 (3.5)	
≥ 350	115 (2.1)	13 (1.0)		3 (1.5)	
Missing	472 (8.4)	79 (6.2)		9 (4.5)	
WHO stage IV at baseline	449 (8.0)	228 (17.9)	<0.0001	45 (22.3)	<0.0001
History of ART (any)	109 (2.0)	8 (0.6)	0.0010	1 (0.5)	0.1369
Pregnant	540 (9.7)	24 (1.9)	<0.0001	2 (1.0)	<0.0001
Peripheral neuropathy	151 (2.7)	81 (6.4)	<0.0001	9 (4.5)	0.1346

P-values are 2-sided by chi-square test compared to "No therapy", † t-test or ‡ Wilcoxon rank sum test. After adjustment for altitude, lower limit of normal hemoglobin is 12.35 g/dl for men, 11.35 g/dl for non-pregnant women, and 10.35 g/dl for pregnant women.

Table 6.2. Hazard ratios and 95% confidence intervals for effect of TB treatment on stavudine substitution by period of TB treatment and duration of co-medication for TB and HIV.

Exposure	N	Crude estimates by duration of co-medication (in months)		
		0-2	3-6	7+
Ongoing TB	**1272**	2.78 (1.61-4.79)	2.24 (1.61-3.12)	1.10 (0.88-1.37)
Recent TB	**202**	5.05 (2.14-11.93)	2.33 (1.14-4.79)	0.95 (0.56-1.63)
Incident TB	**380**	1.26 (0.72-2.22)	1.15 (0.64-2.08)	1.66 (0.67-4.07)

Exposure	N	Weighted estimates by duration of co-medication (in months)		
		0-2	3-6	7+
Ongoing TB	**1272**	2.31 (1.31-4.08)	2.37 (1.58-3.53)	1.24 (0.95-1.62)
Recent TB	**202**	6.54 (2.58-16.57)	1.39 (0.63-3.06)	1.03 (0.55-1.93)
Incident TB	**380**	1.38 (0.70-2.75)	1.13 (0.59-2.15)	1.26 (0.43-3.69)

Table 6.3. Hazard ratios and 95% confidence intervals for outcome of stavudine substitution by duration of TB treatment prior to HAART initiation (sensitivity analysis S6, only among those with ongoing TB treatment).

Length of TB treatment previous to HAART initiation	n	Time period		
		0-2 months	3-6 months	7+ months
15-60 days	**556**	2.64 (1.28-5.42)	2.66 (1.72-4.11)	1.21 (0.83-1.75)
61+ days	**716**	2.15 (1.05-4.40)	2.16 (1.23-3.81)	1.25 (0.88-1.76)

Figure 6.1. Schematic diagram of TB treatment exposure categories: ongoing TB treatment at time of HAART initiation; recent initiation of TB treatment with respect to HAART; incident TB treatment after HAART initiation.

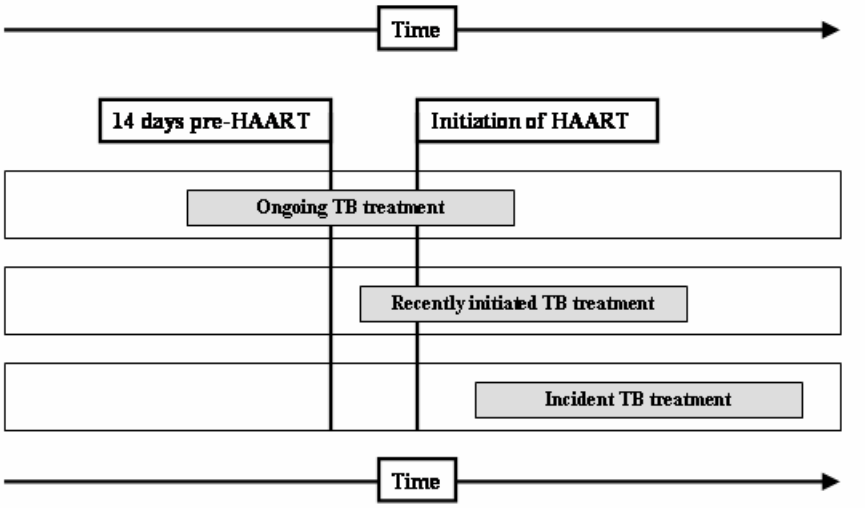
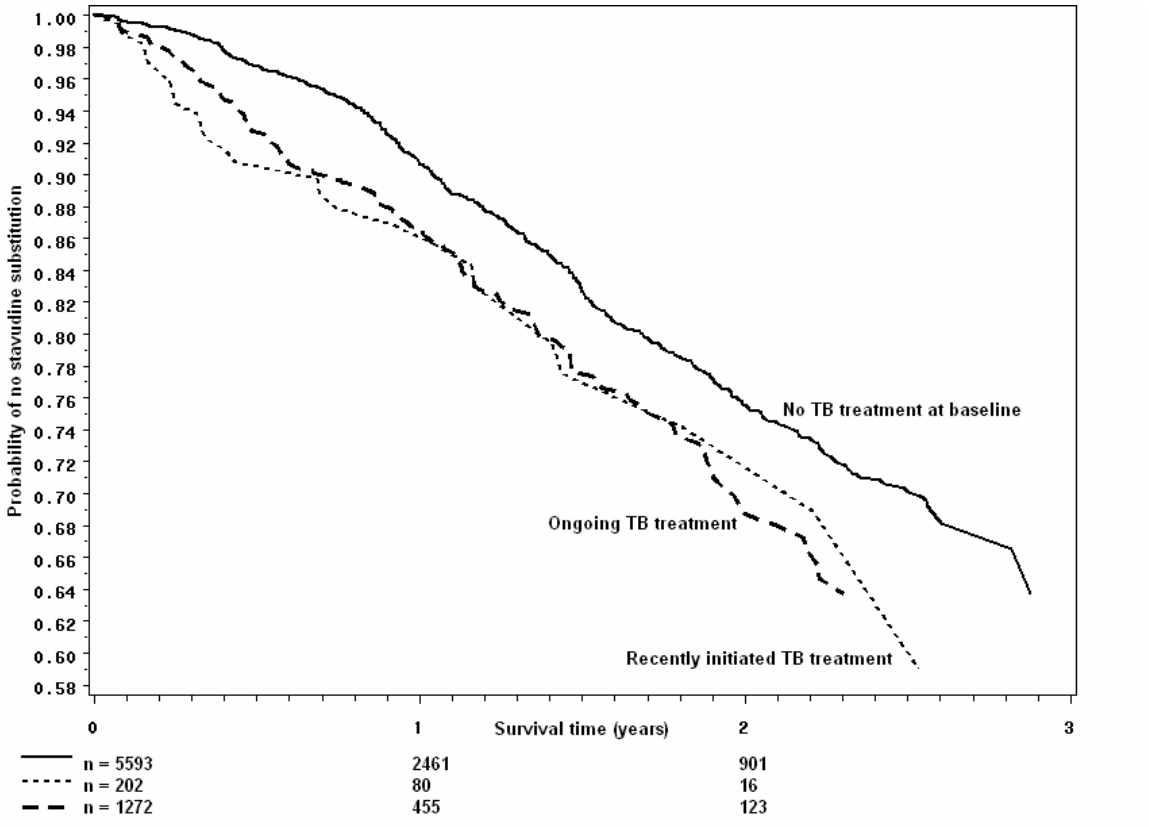


Figure 6.2. Kaplan-Meier analysis of time to d4T substitution by TB treatment status at baseline.



CHAPTER 7

DISCUSSION

OVERVIEW

The goal set by the United Nations is Universal Access to HIV medication by 2010. Achieving this goal would mean initiating approximately 7 million new individuals on HAART over the next two and a half years, and would require a rate of initiation an order of magnitude above what has been achieved so far. The math is overwhelmingly against this goal being achieved.

The wisdom of setting such a goal is debatable. Of course, it is not the first time the WHO has done so. The 3x5 goal (3 million on therapy by 2005) was not reached, but many observers believe that it was the statement – indeed the demand – of that goal which inspired many to roll out access to HAART. And indeed, overall access to HAART has increased geometrically over the past five years, leading to enormous gains in life expectancy among HIV infected individuals throughout the world. The effects of this expanded access have been felt most acutely in sub-Saharan Africa in general and South Africa in particular, where nearly 280,000 individuals, perhaps 30% of those in need, are receiving HAART at present.

The rollout of HAART in South Africa must be considered by almost any measure a great success thus far. Our results from the Themba Lethu Clinic are a

prime example of this success. As we discussed in Chapter III, over three years of follow-up among approximately 7500 patients, virologic suppression remained high, CD4 counts have continued to increase, and survival has been excellent. These results are typical of other reports of cohort in the literature, and taken with these other reports demonstrate that high quality of care is achievable in many situations despite a high rate of initiation of new patients on HAART. Indeed, the results from the Themba Lethu Clinic point towards the future, in that they were achieved with a rate of new patient initiation of 200 per month, or nearly 4 times that reported elsewhere.

At the same time, many large rollout cohorts – including ours – have suffered from significant losses to follow-up, and relatively high rates of drug toxicity and drug substitution. Losses to follow-up may lead us to underestimate the extent of mortality in our cohort, while drug toxicities are both problematic in the short term for reasons of morbidity and sometimes mortality, and in addition may lead to faster virologic failure as toxicities lead to reduced adherence. Worse, the limited antiretroviral drug options available in sub-Saharan Africa (in many cases only a first and second line of HAART) will limit the available options for those same patients who experience drug substitutions for reasons of either virologic failure or toxicity. Going forward, issues of management and durability of HAART and patient retention will both be essential to ensuring that the early successes of the rollout, now demonstrated, are not themselves lost to follow-up.

Another extremely important issue moving forward is the impact of TB in HAART programs. The impact of TB on HIV in this context remains

underappreciated: despite the fact that nearly 17% of patients at Themba Lethu Clinic initiate HAART while receiving ongoing treatment for TB, TB clinics and HIV clinics remain without organizational integration in South Africa and elsewhere. Likewise, the impact of TB therapy on stavudine toxicities is almost entirely unexplored in the literature to date.

STUDY FINDINGS

Our findings on TB highlight several of these key issues. Concerns about the impact of TB even in the presence of HAART persist; in the largest and most rigorous study to date, we have shown that given treatment for both TB and HAART, individuals with TB are not at higher risk of death than others initiating HAART; and what greater risk of death may be observed is due to the greater preponderance of risk factors for death among those with TB, including WHO stage IV disease, low body mass index, and low CD4 cell count. These results will help refocus clinical attention away from TB and towards these key determinants of death, and should help allay fears about the risks of initiating HAART within one month of initiating TB therapy, as we did not see an increased risk among these individuals when considered separately.

Our results on the increased rate of stavudine substitution seen with concomitant therapy for TB point in a different direction. Starting HAART soon after starting TB therapy results in a hazards ratios as large as 6.53 for substitution of stavudine, a first-line HAART drug well known for causing significant toxicity. This

analysis, then, suggests that the longer we can wait after initiation of TB therapy to start HAART, the lower the risk of subsequent d4T substitution (chiefly due to peripheral neuropathy). One caveat in this analysis is that a relatively small number of d4T substitutions occur early in HAART, and so while the hazard *ratio* is large, the absolute risk differences are likely to be substantially smaller. On the other hand, considering that Universal Access proposes to place seven million individuals on HAART in a relatively short period of time, even a small absolute risk difference may prove to have significant public health impact.

Nonetheless, considering that individuals with TB tend to be significantly more immunosuppressed than others, and in light of similar results from other opportunistic infections (Zolopa et al. 2008), there is little doubt that delaying HAART in individuals with TB will lead to unnecessary deaths. Given that our study shows no increased risk of death from starting HAART soon after TB therapy initiation, we feel that in absence of conclusive results from (forthcoming) randomized controlled trials such as ACTG 5221, this study argues strongly for the initiation of HAART no later than two weeks after the initiation of TB therapy.

Nonetheless, there are clearly short and medium term risks to co-administration of TB therapy and stavudine. This is a reason to reconsider stavudine as part of first-line HAART, assuming a less toxic alternative can be found at comparable cost. Clinicians on the ground in South Africa frequently voice a desire for widely available Tenofovir (TDF), which is the best alternative but remains expensive in that setting.

FUTURE DIRECTIONS

Methods moving forward. The larger lesson of this work, however, is how much remains unknown that may be vital for ensuring a high quality of care as the rollout continues and expands. Meeting the challenges posed by an expanding – and ever more quickly expanding – scale-up throughout sub-Saharan Africa will take a great deal of very smart epidemiology. We must ask the right questions and answer those questions rigorously and with the right and most appropriate methods.

With regards to epidemiologic methods, we have made several small advancements in methodology in the course of this dissertation which are worth emphasizing here, as they will serve this work going forward. First, we have developed SAS code which can display non-parametric life-tables from discrete-time hazard models, allowing reweighting not only by baseline treatment weights, but in addition by time-varying inverse probability of censoring weights. This code represents an incremental advance over previous methods (Cole 2004; other) for presenting such data, in that previous methods did not allow for reweighting to compensate for censoring. In fact, the method is relatively simple, relying on the key intuition that for every observed person-month of experience in a discrete time hazard model, there is a single weight, which represents the number of person-months for which the observed person-month must account in the reweighted pseudo-population. The size of the pseudo-population in a given month for the life-table calculation, therefore, is simply the sum of the weights among individuals

observed that month; and the number of outcome events in the same time is simply the sum of the weights among those who had the outcome.

We have not included these curves in previous chapters, because such curves are not easily publishable in the literature until the method of their calculation itself has been published. Nonetheless, the code is included in Appendix 1, and we plan to submit it for publication. We note that we have used life-table calculations rather than the similar but more familiar Kaplan-Meier curves because discrete time hazard models for marginal structural models typically use relatively coarse time units, such as person-months, for which the life-table censoring assumptions are more reasonable. More importantly, as the time-unit of consideration approaches a limit of 0, life-tables converge to Kaplan-Meier estimators (just as discrete time hazard models converge to the proportional hazards model under similar conditions).

Secondly, we believe that we have made a significant methodological contribution in our recognition – touched on briefly in Chapter II – that reweighting (or, for that matter, randomization) cannot be relied upon to control confounding when evaluating a time-varying hazard. This issue was brought to our attention in our analysis in Chapter V, in which we investigated the time-varying effect of TB medication on risk of d4T substitution, looking mainly at a dichotomous time-interaction before and after 3 months of follow-up. The key insight here is that even assuming perfect control of confounding at baseline through reweighting, outcomes which are measured contingent on previous outcomes are subject to selection. In

particular, in our study patients were only at risk of an outcome of d4T substitution after month 2 if they did not have such an event in months 0 to 2.

At first glance, this situation looks like a well-known assumption of hazard models; that they model risk conditional on survival – this is not bias, but instead is assumption and definition. But with deeper inspection, the problem becomes clear: by selecting out those individuals who were at high risk of outcome early, we have changed the composition of individuals at risk of later outcome. As such, control of confounding implemented at the level of the entire population may no longer complete control confounding in this selected, secondary group. This problem is similar to that described in the randomized controlled trials literature on principal stratification, in which analysis of outcomes subject to right-censoring can be biased, but to our knowledge this problem has not yet been described in the literature on observational trials; it has potentially important applications for vast numbers of analyses. We plan to perform simulation studies to confirm this theoretical argument, and publish the findings in the methodological literature.

Future research. Above, we noted that asking the right questions will be essential to ensuring that the rollout maintains a high quality of clinical care. There are several specific areas that I believe are important for epidemiologists to pursue. The issue of patient retention is perhaps more vital than any of these issues, but is a more programmatic than analytic challenge; as such, I will not address it further here.

First, we must improve our management of toxicities and drug substitution. This means, first and foremost, addressing the issue of stavudine. While the ideal

solution at present is universal access to cheap Tenofovir, this is unlikely to happen soon due to issues of cost. In the meanwhile, then, we need to fully delineate the relationship between stavudine and other medications and conditions. In this work I have begun investigating that relationship with regards to TB, the most prevalent opportunistic infection among HIV-positive individuals in South Africa, though much more work remains even with regards to TB, including the impact of TB and TB therapy on particular stavudine toxicities. To the extent that stavudine toxicities can be anticipated, they can be managed proactively, which will reduce mortality and morbidity, as well as increase patient adherence to the drugs and quality of life.

Current clinical practice suggests reducing stavudine dose from 40mg to 30mg in patients with lower BMI, to limit toxicity while maintaining high virologic efficacy. However, little rigorous observational analysis has been performed on the effect of stavudine dose on virologic or drug toxicity outcomes. If stavudine continues to play a key role in initial HAART regimens, we must think in terms of harm reduction; evaluating the outcomes of stavudine dose reductions is a key part of the stavudine harm reduction puzzle. I am currently working with several investigators from Themba Lethu to answer these and related questions. Another related issue is that of TB and virologic suppression, and how TB may interact with different doses of stavudine; and while there is some documentation that TB does not affect rate of initial virologic suppression, there is much less information on how TB may affect long-term risk of virologic failure, and how incident TB interacts with all virologic outcomes.

Another area where epidemiology might help prevent excess toxicities is by creating predictive clinical algorithms of who is at high risk of stavudine (or other) toxicities. However, preliminary analysis, presented in conference in March 2008, suggest that constructing rigorous models of this sort will be quite difficult; risk of stavudine toxicity is predicted strongly by almost no factors besides TB therapy (a result consistent with our finding in Chapter V of little confounding of the TB-d4T substitution relationship). As a result, building predictive models has so far been unsuccessful. However, the application of recursive partitioning algorithms (such as CART) to the problem may yield better predictive models than using parametric models such as logistic regression; this may be an area of future study.

All of these techniques can also help ensure high quality care from midlevel healthcare workers. The more we can simplify first-line HAART regimens and reduce the probability of important toxicities, the easier the job of managing HAART will be for individuals with less training. In general, epidemiology, applied to the right problems, can greatly aid the goals of universal access. This dissertation has made a start, but there is much further to go.

APPENDIX I:

WEIGHT AND FINAL MODELS FOR ANALYSIS IN CHAPTER V

The final structural model (to which the weights were applied) was implemented as a weighted pooled log-binomial (discrete time hazard) model, and specified as

$$\log\left(\Pr(D_k(t)=1 \mid D_k(t-1)=0)\right) = \beta 0 + \beta 1 * X_{0k}$$

where $D_k(t)$ is an indicator of the outcome (death) at time t for patient k, X_{0k} is an indicator variable for the exposure of PTB at time 0 for patient k, β_0 is the intercept, and β_1 is the estimated log-hazard ratio of the effect of PTB on the outcome of death.

In both analyses, the main exposure of baseline PTB does not vary in time, but may affect future risk factors for death: for instance, baseline PTB may affect CD4 reconstitution during follow-up, which may put patients at higher risk of subsequent death. However, such an effect should properly be considered part of the causal action of PTB on mortality; adjustment for such a factor would therefore result in a biased estimate of PTB on mortality. Thus, exposure weights relied only on baseline covariates for this analysis. However, because risk of mortality and LTFU both varied in time and with time-dependent changes in other variables, it was important to consider time-updated covariates in models for censoring weights.

There were four weight models specified for the main analysis of Model 1. The weight component of inverse probability of exposure weights (IPEW, more typically called inverse probability of *treatment* weights(Hernán, Brumback, and Robins 2000)) model (E_W) included terms for gender, ethnicity, employment status, age, history of antiretroviral therapy, history of PTB, and baseline measures of pregnancy, peripheral neuropathy, hemoglobin (adjusted for sex, pregnancy status, and altitude), BMI, CD4 cell count, and WHO stage (IV or other), year of enrollment (measuring years starting at baseline), and whether treatment was initiated after the elimination of all fees for clinical care. The stabilization component of the IPEW (e.g., Appendix A.3.2 in (Hernán and Robins 2006)) included only prevalence of PTB at baseline.

The weight component of the inverse probability of censoring weights (IPCW) model (C_W), included all terms of model E_W as well as a term for the exposure of PTB, terms for month of follow-up and cubic spline terms for month of follow up (with knots at the approximate 5^{th}, 25^{th}, 50^{th}, 75^{th} and 95^{th} percentiles of survival time), and time-updated measures of hemoglobin, CD4 cell count, and BMI. The stabilization component of the IPCW included all terms of Model C_W except for time-updated hemoglobin, CD4, and BMI. Inclusion of interaction terms among key predictors of PTB prevalence did not markedly change re-weighted effect estimates, and so were not included in the final weight models.

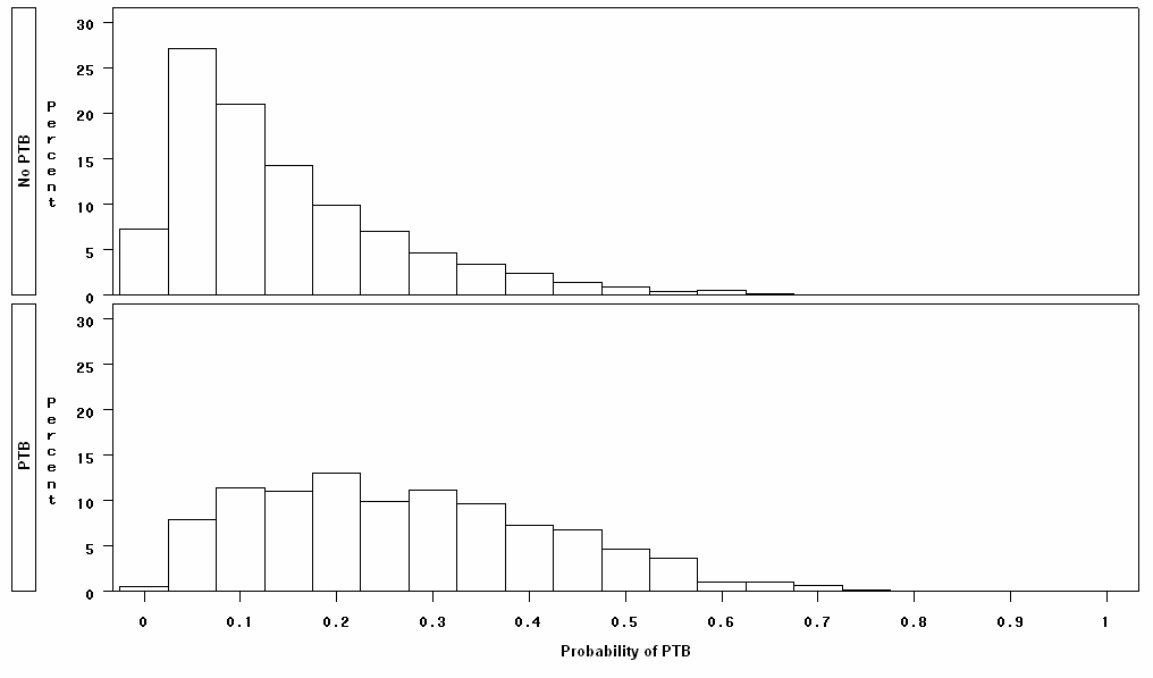
In Model 2, which examined PTB by category of lead-time, the same models were specified, but PTB itself was a five-categorical non-ordinal (nominal) variable, categorized by lead-time (>120 days, 61-120 days, 31-60 days, ≤30 days; and one category of no PTB). Probabilities of being in each category were estimated using polytomous logistic regression.

For both Model 1 and Model 2 we checked the positivity assumption by examining histograms of the conditional probability of exposure for PTB (see Figure A1); we judged that positivity assumption was met in these data, and a similar evaluation for the reduced weight model (S7) yielded similar results. The mean of our stabilized weights was 1.00 in both Model 1, and Model 2; representative values of stabilized and unstabilized weights for Model 1, and stabilized weights for Model 2 and the reduced weight model S7 are shown in Table A1.1.

Table A1.1. Weight distributions for Model 1 (stabilized and unstabilized), Model 2 (stabilized only), and reduced and stabilized Model S7.

	Mean	Median	IQR	Range
Model 1, unstabilized	2.33	1.36	1.21 - 1.77	1.01 - 76.05
Model 1, stabilized	1.00	0.93	0.87 - 1.05	0.19 - 11.32
Model 2, stabilized	1.00	0.94	1.05 - 0.88	0.15 - 9.95
Model S7, stabilized	1.00	0.97	0.90 - 1.05	0.24 - 5.10

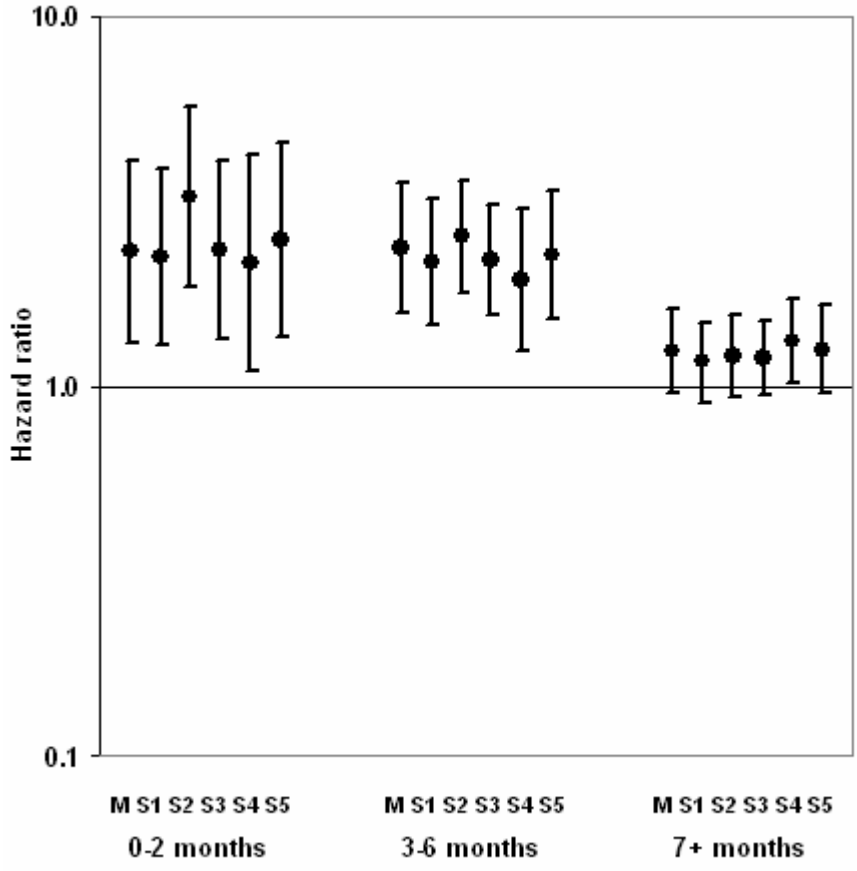
Figure A1.1. Probability of having treated pulmonary tuberculosis (PTB) at baseline in Model 1, by PTB status at baseline (i.e., propensity score curves for PTB).



APPENDIX II:

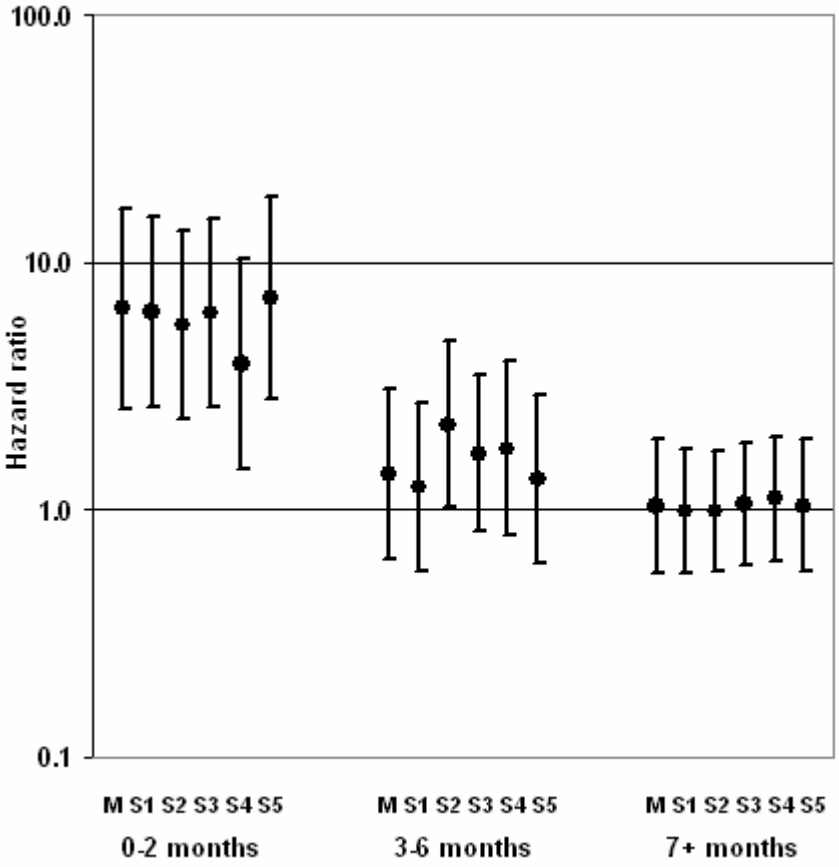
RESULTS OF SENSITIVITY ANALYSES S1-S5 FOR CHAPTER VI.

Figure A2.1. Sensitivity analyses for effect of ongoing TB treatment at time of HAART initiation on risk of stavudine substitution.



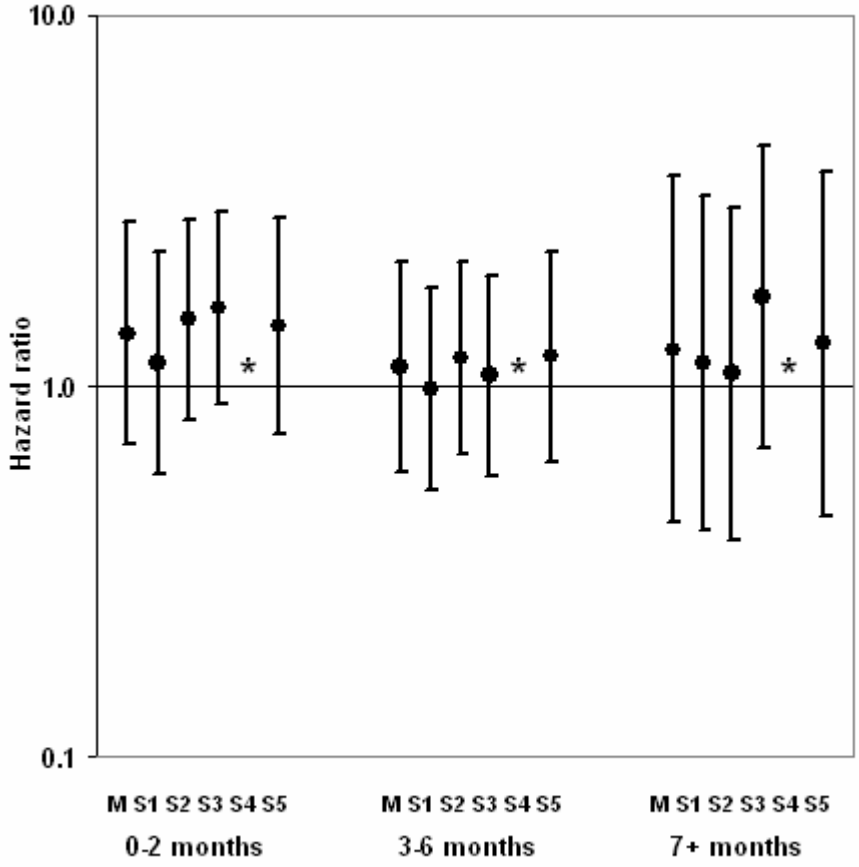
M: main analysis. S1: multi-drug substitution. S2: adjustment for confounding without treatment weights. S3: truncating weights at 1st/99th percentiles. S4: standardized morbidity ratio (effect of the exposure among the exposed). S5: reference group restricted to those without any history of TB.

Figure A2.2. Sensitivity analyses for effect of recent initiation of TB treatment at time of HAART initiation on risk of stavudine substitution.



M: main analysis. S1: multi-drug substitution. S2: adjustment for confounding without treatment weights. S3: truncating weights at 1st/99th percentiles. S4: standardized morbidity ratio (effect of the exposure among the exposed). S5: reference group restricted to those without any history of TB.

Figure A2.3. Sensitivity analyses for effect of incident TB treatment at time of HAART initiation on risk of stavudine substitution. Note that the standardized morbidity ratio (S4) is not well-defined for this exposure.



M: main analysis. S1: multi-drug substitution. S2: adjustment for confounding without treatment weights. S3: truncating weights at 1st/99th percentiles. S4: standardized morbidity ratio (effect of the exposure among the exposed). S5: reference group restricted to those without any history of TB. * Quantity not well-defined.

APPENDIX III:

DETAILS OF MODELS USED FOR TREATMENT, CENSORING, AND COMPETING RISK FOR CHAPTER VI

All hazard ratios were estimated using stabilized inverse probability of treatment and censoring weighted marginal structural models, implemented as weighted pooled log-binomial models (a generalized linear model with a log link and a binomial distribution function). Each weight comprised three factors: a stabilized treatment weight; a stabilized censoring weight; and a stabilized competing risk weight, which modeled the probability of multi-drug substitution including stavudine, but excepting switches to second-line HAART (which is the result of virologic failure and not drug intolerance/toxicity). Because initial analysis showed little change in either estimates or confidence intervals when deaths and switches from first to second line HAART were considered as losses to follow-up, the censoring weights accounted for all three events.

The inclusion of competing risk weights in our overall inverse probability weights has only been described once previously in the literature (Matsuyama and Yamaguchi 2007). However, the intuition behind such an application of inverse probability weights is relatively straightforward, as it just considers multiple censoring categories, modeled separately. Where in other, similar analyses the exposure has been a composite such as “treatment AND not censored”, in this analysis the exposure is simply a larger composite, “treatment AND [not censored OR no competing risk]”.

Covariates considered for inclusion in all models included gender, ethnicity, employment status, age, history of antiretroviral therapy, history of TB; baseline measures of pregnancy, peripheral neuropathy, hemoglobin (adjusted for sex, pregnancy, and altitude), BMI, CD4 cell count, and WHO stage (IV or other), calendar date, and whether treatment was initiated after October 2006 (when all fees for treatment were eliminated). Time varying models (of the exposure of incident TB treatment, censoring, or competing risks) also considered time-updated measures of hemoglobin, CD4 count, BMI, and virologic success (measured only after month 6), and month of follow-up as both a continuous and categorical variable.

The final weight and stabilization treatment models controlled for two exclusive sets of covariates. Set A, included in only the final weight model for treatment, comprised those variables that were confounders of the TB treatment-d4T substitution relationship. Set B, included in both the final weight and stabilization models for treatment, comprised those variables which predicted TB treatment but were only weakly (or non-) predictive of the outcome. The reason for this separation of predictive factors is that the inclusion of covariates which predict exposure but which are not confounders can lead to instabilities in weights, an effect we observed in preliminary weight models. We assigned “true confounder status” and thus membership in Set A or B, using a Cox proportional hazards model for confounding only, using backwards elimination. Our elimination criteria for elimination were a chi-square p-value >0.10 and a change-in-estimate of <0.10 on the log scale.

The final treatment weight model used the following factors to predict probability of TB: gender, age, history of antiretroviral therapy, pregnancy,

peripheral neuropathy at baseline, BMI, year of HAART initiation, and whether clinical care was free or not. In addition, both the weight and stabilization models for treatment contained the following factors: history of TB treatment, ethnicity, employment status, a dichotomous indicator for anemia, WHO stage (IV or otherwise), and categorical CD4 count.

We did not implement such a selection process for models for stabilized censoring or competing risk weights; in both cases, all baseline variables and the exposure were included in stabilizing portions of weights (numerators), while all baseline and time-updated covariates as well as the exposure were included in weight models (denominators) (Hernán, Brumback, and Robins 2000). All weights were estimated using ordinary logistic regression, and final stabilized inverse probability of treatment, censoring, and competing risk weights were calculated by multiplying the three component weights together.

The positivity assumption states that, *given their observed covariates*, there is a non-zero probability that an individual who was being treated for TB could have conceivably not had TB, and vice-versa; this assumption is essentially equivalent to the assumption that there is no complete confounding. We evaluated the positivity assumption by visually examining histograms of multivariate predicted probability of exposure by observed exposure status (a process equivalent to the comparison of propensity score distributions). Because there was substantial overlap in these curves, we judged that the positivity assumption was approximately met in all analyses.

The final structural model (to which the weights were applied) for analyses of recent and ongoing TB treatment was implemented as a weighted pooled log-binomial (discrete time hazard) model, and specified as

$$\log\left(\Pr(S_k(t)=1 \mid S_k(t-1)=0)\right) = \beta_0 + \beta_1 * X_{tk} + \beta_2 * I(3 \le t \le 6) * X_{tk} + \beta_3 * I(t \le 7) * X_{tk} + I(3 \le t \le 6) + I(t \le 7)$$

where $S_k(t)$ is an indicator of the outcome (stavudine substitution) at time t for patient k, X_{tk} is an indicator variable for the exposure of TB treatment at time t for patient k, I(.) represents a 0/1 indicator variable for the truth of the statement (.), β_0 is the intercept, β_1 is the estimated log-hazard ratio of the effect of TB treatment on the outcome of death for months 0-2, $\beta_1+\beta_2$ is the estimated log-hazard ratio of the effect of TB treatment on the outcome of death for months 3-6, and $\beta_1+\beta_3$ is the estimated log-hazard ratio of the effect of TB treatment on the outcome of death for months 7+.

The final structural model for the analysis of the incident TB treatment was similar but omitted the last two indicator terms for month of follow-up, because those terms are assumed in the exposure model for incident TB treatment (as in (Cole et al. 2007)).

REFERENCES

Abraham, OC, A Sivadasan, J Rajan, S Rajkumar, S Pulimood, and D Mathai. 2007. High rates of drug toxicities with generic, fixed dose combination first-line antiretroviral treatment (ART) regimes in south India. In *4th IAS Conference on HIV Pathogenesis, Treatment and Prevention*. Sydney, Australia. Abstract CDB326.

Amoroso, Anthony, R Sheneberger, A Edozien, J Fielder, M Etienne, and K Stafford. 2007. ART-Associated Toxicities Leading to a Switch in Medication: Experience in Uganda, Kenya, and Zambia. In *14th Conference on Retroviruses and Opportunistic Infections*. Los Angeles, CA. Abstract 789.

Anglaret, X., S. Toure, G. Gourvellec, A. Tchehy, L. Zio, M. Zaho, M. C. Kassi, J. Lehou, H. Coulibaly, C. Seyler, T. N'Dri-Yoman, R. Salamon, and G. Chene. 2004. Impact of vital status investigation procedures on estimates of survival in cohorts of HIV-infected patients from Sub-Saharan Africa. *J Acquir Immune Defic Syndr* 35 (3):320-3.

Aurpibul, L., T. Puthanakit, B. Lee, A. Mangklabruks, T. Sirisanthana, and V. Sirisanthana. 2007. Lipodystrophy and metabolic changes in HIV-infected children on non-nucleoside reverse transcriptase inhibitor-based antiretroviral therapy. *Antivir Ther* 12 (8):1247-54.

AVERT.ORG. Accessed 23 August 2007. *AIDS in South Africa: treatment, transmission & the government.* http://www.avert.org/aids-south-africa.htm 2007 [cited Accessed 23 August 2007].

———. Accessed 23 August 2007. *HIV and AIDS in South Africa.* http://www.avert.org/aidssouthafrica.htm 2007 [cited Accessed 23 August 2007].

Badri, M., R. Ehrlich, R. Wood, T. Pulerwitz, and G. Maartens. 2001. Association between tuberculosis and HIV disease progression in a high tuberculosis prevalence area. *Int J Tuberc Lung Dis* 5 (3):225-32.

Badri, M., D. Wilson, and R. Wood. 2002. Effect of highly active antiretroviral therapy on incidence of tuberculosis in South Africa: a cohort study. *Lancet* 359 (9323):2059-64.

Bekker, L. G., L. Myer, C. Orrell, S. Lawn, and R. Wood. 2006. Rapid scale-up of a community-based HIV treatment service: programme performance over 3 consecutive years in Guguletu, South Africa. *S Afr Med J* 96 (4):315-20.

Beutler, E., and J. Waalen. 2006. The definition of anemia: what is the lower limit of normal of the blood hemoglobin concentration? *Blood* 107 (5):1747-50.

Birkus, G., M. J. Hitchcock, and T. Cihlar. 2002. Assessment of mitochondrial toxicity in human cells treated with tenofovir: comparison with other nucleoside reverse transcriptase inhibitors. *Antimicrob Agents Chemother* 46 (3):716-23.

Bisson, G, N Ndwapi, C Rollins, A Avalos, R Gross, S Bellamy, M Mogorosi, H Friedman, D Dickinson, and Tendani Gaolathe. 2007. High Rates of Death among Patients Lost to Follow-up in Botswana's National ART Program: Implications for Monitoring and Evaluation. In *14th Conference on Retroviruses and Opportunistic Infections*. Los Angeles, CA. Abstract 537.

Blower, S., E. Bodine, J. Kahn, and W. McFarland. 2005. The antiretroviral rollout and drug-resistant HIV in Africa: insights from empirical data and theoretical models. *Aids* 19 (1):1-14.

Bolhaar, M. G., and A. S. Karstaedt. 2007. A high incidence of lactic acidosis and symptomatic hyperlactatemia in women receiving highly active antiretroviral therapy in Soweto, South Africa. *Clin Infect Dis* 45 (2):254-60.

Boubaker, K., M. Flepp, P. Sudre, H. Furrer, A. Haensel, B. Hirschel, K. Boggian, J. P. Chave, E. Bernasconi, M. Egger, M. Opravil, M. Rickenbach, P. Francioli, and A. Telenti. 2001. Hyperlactatemia and antiretroviral therapy: the Swiss HIV Cohort Study. *Clin Infect Dis* 33 (11):1931-7.

Boulle, A., C. Orrell, R. Kaplan, G Van Cutsem, M. McNally, K. Hilderbrand, L. Myer, M. Egger, D. Coetzee, G. Maartens, R. Wood, and for the International Epidemiological Databases to Evaluate Aids in Southern Africa (IeDEASA) Collaboration. 2007. Substitutions due to antiretroviral toxicity or contraindication in the first 3 years of antiretroviral therapy in a large South African cohort. *Antivir Ther.* 12 (5):753-60.

Braitstein, P., M. W. Brinkhof, F. Dabis, M. Schechter, A. Boulle, P. Miotti, R. Wood, C. Laurent, E. Sprinz, C. Seyler, D. R. Bangsberg, E. Balestre, J. A. Sterne, M. May, and M. Egger. 2006. Mortality of HIV-1-infected patients in the first year of antiretroviral therapy: comparison between low-income and high-income countries. *Lancet* 367 (9513):817-24.

Breen, R. A., M. C. Lipman, and M. A. Johnson. 2000. Increased incidence of peripheral neuropathy with co-administration of stavudine and isoniazid in HIV-infected individuals. *Aids* 14 (5):615.

Breiman, L. 2001. Statistical Modeling: The Two Cultures. *Statistical Science* 16 (3):199-231.

Byakika-Tusiime, J., J. H. Oyugi, W. A. Tumwikirize, E. T. Katabira, P. N. Mugyenyi, and D. R. Bangsberg. 2005. Adherence to HIV antiretroviral therapy in HIV+ Ugandan patients purchasing therapy. *Int J STD AIDS* 16 (1):38-41.

Calmy, A., L. Pinoges, E. Szumilin, R. Zachariah, N. Ford, and L. Ferradini. 2006. Generic fixed-dose combination antiretroviral treatment in resource-poor settings: multicentric observational cohort. *Aids* 20 (8):1163-9.

Carlin, B. P., and J. S. Hodges. 1999. Hierarchical proportional hazards regression models for highly stratified data. *Biometrics* 55 (4):1162-70.

Carpenter, C. C. 2006. Universal access to antiretroviral therapy: when, not if. *Clin Infect Dis* 42 (2):260-1.

Chen, L. F., J. Hoy, and S. R. Lewin. 2007. Ten years of highly active antiretroviral therapy for HIV infection. *Med J Aust* 186 (3):146-51.

Cheruiyot Bii, S. 2007. Challenges and experiences encountered in the scale up of antiretroviral therapy in the North Rift Region of Kenya. In *4th IAS Conference on HIV Pathogenesis, Treatment and Prevention*. Sydney, Australia. Abstract CDB497.

Chideya, Sekai. 2007. Incidence of Sub-Therapeutic Tuberculosis (TB) Drug Concentrations and Associated Treatment Outcomes among Predominantly HIV-infected TB Patients — Botswana, 1997-2001. In *4th IAS Conference on HIV Pathogenesis, Treatment and Prevention*. Sydney, Australia. Abstract MOAB104.

Chu, H., S. J. Gange, T. E. Yamashita, D. R. Hoover, J. S. Chmiel, J. B. Margolick, and L. P. Jacobson. 2005. Individual variation in CD4 cell count trajectory among human immunodeficiency virus-infected men and women on long-term highly active antiretroviral therapy: an application using a Bayesian random change-point model. *Am J Epidemiol* 162 (8):787-97.

Churchyard, G. J., I. Kleinschmidt, E. L. Corbett, J. Murray, J. Smit, and K. M. De Cock. 2000. Factors associated with an increased case-fatality rate in HIV-infected and non-infected South African gold miners with pulmonary tuberculosis. *Int J Tuberc Lung Dis* 4 (8):705-12.

Clinical HIV Research Unit. 2005. Clinical HIV Research Unit, University of Witwatersrand. http://www.chru.co.za/.

Coetzee, D., A. Boulle, K. Hildebrand, V. Asselman, G. Van Cutsem, and E. Goemaere. 2004. Promoting adherence to antiretroviral therapy: the experience from a primary care setting in Khayelitsha, South Africa. *Aids* 18 Suppl 3:S27-31.

Coetzee, D., K. Hildebrand, A. Boulle, G. Maartens, F. Louis, V. Labatala, H. Reuter, N. Ntwana, and E. Goemaere. 2004. Outcomes after two years of providing antiretroviral treatment in Khayelitsha, South Africa. *Aids* 18 (6):887-95.

Cole, S. R. 2008. Personal communication (April 30 2008).

Cole, S. R., and M. A. Hernán. 2004. Adjusted survival curves with inverse probability weights. *Comput Methods Programs Biomed* 75 (1):45-9.

Cole, S. R., M. A. Hernán, K. Anastos, B. D. Jamieson, and J. M. Robins. 2007. Determining the effect of highly active antiretroviral therapy on changes in human immunodeficiency virus type 1 RNA viral load using a marginal structural left-censored mean model. *Am J Epidemiol* 166 (2):219-27.

Cole, S. R., M. A. Hernán, J. B. Margolick, M. H. Cohen, and J. M. Robins. 2005. Marginal structural models for estimating the effect of highly active antiretroviral therapy initiation on CD4 cell count. *Am J Epidemiol* 162 (5):471-8.

Cole, S. R., M. A. Hernán, J. M. Robins, K. Anastos, J. Chmiel, R. Detels, C. Ervin, J. Feldman, R. Greenblatt, L. Kingsley, S. Lai, M. Young, M. Cohen, and A. Munoz. 2003. Effect of highly active antiretroviral therapy on time to acquired immunodeficiency syndrome or death using marginal structural models. *Am J Epidemiol* 158 (7):687-94.

Cole, S. R., R. Li, K. Anastos, R. Detels, M. Young, J. S. Chmiel, and A. Munoz. 2004. Accounting for leadtime in cohort studies: evaluating when to initiate HIV therapies. *Stat Med* 23 (21):3351-63.

Colebunders, R., M. R. Kamya, J. Laurence, A. Kambugu, H. Byakwaga, P. S. Mwebaze, A. M. Muganga, M. Katwere, and E. Katabira. 2005. First-line antiretroviral therapy in Africa - how evidence-based are our recommendations? *AIDS Rev* 7 (3):148-54.

Colebunders, R., K. R. Moses, J. Laurence, H. M. Shihab, F. Semitala, F. Lutwama, S. Bakeera-Kitaka, L. Lynen, L. Spacek, S. J. Reynolds, T. C. Quinn, B. Viner, and H. Mayanja-Kizza. 2006. A new model to monitor the virological efficacy of antiretroviral treatment in resource-poor countries. *Lancet Infect Dis* 6 (1):53-9.

Collier, A. C., R. W. Coombs, M. A. Fischl, P. R. Skolnik, D. Northfelt, P. Boutin, C. J. Hooper, L. D. Kaplan, P. A. Volberding, L. G. Davis, D. R. Henrard, S. Weller, and L. Corey. 1993. Combination therapy with zidovudine and didanosine compared with zidovudine alone in HIV-1 infection. *Ann Intern Med* 119 (8):786-93.

Cook, J. D., E. Boy, C. Flowers, and C. Daroca Mdel. 2005. The influence of high-altitude living on body iron. *Blood* 106 (4):1441-6.

Cote, J. K., and G. Godin. 2005. Efficacy of interventions in improving adherence to antiretroviral therapy. *Int J STD AIDS* 16 (5):335-43.

Currier, J. 2002. Management of metabolic complications of therapy. *Aids* 16 Suppl 4:S171-6.

Currier, J. S. 2007. Sex differences in antiretroviral therapy toxicity: lactic acidosis, stavudine, and women. *Clin Infect Dis* 45 (2):261-2.

Currier, J. S., and D. V. Havlir. 2005. Complications of HIV disease and antiretroviral therapy. *Top HIV Med* 13 (1):16-23.

D'Agostino, R. B., Jr. 1998. Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. *Stat Med* 17 (19):2265-81.

Dabis, F., E. Balestre, P. Braitstein, P. Miotti, W. G. Brinkhof, M. Schneider, M. Schechter, C. Laurent, A. Boulle, C. Kabugo, G. Capkun, C. Seyler, J. McIntyre, E. Sprinz, D. Bangsberg, S. Van der Borght, and M. Egger. 2005. Cohort Profile: Antiretroviral Therapy in Lower Income Countries (ART-LINC): international collaboration of treatment cohorts. *Int J Epidemiol* 34 (5):979-86.

Davidson, I, H Beardsell, S Mandalia, B Smith, B Gazzard, and M Nelson. 2007. An investigation into the frequency and reasons why patients switch antiretroviral therapy and which antiretrovirals are implicated when switching because of toxicity: The Chelsea switch investigation (CSI) data. In *4th IAS Conference on HIV Pathogenesis, Treatment and Prevention*. Sydney, Australia. Abstract CDB261.

Dean, G. L., S. G. Edwards, N. J. Ives, G. Matthews, E. F. Fox, L. Navaratne, M. Fisher, G. P. Taylor, R. Miller, C. B. Taylor, A. de Ruiter, and A. L. Pozniak. 2002. Treatment of tuberculosis in HIV-infected persons in the era of highly active antiretroviral therapy. *Aids* 16 (1):75-83.

Detels, R., A. Munoz, G. McFarlane, L. A. Kingsley, J. B. Margolick, J. Giorgi, L. K. Schrager, and J. P. Phair. 1998. Effectiveness of potent antiretroviral therapy on time to AIDS and death in men with known HIV infection duration. Multicenter AIDS Cohort Study Investigators. *Jama* 280 (17):1497-503.

Dheda, K., F. C. Lampe, M. A. Johnson, and M. C. Lipman. 2004. Outcome of HIV-associated tuberculosis in the era of highly active antiretroviral therapy. *J Infect Dis* 190 (9):1670-6.

Dieleman, J. P., M. Jambroes, I. C. Gyssens, M. C. Sturkenboom, B. H. Stricker, W. M. Mulder, F. de Wolf, G. J. Weverling, J. M. Lange, P. Reiss, and K. Brinkman. 2002. Determinants of recurrent toxicity-driven switches of highly active antiretroviral therapy. The ATHENA cohort. *Aids* 16 (5):737-45.

Egger, M., A. Boulle, M. Schechter, and P. Miotti. 2005. Antiretroviral therapy in resource-poor settings: scaling up inequalities? *Int J Epidemiol* 34 (3):509-12.

El-Sadr, W, B Elul, W Lo, and D Nash. 2007. Temporal trends in CD4 count at initiation of antiretroviral therapy (ART) and 6 month CD4 response among a large population of HIV-infected individuals receiving HIV care and treatment in 7 sub-Saharan African countries. In *4th IAS Conference on HIV Pathogenesis, Treatment and Prevention*. Sydney, Australia. Abstract WEPEB081.

Etard, J. F., I. Ndiaye, M. Thierry-Mieg, N. F. Gueye, P. M. Gueye, I. Laniece, A. B. Dieng, A. Diouf, C. Laurent, S. Mboup, P. S. Sow, and E. Delaporte. 2006. Mortality and causes of death in adults receiving highly active antiretroviral therapy in Senegal: a 7-year cohort study. *Aids* 20 (8):1181-9.

Evering, T. H., and M. Markowitz. 2008. Raltegravir: an integrase inhibitor for HIV-1. *Expert Opin Investig Drugs* 17 (3):413-22.

Fairall, L. R., M. O. Bachmann, G. M. Louwagie, C. van Vuuren, P. Chikobvu, D. Steyn, G. H. Staniland, V. Timmerman, M. Msimanga, C. J. Seebregts, A. Boulle, R. Nhiwatiwa, E. D. Bateman, M. F. Zwarenstein, and R. D. Chapman. 2008. Effectiveness of antiretroviral treatment in a South African program: a cohort study. *Arch Intern Med* 168 (1):86-93.

Falco, V., D. Rodriguez, E. Ribera, E. Martinez, J. M. Miro, P. Domingo, R. Diazaraque, J. R. Arribas, J. J. Gonzalez-Garcia, F. Montero, L. Sanchez, and A. Pahissa. 2002. Severe nucleoside-associated lactic acidosis in human immunodeficiency virus-infected patients: report of 12 cases and review of the literature. *Clin Infect Dis* 34 (6):838-46.

Ferradini, L., A. Jeannin, L. Pinoges, J. Izopet, D. Odhiambo, L. Mankhambo, G. Karungi, E. Szumilin, S. Balandine, G. Fedida, M. P. Carrieri, B. Spire, N. Ford, J. M. Tassie, P. J. Guerin, and C. Brasher. 2006. Scaling up of highly active antiretroviral therapy in a rural district of Malawi: an effectiveness assessment. *Lancet* 367 (9519):1335-42.

Fischl, M. A., C. B. Parker, C. Pettinelli, M. Wulfsohn, M. S. Hirsch, A. C. Collier, D. Antoniskis, M. Ho, D. D. Richman, E. Fuchs, and et al. 1990. A randomized controlled trial of a reduced daily dose of zidovudine in patients with the acquired immunodeficiency syndrome. The AIDS Clinical Trials Group. *N Engl J Med* 323 (15):1009-14.

Fischl, M. A., D. D. Richman, N. Hansen, A. C. Collier, J. T. Carey, M. F. Para, W. D. Hardy, R. Dolin, W. G. Powderly, J. D. Allan, and et al. 1990. The safety and efficacy of zidovudine (AZT) in the treatment of subjects with mildly symptomatic human immunodeficiency virus type 1 (HIV) infection. A double-blind, placebo-controlled trial. The AIDS Clinical Trials Group. *Ann Intern Med* 112 (10):727-37.

Forna, F., C. A. Liechty, P. Solberg, F. Asiimwe, W. Were, J. Mermin, P. Behumbiize, T. Tong, J. T. Brooks, and P. J. Weidle. 2007. Clinical toxicity of highly active antiretroviral therapy in a home-based AIDS care program in rural Uganda. *J Acquir Immune Defic Syndr* 44 (4):456-62.

Forna, Fatu, D Moore, J Mermin, J Brooks, W Were, K Buchacz, R Downing, C Borkowf, and P Weidle. 2007. Hematologic Changes Associated with Zidovudine after Single-drug Substitution from Stavudine in a Home-based AIDS Care Program in Rural Uganda. In *14th Conference on Retroviruses and Opportunistic Infections*. Los Angeles, CA. Abstract 793.

Garcia-Calleja, J. M., E. Gouws, and P. D. Ghys. 2006. National population based HIV prevalence surveys in sub-Saharan Africa: results and implications for HIV and AIDS estimates. *Sex Transm Infect* 82 Suppl 3:iii64-70.

Geddes, R., S. Knight, M. Y. Moosa, A. Reddi, K. Uebel, and H. Sunpath. 2006. A high incidence of nucleoside reverse transcriptase inhibitor (NRTI)-induced lactic acidosis in HIV-infected patients in a South African context. *S Afr Med J* 96 (8):722-4.

Gilks, C. F., S. Crowley, R. Ekpini, S. Gove, J. Perriens, Y. Souteyrand, D. Sutherland, M. Vitoria, T. Guerma, and K. De Cock. 2006. The WHO public-health approach to antiretroviral treatment against HIV in resource-limited settings. *Lancet* 368 (9534):505-10.

Glynn, R. J., S. Schneeweiss, and T. Sturmer. 2006. Indications for propensity scores and review of their use in pharmacoepidemiology. *Basic Clin Pharmacol Toxicol* 98 (3):253-9.

Hammer, S. M., K. E. Squires, M. D. Hughes, J. M. Grimes, L. M. Demeter, J. S. Currier, J. J. Eron, Jr., J. E. Feinberg, H. H. Balfour, Jr., L. R. Deyton, J. A. Chodakewitz, and M. A. Fischl. 1997. A controlled trial of two nucleoside analogues plus indinavir in persons with human immunodeficiency virus infection and CD4 cell counts of 200 per cubic millimeter or less. AIDS Clinical Trials Group 320 Study Team. *N Engl J Med* 337 (11):725-33.

Harrell, Frank E. *RCSPLINE Macro*. Department of Biostatistics, Vanderbilt University (http://biostat.mc.vanderbilt.edu/twiki/bin/view/Main/SasMacros) 2004.

Harries, A. D., E. J. Schouten, S. D. Makombe, E. Libamba, H. N. Neufville, E. Some, G. Kadewere, and D. Lungu. 2007. Ensuring uninterrupted supplies of antiretroviral drugs in resource-poor settings: an example from Malawi. *Bull World Health Organ* 85 (2):152-5.

Hawkins, C., C. Achenbach, W. Fryda, D. Ngare, and R. Murphy. 2007. Antiretroviral durability and tolerability in HIV-infected adults living in urban Kenya. *J Acquir Immune Defic Syndr* 45 (3):304-10.

Heath, K. V., J. S. Montaner, G. Bondy, J. Singer, M. V. O'Shaughnessy, and R. S. Hogg. 2003. Emerging drug toxicities of highly active antiretroviral therapy for human immunodeficiency virus (HIV) infection. *Curr Drug Targets* 4 (1):13-22.

Hernán, M. A., B. A. Brumback, and J. M. Robins. 2002. Estimating the causal effect of zidovudine on CD4 count with a marginal structural model for repeated measures. *Stat Med* 21 (12):1689-709.

Hernan, M. A., B. Brumback, and J. M. Robins. 2000. Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. *Epidemiology* 11 (5):561-70.

Hernán, M. A., B. Brumback, and J. M. Robins. 2000. Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. *Epidemiology* 11 (5):561-70.

Hernán, M. A., and J. M. Robins. 2006. Estimating causal effects from epidemiological data. *J Epidemiol Community Health* 60 (7):578-86.

Hing, K, O Koole, M Haverkamp, D Sculier, S Thai, and L Lynen. 2007. Reasons for switching antiretroviral therapy in an HIV cohort in Phnom Penh, Cambodia. In *4th IAS Conference on HIV Pathogenesis, Treatment and Prevention*. Sydney, Australia. Abstract WEPEB058.

Ho, D. D., and Y. Huang. 2002. The HIV-1 vaccine race. *Cell* 110 (2):135-8.

Ho, D. D., A. U. Neumann, A. S. Perelson, W. Chen, J. M. Leonard, and M. Markowitz. 1995. Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. *Nature* 373 (6510):123-6.

Hoffmann, C. J., and C. L. Thio. 2007. Clinical implications of HIV and hepatitis B co-infection in Asia and Africa. *Lancet Infect Dis* 7 (6):402-9.

Janeway, Charles, Paul Travers, Mark Walport, and Mark J. Shlomchik. 2005. *Immunobiology: the Immune System in Health and Disease, Sixth Edition*. New York, NY: Garland Science Publishing.

Kapp, C. 2007. South Africa unveils new 5-year HIV/AIDS plan. *Lancet* 369 (9573):1589-90.

Kocholla, L, M Wangai, N Kusu, J Maundu, and M Thuo. 2007. Reasons for switching highly active antiretroviral therapy (HAART) regimens among HIV/AIDS patients in low-resource settings: the case of Mbagathi District Hospital, Kenya. In *4th IAS Conference on HIV Pathogenesis, Treatment and Prevention*. Sydney, Australia. Abstract MOPEB015.

Koenig, S. P., F. Leandre, and P. E. Farmer. 2004. Scaling-up HIV treatment programmes in resource-limited settings: the rural Haiti experience. *Aids* 18 Suppl 3:S21-5.

Kumarasamy, N., S. Vallabhaneni, A. J. Cecelia, T. Yepthomi, P. Balakrishnan, S. Saghayam, T. P. Flanigan, C. C. Carpenter, S. Solomon, and K. H. Mayer. 2006. Reasons for modification of generic highly active antiretroviral therapeutic regimens among patients in southern India. *J Acquir Immune Defic Syndr* 41 (1):53-8.

Laurent, C., N. Diakhate, N. F. Gueye, M. A. Toure, P. S. Sow, M. A. Faye, M. Gueye, I. Laniece, C. Toure Kane, F. Liegeois, L. Vergne, S. Mboup, S. Badiane, I. Ndoye, and E. Delaporte. 2002. The Senegalese government's highly active antiretroviral therapy initiative: an 18-month follow-up study. *Aids* 16 (10):1363-70.

Lawn, S. D., M. Badri, and R. Wood. 2005. Tuberculosis among HIV-infected patients receiving HAART: long term incidence and risk factors in a South African cohort. *Aids* 19 (18):2109-16.

Lawn, S. D., L. Myer, L. G. Bekker, and R. Wood. 2006. Burden of tuberculosis in an antiretroviral treatment programme in sub-Saharan Africa: impact on treatment outcomes and implications for tuberculosis control. *Aids* 20 (12):1605-12.

———. 2006. CD4 cell count recovery among HIV-infected patients with very advanced immunodeficiency commencing antiretroviral treatment in sub-Saharan Africa. *BMC Infect Dis* 6:59.

Lawn, S. D., L. Myer, G. Harling, C. Orrell, L. G. Bekker, and R. Wood. 2006. Determinants of mortality and nondeath losses from an antiretroviral treatment service in South Africa: implications for program evaluation. *Clin Infect Dis* 43 (6):770-6.

Lawn, S. D., L. Myer, C. Orrell, L. G. Bekker, and R. Wood. 2005. Early mortality among adults accessing a community-based antiretroviral service in South Africa: implications for programme design. *Aids* 19 (18):2141-8.

Lawn, S. D., R. J. Wilkinson, M. C. Lipman, and R. Wood. 2008. Immune reconstitution and "unmasking" of tuberculosis during antiretroviral therapy. *Am J Respir Crit Care Med* 177 (7):680-5.

Lawn, S.D., D.J. Edwards, and R. Wood. 2007. Concurrent drug therapy for tuberculosis and HIV infection in resource-limited settings: present status and future prospects. *Future HIV Ther.* 1 (4):387-98.

Lawn, Stephen. 2008. Concurrent Management of TB and HIV. In *15th Conference on Retroviruses and Opportunistic Infections*. Boston, MA. Abstract 11.

Lehrman, G., I. B. Hogue, S. Palmer, C. Jennings, C. A. Spina, A. Wiegand, A. L. Landay, R. W. Coombs, D. D. Richman, J. W. Mellors, J. M. Coffin, R. J. Bosch, and D. M. Margolis. 2005. Depletion of latent HIV-1 infection in vivo: a proof-of-concept study. *Lancet* 366 (9485):549-55.

Libamba, E., S. Makombe, A. D. Harries, R. Chimzizi, F. M. Salaniponi, E. J. Schouten, and R. Mpazanje. 2005. Scaling up antiretroviral therapy in Africa: learning from tuberculosis control programmes--the case of Malawi. *Int J Tuberc Lung Dis* 9 (10):1062-71.

Libamba, E., S. Makombe, E. Mhango, O. de Ascurra Teck, E. Limbambala, E. J. Schouten, and A. D. Harries. 2006. Supervision, monitoring and evaluation of nationwide scale-up of antiretroviral therapy in Malawi. *Bull World Health Organ* 84 (4):320-6.

Llibre, J. M., P. Domingo, R. Palacios, J. Santos, M. J. Perez-Elias, R. Sanchez-de la Rosa, C. Miralles, A. Antela, and S. Moreno. 2006. Sustained improvement of dyslipidaemia in HAART-treated patients replacing stavudine with tenofovir. *Aids* 20 (10):1407-14.

Lonergan, J. T., R. E. Barber, and W. C. Mathews. 2003. Safety and efficacy of switching to alternative nucleoside analogues following symptomatic hyperlactatemia and lactic acidosis. *Aids* 17 (17):2495-9.

Lonergan, J. T., G. A. McComsey, R. L. Fisher, P. Shalit, T. M. File, Jr., D. J. Ward, V. C. Williams, S. M. Hessenthaler, L. Lindsey, and J. E. Hernandez. 2004. Lack of recurrence of hyperlactatemia in HIV-infected patients switched from stavudine to abacavir or zidovudine. *J Acquir Immune Defic Syndr* 36 (4):935-42.

Lopez-Gatell, H., S. R. Cole, N. A. Hessol, A. L. French, R. M. Greenblatt, S. Landesman, S. Preston-Martin, and K. Anastos. 2007. Effect of tuberculosis on the survival of women infected with human immunodeficiency virus. *Am J Epidemiol* 165 (10):1134-42.

Madruga, JV, I Cassetti, E Koenig, A Etzel, Y Zhou, A Cheng, J Enejosa, and 903E Study Team. 2007. Six year safety and efficacy of tenofovir DF (TDF) in combination with lamivudine (3TC) and efavirenz (EFV) in antiretroviral-naïve patients. In *4th IAS Conference on HIV Pathogenesis, Treatment and Prevention*. Sydney, Australia. Abstract WEPEB030.

Malope, B. I., R. M. Pfeiffer, G. Mbisa, L. Stein, E. M. Ratshikhopha, D. L. O'Connell, F. Sitas, P. MacPhail, and D. Whitby. 2007. Transmission of Kaposi sarcoma-associated herpesvirus between mothers and children in a South African population. *J Acquir Immune Defic Syndr* 44 (3):351-5.

Manosuthi, W., S. Chottanapand, S. Thongyen, A. Chaovavanich, and S. Sungkanuparph. 2006. Survival rate and risk factors of mortality among HIV/tuberculosis-coinfected patients with and without antiretroviral therapy. *J Acquir Immune Defic Syndr* 43 (1):42-6.

Manosuthi, W., W. Mankatitham, A. Lueangniyomkul, S. Chimsuntorn, and S. Sungkanuparph. 2008. Standard-dose efavirenz vs. standard-dose nevirapine in antiretroviral regimens among HIV-1 and tuberculosis co-infected patients who received rifampicin. *HIV Med* 9 (5):294-9.

Manosuthi, W., S. Sungkanuparph, A. Thakkinstian, A. Vibhagool, S. Kiertiburanakul, S. Rattanasiri, W. Prasithsirikul, J. Sankote, A. Mahanontharit, and K. Ruxrungtham. 2005. Efavirenz levels and 24-week efficacy in HIV-infected patients with tuberculosis receiving highly active antiretroviral therapy and rifampicin. *Aids* 19 (14):1481-6.

Marceau, G., V. Sapin, C. Jacomet, S. Ughetto, L. Cormerais, C. Regagnon, B. Dastugue, H. Peigue-Lafeuille, J. Beytout, and H. Laurichesse. 2003. Frequency, risk factors, and outcome of hyperlactatemia in HIV-positive persons: implications for the management of treated patients. *Clin Chem* 49 (7):1154-62.

Marlink, Richard, M Bulterys, T Ellerbrock, R McCullough, T Neluheni, M Nolan, J Stringer, A Tanoh, D Tindyebwa, and S Wiktor. 2007. Project HEART: Initiating Care and ART for >134,000 People in 28 Months in 4 African Countries. In *14th Conference on Retroviruses and Opportunistic Infections*. Los Angeles, CA. Abstract 540.

Martin, J, N Musinguzi, N Emenyonu, A Ellman, B Mwebesa, S Deeks, C Yiannoutsos, D Bangsberg, and East Africa IeDEA Consortium. 2007.

Towards more valid estimation of antiretroviral treatment outcomes in resource-limited settings: feasibility of finding and determining outcomes in patients who have become lost-to-follow-up. In *4th IAS Conference on HIV Pathogenesis, Treatment and Prevention*. Sydney, Australia. Abstract CDB284.

Matsuyama, Y., and T. Yamaguchi. 2007. Estimation of the marginal survival time in the presence of dependent competing risks using inverse probability of censoring weighted (IPCW) methods. *Pharm Stat*.

McCaffrey, D. F., G. Ridgeway, and A. R. Morral. 2004. Propensity score estimation with boosted regression for evaluating causal effects in observational studies. *Psychol Methods* 9 (4):403-25.

McCarthy, E. A., M. E. O'Brien, and W. R. Rodriguez. 2006. Training and HIV-treatment scale-up: establishing an implementation research agenda. *PLoS Med* 3 (7):e304.

McComsey, G. A., D. M. Paulsen, J. T. Lonergan, S. M. Hessenthaler, C. L. Hoppel, V. C. Williams, R. L. Fisher, C. L. Cherry, C. White-Owen, K. A. Thompson, S. T. Ross, J. E. Hernandez, and L. L. Ross. 2005. Improvements in lipoatrophy, mitochondrial DNA levels and fat apoptosis after replacing stavudine with abacavir or zidovudine. *Aids* 19 (1):15-23.

McIlleron, H., G. Meintjes, W. J. Burman, and G. Maartens. 2007. Complications of antiretroviral therapy in patients with tuberculosis: drug interactions, toxicity, and immune reconstitution inflammatory syndrome. *J Infect Dis* 196 Suppl 1:S63-75.

Meloni, S, Ernest Ekong, D Onwujekwe, C Okany, I Adewole, C Idehen, W Gaschau, J Idoko, R Murphy, P Kanki, and APIN Plus/Harvard PEPFAR Team. 2007. Evaluation of ART in the Harvard PEPFAR Treatment Program in Nigeria. In *14th Conference on Retroviruses and Opportunistic Infections*. Los Angeles, CA. Abstract 544.

Mermin, J., J. Lule, J. P. Ekwaru, S. Malamba, R. Downing, R. Ransom, F. Kaharuza, D. Culver, F. Kizito, R. Bunnell, A. Kigozi, D. Nakanjako, W. Wafula, and R. Quick. 2004. Effect of co-trimoxazole prophylaxis on morbidity, mortality, CD4-cell count, and viral load in HIV infection in rural Uganda. *Lancet* 364 (9443):1428-34.

Meya, D. B., and K. P. McAdam. 2007. The TB pandemic: an old problem seeking new solutions. *J Intern Med* 261 (4):309-29.

Mocroft, A., B. Ledergerber, C. Katlama, O. Kirk, P. Reiss, A. d'Arminio Monforte, B. Knysz, M. Dietrich, A. N. Phillips, and J. D. Lundgren. 2003. Decline in the

AIDS and death rates in the EuroSIDA study: an observational study. *Lancet* 362 (9377):22-9.

Montaner, J. S., H. C. Cote, M. Harris, R. S. Hogg, B. Yip, J. W. Chan, P. R. Harrigan, and M. V. O'Shaughnessy. 2003. Mitochondrial toxicity in the era of HAART: evaluating venous lactate and peripheral blood mitochondrial DNA in HIV-infected patients taking antiretroviral therapy. *J Acquir Immune Defic Syndr* 34 Suppl 1:S85-90.

Moore, D., C. Liechty, P. Ekwaru, W. Were, G. Mwima, P. Solberg, G. Rutherford, and J. Mermin. 2007. Prevalence, incidence and mortality associated with tuberculosis in HIV-infected patients initiating antiretroviral therapy in rural Uganda. *Aids* 21 (6):713-9.

Moore, R. D., W. M. Wong, J. C. Keruly, and J. C. McArthur. 2000. Incidence of neuropathy in HIV-infected patients on monotherapy versus those on combination therapy with didanosine, stavudine and hydroxyurea. *Aids* 14 (3):273-8.

Mosoko, J, W Akam, P Weidle, J Brooks, A Aweh, M Ebad, T Kinge, C Vitek, and P Raghunathan. 2007. Survival and Adherence to ART in an Era of Decreasing Drug Cost in Limbe, Cameroon. In *14th Conference on Retroviruses and Opportunistic Infections*. Los Angeles, CA. Abstract 536.

Moyle, G. J., D. Datta, S. Mandalia, J. Morlese, D. Asboe, and B. G. Gazzard. 2002. Hyperlactataemia and lactic acidosis during antiretroviral therapy: relevance, reproducibility and possible risk factors. *Aids* 16 (10):1341-9.

Moyle, G. J., and M. Sadler. 1998. Peripheral neuropathy with nucleoside antiretrovirals: risk factors, incidence and management. *Drug Saf* 19 (6):481-94.

Mukadi, Y. D., D. Maher, and A. Harries. 2001. Tuberculosis case fatality rates in high HIV prevalence populations in sub-Saharan Africa. *Aids* 15 (2):143-52.

Murdoch, D. M., W. D. Venter, C. Feldman, and A. Van Rie. 2008. Incidence and risk factors for the immune reconstitution inflammatory syndrome in HIV patients in South Africa: a prospective study. *Aids* 22 (5):601-10.

Murdoch, D. M., W. D. Venter, A. Van Rie, and C. Feldman. 2007. Immune reconstitution inflammatory syndrome (IRIS): review of common infectious manifestations and treatment options. *AIDS Res Ther* 4:9.

Muthén, Bengt, and Katherine Masyn. 2005. Discrete-Time Survival Mixture Analysis. *Journal of Educational and Behavioral Statistics* 30 (1):27-58.

Mutimura, E., A. Stewart, P. Rheeder, and N. J. Crowther. 2007. Metabolic function and the prevalence of lipodystrophy in a population of HIV-infected African subjects receiving highly active antiretroviral therapy. *J Acquir Immune Defic Syndr* 46 (4):451-5.

Muula, A. S. 2004. Ethical and programmatic challenges in antiretroviral scaling-up in Malawi: challenges in meeting the World Health Organization's "Treating 3 million by 2005" initiative goals. *Croat Med J* 45 (4):415-21.

Nattrass, N. 2006. South Africa's "rollout" of highly active antiretroviral therapy: a critical assessment. *J Acquir Immune Defic Syndr* 43 (5):618-23.

Navas, E., P. Martin-Davila, L. Moreno, V. Pintado, J. L. Casado, J. Fortun, M. J. Perez-Elias, E. Gomez-Mampaso, and S. Moreno. 2002. Paradoxical reactions of tuberculosis in patients with the acquired immunodeficiency syndrome who are treated with highly active antiretroviral therapy. *Arch Intern Med* 162 (1):97-9.

Ncayiyana, D. J. 2004. Antiretroviral treatment roll-out policy is in place--now for the implementation challenges. *S Afr Med J* 94 (2):61.

Nunn, P., R. Brindle, L. Carpenter, J. Odhiambo, K. Wasunna, R. Newnham, W. Githui, S. Gathua, M. Omwega, and K. McAdam. 1992. Cohort study of human immunodeficiency virus infection in patients with tuberculosis in Nairobi, Kenya. Analysis of early (6-month) mortality. *Am Rev Respir Dis* 146 (4):849-54.

Orrell, C. 2005. Antiretroviral adherence in a resource-poor setting. *Curr HIV/AIDS Rep* 2 (4):171-6.

Orrell, C., D. R. Bangsberg, M. Badri, and R. Wood. 2003. Adherence is not a barrier to successful antiretroviral therapy in South Africa. *Aids* 17 (9):1369-75.

Osler, Meg, D Stead, K Rebe, A Boulle, and G Meintjes. 2007. Severe Hyperlactatemia Complicating ART with Stavudine First-line Therapy in South Africa: Incidence, Risk Factors, and Outcomes. In *14th Conference on Retroviruses and Opportunistic Infections*. Los Angeles, CA. Abstract 792.

Oyugi, J. H., J. Byakika-Tusiime, E. D. Charlebois, C. Kityo, R. Mugerwa, P. Mugyenyi, and D. R. Bangsberg. 2004. Multiple validated measures of adherence indicate high levels of adherence to generic HIV antiretroviral therapy in a resource-limited setting. *J Acquir Immune Defic Syndr* 36 (5):1100-2.

Parienti, J. J., V. Massari, D. Descamps, A. Vabret, E. Bouvet, B. Larouze, and R. Verdon. 2004. Predictors of virologic failure and resistance in HIV-infected

patients treated with nevirapine- or efavirenz-based antiretroviral therapy. *Clin Infect Dis* 38 (9):1311-6.

Parienti, J. J., V. Massari, D. Rey, P. Poubeau, and R. Verdon. 2007. Efavirenz to nevirapine switch in HIV-1-infected patients with dyslipidemia: a randomized, controlled study. *Clin Infect Dis* 45 (2):263-6.

Patel, A., K. Patel, J. Patel, N. Shah, B. Patel, and S. Rani. 2004. Safety and antiretroviral effectiveness of concomitant use of rifampicin and efavirenz for antiretroviral-naive patients in India who are coinfected with tuberculosis and HIV-1. *J Acquir Immune Defic Syndr* 37 (1):1166-9.

Patel, KK, AK Patel, E Naik, R Ranjan, JK Patel, K Tash, and J Sinnott. 2007. TB co-infection treated at onset of therapy does not affect long-term risk of treatment failure among HIV-1 patients initiating EFV-based combination antiretroviral treatment (cART). In *4th IAS Conference on HIV Pathogenesis, Treatment and Prevention*. Sydney, Australia. Abstract MOAB103.

Paterson, D. L., S. Swindells, J. Mohr, M. Brester, E. N. Vergis, C. Squier, M. M. Wagener, and N. Singh. 2000. Adherence to protease inhibitor therapy and outcomes in patients with HIV infection. *Ann Intern Med* 133 (1):21-30.

Pettifor, A. E., H. V. Rees, I. Kleinschmidt, A. E. Steffenson, C. MacPhail, L. Hlongwa-Madikizela, K. Vermaak, and N. S. Padian. 2005. Young people's sexual health in South Africa: HIV prevalence and sexual behaviors from a nationally representative household survey. *Aids* 19 (14):1525-34.

Pinti, M., P. Salomoni, and A. Cossarizza. 2006. Anti-HIV drugs and the mitochondria. *Biochim Biophys Acta* 1757 (5-6):700-7.

Portsmouth, S., J. Stebbing, and B. Gazzard. 2003. Current treatment of HIV infection. *Curr Top Med Chem* 3 (13):1458-66.

PubMed. http://www.ncbi.nlm.nih.gov. *Accessed 03 May 2008.* [cited 03 May 2008.

Republic of South Africa Department of Health. 2004. National Antiretroviral Treatment Guidelines. In *HIV and AIDS Policy Guideline*.

Right To Care. *Right To Care website.* 2005 [cited February 6, 2006.

Robins, J. M., M. A. Hernán, and B. Brumback. 2000. Marginal structural models and causal inference in epidemiology. *Epidemiology* 11 (5):550-60.

Rockstroh, J. K., and U. Spengler. 2004. HIV and hepatitis C virus co-infection. *Lancet Infect Dis* 4 (7):437-44.

Ruiz-Arguelles, G. J. 2006. Altitude above sea level as a variable for definition of anemia. *Blood* 108 (6):2131; author reply 2131-2.

Saag, M. 2003. Is it time to proactively switch successful antiretroviral therapy? Carefully check your SWATCH. *Ann Intern Med* 139 (2):148-9.

Saarto, T., and P. J. Wiffen. 2007. Antidepressants for neuropathic pain. *Cochrane Database Syst Rev* (4):CD005454.

Sato, T., and Y. Matsuyama. 2003. Marginal structural models as a tool for standardization. *Epidemiology* 14 (6):680-6.

Schim van der Loeff, M. F., S. Jaffar, A. A. Aveika, S. Sabally, T. Corrah, E. Harding, A. Alabi, A. Bayang, K. Ariyoshi, and H. C. Whittle. 2002. Mortality of HIV-1, HIV-2 and HIV-1/HIV-2 dually infected patients in a clinic-based cohort in The Gambia. *Aids* 16 (13):1775-83.

Severe, P., P. Leger, M. Charles, F. Noel, G. Bonhomme, G. Bois, E. George, S. Kenel-Pierre, P. F. Wright, R. Gulick, W. D. Johnson, Jr., J. W. Pape, and D. W. Fitzgerald. 2005. Antiretroviral therapy in a thousand patients with AIDS in Haiti. *N Engl J Med* 353 (22):2325-34.

Shepherd, B. E., P. B. Gilbert, Y. Jemiai, and A. Rotnitzky. 2006. Sensitivity analyses comparing outcomes only existing in a subset selected post-randomization, conditional on covariates, with application to HIV vaccine trials. *Biometrics* 62 (2):332-42.

Shepherd, J, A Habib, M Charurat, P Mondal, K Falajayo, A Abimiku, P Dakum, J Farley, W Blattner, and ACTION Project Study Group. 2007. Efficacy and tolerability of alternative first-line antiretroviral drug regimens in Nigeria. In *4th IAS Conference on HIV Pathogenesis, Treatment and Prevention*. Sydney, Australia. Abstract WEPEB060.

Smith, T.C., and S.P. Novella. 2007. HIV Denial in the Internet Era. *PLoS Med* 4 (8):e256.

South African National AIDS Council. *HIV and AIDS and STI Strategic Plan for South Africa, 2007-2011*. http://www.doh.gov.za/docs/misc/stratplan-f.html 2007 [cited 14 September 2007.

Stafford, K, M Etienne, O Aina, S Lewin, A Bositis, C Bositis, A Doherty, R Sheneberger, and A Amoroso. 2007. Time to event similarities between deceased patients and patients lost to follow up after initiating antiretroviral therapy (ART)in Zambia. In *4th IAS Conference on HIV Pathogenesis, Treatment and Prevention*. Sydney, Australia. Abstract CDB017.

Stebbing, J., and G. Moyle. 2003. The clades of HIV: their origins and clinical significance. *AIDS Rev* 5 (4):205-13.

Stringer, J. S., I. Zulu, J. Levy, E. M. Stringer, A. Mwango, B. H. Chi, V. Mtonga, S. Reid, R. A. Cantrell, M. Bulterys, M. S. Saag, R. G. Marlink, A. Mwinga, T. V. Ellerbrock, and M. Sinkala. 2006. Rapid scale-up of antiretroviral therapy at primary care sites in Zambia: feasibility and early outcomes. *Jama* 296 (7):782-93.

Subbaraman, R., S. K. Chaguturu, K. H. Mayer, T. P. Flanigan, and N. Kumarasamy. 2007. Adverse effects of highly active antiretroviral therapy in developing countries. *Clin Infect Dis* 45 (8):1093-101.

The Global Fund. 2007. The Global Fund To Fight AIDS, Tuberculosis, and Malaria. *http://www.theglobalfund.org/en/ (accessed 14 September 2007)*.

Tronchet, J. M., and M. Seman. 2003. Nonnucleoside inhibitors of HIV-1 reverse transcriptase: from the biology of reverse transcription to molecular design. *Curr Top Med Chem* 3 (13):1496-511.

Tsiouris, S. J., D. Coetzee, P. L. Toro, J. Austin, Z. Stein, and W. El-Sadr. 2006. Sensitivity analysis and potential uses of a novel gamma interferon release assay for diagnosis of tuberculosis. *J Clin Microbiol* 44 (8):2844-50.

UNAIDS. *South Africa: Country HIV and AIDS estimates, end 2003* 2006 [cited 1 February 2006.

UNAIDS/WHO. *AIDS Epidemic Update 2005* 2006 [cited 1 February 2006.

———. *AIDS Epidemic Update 2007* 2007 [cited 7 January 2007.

United Nations Special Assembly. 2001. "Global Crisis - Global Action", Declaration of Commitment on HIV/AIDS.

United States of America Department of Health and Human Services. 2005. Guidelines for the Use of Antiretroviral Agents in HIV-1-Infected Adults and Adolescents.

USAID. 2005. The President's Emergency Plan for AIDS Relief. *http://www.usaid.gov/our_work/global_health/aids/pepfar.html*.

Van Cutsem, G, K Hilderbrand, D Coetzee, E Goemaere, and A Boulle. 2007. Clinical Outcomes and Emerging Challenges after 5 Years of ART in a South African Township. In *14th Conference on Retroviruses and Opportunistic Infections*. Los Angeles, CA. Abstract 535.

Van Damme, W., K. Kober, and M. Laga. 2006. The real challenges for scaling up ART in sub-Saharan Africa. *Aids* 20 (5):653-6.

van der Sande, M. A., M. F. Schim van der Loeff, R. C. Bennett, M. Dowling, A. A. Aveika, T. O. Togun, S. Sabally, D. Jeffries, R. A. Adegbola, R. Sarge-Njie, A. Jaye, T. Corrah, S. McConkey, and H. C. Whittle. 2004. Incidence of tuberculosis and survival after its diagnosis in patients infected with HIV-1 and HIV-2. *Aids* 18 (14):1933-41.

Van Rie, A., and D. Enarson. 2006. XDR tuberculosis: an indicator of public-health negligence. *Lancet* 368 (9547):1554-6.

Villamor, E., F. Mugusi, W. Urassa, R. J. Bosch, E. Saathoff, K. Matsumoto, S. N. Meydani, and W. W. Fawzi. 2008. A Trial of the Effect of Micronutrient Supplementation on Treatment Outcome, T Cell Counts, Morbidity, and Mortality in Adults with Pulmonary Tuberculosis. *J Infect Dis* 197 (11):1499-1505.

Weinberger, Steven E. 2004. *Principles of Pulmonary Medicine, Fourth Edition*. Philadelphia: Saunders/Elsevier.

Wester, C. W., O. A. Okezie, A. M. Thomas, H. Bussmann, S. Moyo, T. Muzenda, J. Makhema, E. van Widenfelt, R. Musonda, V. Novitsky, T. Gaolathe, N. Ndwapi, M. Essex, D. R. Kuritzkes, V. Degruttola, and R. G. Marlink. 2007. Higher-Than-Expected Rates of Lactic Acidosis Among: Preliminary Results from a Large Randomized Clinical Trial. *J Acquir Immune Defic Syndr*.

Whalen, C. C., P. Nsubuga, A. Okwera, J. L. Johnson, D. L. Hom, N. L. Michael, R. D. Mugerwa, and J. J. Ellner. 2000. Impact of pulmonary tuberculosis on survival of HIV-infected adults: a prospective epidemiologic study in Uganda. *Aids* 14 (9):1219-28.

White, A. J. 2001. Mitochondrial toxicity and HIV therapy. *Sex Transm Infect* 77 (3):158-73.

WHO. *Iron Deficiency Anaemia; Assessment, Prevention, and Control: A guide for programme managers* 2001 [cited.

———. *TB/HIV: A Clinical Manual. Second edition.* 2004 [cited.

———. *The 3x5 Initiative* 2005 [cited.

———. *Interim WHO Clinical Staging of HIV/AIDS and HIV/AIDS Case Definitions for Surveillance* 2005 [cited.

———. *Antiretroviral Therapy for HIV Infection in Adults and Adolescents: Recommendations for a public health approach. 2006 revision.* 2006 [cited.

———. *Universal Access by 2010* 2006 [cited.

———. *Malaria and HIV/AIDS* 2007 [cited.

———. *TOWARDS UNIVERSAL ACCESS. Scaling up priority HIV/AIDS interventions in the health sector. Progress Report, April 2007.* 2007 [cited.

———. *Tuberculosis (Fact sheet 104).* 2007 [cited.

———. *Tuberculosis Care with TB-HIV Co-management: Integrated Management of Adolescent and Adult Illness (IMAI)* 2007 [cited.

———. *WHO Report 2007: Global Tuberculosis Control. Surveillance, Planning, Financing* 2007 [cited.

———. *BMI classification* 2008 [cited.

———. *WHO Report 2008. Global tuberculosis control - surveillance, planning, financing* 2008 [cited.

Wikipedia. *Johannesburg* 2008.

Wohl, D. A., G. McComsey, P. Tebas, T. T. Brown, M. J. Glesby, D. Reeds, C. Shikuma, K. Mulligan, M. Dube, D. Wininger, J. Huang, M. Revuelta, J. Currier, S. Swindells, C. Fichtenbaum, M. Basar, M. Tungsiripat, W. Meyer, J. Weihe, and C. Wanke. 2006. Current concepts in the diagnosis and management of metabolic complications of HIV infection and its therapy. *Clin Infect Dis* 43 (5):645-53.

Wohl, D. A., C. D. Pilcher, S. Evans, M. Revuelta, G. McComsey, Y. Yang, R. Zackin, B. Alston, S. Welch, M. Basar, A. Kashuba, P. Kondo, A. Martinez, J. Giardini, J. Quinn, M. Littles, H. Wingfield, and S. L. Koletar. 2004. Absence of sustained hyperlactatemia in HIV-infected patients with risk factors for mitochondrial toxicity. *J Acquir Immune Defic Syndr* 35 (3):274-8.

Wools-Kaloustian, K., S. Kimaiyo, L. Diero, A. Siika, J. Sidle, C. T. Yiannoutsos, B. Musick, R. Einterz, K. H. Fife, and W. M. Tierney. 2006. Viability and effectiveness of large-scale HIV treatment initiatives in sub-Saharan Africa: experience from western Kenya. *Aids* 20 (1):41-8.

Wynn, G. H., M. J. Zapor, B. H. Smith, G. Wortmann, J. R. Oesterheld, S. C. Armstrong, and K. L. Cozza. 2004. Antiretrovirals, part 1: overview, history, and focus on protease inhibitors. *Psychosomatics* 45 (3):262-70.

Zachariah, R., M. Fitzgerald, M. Massaquoi, O. Pasulani, L. Arnould, S. Makombe, and A. D. Harries. 2006. Risk factors for high early mortality in patients on antiretroviral treatment in a rural district of Malawi. *Aids* 20 (18):2355-60.

Zachariah, R., R. Teck, O. Ascurra, P. Gomani, M. Manzi, P. Humblet, P. Nunn, F. M. Salaniponi, and A. D. Harries. 2005. Can we get more HIV-positive tuberculosis patients on antiretroviral treatment in a rural district of Malawi? *Int J Tuberc Lung Dis* 9 (3):238-47.

Zapor, M. J., K. L. Cozza, G. H. Wynn, G. W. Wortmann, and S. C. Armstrong. 2004. Antiretrovirals, Part II: focus on non-protease inhibitor antiretrovirals (NRTIs, NNRTIs, and fusion inhibitors). *Psychosomatics* 45 (6):524-35.

Zolopa, Andrew, J Andersen, L Komarow, A Sanchez, C Suckow, I Sanne, E Hogg, W Powderly, and and ACTG A5164 Study Team. 2008. Immediate vs Deferred ART in the Setting of Acute AIDS-related Opportunistic Infection: Final Results of a Randomized Strategy Trial, ACTG A5164. In *15th Conference on Retroviruses and Opportunistic Infections*. Boston, MA. Paper 142.