The Effect of Antiretroviral Treatment on HIV Associated Tuberculosis Incidence and Outcomes in the Free State Province, South Africa

Venessa Timmerman
(TMMVAN001)

Supervisors
Dr. LR Fairall
Prof. MO Bachmann

Thesis Presented for the Degree of
DOCTOR OF PHILOSOPHY
in the Department of Medicine, Faculty of Health Sciences
UNIVERSITY OF CAPE TOWN
February 2013

ABSTRACT

Background: Of people co-infected with HIV and tuberculosis (TB) globally, 79% are in Sub-Saharan Africa. While scale-up of antiretroviral treatment (ART) is reducing mortality in the region, there is little direct evidence of the magnitude of its effect on mortality in people co-infected with TB, or on TB treatment outcomes. There is little evidence on the effect of ART on TB incidence in HIV positive people in South Africa, which has most HIV-TB cases in the region.

Methods: The study population was a cohort of all patients aged ≥16 years enrolled with the public-sector ART programme in the Free State province of South Africa during the first six years (2004-2010). Cohort data were linked between the HIV and TB programmes, laboratory and hospital services and the national population register. Various multiple regression models estimated the effects of ART on time to TB and incidence of TB episodes in all HIV-positive patients (n=74 075) and, in subsets of co-infected patients, estimated the effects of ART on mortality (n=10 498) and TB treatment outcomes (n=9 276). The analyses compared outcomes during person-time with and without ART, adjusting for patient-level confounders, with multiple imputation of missing CD4 counts and weights.

Results: ART was independently associated with a lower hazard of time to TB (sub-distribution hazard ratio 0.64, 95% confidence interval (CI) 0.61-0.67), lower incidence of TB episodes (incidence rate ratio 0.66, 95% CI 0.64-0.69), lower hazard of time to death after TB diagnosis (hazard ratio 0.53, 95% CI 0.47-0.61), and higher odds of successful TB treatment outcome (odds ratio 3.25, 95% CI 2.62-4.06).

Conclusion: ART was effective at reducing the risk of TB in HIV-positive adults and is thus an important part of TB control. It also reduced mortality in co-infected adults, supporting recent policy changes to make it available to all irrespective of CD4 counts. However more research is needed to understand better the causes of high mortality in co-infected patients.

V. Timmerman February 2013

AUTHOR'S DECLARATION

I, Venessa Timmerman, hereby declare that the work on which this thesis is based is my original work (except where acknowledgements indicate otherwise) and that neither the whole work nor any part of it has been, is being, or is to be submitted for another degree in this or any other university.

I empower the university to reproduce for the purpose of research either the whole or any portion of the contents in any manner whatsoever.

Signature: ……

Date: …….

ACKNOWLEDGEMENTS

I wish to express my sincere appreciation to Dr. Lara Fairall and Professor Max Bachmann, my research supervisors, for their patient guidance and enthusiastic encouragement, valuable and constructive suggestions during the planning and development of this research work, and most importantly their willingness to offer their time so generously. Max, a special word of gratitude for your effort in designing and implementing the sophisticated statistics presented in this thesis and Lara, for all your guidance and support from initial writing stages through final editing.

My greatest appreciation to Professor Eric Bateman for Dr. Beverley Draper for their advice and assistance during final reviewing. I wish to thank various people for their contribution and proof reading of the thesis, Dr. Naomi Folb, Dr. Marlene Fourie, Dr. Marina Muller.

Valuable gratitude is expressed for contributions made by staff members from the Free State Department of Health at the time, Ms. Sonja van der Merwe (TB programme manager) and Dr. Eduan Kotze (former head of Information technology unit)

Finally, I wish to thank my parents, brother Conrad and husband Ben for their support, continuous belief, and encouragement. Ben, for all your help during final layout and formatting of the document.

I want to thank the Lord for the privilege and strength to have completed this lifelong dream.

This thesis is dedicated to my mother, Hetsie Timmerman. Thank you for all your encouragement and prayers!

CONTENTS

TABLES

FIGURES

ABBREVIATIONS AND DEFINITIONS

AIDS	Acquired Immunodeficiency Syndrome
ART	AntiRetroviral Treatment
ART TB	Patient recorded as being on TB treatment during HIV programme visit
AZT	Azidovudine (formerly called azidothymidine)
CD4	Cluster of Differentiation 4 (glycoprotein found on the surface of immune cells)
CI	Confidence Interval
CRR	Competing Risks Regression
ddI	Didanosine
DOB	Date of Birth
DOH	National Department of Health
DOTS	Directly Observed Treatment Strategy (the basic package that underpins the Stop TB Strategy)
DVD	Digital Versatile Disk
ETR	Electronic TB Register
ETR-TB	TB information with electronic TB register as source
HIV	Human Immunodeficiency Virus
HIV-TB	TB information with HIV programme database as source
HR	Hazard Ratio
IeDEA	International epidemiological Databases to Evaluate AIDS
IPT	Isoniazid Preventive Therapy
IR	Incidence Rate
IRR	Incidence Rate Ratio
IRIS	Immune Reconstitution Inflammatory Syndrome
MDG	Millennium Development Goal
NDOH	National Department of Health
no-ART	Follow-up time during the period of no-ART or pre-ART treatment
NSP	National Strategic Plan
OR	Odds Ratio
PALSA PLUS	Practical Approach to Lung Health in South Africa
RHZE	Rifampicin+Isoniazid+Pyrazinamide+Ethambutol

RNA	Ribonucleic acid
SACEMA	South African Centre for Epidemiological Modelling and Analysis
SA ID	South African Identification Number (ID)
SHR	Sub-hazard Ratio
SQL	Structured Query Language
TAC	Treatment Action Campaign
TST	Tuberculin Skin Test
TB	Tuberculosis
VL	Viral Load
WHO	World Health Organisation

1. INTRODUCTION

1.1. Background: The "Syndemic" of infection with HIV and tuberculosis in South Africa

The work reported in this thesis focuses on the "syndemic" of Tuberculosis (TB) and human immunodeficiency virus (HIV) infection in South Africa, a health disaster of major proportions that over the last three decades has dominated the medical agendas for the countries of Sub-Saharan Africa and overwhelmed the health resources of these countries. The term "syndemic" highlights the fact that their combined negative effects on the health of nations is not additive but synergistic. The presence of HIV has reversed gains made in combating TB over the previous half century (since the introduction of effective TB therapy) and resulted in dramatic increases in the incidence and mortality from this disease. TB, in turn, has greatly increased the pace of development, morbidity and mortality of infection with HIV. The mechanisms of interaction of these "terrible twins" have been extensively studied. But, perhaps more importantly, their interaction has demanded a scrutiny of TB control programmes, and the need for reorganisation of health services at every level to jointly address these problems through systems of integrated care. Central to such programmes is understanding the impact of effective treatment of each on the prognosis of patients with co-infection, and on the spread of disease in communities. This thesis provides insights on these and other questions through analysis of a dataset from carefully documented cohorts of patients receiving care for TB or HIV (or both) in the state-run health services in the Free State province of South Africa for a 6-year period, from the time of the antiretroviral (ART) roll-out (2004). However, in order to appreciate the importance of this work, it is appropriate to review the nature, size and historical health system responses to this "syndemic".

1.1.1. HIV in Sub-Saharan Africa.

As the world enters the 30th year of the Acquired Immunodeficiency Syndrome (AIDS) epidemic, figures are still alarmingly high; 34 million people were living with HIV at the end of 2011, an overall global prevalence of 0.8%, and 2.5 million people were newly infected with HIV during that year (UNAIDS Global report 2012). In recent years, most high burden countries have seen a decline in new HIV infections, but in the Middle East, North Africa, Eastern Europe and Central Asia an increase in HIV incidence has been reported. The rapid large-scale rollout of

antiretroviral therapy (ART) has saved an estimated 14 million life years in low and middle income countries. By 2011, more than 8 million HIV positive patients had started ART and 1.7 million deaths from AIDS related causes were reported globally - 24 % fewer than in 2005 (UNAIDS Global report 2012).

Sub-Saharan Africa is the most heavily affected region, representing 69% of all HIV positive people (23.5 million in 2011), 92% of HIV positive pregnant women, and 90% of children diagnosed with HIV. Concerted efforts to prevent infection has led to a 25% decline in new HIV infections between 2001 and 2011, which numbered 1.8 million in 2011 (Figure 1). ART coverage has also improved substantially since 2005 (Figure 2). An estimated 56% of eligible patients were receiving ART by 2011, ranging from countries who achieved ART initiation in more than 80% of eligible patients (including five countries in the Sub-Saharan Africa Botswana, Namibia, Rwanda, Swaziland, Zambia) to suboptimal performers (40% to 79% ART coverage for other Africa countries like Lesotho, Mosambique, Malawi, Uganda, South Africa and Democratic Republic of Congo). Although 70% of all people dying from AIDS-related causes are from the Sub-Saharan region, deaths declined by 32%, from 3.2 million in 2001 to 1.2 million in 2011 (UNAIDS Global report 2012).

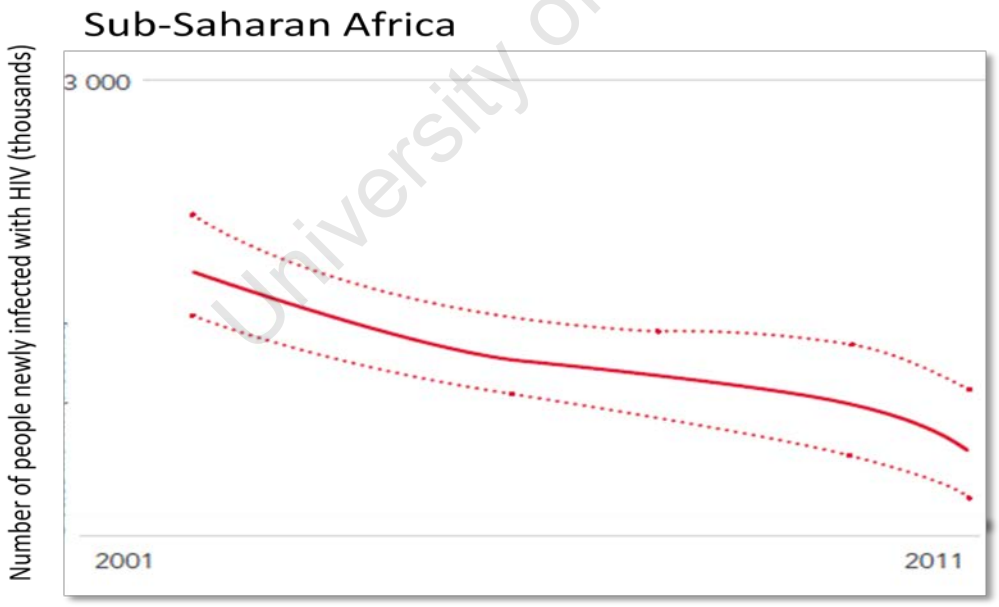

Figure 1: Number of people newly infected with HIV, 2001–2011 (UNAIDS Global report, 2012)

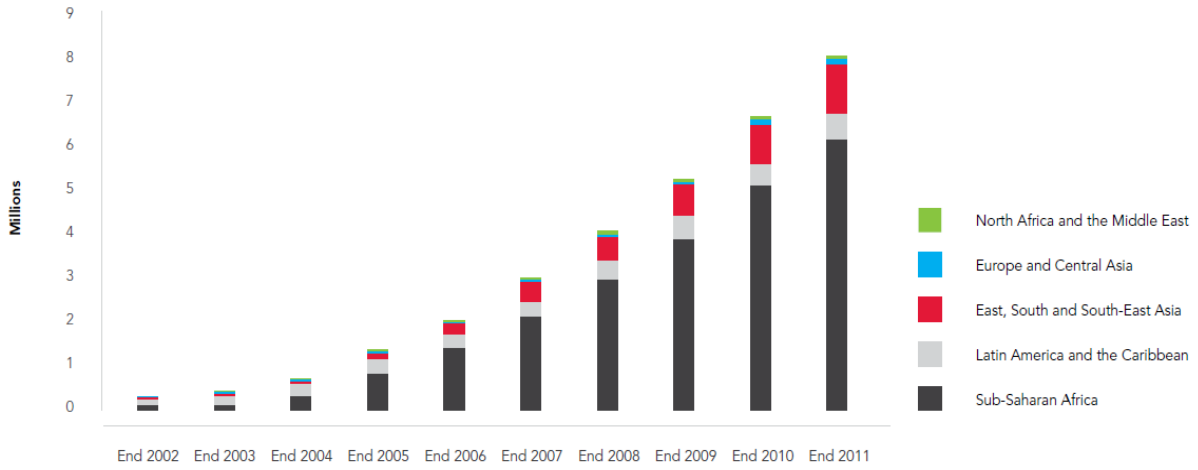

Figure 2: Number of people receiving antiretroviral therapy in low- and middle-income countries, by region, 2002–2011 (UNAIDS Global report, 2012)

1.1.2. HIV/AIDS and South Africa

South Africa is one of the most affected countries worldwide. In 2011, it accounted for 16% (5.36 million) of the worldwide HIV burden while its population represented only 0.7% of the world's population (UNAIDS global report 2012; Stats SA July 2007). Fifty eight percent (approximately 3 million) of affected persons were women and 6% were children, aged 0-14 years (334 000). The national prevalence was estimated in 2011 by an UNAIDS SPECTRUM model to be 17.3% in the general adult population (15-49 years), while the National Antenatal Sentinel Survey Report of 2012 estimated prevalence at 29% among antenatal women, with significant regional differences ranging from 4.75% in the Western Cape to 24.1% in KwaZulu-Natal. In the Free State Province, from where data for this thesis were drawn, the antenatal prevalence in 2011 was found to be 19.58%. (National Department of Health (NDOH) 2012 Antenatal Sentinel Survey Report; UNGASS Global AIDS Response Progress Response 2012: Republic of South Africa). Although the HIV prevalence rate has plateaued, it is at a high level of nearly 30% and the HIV epidemic is continuing. It was estimated that the number of new infections in 2011 was 1.43%, which is slightly lower than the 1.63% new infections in 2008. HIV incidence is exceptionally high compared to other Sub-Saharan countries. (National Department of Health 2011 National Antenatal prevalence survey; NDOH National Strategic Plan 2012-2016; UNAIDS HIV epidemic report. 2010).

South Africa has implemented one of the world's largest HIV treatment programmes during the past few years, with treatment access scaled-up by 75% between 2009 and 2011 as displayed in Figure 3 (UNAIDS global report, 2012). This rapid scale up made it possible to treat 2 million

people by October 2012, which is 80% of all those in need of treatment, reaching the target of universal access to treatment (Irin News 9 October 2012, Report 94365: 'SOUTH AFRICA: Revamped AIDS council makes its debut'). One of the benefits of good ART coverage in communities was shown by a prospective cohort study from Hlabisa sub-district in KwaZulu-Natal which found that the risk of acquiring HIV infection was 38% lower (aHR = 0.62, CI: 0.50-0.76) in communities with high ART coverage (30-40%) compared to low ART coverage (<10%) communities (Tanser F et al., 2013). Mortality is also starting to decrease due to improved ART coverage. An estimated 270 190 South Africans died of AIDS related causes in 2011, placing South Africa in the 25-49% death reduction category of deaths that occurred between 2005 and 2011, and increased the country's life expectancy by five years. (UNAIDS global report 2012; Mayosi BM *et al*., 2012).

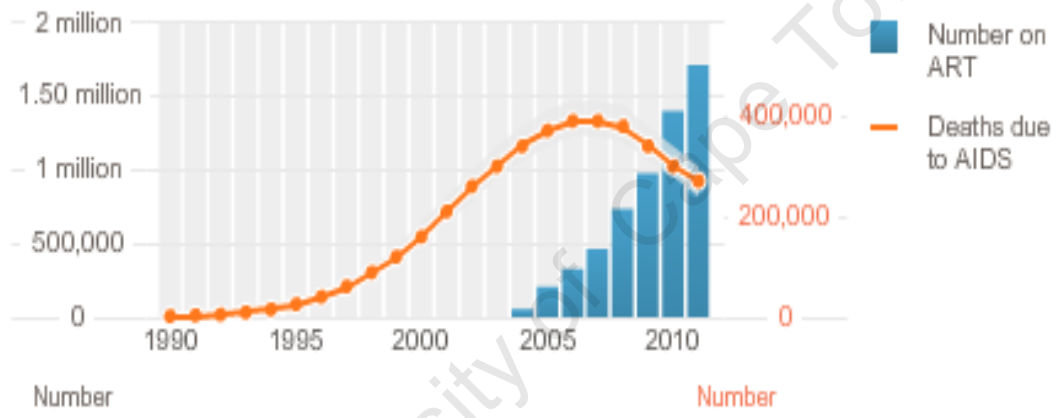

Figure 3: ART scale-up vs. AIDS deaths for South Africa (From UNAIDS 2012 country specific report)

1.1.3. Expansion of ART services in South Africa

Staging of HIV infection and patients' need for ART is based on both clinical manifestations of infection, and an assessment of state of immuno-compromisation, for which measurement of CD4 lymphocytes is the most useful marker.

In November 2003 the South African cabinet approved a national plan for HIV/AIDS prevention, care and treatment. ART guidelines were based on the World Health Organisation (WHO) recommendations (National ART Guidelines, National Department of Health 2004; National Department of Health Strategic Plan 2000-2005). Adults were eligible for ART when their CD4 count fell below 200 cells/µL and/or they presented with WHO Stage 4 disease. The cabinet originally estimated that 53 000 people would be using ART by the end of March 2004. The

target was revised in March 2005, owing to delays in the rollout. In July 2004, some eight months after the rollout started, the Treatment Action Campaign (TAC) organised a symposium to establish a review of the current state of the rollout in South Africa and strategies for increasing access to treatment. It was established that only 6 000 patients were on ART (of the estimated 53 000 who required it). The major reasons for the slow progress in the early phases of the roll out included shortage of qualified doctors, nurses, pharmacists, and counsellors (human resources problem), high prices of drugs and difficulties with their procurement, lack of transparency of government in informing stake-holders of issues requiring attention, logistical problems for example adequate safe storage for drugs, and lack of electronic systems at sites for monitoring treatment and outcomes (Schneider H *et al*., 2006; Steyn F *et al*., 2006).

The HIV treatment programme was also affected by structural problems within the government health sector; care in clinics being organized in "vertical" programmes, and overly strict accreditation system for clinics and hospitals to provide ART. Consequently HIV treatment and care was only accessible at certain clinics, and it was not integrated into primary health care facilities (referred to as horizontal or integrated approach) and only a small proportion of doctors were authorised to prescribe ART. HIV positive patients had to be referred to an ART facility where a registered ART nurse or doctor would assess their eligibility for ART. This vertical model placed pressure on an already straining health care system, diverting staff from other primary care facilities (Steyn F. *et al*., 2006), and causing heavy additional workloads since patients on ART need close monitoring and frequent visits.

Between 2007 and 2011, a revised National Strategic Plan for HIV/AIDS and sexually transmitted infections (National Strategic plan 2007, released by National Department of Health on 14 March 2007) was in operation. Three specific goals were to reduce the number of HIV infections, to reduce the impact of the epidemic on individuals, families, and communities by expanding access to individuals for treatment care and support, and to reduce mother-to-child transmission. In addition, in 2010, the new national minister of health, Dr. Aaron Motsoaledi, usher in several key changes to further accelerate access to treatment including, a mass campaign to test 15 million South Africans for HIV, the scrapping of stringent accreditation criteria for sites providing treatment, the extension of prescribing authority to trained nurses, and the revision of eligibility criteria for starting treatment. These revised criteria made it possible for all TB patients and pregnant women with a CD4 count of less than 350 cells/µl to receive ART. In 2011 eligibility criteria were expanded again, making provision for ART for TB

patients regardless of CD4 count and raising the CD4 count threshold to350 cells/µl for all other HIV positive adults.

Despite all these efforts, the targets for the number of people on ART have not been met during the first few years of the programme. After the initial stumbling blocks the HIV programme had to overcome, the programme is still faced with a number of significant limitations; high staff vacancies in clinics, inadequate budget allocation for primary care, the vertical service model (Uebel KE *et al*., 2010), and long patient waiting times for initiation of ART (Ingle SM *et al*., 2010). The slow pace of rollout has not kept pace with targets or the demand for ART resulting in accumulation of the number of patients waiting for ART initiation. This is illustrated by the figures from the Free State Province as seen in the graph below (Figure 4).

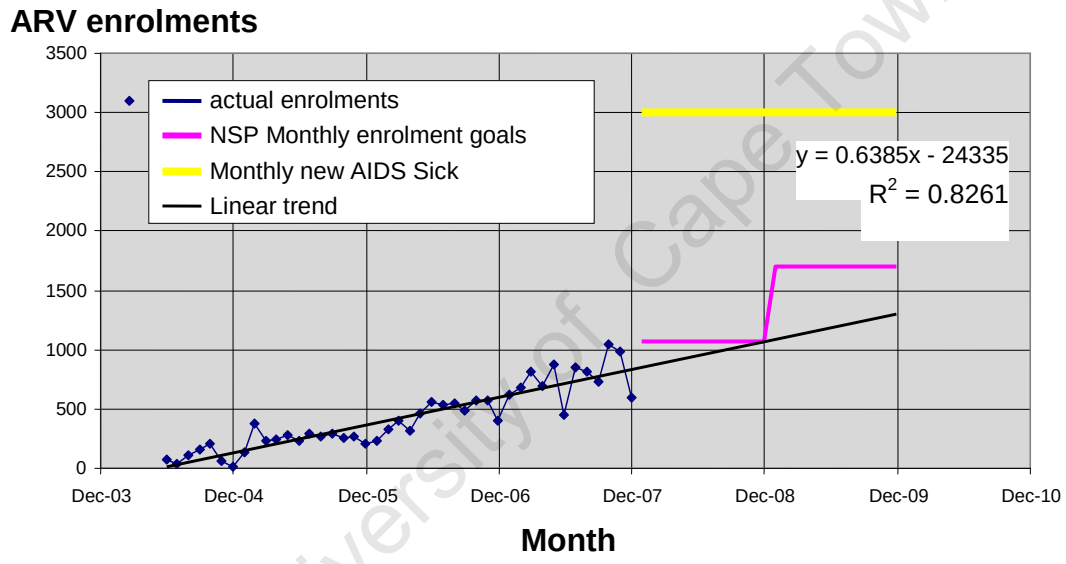

Figure 4: Monthly new enrolments (2004-2007) plotted against need and goals as envisaged in the National Strategic Plan 2008 and 2009 (From Uebel KE *et al*., 2010)

A new National Strategic Plan (NSP) for the period 2012-2016 was released in November 2011 (NDOH: National Strategic Plan 2012-2016) and raised the treatment target from 80% to 90% of those that need ART, addressing Target 6B (Achieve universal access to treatment for HIV/AIDS for all those who need it by 2015) of the Millennium Development Goals number 6 (WHO Millennium Development Goals, 2012, see Figure 5).

The Millennium Development Goals (MDGs) are eight international development goals that were officially established following the Millennium Summit of the United Nations in 2000, whereby all 193 United Nations member states have agreed to achieve these goals by the year 2015. MDG number 6 is to **Combat HIV/AIDS, malaria, and other diseases.**

The targets set for 2015 by this goal include:

Target 6A: Halt and begin to reverse the spread of HIV/AIDS
Target 6B: Achieve, by 2010, universal access to treatment for HIV/AIDS for all those in need.
Target 6C: Halt and begin to reverse the incidence of malaria and other major diseases

Stop TB partnership targets set for 2015 and 2050:

By 2015: Reduce prevalence and death rates by 50%,compared with their levels in 1990
By 2050: Reduce the global incidence of active TB cases to <1 case per 1 million population per year

WHO Health Strategy on HIV launched in 2010:

- Optimise HIV prevention, treatment and care outcomes.
- Leverage broader health outcomes through HIV responses: Strengthen links between HIV programmes and other health programmes
- Build stronger and sustainable systems.
- Reduce vulnerabilities and remove structural barriers to accessing services.

WHO 2015 Targets

- **Reduce new infections:** reduce by 50% the percentage of young people aged 15–24 years who are infected (compared with a 2009 baseline)
- **Eliminate new HIV infections in children:** reduce new HIV infections in children by 90% (compared with a 2009 baseline)
- **Reduce HIV-related mortality:** reduce HIV-related deaths by 25% (compared with a 2009 baseline)
- **Reduce tuberculosis-related mortality:** reduce tuberculosis deaths by 50% (compared with a 2004 baseline).

Source: WHO Global TB report, 2012; UN Millennium goals; WHO Health Strategy

Figure 5: Millennium Development Goals (MDG) and the details of MDG 6, Stop TB targets, and WHO health strategy and targets launched in 2010.

In a further attempt to address target 6A, which is to have reduce mortality related to HIV by 25%, and halted and reversed the spread of HIV/AIDS by 2015, some of the National Strategic plan's goals include the following: to reduce new HIV infections in children by 90% by 2016, improve life expectancy by five years for men and women by 2016. Some of the key interventions proposed in the NSP 2012-2016 include: HIV testing and TB screening for all

South Africans (12 and older) on an annual basis for those with unknown status, or previously tested HIV negative (improved case finding), early diagnosis of HIV and TB, early treatment for HIV and TB, and taking biomedical and behavioural prevention interventions.

The 2012-2016 NSP recognised that the centralised ART model, where only selected facilities provided ART, made it difficult to refer co-infected patients from decentralised TB services when they required concomitant treatment for both conditions. Since 2010 the Department of Health started to decentralise ART initiation to primary health care, which was also supported by a randomised trial in the Free State to assess task shifting of ART prescription to nurses (Fairall LR *et al*., 2012, Uebel KE *et al*., 2010; DOH NSP 2012-2016)

1.2. Tuberculosis in South Africa and co-infection with HIV

WHO reports major progress in the reduction of TB cases and deaths in the past two decades (WHO Global TB report, 2012 and WHO TB fact sheet, 2012). Although the MDG to halt and reverse the TB epidemic by 2015 has been achieved (new cases of TB have been falling and fell at a rate of 2.2% in 2011), they caution that the global burden of TB remains enormous. It is estimated that there were 8.7 million people who fell ill with TB in 2011, of which 13% (1.1 million) were co-infected with HIV and 5.8 million were newly diagnosed cases (two thirds). It remains the second highest global killer due to a single infectious agent (second to HIV/AIDS), with an estimated 1.4 million deaths reported in 2011, 990 000 deaths among HIV-negative and 430 000 among HIV-positive patients. Geographically the burden of TB is highest in Asia and Africa. Approximately one quarter of the world's cases occur in the African region (Figure 6), and Africa shows the highest rates of cases and deaths relative to population, presumably as part of its HIV epidemic.

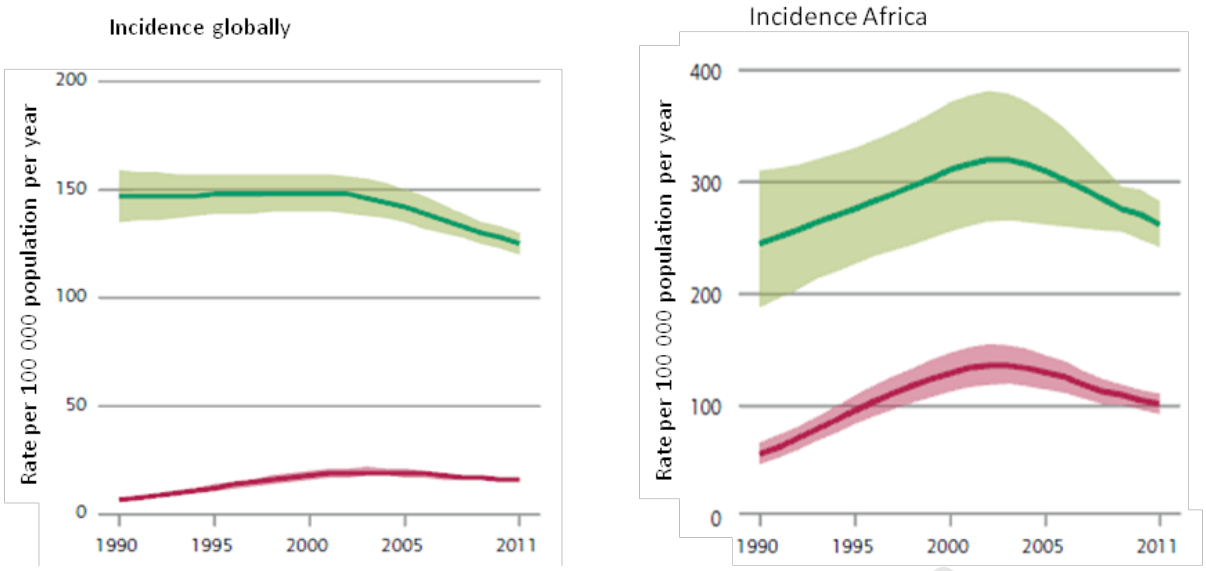

Figure 6: Global trends in estimated rates of TB incidence. Global trends in estimated incidence rate including HIV-positive TB (Green) and estimated incidence rate of HIV-postive TB (Red). From WHO Global TB report, 2012.

The risk of developing TB is estimated to be 20-37 times greater in people living with HIV than among those without HIV infection (WHO TB and HIV). It is well-known that HIV positive patients have a dramatically increased risk of TB infection, either as a result of primary progression or reactivation of latent infection (Manosuthi W *et al.*, 2006) although the importance of re-infection in high prevalence settings is now recognised. The association between HIV and TB has been shown to be strongly associated with CD4 cell recovery during ART in South Africa (Lawn SD *et al.*, 2009a). The incidence of TB is markedly reduced in persons receiving ART (Girardi E *et al.*, 2012), and delaying ART increases the risk of both TB and also other opportunistic infections. The HPTN 052 trial done in sero-discordant couples to evaluate HIV-1 transmission to a HIV negative partner, reported that early ART (started when CD4 count was between 350 and 500) vs. delayed ART (CD4<250) reduced the risk by 89% (HR 0.11 CI: 0.04-0.32), and early ART was also associated with a 41% (HR 0.59 CI: 0.40-0.88) reduction in clinical events (mainly due to extra-pulmonary TB) in the HIV-infected partner (Cohen MS *et al.*, 2011). A mathematical model based on data from nine countries in Sub-Saharan Africa showed that the earlier ART is started after sero-conversion the more TB incidence will be reduced, e.g. when ART is started 5 years after HIV sero-conversion TB incidence in 2050 will be reduced by 66% while ART started 1 year after HIV sero-conversion will reduce TB incidence by 97% (Williams BG *et al.*, 2010). The proportion of TB patients co-infected with TB was overall 39% for the Sub-Saharan Africa region, contributing 79% of global TB-HIV co-infection cases due to its high HIV prevalence (WHO TB fact sheet, 2012).

The HIV epidemic is stabilising in terms of the number of new cases reported, and TB incidence rate was 3.1% lower in 2011 than in 2010 for the Africa region. However, the decline is still slow. This is attributable to the fact that the risk of TB is high even in the early stages of HIV infection, when the risk of other opportunistic infections is still low (i.e. CD4 counts of 350 to 500 cells/µl) (Houlihan CF *et al*., 2009). Peak ages for TB are the same as for HIV; women in their twenties and men in their thirties (Lawn S.D. *et al*, 2006). Patients contracting both TB and HIV (TB-HIV co-infection) are a major concern for South Africa with an incidence rate of 650 TB cases/100 000 for co-infected HIV-TB patients in 2011 (WHO TB South Africa, 2012).

In 2006 WHO ranked South Africa's TB burden as fourth highest in the world (WHO Global TB Report, 2009). By 2011 it still remained among the top five countries with the largest number of incident cases (WHO Global TB report, 2012). The TB notification rate increased fourfold between 1986 and 2006 from 163 cases/100 000 population to 628 cases/100 000 population (NDOH Report 2007: TB Strategic Plan for South Africa, 2007–2011). By 2011, over 900 cases per 100 000 population were reported (WHO TB South Africa, 2012). This dramatic increase in South Africa's TB caseload is shown in Figure 7.

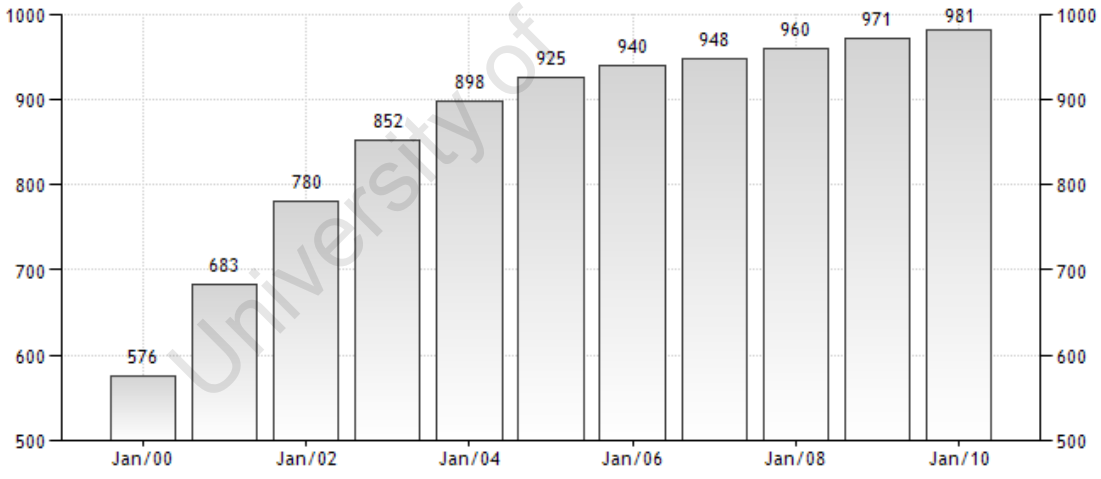

Figure 7: Incidence of TB (per hundred thousand people) in South Africa (from World Bank report, 2012)

In 2011, new cases of TB in South Africa approximated 325 321 (343 715 is the total) of which only 40% were smear positive cases (Table 1). The relatively low percentage of smear positive patients can potentially be attributed to co-infection of TB and HIV due to the atypical presentation of TB in co-infected patients (Jayasundera CI *et al*., 1993; Mahajan V *et al*., 2008).

Table 1: TB case notifications 2011 from the WHO global TB report, 2012

TB case notifications 2011					
New cases		(%)	Retreatment cases		(%)
Smear-positive	129 770	(40)	Relapse	18 394	(40)
Smear-negative	70 341	(22)	Treatment after failure	2 578	(6)
Smear-unknown / not done	77 925	(24)	Treatment after default	4 642	(10)
Extrapulmonary	47 285	(15)	Other	20 301	(44)
Other	0	(0)			
Total new	**325 321**		**Total retreatment**	**45 915**	
Other (history unknown)	18 738				
Total new and relapse	**343 715**		**Total cases notified**	**389 974**	

Due to concerns about the high pill burden, IRIS (immune reconstitution inflammatory syndrome), drug-drug interactions and co-toxicities (Dean GL et al., 2002; Niemie M et al., 2003; Manosuthi W et al., 2010), earlier guidelines recommended that co-infected adults complete their TB treatment before starting ART (National Department of Health Antiretroviral Treatment Guidelines 2004), often resulting in delayed referral to ART programmes. This delay left the large proportion of TB patients co-infected with HIV in the TB programme, without ready access to ART (66% not receiving ART, Table 2), contributing to mortality.

The new NSP (2012-2016) aims at addressing the problem of delayed and small numbers of co-infected patients initiated on ART in support of MDG 6 - Target 6c. The NSP has set the country's targets to increase TB cure rate to 85% by 2016, to reduce TB infections to 2010 levels by 2016, and to reduce mortality due to TB by 50%. Interventions are planned to improve early diagnosis and treatment of TB and reducing vulnerability to TB infection. These interventions to prevent TB infection through intensified case-finding of TB are critical to have reduced and reversed the spread of TB and TB related mortality by 50% in 2015 (WHO Global Health Observatory). Restructuring of the primary health care service platform is a priority to ensure that HIV and TB services are fully integrated at all levels of responsibility and implementation, and that the TB and HIV programmes are incorporated integrated development plans (NDOH NSP 2012-2016).

Table 2: Co-infected HIV/TB patients during 2011 from WHO global TB report, 2012

TB/HIV 2011	Number	(%)
TB patients with known HIV status	323 440	(83)
HIV-positive TB patients	211 800	(65)
HIV-positive TB patients on co-trimoxazole preventive therapy (CPT)	161 298	(76)
HIV-positive TB patients on antiretroviral therapy (ART)	92 376	(44)
HIV-positive people screened for TB	1 256 212	
HIV-positive people provided with IPT	372 994	

Despite all the difficulties in implementing ART and delayed treatment in co-infected HIV-TB patients, the WHO Global TB report shows that since 2002, ART coverage has increased, incidence and prevalence have stabilised, mortality is stabilising (unfortunately excluding figures on HIV positive patients with TB), and treatment success rate has increased (Figure 8).

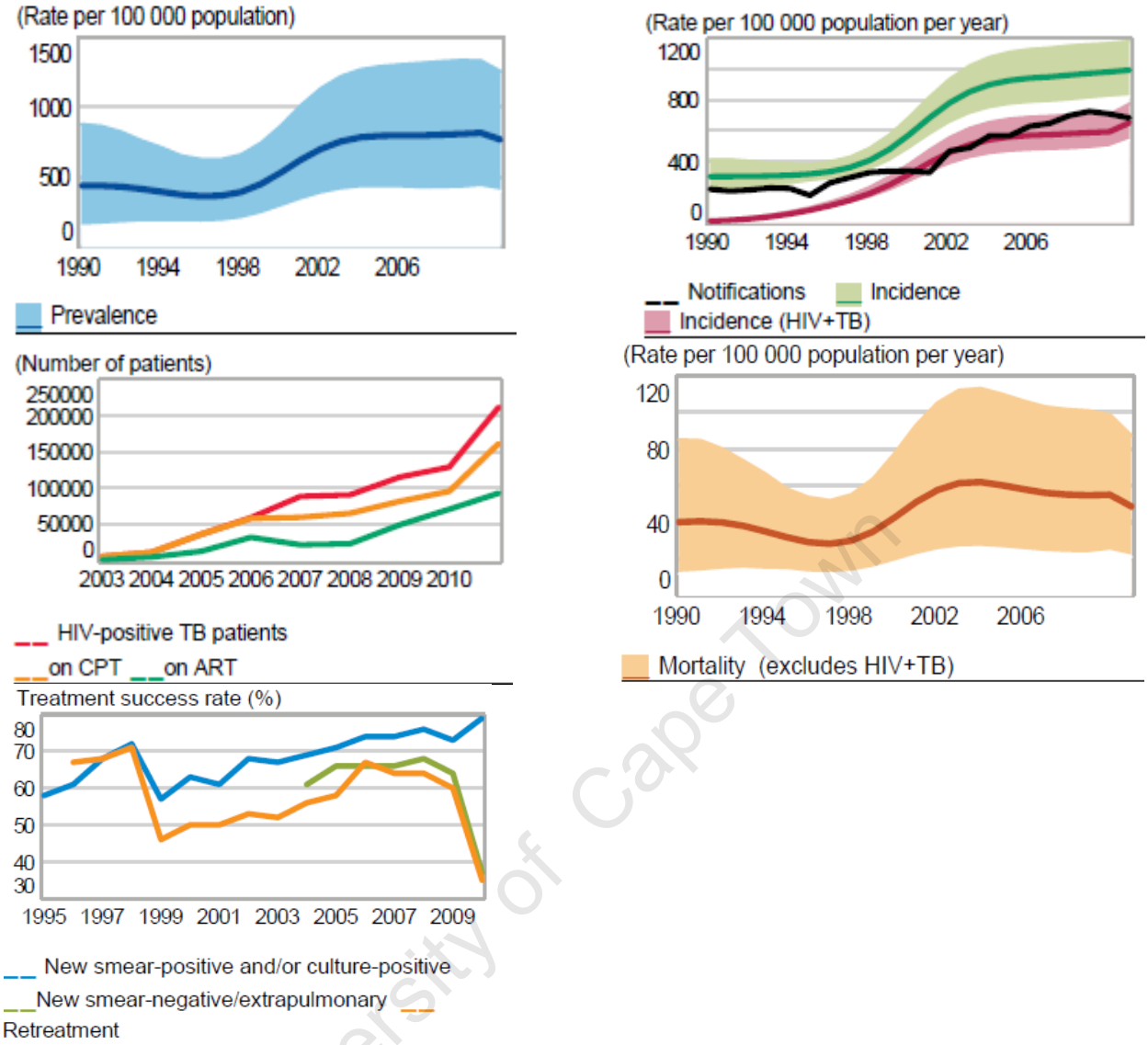

Figure 8: Graphical presentation of ART coverage, incidence and prevalence of TB, mortality (excluding HIV positive and TB), and treatment success in South Africa during 2011 (from WHO Country specific report, 2012)

1.3. Management of HIV and TB

Despite the potential availability of affordable medications, there remains a large number of avoidable deaths and co-morbidity in co-infected HIV-TB patients, especially in Africa (UNAIDS global report, 2010). Only 42% of the nine million people in need of ART treatment globally were receiving it in 2010 (UNAIDS, Aids at 30, 2011). WHO has recognised that the DOT (directly observed treatment) strategy alone is not able to reduce the prevalence and incidence of TB in high HIV burden settings (De Cock KM and Chaisson RE, 1999), and in 2008 recommended the Three I's strategy: intensified case-finding, isoniazid preventive therapy (IPT) and TB infection control (WHO for IPT, 2011).

Since 1993 WHO has recommended the use of isoniazid as a therapy for preventing TB in adults with HIV (WHO preventative therapy policy, 1998). Yet, only a small proportion of patients received IPT according to the WHO Global TB control report, 2012. South Africa was the top country in prescribing IPT during 2011, with 373 000 HIV positive patients without active TB receiving IPT (WHO global TB report, 2012). The reasons for poor implementation of IPT are complex (Lawn SD et al., 2010). IPT should only be prescribed in patients without active TB therefore one of the main reasons for the low uptake of IPT is the process of IPT screening, tuberculin skin test (TST), to assess eligibility for IPT (Kerkhoff AD et al., 2012). The WHO revised their IPT guidelines in 2011 and made an explicit recommendation that TST is not a requirement for initiation of IPT in HIV positive people (WHO for IPT, 2011). A meta-analysis of placebo-controlled trials reported that IPT was effective in reducing the risk of active TB by 32% (RR 0.68 CI: 0.54-0.85), however IPT was not associated with a reduction in all-cause mortality (Akolo C et al., 2010). Observational studies (Fenner L et al., 2011; Charalambous S et al., 2010; Golub JE et al., 2009) and a clinical trial (Durovni B et al., 2010) have shown that the risk of TB and mortality was reduced with the combined effect of IPT and ART, while the results from two ongoing trials are eagerly awaited (NCT00495651 and NCT00463086) to assess the synergy between ART and IPT.

Cotrimoxazole can be used as a prophylactic treatment due to its wide range of action against common infectious agents. It is cheap and widely available but there are concerns about the possibility of increased drug resistance since it is shares a component with drugs commonly used for Malaria. Since patients co-infected with TB-HIV have higher death rates than those infected with either TB or HIV on its own (Colebunders RL *et al*., 1989; De Cock KM *et al*., 1992; Perriens JH *et al*., 1995), information on the effect of cotrimoxazole prophylaxis in patients with both infections are needed. A Cochrane review published in 2009 (Grimwade K and Swindler GH, 2009) identified four placebo controlled trials and reported that cotrimoxazole appears to be effective in preventing death and illnesses in HIV infected patients. However, insufficient evidence existed in people without TB and they recommended that studies needs to be conducted to evaluate the cost-effectiveness of this intervention. A systematic review and meta analysis by Suthar and colleagues (2012a) found that cotrimoxazole therapy, an additional preventative therapy, also increased survival significantly in HIV-infected adults on ART, not limited to co-infected HIV-TB patients. There was no publication bias but high heterogeneity in study results, and most of the studies followed patients for less than one year. Therefore more research is needed to determine the optimal duration of treatment, which is currently evaluated

by a trial in Uganda, testing the hypothesis of stopping cotrimoxazole in adults on ART will not increase morbidity or mortality and decrease haematological adverse events (Suthar AB *et al*., 2012a; Trial ISRCTN44723643).

1.3.1. The effect of ART on TB incidence in people with HIV

The most comprehensive meta-analysis to date examining the impact of ART on TB incidence included eleven studies from developing country settings, of which four were from Sub-Saharan Africa, and one from a combination of regions including Sub-Saharan Africa (Suthar AB *et al*., 2012b). The main finding of this analysis was that ART is strongly associated with reducing risk of TB in adults living with HIV (HR 0.35, 95% CI 0.28 to 0.44). The meta-analysis was also the first to examine the effect of ART on TB incidence in categories defined by the CD4 count at the time of ART initiation. It reported dramatic reductions in all CD4 strata: HR 0.16, (95% CI 0.07 to 0.36) in patients with CD4 less than 200 cells/µL at initiation; HR 0.34, 95% (CI 0.19 to 0.60) in patients with CD4 between 200 and 350 cells/µL at initiation, and in those with CD4 counts of more than 350 cells/µL at initiation HR 0.43 (95% CI 0.30 to 0.63). This meta-analysis is important because it highlights the role ART could play in TB control in high HIV burden settings. Further, the finding of a preventative effect of ART on TB incidence in patients with CD4 counts of more than 350 cells/µl lends weight to arguments supporting earlier initiation of ART.

One multi-cohort study that missed the date cut-off for inclusion in the Suthar review, analysed the incidence of TB during pre-ART and ART periods from eight large HIV programmes in six countries in the Sub-Saharan region (Nicholas S *et al*., 2011). The incidence rates reported were consistent with other studies which also showed that the incidence of TB decreases with starting ART (Brinkhof MW *et al*., 2007; Moore D *et al*., 2007). These findings were consistent with previously published studies that found that immuno-compromised patients were more at risk (Lawn SD *et al*., 2005a & 2006). Importantly, a study by Pettit and colleagues (2011) in low TB incidence setting found that ART reduced the risk of TB, especially when ART was initiated during first six months of ART compared to later. However, after they adjusted for the most recent CD4 count with marginal structural models, the risk of TB did not remain elevated during the first six months of ART compared to not receiving ART, as it was in the crude model. This suggested that the elevated risk of TB during the six month period was related (at least in part) to low CD4 cell counts (Pettit AC *et al*., 2011).

In a collaborative study between researchers in Europe and the United States, data from 12 cohorts were evaluated for the effect of ART on TB incidence in high income countries (HIV Causal Collaboration, 2012). A total of 65 112 patients started ART between 1997 and 2006 of which 712 (1.09%) were diagnosed with TB (the incidence rate was 0.3/100 py). Marginal structural models were used to compare patients receiving ART and those not yet receiving ART. ART was effective in reducing the risk of TB by 44% (HR 0.56 CI: 0.44-0.72) when compared to patients not receiving ART. However, the risk of TB was not reduced by ART in patients with a CD4<50 cells/µl.

Data from a South African cohort also showed dramatic reductions in the risk of TB associated with starting ART (Fairall LR *et al*., 2008). In that study, patients (n=14 267) from the Free State HIV programme were followed up for 20 months by the end of 2005. The incidence of TB 15.08/100 py for patients with a CD4 count below 200 cells/µl and not receiving ART and 5.85/100 py for those receiving ART. Crude incidence rate ratios indicated that ART was highly effective in preventing TB in the CD4<200 cells/µl stratum, IRR 0.39 CI: 0.31-0.48, while the overall crude incidence rate ratio associated with ART was IRR 0.61 CI: 0.49-0.75. Using a marginal structural model and adjusting for baseline variables and time-varying covariates (months on ART, CD4 and weight results), the risk of developing TB was reduced by 39% (HR 0.61 CI: 0.46-0.81) overall, and by 33% in the CD4<200 cells/µl stratum (HR 0.67 CI: 0.50-0.89).

There is consistent evidence from observational studies that TB incidence during the first few months of ART is higher than in later years of treatment (Girarde E *et al*., 2005; Lawn SD *et al*., 2005a; Bonnet MM *et al*., 2006; Brinkhof MW *et al*., 2007; Rajesekaran S *et al*., 2009; Dembele M *et al*., 2010; Van Rie A *et al*., 2011). Higher incidence during the first few months of treatment has been attributed to Immune Reconstitution Inflammatory Syndrome (IRIS) and the unmasking of subclinical TB during initial restoration of the immune system (Dhasmana DJ *et al*., 2008; Meintjies G *et al*., 2008; Lawn SD *et al*., 2008 and 2009c), and also the increased risk of infection before the CD4 count rises.

While the risk of TB declines with cumulative time on ART, it never reaches the normal background rate of TB. A cohort study at Thembalethu clinic in Johannesburg assessed TB incidence in patients receiving ART between 2004 and 2007 (Van Rie A *et al*., 2011). The cumulative probability of TB on ART was a 10% risk of TB by the end of the fourth year, whereas the estimated lifetime risk of TB in HIV uninfected individuals is only 10% (Styblo K,

1985). This is consistent with other studies which also found that TB incidence rate in ART patients stabilises above the estimated background levels of TB incidence (Gupta A *et al*., 2012, Hermans SM *et al*., 2010; Eshun-Wilson I *et al*., 2012), suggesting that ART is not able to restore the immune system completely (Jones JL *et al*., 2000; Brinkhof MW *et al*., 2007; Lawn SD *et al* 2009c; Lawn SD *et al*., 2005).

While many studies have by now reported an association between ART and a reduction in TB incidence, relatively few have included a comparison of ART (3 drugs) vs. no drugs, which is necessary for a more reliable estimate of the effect of ART on TB incidence. The absence of a comparator group is a serious limitation of many studies, and is the reason given for why as many of 40 studies, or two thirds of studies identified, were not included in the Suthar meta-analysis (Suthar AB *et al*., 2012b). Further limitations in these analyses include failure to account for the competing risk of death and the effect of ART on multiple TB episodes.

1.3.2. Effect of ART on mortality in people with HIV and TB

Several studies have shown that HIV positive patients are more likely to die from TB than HIV uninfected TB patients (El Soni AI *et al*., 2002; Sarder N *et al*., 2006; Moolphate S *et al*., 2011). A recent study in Ethiopia evaluating the effect of HIV status on survival in TB patients attending a health centre between 2006 and 2010, confirmed that co-infected HIV positive TB patients were 60% more likely to die than un-infected HIV patients with TB (Shaweno D *et al*., 2012). A significant difference in the death rate was seen during the continuation phase, 7.5 / 1 000 person months for HIV positive vs. 1.7 / 1 000 person months for HIV negative patients (p=0.003). It was hypothesized that late presentation of patients might be a critical factor for early death in this cohort instead of just HIV infection itself since no difference in death was seen during the intensive phase of TB treatment.

TB is a leading reported cause of death among HIV positive patients worldwide (WHO Interim policy 2004; Lopez-Gatell H *et al*., 2008; Saraceni V *et al*., 2008; Cain KP *et al*., 2007; Etard JF *et al*., 2006). The impact of TB on HIV associated mortality may even be underestimated. The atypical clinical presentations of TB in HIV positive patients results in diagnostic challenges and delays in diagnosis (Jayasundera CI *et al*., 1993; Mohan A and Sharma SK , 2008). A high frequency of undiagnosed disseminated TB at post mortem in HIV patients has been reported in several studies (Lucas SB *et al*., 1993; Rana FS *et al*., 2000; Ansari NA *et al*., 2002; Cohen T *et al*., 2010a; Wong EB *et al*., 2010).

Cohorts from both high, low and middle income countries suggest that ART is associated with reduced the risk of mortality in co-infected HIV-TB. (Lawn SD *et al.*, 2009b). In the Netherlands co-infection of HIV and TB was associated with increased risk of death and that the decrease in mortality was due the more widespread use of ART (Haar CH *et al.*, 2007). In Spain mortality was also shown to be reduced in co-infected TB patients who received ART (Velasco M *et al.*, 2009). Some studies in Malawi (Makombe SD *et al.*, 2007; Bong CM *et al.*, 2007) and Thailand (Manosuthi W *et al.*, 2006) have shown improved mortality in TB patients on ART compared with non-TB patients (receiving ART). Surprisingly, a study from Malawi (Makombe D *et al.*, 2007) showed that the six month cohort outcome of co-infected HIV-TB patients on ART had better survival than HIV positive patients on ART without TB (77% vs. 71%, $p<0.001$).

In a study from Thailand, the impact of ART on the survival of HIV positive TB patients compared patients who received ART during TB treatment with those who did not. (Sanguanwonse N *et al.*, 2008). There was an 83% risk reduction (HR 0.17 CI: 0.12-0.24) in mortality. Since this study required a large sample size (n=1 269) patients from health care facilities across the country in both public and private health facilities in 5 provinces were included. Sophisticated analyses (propensity score analysis) were used to adjust for an important survivor bias, i.e. physicians' potentially avoiding commencement of ART in patients with high risk of death (Akksilp S *et al.*, 2007). A smaller study from Thailand published a year earlier by Akksilp *et al* (2007) involved patients from the same setting in the population-based TB surveillance system in one province only (n=329). ART given before or during TB treatment (compared with no ART) reduced deaths by 80% (OR 0.2 CI: 0.1-0.5). Both of these studies monitored death during the course of TB treatment and excluded patients who defaulted or transferred as the vital status of such patients was not known.

Although these studies suggest a strong association between ART and reduction in mortality in co-infected patients, early treatment guidelines advised that ART treatment be delayed. Reasons included concerns about adherence (the high pill burden), (Dean GL *et al.*, 2002; Manosuthi W *et al.*, 2010), potential pharmacologic interactions or drug-drug interactions (Niemie M *et al.*, 2003), higher likelihood of side-effects or drug toxicities, and the risk of IRIS precipitated by mycobacterial products from dying organisms evoking an overwhelming inflammatory response, triggered by a rapidly recovering immune system (Cohen K *et al.*, 2010b; Kwara A *et al.*, 2010; McIlleron H *et al.*, 2007; Lawn SD *et al.*, 2007a; Shelburne SA *et al.*, 2005). However, although real, these risks need to balanced with the risks associated with delaying ART too long (Lawn

SD *et al*., 2005c and 2007b,c; Violari A *et al*., 2008; Basset IV *et al*., 2009; Abdool Karim SS *et al*., 2011).

During the past decade several studies, clinical trials and observational studies have been undertaken to establish the optimal time to commence ART during TB treatment in co-infected patients (Manosuthi W *et al*., 2006; Dheda K *et al*., 2004; Lawn SD *et al*., 2006 and 2011a). Some observational studies found that early initiation of ART initiation was not associated with better survival (Andia I *et al*., 2010; Yotebien M *et al*, 2010), but a study from five hospitals in Spain which compared survival when ART was administered within two months of TB treatment with ART administered after two months of TB treatment suggested benefit from earlier administration (Velasco M *et al*., 2009). Only nine percent (9%) of patients died in the early ART group compared with 20% of patients in the later ART group, a reduction in the risk of death of 63% (HR 0.37 CI: 0.17-0.66). However, results from observational studies like this must be regarded with caution as they do not account for deaths in patients before they started ART or biases inherent in ART allocation (Lawn SD *et al*., 2011a).

Fortunately this question has been the subject of five randomised trials the results of which have most recently been summarised by Naidoo and colleagues (2013, Table 3). Overall, the trials found that only severely immuno-compromised patients (CD4<50 cells/µl) experienced a survival benefit with early ART, except for patients with TB meningitis. Whether ART was initiated within two months or after two months of TB treatment, no difference in AIDS or mortality was found. The Camelia trial reported an advantage of starting ART early (after two weeks of TB) compared to starting ART later (after eight weeks) in patients with a CD4 count below 200cells/µl (Blanc FX *et al*., 2010). Interestingly although the risk and severity of IRIS was increased with early ART, the overall IRIS-associated mortality was lower.

Table 3: Summary of randomised trials evaluating the optimal to start ART in co-infected HIV-TB patients.

Study	Country and Enrolment	Sample size and basic cohort description	Comparison	Outcome
SAPIT Abdool Karim S *et al*., 2010	South Africa (Durban), one facility Enrolment: 28 June 2005 to 11 July 2008	n=429 HIV positive Smear positive pulmonary TB >= 18 years Median CD4: 150 Median age: 34 years Males: 48%	Early arm: started ART within 4 weeks after start of TB (n=214) Late arm: started ART during first 4 weeks of continuation phase of TB (n=215)	Early ART for patients with CD4<50 increased AIDS free survival. Deferred ART in patients with CD4>50 reduced risk of IRIS without increasing risk of AIDS or death.

Study	Country and Enrolment	Sample size and basic cohort description	Comparison	Outcome
STRIDE (ACTG5221) Havlir DV *et al*., 2011	4 Continents, 26 sites Enrolment: Sep 2006 to Aug 2009	n=809 HIV positive suspected TB >= 18 years Median CD4: 77 cells/µl	Early arm: started ART within 2 weeks of TB treatment (n=332) Late arm: started ART after 8-12 weeks of TB treatment (n=329)	Overall, early ART did not reduce AIDS-defining illnesses and death compared to early ART. For CD4<50, early ART had 42% less AIDS defining illnesses and death compared to late ART (p=0.02).
CAMELIA Blanc FX *et al*., 2010	Cambodia, 5 hospitals Enrolment: 31 Jan 2006 to 27 May 2009	n=661 HIV positive Newly diagnosed acid-fast bacilli positive TB >= 18 years Median CD4: 25 Median age: 35 years Males: 64%	Early arm: started ART 2 weeks after start of TB treatment (n=332) Late arm: started ART after 8 weeks of TB treatment (n=329)	Early ART significantly improved survival in patients with CD4<=200 HR 0.62 CI: 0.44-0.86) IRIS significantly increased in early ARM (HR 2.51; CI: 1.78-3.59) Median gain in CD4 was 144 cells/µl irrespective of arm
OXTREC 023-04	Vietnam	n=253 HIV positive TB meningitis Mean CD4: 41	Early arm: Immediate, within 7 days. Late arm: Deferred within 2 months	High overall mortality No difference in time to new AIDS event or death (HR 1.16 CI: 0.87-1.55 p=5.31)
Degu WA *et al*., 2012	Ethiopia	n=512 HIV positive TB Median CD4: 78 cells/µl	ART after TB was started: Arm: 1 week Arm 2: 4 weeks Arm 3: 8 weeks after	Overall incidence of mortality: 1 week: 36.4/100 py 4 weeks: 26/100 py 8 weeks: 20.8/100py (p=0.4) Relative risk of mortality CD4<=50 vs. CD4>50 Wk 1: RR 1.5 CI: 0.6-0.9 Wk 4: RR 2.0 CI: 0.7-5.2 Wk 8: RR 2.9 CI: 0.8-9.9 (NS) Early ART did not reduce risk of mortality and rate of AIDS-defining illnesses, rate of IRIS was increased. Better survival trend seen in patients with CD4<50 cells/µl.
CD4: cells/µl				

The reasons for and mechanisms underlying the excess mortality in co-infected patients remains controversial. A 2010 review on the effect of TB on mortality in HIV positive people concluded

that there is insufficient data to support the causal effect of TB on mortality in co-infected people on ART (Straetemans M *et al*., 2010). The role of advanced immuno-suppression warrants further research (Lawn SD *et al*., 2009b). While it has been shown that TB is strongly associated with mortality in co-infected patients, a direct causal effect has not yet been established as an independent predictor of mortality. Mortality data from ART services in Uganda, Malawi and South Africa have confirmed an associated risk with low baseline CD4 cell counts, increasing age, body mass index, and clinical staging (Moore D *et al*., 2007; Zacharia R *et al*., 2006, Westreich D et *al*., 2009) .

1.3.3. Effect of ART on TB treatment outcomes

Despite long-established TB reporting systems, there are few studies evaluating the effect of ART on TB treatment outcomes. A recently published study from Malawi (Kanyerere HS *et al*., 2012) evaluated outcomes in 1 190 co-infected patients in whom cotrimoxazole was initiated in 88% and ART in 58%. ART was initiated within the first two months in 61% of patients, even though the national guidelines during the study period recommended that ART should be started during the continuation phase of TB treatment. A significant improvement in successful (cured/completed) TB treatment outcomes was found for patients who started ART early (<2 months) compared to late ART (>=2 months); the likelihood of a successful outcome was 60% higher (RR 1.6 CI: 1.4-1.8). They also reported that the risk of death was reduced by 75% (RR 0.25 CI: 0.19-0.35). No comparison was made between patients who received ART and those who did not.

A multi-centre study in Italy evaluated TB treatment outcomes of co-infected patients and the effect of ART on death as recorded in a TB register (Girardi E *et al*., 2012). The use of ART during TB treatment was associated with a reduction in the hazard of death (IRR 0.12 CI: 0.06-0.30) while lower CD4 cell count, age more than 40 years and MDR increased the risk of death.

A study by Hermans and co-workers (2012) showed that in an integrated HIV-TB setting, successful TB treatment outcomes (cured or completed) increased from 62% to 68% and unfavourable outcomes (died and defaulted) decreased from 33% to 25%. More patients in the CD4<100 cells/µl category started ART during TB treatment in 2009 (94% vs. 78%, p<0.001), and the median time between TB treatment and initiation of ART decreased from 103 days to 45 days. Patient's data from 2007 (n=346, before integrated service started) were compared with data from 2009 (n=366, after integration). In addition more patients were diagnosed with extra-

pulmonary TB and started on ART in 2009. The results of this supported integration of TB and HIV services (Hermans SM *et al*., 2012). Similar conclusions were reached by studies in Cape Town which found that a non-integrated HIV-TB service was associated with a delay in initiating ART in co-infected patients (Kerschberger B *et al*., 2012; Lawn SD *et al*., 2011b).

With small sample size studies it is possible to collect data manually from non-integrated HIV and TB services and to combine the data in a research database. Often these research facilities are equipped for such a data collection process with specific data collection forms since principal investigators are guiding the process for research purposes. For a province wide programme surveillance dataset, like the Free State province, collecting HIV and TB information from more than 200 facilities spread across the province was not feasible and another approach had to be taken.

1.4. Data linkage

Access to accurate, complete, timely patient - and organisational data is needed for research on current health systems services and problems (Roos LL *et al*., 1985; Qauad MG *et al*., 2009). Health datasets that would provide insight to answering questions like these are often collected by diverse groups (researchers, analysts) in the public and private sector. It takes skilled researches months or years to obtain and link these datasets to extract meaningful information (Bradley CJ *et al*., 2010). However, should it be possible to link these datasets, it provides a unique source to perform epidemiological studies, which would otherwise not be possible with single source data (Tromp M *et al*., 2010). Weaknesses in the recording, reporting and accurate measurement of HIV and TB control programmes have limited the use of observational data for research purposes due to the separate treatment and data collection systems.

This data can be greatly enriched by linking this reporting system with other sources of population data, making it suitable for interrogation to address important questions facing decision-makers in terms of treatment and prevention strategies and priorities and resource allocation, (Johnson WD *et al*., 2008; Heunis C *et al*., 2009).

Electronic data linkage has been in use in a myriad of fields over the last 40 years. The need for complete and longitudinal datasets on patients to make better informed decisions to improve quality of care has been a critical factor in healthcare (Rogot M *et al*., 1983; Zingmond DS *et al*., 2004; Parker DP *et al*., 2008). Furthermore, by using previously collected data it saves time,

money, and the burden of collecting data once more, therefore driving the need for linkage strategies (Bradley CJ *et al*., 2010).

Two methods are available for data linkage, deterministic and probabilistic linkage. Initially deterministic linkage was more common, however with increased computing power probabilistic matching gained popularity. In deterministic record linkage a predefined subset of linking variables have to agree to consider a pair as a link. With probabilistic record linkage weights for agreement (rewards) or disagreement (penalty) are estimated for each variable based on the difference in probability that a variable agrees among matches and non-matches (Tromp M *et al*., 2010).

A link is established when two or more records in a set of databases represent the same individual and could be combined to form one record in a new database. False-positive links are records incorrectly linked, false-negatives are difficult to determine; these refer to records that should have been matched but they were not matched. A true match is a correctly matched record. The aim is to maximise true matches and minimise false matches. There is always a trade-off to consider, to obtain more matches it is likely that there will be more mismatches (Hser YI *et al*., 2008).

While trials have since provided compelling evidence supporting early ART for co-infected patients (Naido K *et al*., 2013), they represent small patient populations recruited from research-augmented programmes, and so findings may not be generalisable to large public sector populations where ART and TB treatment are provided at scale. However, observational cohort data can be used to support or refute RCT findings in these populations, provided that appropriate statistical methods are used to minimise potential biases associated with observational data (Franke MF *et al*. 2011; Lawn SD *et al*. 2011). This is particularly true of programmes followed for several years during which policy and clinician enthusiasm for treating co-infected patients with ART changed, permitting comparisons to be made between patients who received treatment earlier or later, or not at all.

In the Free State Province of South Africa systems were in place to support and manage large scale routine TB and HIV programme data. However, these datasets were not linked and could consequently not being used to evaluate the scale of the problem of co-infection in the Free State province, to describe the heterogeneity of the epidemic, to identify populations at greatest risk of co-infection, establish the factors driving the co-epidemic, and the impact of effective treatment

on the prognosis of these patients. This study has created a unique linkage that provides the opportunity to interrogate the dataset to evaluate the effectiveness of ART over a long term in a public health service setting, in preventing TB in HIV positive patients, reducing the risk of mortality in these co-infected patients, and the impact thereof on TB treatment outcomes. It is estimated that ART's effectiveness is overestimated in smaller controlled settings, therefore this study aims to evaluate the effect of ART, compared to patients not receiving ART, in a heterogeneous setting with a large sample size, appropriately adjusted for confounding to reduce the effect of bias and to provide an epidemiological evidence base for the formulation of evidence-informed target and prevention strategies.

First I will describe the process of linking the HIV and TB programme datasets to establish a data platform from which sample populations were identified to evaluate the effect of ART on incidence of TB in this HIV positive population, how ART affected the risk of mortality in this HIV-TB co-infected population and their outcomes.

2. METHODS

The purpose of this study was to evaluate the effectiveness of ART on TB incidence and outcomes in co-infected adults in the Free State Province's public sector. The ART and TB treatment programmes were managed separately, therefore it was important to link the Free State public sector HIV programme database with the provincial TB register to obtain an all-inclusive dataset to support aims listed below.

2.1. Aims

1. Estimate the effect of ART on incidence of TB disease (time to first TB episode[1], and rate of recurrent TB episodes) in HIV positive patients.
2. Estimate the effect of ART and cotrimoxazole on mortality (time to death) in co-infected HIV-TB patients.
3. Estimate the effect of ART on TB treatment outcomes in co-infected HIV-TB patients.

2.2. Study population

The study population was all HIV positive people, 16 years and older, enrolled in the Free State Comprehensive Care, Management and Treatment of HIV/AIDS Programme from the start of the programme in May 2004 until July 2010, when final extracts of the relevant datasets were drawn and closed for analyses. This HIV cohort was linked to additional data sources to enrich the dataset and reduce the amount of missing baseline and outcome data. Patients between the ages of 16 and 18 were included since they received standard adult treatments. Patient populations are defined separately for each analysis.

2.3. Setting and data collection

2.3.1. HIV Treatment Programme

The Free State Department of Health opened their first ART facilities in May 2004 after the South African government made ART available in the public sector. At the start of the programme, ART was provided at 25 facilities, five facilities per district, and scaled up during

[1] A TB episode was defined as the day when TB treatment was started until the date it was completed, stopped (transferred), or the patient became lost to follow-up (defaulter). Due to a lack of TB diagnosis date it was assumed that TB treatment started soon after TB diagnosis or on the same date.

consecutive years such that by the close of the study healthcare was provided at 21 assessment clinics, 12 treatment sites and 13 combined assessment/treatment sites throughout the five districts in the Province. Assessment sites (clinics) were staffed by nurses while doctors were mostly based at treatment sites (hospitals); this required the patient to travel between sites since different services were provided at each site. Some facilities operated as a combined site when nurses and doctors (permanent or visiting) were based at the same premises (Fairall LR *et al*. 2008, Ingle SM *et al*. 2010). See Figure 9 for a map of the Free State Province with districts and ART facilities.

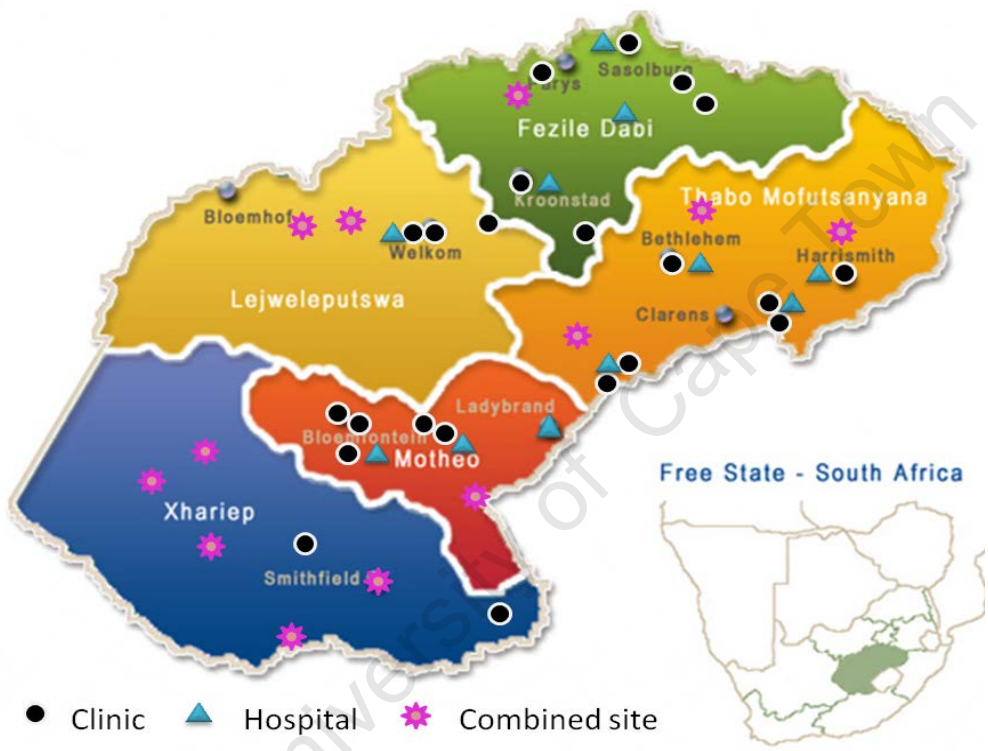

Figure 9: Map of Free State Province displaying clinic (assessment sites), hospitals (treatment sites) and combined treatment-assessment sites at the close of the study, July 2010.

Patients testing positive for HIV at any primary care facility were sent to an ART facility, typically a clinic (assessment site) and triaged according to their CD4 cell count. Patients with a CD4 cell count above 500 were monitored annually. When the HIV programme was initiated in 2004, patients with a CD4 cell count between 200 and 500 were monitored six monthly and others with a CD4 cell count below 200 or clinical AIDS were referred to a doctor at a hospital (treatment site), or later to a doctor visiting rural assessment sites on a sessional basis. Eligible patients would then undergo a treatment readiness program for three weeks before the first line ART regimen comprising stavudine (d4T), lamivudine (3TC), and efavirenz or nevirapine was initiated. Patients receiving ART were monitored on a monthly basis at assessment sites (nurse

visit) and six monthly doctor visits at treatment sites. Monthly monitoring included weight measurement, family planning, screening for TB and side effects, and monitoring ART adherence. Blood samples were taken every six months to measure a patient's CD4 cell count and viral load (HIV viral RNA copies/ml) (South African National Department of Health (NDOH) National Antiretroviral Treatment Guidelines; 2004). Second line treatment was reserved for patients who failed first line treatment and comprised zidovudine (AZT), didanosine (ddI) and lopinovir.

Clinical policy regarding the CD4 cell count eligibility criteria for ART and timing of initiation in co-infected HIV-TB patients evolved as the programme expanded, WHO (World Health Organisation) revised guidelines in 2009 and more data became available, suggesting a risk benefit ratio in favour of starting ART earlier (Table 4). At the outset of the programme patients were started on ART when their CD4 cell count was 200 cells/µl or less whereas currently patients are started when their CD4 cell count is below 350 cells/µl. Table 4 displays ART eligibility criteria policies implemented by the Free State during the study period (May 2004 to July 2010).

Table 4: Eligibility criteria for starting ART and in patients diagnosed with TB. Source: PALSA PLUS [a] guidelines implemented in the Free State Province based on policy guidelines published by the Free State – and National Department of Health.

Year	Eligibility for ART (adult criteria only)	Timing of ART when patient has TB (adult criteria)
2004	Adults and pregnant women: CD4 [b] ≤ 200 and/or WHO [c] Stage 4	Complete standard TB treatment (6 months) before starting ART.
2006	Pregnant women: CD4 ≤ 350 and/or WHO Stage 4 Men and non-pregnant women: CD4 ≤ 200 and/or WHO Stage 4	Dependent on CD4 cell count measured during TB episode: CD4 < 50: Start ART as soon as possible. CD4 50-200: Defer ART until the end of the intensive phase of TB treatment (2 months). CD4 > 200: Repeat CD4 test and re-evaluate eligibility for ART at the end of TB treatment.
2009	Unchanged	Start ART within 2 to 8 weeks of starting TB treatment, and as early as possible in those with CD4 ≤ 50.
2010	TB and pregnant women: CD4 ≤ 350 and/or Stage 4 Men and non-pregnant women: CD4 ≤ 200 and/or Stage 4	Unchanged
2011	CD4 ≤ 350 and/or Stage 4 (regardless of whether pregnant or TB)	Unchanged

a PALSA PLUS: Practical Approach to Lung Health and HIV/AIDS in South Africa

b CD4: CD4 cell count in cells/µl

c WHO Staging

At the start of the programme in 2004, guidelines recommended that ART be deferred in co-infected TB patients until TB treatment was completed. By 2006, co-infected patients with a CD4≤50 cells/µl were referred for urgent ART work-up. More recently doctors have tended to start ART patients with very low CD4 counts or who are clinically ill, as soon as they have been stabilised on TB treatment. On World AIDS Day 2009, the South African Department of Health expanded eligibility criteria to include pregnant women or those with TB and a CD4 cell count between 201 and 350 cells/µl (South African National Department of Health Government speech 2009). In 2011, this was again expanded to include all HIV positive patients with a CD4 cell count of 350 cells per µl or less (South African National Department of Health, 2011).

2.3.2. HIV Programme data collection

Nine structured paper clinical records were designed for clinicians and nurses to record consultations before and after ART was started (see Appendix B for HIV Programme clinical visit forms). Figure 10 presents a schematic diagram of HIV programme visits. There was a form for each type of visit for example voluntary counselling and testing (VCT) visit, drug readiness training visit etc. For patients on ART only two forms were used: a nurse follow-up and a doctor follow-up form. Data clerks were appointed to capture completed clinical records into the Province's computerised patient information system. At some facilities, additional computers were installed for health care personnel to capture visit information directly on the system during a consultation. Captured forms (visit data) were uploaded to a central database server in Bloemfontein once a week.

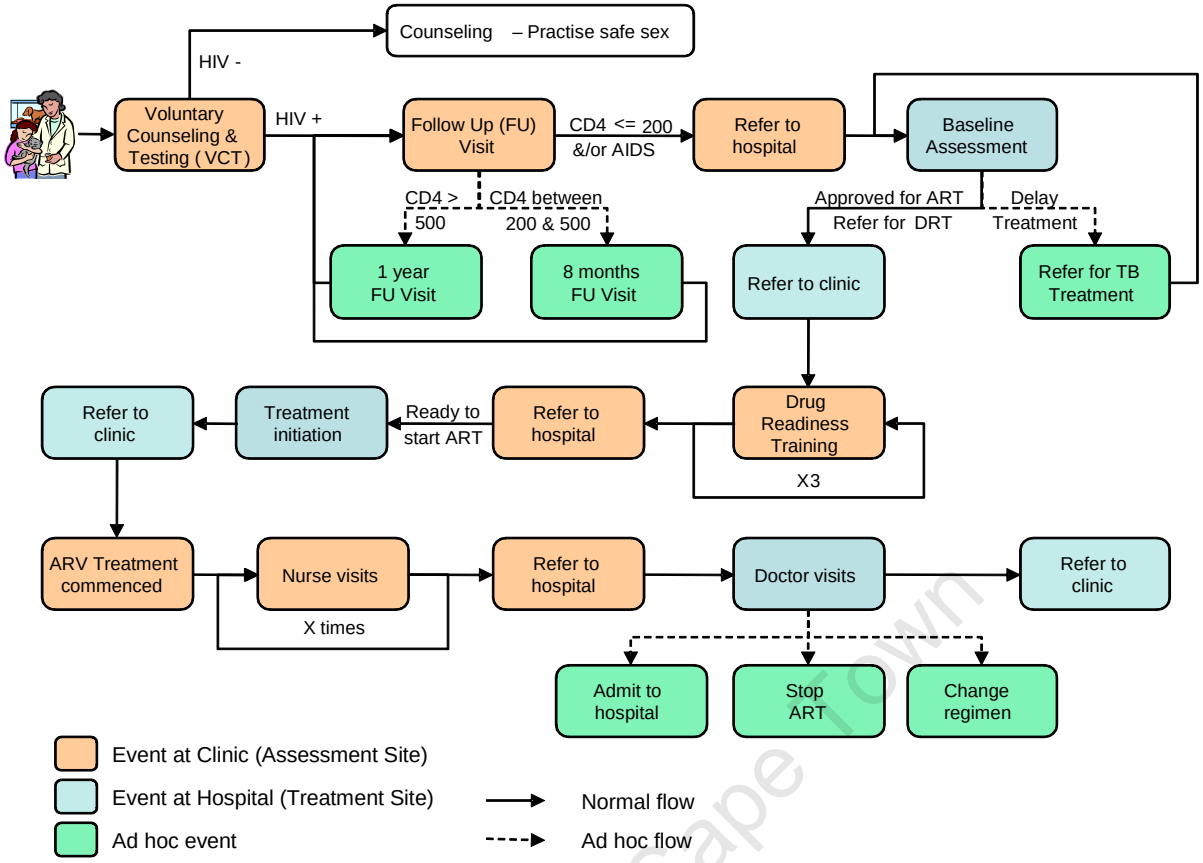

Figure 10: HIV programme visits in the Free State Province.

Visit information was saved in a database (data warehouse) combining information from the different facilities, which was designed to facilitate reporting, run quality routines for data cleaning, support monitoring and evaluation of the HIV programme, and longitudinal data analysis of the HIV programme before and after ART initiation (Fairall LF *et al*, 2008; Louwagie G *et al*. 2007; Kotze JE *et al*, 2010). The HIV database and data loading processes were developed by a team (of which the candidate was part) at the Medical Research Council during 2004-2005 and handed over to the Free State Department of Health (DOH) in December 2005. A Microsoft SQL (Structured Query Language) Server 2000™ was used to deploy the HIV database on in the Free State. The Free State DOH information technology manager incorporated the HIV database in a data warehouse at the Department, which included human resource data, hospitalisation data, laboratory data, and death population registry data. The Free State Department of Health has linked hospital and laboratory data monthly with the HIV programme data since June 2009 and incorporated death registry information from the Department of Home Affairs.

The death linkage was done monthly by the Department of Home Affairs based on a list of South African identification numbers (SA ID) exported from the HIV programme database. A date of death was provided by the Department of Home Affairs based on the SA ID match and imported back into the data warehouse. The cause of death was not available. It is estimated that the national death population register captures over 90% of adult deaths in South Africa (Statstics SA, 2003; Joubert G et al., 2007). ID numbers were considered to be valid if the DOB corresponded with the first six digits of the ID number and if the ID number appeared on the Department of Home Affairs' database. A check was implemented at the point of data entry at facilities in the Free State to alert data clerks or healthcare workers if the ID number appeared to be incorrect based on the algorithm used by Department of Home Affairs to create ID numbers or if a patient with the same ID number existed on the system. A concerted effort was made by the Free State Department of Health and clinics to obtain ID numbers for patients to improve linkage with the population register and support monitoring of the programme (Kotze JE et al, 2007).

The Free State HIV database was unique since it contained visit data, both before and after ART was started for all ART facilities in the Province. However, TB information collected by the HIV programme was limited, and included TB status at the time of each visit (Figure 11). This recorded whether the patient was receiving TB treatment, had symptoms suggestive of TB, or was neither on treatment nor symptomatic. Categories were mutually exclusive. TB information not recorded by the HIV programme included TB treatment start date, TB classification (pulmonary or extra-pulmonary), culture results, sputum conversion information or TB treatment outcomes.

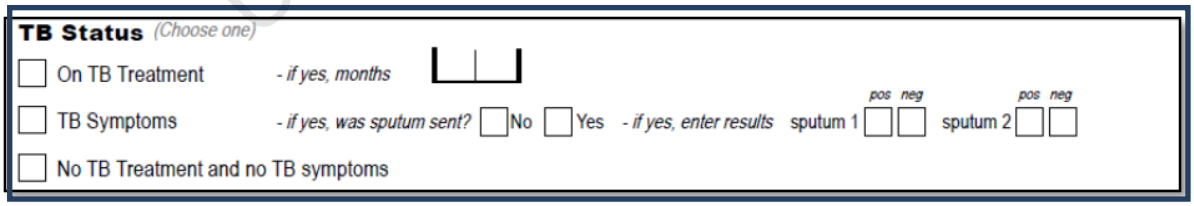

Figure 11: Example of TB information collected during HIV programme visits.

2.3.3. TB Programme

Patients were referred to TB facilities from primary care clinics or ART facilities. There were 328 facilities providing TB care at the close of the study, 91 of the facilities were mobile facilities and 237 fixed band facilities located within the primary care clinic. The WHO and

South African National TB programme defined clinical algorithms for the diagnosis of TB (SA National Dept of Health, 2007,2008; WHO 2008). Diagnosis of pulmonary TB is confirmed by a positive sputum smear for acid fast bacilli while smear negative patients are diagnosed by a medical doctor considering all available diagnostic information. Treatment is divided in two phases, a two to three month intensive phase followed by a four month continuation phase. A new TB treatment episode is treated with a fixed combination of -rifampicin, isoniazid, pyrazinamide, and ethambutol (RHZE) during the intensive phase, and rifampicin and isoniazid during the continuation phase. Depending on the sputum smear outcome after two months of treatment, intensive phase treatment is continued for an additional month (smear positive) or the patient can start with continuation phase drugs, (rifampicin and isoniazid), for four to seven months (National TB Strategic Plan, 2007-2011).

TB treatment outcomes are recorded based on WHO definitions. A patient is *cured* when sputum converted from an initial positive result to smear negative in the last month of treatment and at least one previous occasion; *Failed* is when a patient was initially smear positive and remained smear positive five months later; *Completed* is when patient a completed treatment but did not meet the criteria for cure or failure; *Defaulted* is when a patient's treatment was interrupted for two consecutive months or more; S*uccess* is a combination of cure and completed. Other outcomes include death, moved, transferred out, or not evaluated. TB episodes are also classified as a new or retreatment case. An episode was classified as *new* if the patient never had treatment for TB before or received TB treatment for less than one month. A *Retreatment* case was when a patient received at least one month or more TB treatment in the past or when a patient was returning after default or retreated after treatment failure (WHO treatment guidelines 2008, 2009, 2011).

2.3.4. TB Programme data collection

TB treatment information was recorded on a paper-based TB register at each TB clinic. Information was collected on four copies. The first copy was submitted to the sub-district office when the TB initiation section was completed for each record on the page or at the end of a quarter. A data capturer (or health care worker) captured the data into a standalone (not connected to a central server database) electronic TB register (ETR) at sub-district level. The ETR software created an encrypted dispatch file with newly added and/or updated data, which was sent to a district manager who collated and integrated all the sub-district data in a district database. The same process was repeated to collate data from district to provincial level and

then from provincial to national level. Baseline characteristics recorded at diagnosis included age, sex, smear results (positive, negative or no smear), new or retreatment case, and anatomical site of disease (pulmonary, extra-pulmonary or both). From January 2007, SA ID numbers and date of birth (DOB) were entered on the ETR, however the data are still limited (Table 5 and Table 6). From a total of 111 930 records in the ETR 2007-2010, only 3.5% (3 954 / 111 930) records contained a SA ID number with 13 characters, which could be used for linkage. Only 1.1 % (1 194 / 111 930) of TB records with a valid SA ID number were linked to HIV database patients.

Table 5: ID numbers recorded in ETR between 2007 and 2010.

Year TB registered	Number of Records N	Valid ID number (13 characters) n (%)	Linked to HIV database n (%)
2007	28 715	1 (0%)	0%
2008	32 299	921 (2.9%)	214 (0.7%)
2009	32 985	2 036 (6.2%)	670 (2.03%)
2010	17 931	996 (5.6%)	310 (1.7%)
	111 930	3 954 (3.5%)	1 194 (1.1%)

Overall, DOBs were entered for 35.1% (70 220 / 199 998) of records in the ETR. Initially only 4.2% of records had a DOB in 2007 but it increased to 98.7% by June 2010 when the data extract were exported (Table 6).

Table 6: DOBs entered in ETR between 2007 and 2010.

Year TB registered	Number of Records N	DOB entered n (%)
2007	28 715	1 198 (4.2%)
2008	32 299	19 160 (59.3%)
2009	32 985	32 161 (97.5%)
2010	17 931	17 701 (98.7%)
	111 930	70 220 (62.7%)

Since 2003 the HIV status of a TB patient was recorded in the ETR (Table 7). However, the completeness of this variable ranged between 25% in 2003 to 46% in 2009 and the percentage of records with a HIV positive status ranged between 2.4% in 2003 and 3.8% in 2009, suggesting substantial under-reporting.

Table 7: HIV status entered in ETR between 2003 and 2010.

Year TB registered	Number of Records N	HIV status recorded n (%)	HIV positive n (%)
2003	16 395	4 111 (25.1%)	394 (2.4%)
2004	21 246	6 982 (32.8%)	661 (3.1%)
2005	23 288	7 042 (30.2%)	727 (3.1%)
2006	27 139	4 903 (18.1%)	504 (1.9%)
2007	28 715	5 713 (19.9%)	831 (2.9%)
2008	32 299	15 159 (46.9%)	1 239 (3.8%)
2009	32 985	15 058 (45.7%)	1 256 (3.8%)
2010	17 931	5 466 (30.5%)	666 (3.7%)
	199 998	64 434 (32.2%)	6 278 (3.1%)

The ETR software was expanded in 2003 to capture more HIV related information. Table 8 shows the HIV related variables, the total number of records registered per year and how many records contained data entered for the specific variables. Given that public sector ART was implemented in May 2004, either HIV related information was entered more regularly since 2008 or the clinical profile of TB patients started to change, explaining the difference in the completeness of the variables for example 26% of records had a CD4 cell count entered in 2008 compared to <1% before 2008 and 46% of records had an indication of a patient receiving cotrimoxazole therapy in 2008 whereas less than 10% of records recorded cotrimoxazole therapy it in the preceding years.

Table 8: HIV related information entered in ETR per year of TB registration.

Year TB reg.	Number of records	CD4 between 1 and 2 000	On cotrimoxazole treatment (Yes,No, Positive, Negative, Unknown) [a]	HIV test date	On ART (Yes,No,Not eligible) *	On ART at start of TB (Yes)
2003	16 395	0 (0%)	878 (5.4%)	3 (0.02%)	1 (0.01%)	4 (0.02%)
2004	21 246	0 (0%)	1013 (4.8%)	15 (0.07%)	7 (0.03%)	0 (0%)
2005	23 288	1 (0%)	1325 (5.7%)	264 (1.1%)	212 (0.9%)	7 (0.03%)
2006	27 139	7 (0.03%)	1771 (6.5%)	837 (3.1%)	624 (2.3%)	9 (0.03%)
2007	28 715	23 (0.08%)	2712 (9.4%)	1729 (6.02%)	724 (2.5%)	95 (0.3%)
2008	32 299	3 172 (9.8%)	14780 (45.8%)	56 (0.2%)	6329 (19.6%)	1231 (3.8%)
2009	32 985	8 719 (26.4%)	24016 (72.8%)	0 (0%)	11961 (36.3%)	1161 (3.5%)
2010	17 931	5 333 (29.7%)	11509 (64.2%)	0 (0%)	7021 (39.2%)	992 (5.5%)
	199 998	17 255 (8.6%)	58 004 (29.0%)	2 904 (1.5%)	26 879 (13.4%)	3 499 (1.8%)

a Answers recorded in database displayed in brackets.

The WHO recommended a reporting format as a key management tool to evaluate effectiveness of a national TB control programme. This reporting format is based on cohort reporting, for example reporting will be for patients who were diagnosed with TB during the first quarter of the year and that particular group's (cohort's) outcomes, the second quarter's patients and their outcomes etc. This approach does not emphasise or report specifically on patients with multiple TB episodes. Earlier versions of the ETR were also not designed to monitor patients with multiple TB episodes. Reporting was structured around episode outcomes, not per patient outcomes due to the lack of a unique identifier. However, multiple TB episode entries per patient do exist in the provincial ETR, either from multiple treatment episodes at the same facility or from episode information from different facilities. The ETR database was designed such that each record in the TB register reflects one treatment episode, and a patient can have more than one record (TB episode). The inclusion of more personal identifiers recently will allow for identification of multiple TB episodes per patient in future.

2.3.5. Data security and patient confidentiality

Data security and privacy for the HIV programme and TB programme data are managed by the Free State Department of Health (DOH) with secure network connections and security policies. Password protected TB data files were obtained by the candidate on three occasions with permission (Dec 2007, Jun 2008, Aug 2010). Research copies of the HIV database were either collected in person at the Department of Health in the Free State or copied to a digital versatile disk (DVD) and sent by registered post. The HIV database was encrypted with a public-private key cryptography method, which uses a pair of keys for encryption and decryption. A public key, which encrypts data, and a corresponding private, or secret key for decryption. Only the candidate and the Information Technology manager from the Free State DOH had access to these encryption keys. The databases are hosted on a secure server at the University of Cape Town Lung Institute. The candidate is the only person with administrator rights for the server hosting the databases and the SQL databases containing the data. Database backups are created each night and stored on the server. Backups from the server are copied to an external drive bi-monthly and stored off-site in a secure location. Personal identifiers were included in the data sources since they were required for the linkage process. Only anonymised data (name and SA ID numbers removed) were shared with supervisors, and all facility information was coded as categorical variables. Only the candidate had full access to personal identifiers from the HIV programme and TB register data.

2.4. Data linkage of Free State public sector ART Programme database with provincial TB register

The provincial databases for the ART and TB programmes existed in parallel, but no substantive attempt had been made to link patient information between the two databases before this thesis. Completeness of variable rates of person identifiers, particularly in the ETR, proved challenging and required that the candidate explore a range of methods for data linkage, described below.

2.4.1. Types of data linkage

Deterministic data linkage

With deterministic linkage a single or set of variables is used in a series of steps to link two datasets and it requires an exact match on the variable(s) to consider it a match. It does not draw on a statistical method. A common identifier between two datasets in order to link them for example social security number (USA) or national identification number (South Africa) is ideal and provides the best match. However, these are not always available. Other variables used include name, surname, sex, DOB, addresses, area codes and any other discriminating variables pertaining to the datasets (Qayad MG *et al*., 2009). Rules are developed to specify which variables and in which combination the variables need to match exactly before a linkage is made. For example, if six variables are available (name, surname, DOB, gender, facility, and street address) and three out of the six variables should be matched before a linkage is made, it is important to specify which three variables (combination) must match. An exact match on surname, name, and DOB could be considered a better match than gender, DOB and facility. Experimenting with rule-variable combinations is necessary to identify the best rules for a particular dataset.

Deterministic linkage is particularly useful when the scrutinised outcome is very critical and incorrectly linked outcomes could cause problems, for example legal issues, fatalities, treatment outcomes, and blood results. It provides a high level of specificity, but at the cost of sensitivity (Hser YI *et al*., 2008; Tromp M *et al*., 2010). Data errors or missing data will hamper deterministic linkage and the magnitude of under-linkage depends heavily on the accuracy of the linking variables (Winkler WE, 2006; Qayad MG *et al*., 2009).

In addition to linking different datasets, computerised record linking is very useful for identifying duplicates within a dataset. Often the majority of work (estimated at 90%) in data warehouses is associated with eliminating (cleaning) duplicates (Fayad U *et al*., 2000) which is costly (Thornton SN, 2005). Identifying duplicates in such a way is often referred to as one-file matching where the same file is linked to itself to identify the duplicates (Kotze E *et al*., 2010). Duplicate records may also put the safety and wellbeing of patients at risk due to information captured under a different record (patient name mistyped for instance), which is not available at the time of decision making (Stiell A *et al*., 2003).

Probabilistic data linkage

Newcombe and Kennedy (1962) introduced probabilistic record linkage concept and in 1969 Fellegi and Sunter formalised a mathematical method which is the basis for probabilistic matching. A mathematical criterion or probability is used to separate matches from non-matches (Felligi IF *et al*., 1969).

Statistical methods have been developed to improve the accuracy and efficiency of the original work specifically to compensate for data quality problems by adjusting weights or scores used for deciding whether a pair of records is a match (Jaro MA, 1989 & 1995; Winkler WE, 1993). In order to decide on probabilities *a priori* knowledge of the variables is helpful, i.e. distribution characteristics and information gathered from working with these variables. In the absence of knowledge about the dataset and variables, arbitrary values are chosen and with trial-and-error better estimate for probabilities and thresholds could be developed. This process can significantly increase the energy, time and cost expenditure (Hser YI *et al*., 2008).

Matching relies heavily on computerised routines; however human judgment should never be completely omitted for either method of linkage. Matching records is not always an unambiguous choice between a match and a non-match. It is often the combination of available data at hand and human judgement which helps to make the decision during review processes (Hser YI *et al*., 2008; Bradley CJ *et al*., 2010).

In South Africa, some attempts at linkage have been made to strengthen reporting systems in the government-run public health service. For example, the Free State Department of Health implemented a computerised data collection system in all public sector ART facilities and set up a system that links provincial ART programme data with the national death population register

using national ID numbers (Fairall LR *et al*. 2008). Similar linkages with a national population register have been set up at some sentinel and research sites (Boulle A *et al*. 2010, Manoshuthi W *et al*. 2006a; Van Rie A *et al*. 2011). The Free State Department of Health has also linked their HIV programme data with admission data from government-run hospitals in the province and with the state laboratories to improve data availability for clinical decision making and reporting purposes (Kotze JE *et al*, 2007).

For research purposes of this thesis the candidate linked the provincial HIV database and ETR using Microsoft SQL server 2005™ for data cleaning and linkage of the two datasets. A database copy of the Free State HIV database was provided to the candidate after integration of laboratory, hospitalisation and death registry data and the provincial TB manager extracted data from the TB register in excel format for the purpose of this project.

2.4.2. Overview of basic steps for deterministic data linkage

Before any linkage could start a comprehensive data cleaning effort was needed to create variables that would be useful for the data linkage process and to increase specificity. These tasks included removal of non-alphabetical characters, standardising the format of names, and consolidating age data using multiple sources (for example DOB, age at registration) to improve comparability. SQL server's text and string manipulation functions were studied to understand the programmability potential of SQL server to define rigid and more flexible rules for linkage. A rule matrix (cross tab) was designed to identify rules which could be programmed using a combination of linking variables. Reviewing of linked output results was crucial to develop a sense of the accuracy of a linking rule and its feasibility, as well as reviewing multiple linkages for the same person. Overlapping TB episodes, generated by combining TB treatment information from both the ART and TB programmes, then had to be reviewed creating a coherent history of TB episodes when multiple episodes were identified. The steps are now described in detail in sections 2.4.3 to 2.4.7.

2.4.3. Data preparation prior to linkage

The provincial TB manager of the Free State Province provided data in EXCEL spreadsheet format, extracted from the TB register, for each year from 2003 to 2010. Older versions of EXCEL have a maximum row set of 65 569 resulting in some loss of data during data exports,

hence each year was extracted separately and merged into one table (TB register source file) with SQL server.

The final HIV dataset was updated with mortality data on 3 Aug 2010 and the latest TB treatment date registered on the ETR was 13 June 2010. The death population registry database was current at the time of linkage with the most recent death date reported in the extract being 2 Aug 2010 on an extract provided by the Department of Home Affairs on the following day. Table 9 summarises the key variables used for the linkage, and the availability of these in each dataset.

Table 9: Completeness of variables in HIV programme database and ETR database.

Variable	HIV database N (%)	ETR N (%)
Total no of records in database	97 476	199 998
Full name	78 338 (80.0%)	198 681 (99.3%)
Surname	97 445 (99.9%)	199 977 (99.9%)
DOB	97 430 (99.9%)	70 220 (35.1%)
Sex	97 476 (100.0%)	199 998 (100.0%)
Age at enrolment	97 476 (100.0%)	199 998 (100.0%)
SA ID number	76 226 (78.0%)	3 954 (0.02%)
Facility	97 455 (99.9%)	199 998 (100.0%)
District	97 455 (99.9%)	199 998 (100.0%)
Is the patient on TB treatment?	12 336 (12.6%)	199 997 (100.0%)
Address (more than 4 digits long)	93 766 (96.0%)	176586 (88.3%)
HIV status	97 476 (100.0%)	64 434 (32.2%)
ART status:		
Recorded (Yes/No)	97 467 (100.0%)	26 879 (13.4%)
On ART treatment (Yes)	52 119 (54.0%)	4621 (17.0%)
ART start date available	52 119 (100.0%)	277 (0.06%)

Personal identifiers like name, surname, sex, facility, age and address were fairly complete and could be used for linkage. However, DOB, South African ID number and HIV information were less complete on the TB register. Linkage then proceeded as follows:

a) Preparation of source datasets..

- Copies of the datasets were made before any changes/updates were performed. A copy of the variables, which were used during the linking process, was made in each dataset (name, surname, district, DOB, ID number, facility, and district).

- A unique row number was created for each record in the TB register database. A unique patient identifier (ZU number) existed in the HIV database for each record.
- Before linkage was performed it was necessary to make sure that the data types of the variables to be compared were of the same type for example ID number was a float in one set and a text variable in the other. The float variable was converted to a text variable in this instance.
- Additional columns were added to the HIV and TB dataset to use as markers (cross referencing) to indicate which records have been matched.

b) Identification and merging of duplicate patients.

Duplicate patients (same person but with two unique identifiers) in the HIV database were identified and merged to obtain one unique identifier. The duplicates were identified by sorting the HIV dataset on surname and DOB and counting patients with the same surname and DOB. In cases where more than one record were found, entries were manually reviewed to determine if it was the same patient based on additional information (name, address, facility, gender, ID number). Occasionally data capturers in the HIV programme registered a patient twice, creating two unique patient identifiers for the same patient. Either the ID number was mistyped or an ID number was not entered for one of the patient registrations. The HIV programme data collection software used ID numbers to warn the data capturer if they were about to register a duplicate patient. For data linkage (and reporting) purposes, it was necessary to identify these duplicate patient identifiers and to eliminate them. One unique identifier (study number) was selected and used to overwrite the second (or sometimes a third duplicate) study number such that all visit information and blood result records could be linked to the chosen primary identifier in the database.

c) Cleaning of name variables.

Names and surnames needed extensive cleaning for the linkage rules to perform best since deterministic linkage is based on exact matching. Variables had to be spelled exactly the same before a computer regarded it as an exact match. Database matches are very specific; a double space between two words (“Joan Smith”) would not be matched with the same name with a single space between the words (“Joan Smith”).

Name variables were cleaned as follows:

1. Copies of FIRSTNAME and SURNAME columns were created before any changes were made to the data variables themselves.
2. Spaces or atypical characters at the beginning of each name or surname were removed.
3. Zeros “0” were replaced with “o” (alphabet letter o)
4. Numbers or atypical characters within names or surnames were removed.
5. Names with a length of three or less characters: full stops were removed and consecutive initials were created.
6. Full stops in names with a length of six or less and with two or more full stops in it were cleaned.
7. Commas were replaced with a space.
8. Surnames repeating within the NAME field were removed, for example, SURNAME: ‘Smith’ NAME: ‘Joe Smith’ > NAME: ‘Joe’.
9. Spaces at the beginning of a name (first character) were removed.
10. Variables where surnames were removed and the comma replaced by a space (step 7 and 8), the extra space had to be removed, for example;
    ‘Smith, Joe’ > ‘Smith Joe’ (step 7)
    > ‘ Joe’ after step 8 (surname removed)
    > ‘Joe’ after repeating step 9 twice
11. Extra spaces between names were removed for example ‘Monku A T’ > ‘Monku AT’
12. A list of distinct names and surnames were prepared for manual review, to ensure all atypical characters and spaces were removed and to determine if any rules had to be repeated.
13. Spaces and full stops were removed from initials columns.
14. A surname count variable was added to each dataset to display the frequency of the surname in the dataset.

d) Cleaning of facility information

Records on the ETR had different ways of referring to the same facility for example “FS BATHO CLINIC” and “BATHO CLINIC”. All facility names were standardised by removing FS prefixes and duplicate names were merged to one name if there were small spelling differences. The HIV database’s facility and district names were standardised during the data loading process therefore no cleaning was necessary for it and each facility had a unique site-ID.

Two extra columns were added to the TB table, DISTRICT and SITE_ID. Districts were assigned to all the facilities in the TB dataset and where it was possible to map a TB facility to one with an ART service (TB and ART facility on the same premises or TB and ART facilities servicing the same town), the same site-ID from an ART facility was assigned to the TB facility.

e) Calculating patient age

As of 2007, the TB programme collected DOB on their electronic system. Before this, only the age of the patient at the start of TB treatment had been recorded. For each of these records the patient's age at the time of creating the dataset was added by subtracting two dates, the year in which the TB dataset was extracted (2010) less the year when TB treatment was started, and added to the age at registration. Figure 12 displays two examples of how this variable was calculated.

TBRegID (unique row identifier)	**AGE (at registration)**	**TB treatment date**	**CURRENT_AGE (new calculated field)**
131	30	2004/03/12	36
132	28	2007/02/10	31

Figure 12: Example of variables and how CURRENT_AGE was calculated.

This calculated variable was compared with an age calculated from documented DOB in the HIV programme database. A difference of one to two years between the two variables were allowed (Age ± 1 or Age ± 2) during the matching process.

2.4.4. Software used to execute linkage

Microsoft SQL server™ has standard build-in functions which can be used with database programming. Text string manipulation functions were used with the programming of the linkage rules. An easy match is when first names match exactly like "Marie" = "Marie". In some cases two first names were recorded on a dataset, "Marie Liouse" and "Liouse". Some linking rules were programmed to search for the word "Liouse" in "Marie Liouse". The same principle was applied to search for a piece of an address (first 6 letters) within the address field of the other dataset. A small typing error would result in a mismatch for example "Liouse" is not the same as "Louise" therefore a SQL soundex function was used that would identify "Liouse" and "Louise" as a match.

Initials were inconsistently recorded, for example “LR” or “RL”. A SQL function was used to swap the initials from one of the variables and to test it for an exact match with the other variable. The month and day of a DOB were often swapped between datasets. The month and day of each DOB variable were extracted separately and compared with the other date parts of the DOB variable, for example 1956-03-09 and 1956-09-03.

A frequency count of each surname in the TB and HIV dataset was calculated and stored in a surname count variable in the separate datasets. The surname count variable was used during the manual review process. If there was any doubt regarding a linkage, and the surname was a very popular surname with a high surname count, for example “Mokoena” with 3 081 repeats of the surname on the HIV database, and 2008 repeats on the TB dataset, the linkage would be discarded. However, if it was a scarce surname like “Matsi” with only three appearances on the HIV database and five appearances on the TB dataset, it would be used in favour of the linkage after considering the other matching variables.

2.4.5. Linkage and rule matrix

A cross-tab/rule matrix (Table 10) was developed during pilot linkage projects of the datasets to identify “linking rules” based on combinations of variables from the two separate datasets, which could be used to match patients. The table was also used to keep a record of rules which have been programmed. Table 10 displays variables used for linkage and a unique rule number assigned to each combination of variables used for the linkage and rules. The unique rule number in a cell represents the combination of variables intersecting in that particular cell and if any additional restrictions besides the variable combinations were used. Colour coding and indentation indicate if the matching criteria applies to variables on a level below. All cells below the dark orange “DOB the same” also matched on DOB unless otherwise indicated in the cell. All of the records have an exact match on sex or surname (basic condition). For example, Rule 1-1a matched on the same surname, sex, DOB, facility and first name. An additional rule was used in some cases called “HIV TB”, which means a patient was identified with TB on the HIV database. This criteria was used during the linkage process when name variables did not match perfectly, but there was an indication in the HIV programme that the patient was undergoing TB treatment, making the linkage more likely.

Table 10: Rule matrix used during linkage process.

Basic Condition: Exact match on Surname and Sex		**Same first name**	**Name sounds alike**	**Matching initials (either one or both of the first names are missing)**	
	Rule Number			One initial	Two or more initials
DOB the same					
Same Facility	1	1a	1g		
Same District	1	1h- HIV TB 1n- Day and month swap in DOB	1m-HIV TB	1k - HIV TB & TB pat must have ID number	
Not same District	1	1h- HIV TB 1n- Day and month swap in DOB	1m- HIV TB		
Age +- 1					
Same Facility	1	1b	1q	1i	1c 1e-TB patient must have ID number
Same District	1		1p- Soundex (Surname) and Part of address match		1d 1j- IDs match
Not same District	1				
Age +- 2					
Same Facility	2	1b	1q	1i	1c 1e-TB pat must have ID number
Same District	2		1p- Soundex (Surname) and Part of address match		1d 1j- IDs match
Not same District	2				

These rules were programmed with SQL. It was recorded on both the HIV and TB dataset if a patient or TB record was linked; the unique identifier from each source was copied to the other source (cross referencing). After a record from the ETR was matched and linked to a patient on the HIV database, it was excluded from further matching. Therefore, it was important to execute rules with the best possibility of identifying a match between a patient on the HIV database with a patient on the TB database first, for example linking on exactly the same first name, surname, DOB, ID number and facility was a definite match. If a very certain match was found, there was no need to try and link the record with less strict criteria, therefore, it was removed from the set

to match. Once a match was found for an HIV database patient, the patient was not excluded from further matching since the patient could be matched with a different rule to a second/third TB record.

A database table (TB-Linkages) was created to keep record of HIV database patients and TB records being matched as well as all the variables from both sources used for the linkage and the particular linkage rule. Each rule had a unique number such that a particular linkage could be traced back to the rule in case it was found later that a specific rule did not link patients accurately enough. All the records linked (linkages) with such a rule could then be reversed/cancelled. Table 11 displays a few variables and fictitious data to illustrate the TB-linkages table.

Table 11: Example of some variables in TB-Linkages table after execution of linkage rules.

TBRegID	HIV_SUBJECT	Rule	Use	HIV_NAME	TB_NAME	HIV_SURNAME	TB_SURNAME	HIV CURRENT_AGE	TBCURRENT_AGE
235	ZU00023	2-1a	Y	Venessa	Vanessa	Smith	Smit	29	30
1025	ZU00204	1-1d	Y	Ben	Benjamin	Coetzee	Coetzee	35	35
120158	ZU00999	1-1h	N	John	Johanna	Mokoena	Mokoena	29	36
21025	ZU00023	2-1b	Y	Venessa	Venessa	Smith	Smit	30	30

2.4.6. Review of linkage output

Record (patient) linkages were reviewed manually after each rule was executed to ensure that the linkages were accurate. A binary variable (USE) was added to the TB-Linkages table and used to identify/indicate a mismatch and to "undo" the link by removing the cross referencing variables on the TB Register and HIV table.

If it was noted that a rule returned ambiguous matches, the rule was discarded, linkages from the particular rule were deleted from the TB-Linkages table and the cross referencing on the separate data sources were removed.

Patients known to have TB according to the HIV programme, but who were not yet linked to a record on the TB register, were manually searched for on the ETR after the programmed rules were executed. This process helped to identify more rules which could be programmed and identified matches that had too many variances for a deterministic computer linkage method. Typing errors in name, surname, ID number and sex (for the same person) would not result in a computer linkage but using human judgement the patient could be linked.

The rules recognised that a patient may have multiple TB records or episodes. Each time a patient enrolled for a TB treatment episode lasting between six and nine months, a record was entered in the ETR. Therefore more than one TB record can be linked to an HIV patient due to the multiple TB episodes, see HIV_SUBJECT = ZU00023 in Table 11. Additional information from the TB register such as the TB treatment start date and treatment end date as well as TB outcomes (died, failed, defaulted etc.) were often used to clarify/confirm multiple linkages to the same subject and to determine if the multiple episodes made logical sense. There should not be an overlap between TB episode start dates. A patient can only start a second episode after he/she completed the first episode. Also, the first record/episode may not have "Died" as an outcome and a second TB episode recorded a year later with a different outcome (cured). If there was any uncertainty regarding a linkage, the linkage was discarded.

Table 12 displays a summary of the programmed rules, how many TB records were linked by the rule and how many TB linkages were removed after the manual review process. Only Rule group 1 is displayed in Table 12. Rule group 2 used the same combinations of matching variables except for age ± 2 years. In total 52 rules were programmed.

Table 12: Linkage results

Rule group	Rule No.	No. linked	No. Discarded
1	1a	2205	10
1	1b	4027	130
1	1c	227	2
1	1d	541	
1	1da	68	
1	1db	357	1
1	1e	43	
1	1Ga	723	
1	1Gb	634	1
1	1Gc	52	

Rule group	Rule No.	No. linked	No. Discarded
1	1Gd	209	2
1	1Ge	94	
1	1Gf	1340	3
1	1Gg	802	7
1	1Gh	630	6
1	1Gi	833	27
1	1Gj	1713	47
1	1H	130	
1	1I	63	
1	1j	23	
1	1k	31	1
1	1L	17	
1	1M	49	
1	1Ma	67	
1	1N	4	
1	1P	291	

After the linkage process was completed, 14.2% (24 279 / 199 998) records of the TB register records were linked to an HIV database patient. The reviewing process discarded 281 linkages, resulting in 20.4% (18 438 / 90 367) of HIV database patients linked to a record in the TB register, containing TB treatment episode data between 2003 and 2010. Seventy three percent (73%, 8 730 / 11 896) of HIV patients known to have TB, as recorded by the HIV programme, were linked to the TB register, while the other 27% (3166 / 11 896) of HIV patients could not be linked to any record on the ETR after programmed rules and manual searches. This likely indicates under-reporting of TB cases in the ETR, as reported by other studies (Van Hest NAH. *et al*., 2006; Melosini L *et al.*, 2012) but may reflect a sensitivity of 73% for the record linkage process An additional 9 708 HIV patients were identified as co-infected with TB, after the linkage to the ETR, which was not recorded at any HIV programme visit. In total, 24% (18 438 + 3 166) / 90 367) of HIV patients were known to have TB, either recorded by ETR (45%, 9 708 / 21 604), HIV database (14.6%, 3 166 / 21 604), or both (40.4%, 8 730 / 21 604). See Figure 13 for a diagram representing the linkage results.

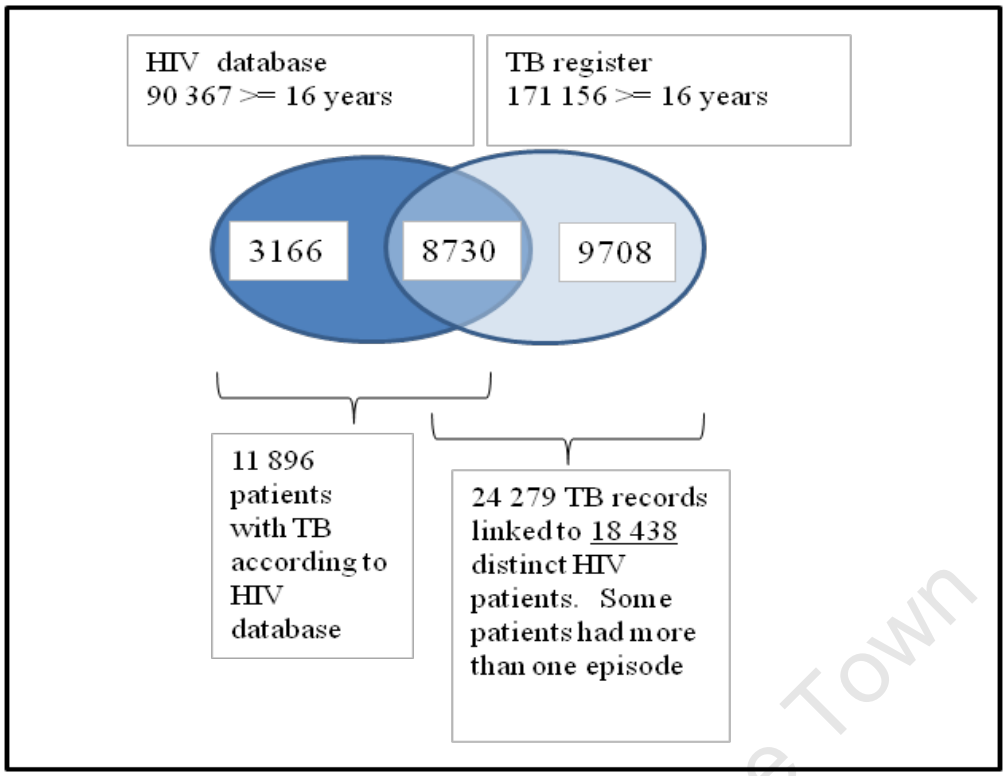

Figure 13: TB linkage results.

TB episodes were then arranged logically, especially for the 8 730 patients with TB information from both the HIV programme (HIV-TB) and electronic TB register (ETR-TB). This was done to make sure that the recorded HIV-TB episode occurred during the same time period the ETR has a recorded TB episode for the patient. If not, the HIV-TB episode could be a different TB episode not matched or recorded on the TB register. An "estimated TB episode" had to be used for patients with HIV programme TB information since no TB treatment start or end date was available. Section 2.4.7 describes how these HIV-TB episodes dates were created.

2.4.7. Refinement of TB episode data

TB data collected in the HIV programme was limited (Figure 11). This information was not accurate enough to calculate an exact TB start date, therefore an estimated TB start date was calculated from the HIV programme visit data. HIV programme visit records, indicating that the patient was on TB treatment, were retrieved and sorted by patient and visit date (consultation date) as indicated on the form. The first date in the sequence was selected, 12 months were added to it and recorded as a TB episode. The first record with TB information after this 12 month period was taken as the start date for a new TB episode etc. See Figure 14 for a schematic representation of the process described. Each blue cross on the timeline represents an

HIV visit form where "On TB treatment" was selected by the healthcare worker. Although a standard TB episode is usually between 6-9 months, a 12-month period was used for an HIV programme TB episode[1].

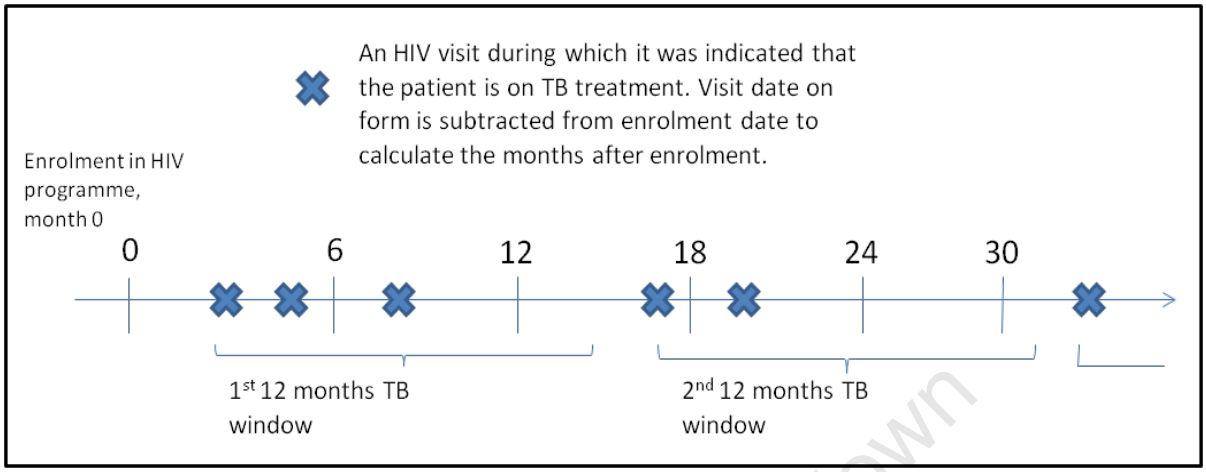

Figure 14: A schematic representation to indicate how HIV programme TB episodes were calculated.

Linked TB register data were also cleaned to identify episodes to be used for analyses. On occasion it was found that two episodes were registered for a patient within three months of each other. The first episode's outcome would be moved/not evaluated and the second episode would be cured/completed/died. The first episode would be discarded and the second one would be used. If the TB start date of second episode was within three months of the first episode, the earliest start date would be used to replace the start date of the episode selected.

The ETR episodes (ETR-TB) and HIV programme TB episodes (HIV-TB) were combined after being cleaned separately. Records were sorted on TB episode start date and numbered consecutively from '1' to 'N' per patient. If any HIV-TB treatment start date was within a six month period of an ETR-TB treatment start date, the HIV-TB episode was discarded, since the TB treatment episode dates in the ETR were more accurately recorded (Figure 15). After the HIV-TB episodes were removed, the data were sorted and numbered again. TB episodes within six months of the first episode would get the same episode number and were reviewed manually to decide if it was referring to the same episode or if it was a new episode. If it was evident from

[1] Only 100 patients were recorded as having multi-drug resistance TB (MDR TB) on ETR. The maximum treatment duration recorded for these patients in the ETR was 9 months and a failed treatment outcome was recorded for all of them.

the data that the patient was treated for a second time, the episode number was changed to '2', for example;

TB Start: 15 Feb 2009, TB Stop: 8 Aug 2009, TB outcome: Completed > Episode 1.

TB Start: 10 Aug 2009, TB Stop: 10 Feb 2010, TB outcome Died > Episode 2

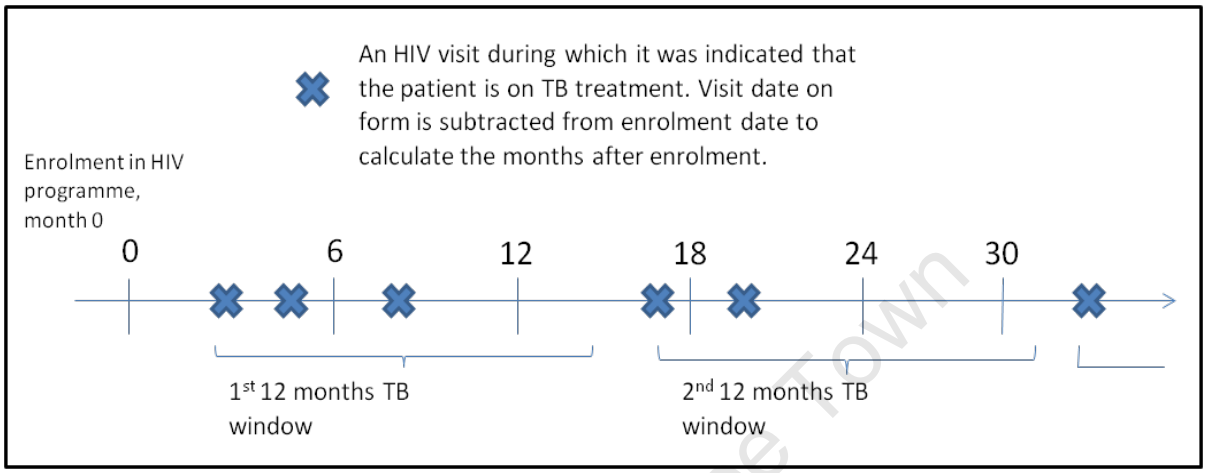

Figure 15: Diagram indicating how TB episodes were selected after ETR and MT TB merge.

Patients with more than 4 episodes were reviewed manually to confirm that the linkages were correct and any incorrect linkages were removed. After the data cleaning and TB episode selections were completed, the final record set was created which was used as the starting dataset for study populations in section 2.5 (Aim 1), 2.6 (Aim 2), and 2.7 (Aim 3).

2.5. Effect of ART on incidence of TB disease in HIV-positive patients (Aim 1)

This analyses was designed to include all of the HIV database patients (≥16 years) who were HIV-positive and enrolled between 2004 and 2010. This provided the opportunity to evaluate the effect of ART on incidence of TB disease (time to first episode of TB) for two groups of patients in this high risk cohort, those who did not receive ART by the end of follow-up and those who had. It was also possible to evaluate the effect of ART on multiple episodes of TB and the effect of the linkage with the ETR.

2.5.1. Specific aims

1. To estimate TB treatment episode incidence rates for the overall group, and crude incidence rate ratios comparing follow-up periods during which patients received ART with follow-up periods during which patients did not receive ART.
2. To estimate the effect of initiating ART compared with not initiating ART on time until the first episode of TB treatment, taking into account the competing risk of death.
3. To perform a sensitivity analysis estimating the effect of initiating ART compared with not initiating ART on time until the first episode of TB, but limited to TB information collected by the HIV programme, thereby assessing the effect of the linkage with the ETR on results.
4. To estimate the effect of receiving ART on incidence rates of multiple TB episodes, i.e. ART initiated at different time points (before, during or after TB treatment episodes)

2.5.2. Study population

This study was conducted among all HIV positive adults enrolled in the Free State Province HIV programme at public sector facilities between May 2004 and June 2010. This dataset included patients who have and have not received ART as well as patients with and without TB. TB case information was determined either from the ETR using standard treatment definitions (section 2.3.3) or from TB information routinely collected during HIV programme visits (section 2.3.2).

Inclusion criteria:

- Patients enrolled in HIV programme between May 2004 and June 2010.
- Patients 16 years or older at enrolment in the HIV programme.

Exclusion criteria:

- HIV negative patients
- Patients starting ART before public sector HIV programme rollout.
- Patients without a valid SA ID number and one visit to the HIV programme in whom it was thus not possible to link to the population register to evaluate time at risk of TB.
- Patients with no visit information recorded on the HIV programme.

- Patients with discrepancies in the sequencing of key dates
- Patients who died on the date of enrolling in the HIV programme, and thus had not follow-up time to contribute to the analysis.

2.5.3. Design

This was a retrospective cohort study to estimate the effect of ART on incident TB in Free State public sector HIV programme patients.

2.5.4. Data variables

Individual patient level data from the HIV programme data were analysed. TB patients were identified from either HIV programme (HIV-TB) or from ETR-TB (section 2.4.7). Patients with incident TB were those with a TB treatment start date after their enrolment in the HIV programme. History of TB was defined as TB episodes before enrolling in the HIV programme or being on TB treatment at the time of enrolment (prevalent TB).

Follow-up time, also referred to as person-time or time at risk, was divided in two records, follow-up time during the period of no ART treatment (no-ART) or also known as pre-ART, and follow-up time when the patient received ART (ART). The no-ART period was from enrolment in the HIV programme until one day before ART was initiated and the ART period was from the date of ART initiation to the censor date. Patients on ART contributed person-time to both no-ART and ART records. Follow-up time (time at risk for TB) was defined as time from HIV programme enrolment to the first TB diagnosis or censoring event (date) and this follow-up time was divided in a no-ART and ART record.

Follow-up time was censored as follows:

a) Death date for patients who died.
b) Patients with a valid South African ID number (including lost to follow up patients) and not known to be dead when the last linkage with the death population register was done (2010/08/03), were assumed to be alive on this date, and so have the date of linkage (2010/08/03) as their censor date[1]

[1] The death population register was considered to be comprehensive and up to date. The extract included deceased patients with death dates as recent as the week before the date when the death extract was prepared

c) For patients without a valid South African ID number and not known to be dead, the date of their most recent visit to any health facility for example hospital visit, laboratory test date, HIV programme visit or TB treatment completion date was used.

Patients who were lost to follow up to the ART programme might have moved, had their care transferred or had a visit at other healthcare facilities after their last ART programme visit. Linkage of the different health system data sources provided more follow-up information (time) than using only the ART programme database. In the absence of a valid South African ID number, using their last health system contact date gave them maximum follow-up time.

Outcome variable

The outcome was a diagnosis of TB either recorded during an HIV programme visit (HIV-TB) or from linkage to the electronic TB register (ETR-TB). A binary variable (TB) was used to indicate if a patient had TB during the no-ART or ART record.

Explanatory variables

The primary explanatory variable for incidence of TB was whether or not the patient received ART by the end of follow-up. Other explanatory variables included in the regression models were history of TB at enrolment, sex, age, and district. A binary variable (PREPOSTART) was created to classify no-ART and ART follow-up time records. For no-ART records, PREPOSTART was set to '0' and for ART records PREPOSTART was set to '1'. Age was modelled as cubic splines with knots at the 25th, 50th, and 75th percentiles to optimise adjustment for confounding without assuming a linear relationship (Fairall LR *et al*., 2008). The respective cut-points were 29, 36 and 43 years.

A binary variable (DIED) was coded for use in the competing risks analysis. In this analysis the assumption was that TB and death compete with each other for the outcome. This is simply because patients who die cannot be diagnosed with TB. DIED was set to '1' if a patient died during the no-ART or ART record and had no TB prior to death.

CD4 cell count, weight and cotrimoxazole prescription were also included in the regression models.

---

by the Department of Home Affairs (2010/08/03). The extracts were done on a monthly basis for the Free State Department of Health.

A CD4 cell count and weight documented closest to enrolment in the HIV programme were used for no-ART records, and a CD4 cell count and weight documented closest to the date when ART was started were used for ART records. The no-ART CD4 cell count and weight values were from the period six months before the patient enrolled in the HIV programme and six months after the enrolment date but closest to the enrolment date. A six month window before and after enrolment in the HIV programme was used since a CD4 cell count was a key factor for referring the patient to the HIV programme and for monitoring the patient's status before ART could be initiated. For ART records, values between six months before ART was started and fourteen days thereafter were searched to identify a value closest to the date when ART was started.

Categorical variables for CD4 cell count and weight were created for no-ART and ART records. CD4 cell count categories were CD4 0-49, CD4 50-99, CD4 100-199, CD4 200-349, and CD4 350-2000 cells/µl. The 25%, 50%, and 75% percentiles were used to define the weight categories. Weight categories were thus 0-49 kg, 50-56 kg, 57-65 kg, and 66-160 kg for this population.

Cotrimoxazole prescription was coded as a binary variable (COTRI). It was assumed that once cotrimoxazole was prescribed, the patient remained on it and the COTRI variable was set to '1'. If no record of cotrimoxazole prescription was found, it was assumed that the patient did not receive cotrimoxazole and the variable was set to '0'. In the case of patients with two records, if a patient was on cotrimoxazole during no-ART follow-up time, COTRI was set to '1' for both the no-ART and ART record. However, if cotrimoxazole was only prescribed during ART follow-up time, COTRI was set to 0 for the no-ART record and '1' for the ART record.

2.5.5. Statistical analysis

Baseline demographics were characterised using standard descriptive statistics. Categorical variables were described as number (%) and continuous variables as mean (standard deviation) or median (inter quartile range). Binary outcome variables were compared with Pearson chi-squared test and continuous outcomes with hierarchical linear regression. A complex survey design was used since the data had a multi-level structure. Outcomes tend to be concentrated within clinics within districts. It was assumed that patients who started ART remained on therapy through to the end of follow-up.

TB incidence rate was defined as the number of TB cases occurring per 100 person-years observation at risk for either the no-ART group, on ART group or overall.

When multiple TB episodes are taken into account, person-time accrued during TB episodes (either prevalent TB at enrolment or during TB treatment after enrolment) was excluded from the denominator (person time at risk) when TB incidence rates were calculated. Figure 16 gives a schematic representation of how follow-up time was counted; the red lines on the time line was added for no-ART follow-up time, and the green lines represent follow-up time counted during ART. Unadjusted TB incidence rates and incidence rate ratios were calculated under a Poisson distribution and expressed as number of events per 100 person-years.

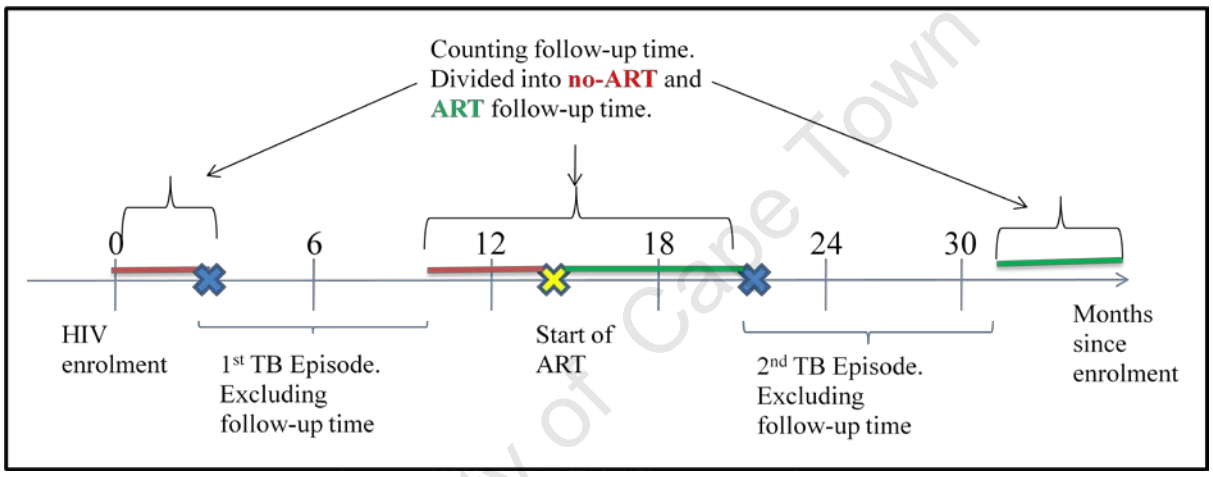

Figure 16: Schematic representation of how follow-up time was counted and divided between no-ART (pre-ART) and ART follow-up for analysis of multiple TB episodes.

Censoring was described under section 2.5.4. For analyses of time to first TB episode, each person could contribute a maximum of one case.

The primary analysis, a competing risks proportional hazards regression model was used to estimate adjusted associations of patient characteristics with rates of incident TB (Fine JP et al. 1999). Time from enrolment to TB diagnosis was analysed, with death as a competing risk. Sub distribution hazards (SHR) were estimated for these competing events and can be interpreted similarly to hazard ratios estimated in standard Cox models, but account for the hazard of the competing event (Ingle SM et al. 2010b). Andersen and colleagues (2012) summarise the key feature of competing risks regression as being "that the one-to-one correspondence between cause-specific hazard and cumulative incidence, between rate and risk, is lost. This fact has two important implications. First, the naive Kaplan–Meier that takes the competing events as censored observations, is biased. Secondly, the way in which covariates are associated with the

cause-specific hazards may not coincide with the way these covariates are associated with the cumulative incidence." (Andersen PK et al., 2012). A Marginal Structural Model (MSM) could not adequately account for the competing risk of death, that is, they could not distinguish between censorship due to death, censorship due to incident TB and censorship due to end of follow-up for other reasons.

The ART facility variable was excluded from competing risks model due to the large dataset and the impact of the facility variable on computing time and performance of the statistical software. Also cotrimoxazole prescription was excluded from the model since we know that cotrimoxazole does not prevent TB, although it prevents death and is highly correlated with ART prescription (Fairall LR et al., 2008; Grimwade K et al., 2005). Multiple imputation was used to impute missing CD4 cell counts and weights (Royston P, 2005).

Stata offers procedures for creating and analysing multiple imputed datasets for incomplete multivariate data. Multiple imputation provides a useful strategy for dealing with datasets with missing values. Instead of filling in a single value for each missing value, Rubin's (1987) multiple imputation procedure replaces each missing value with a set of plausible values that represent the uncertainty about the right value to impute. These multiple imputed datasets are then analysed by using standard procedures for complete data and combining the results from these analysis. No matter which complete-data analysis is used, the process of combining results from different imputed datasets is essentially the same. This results in statistically valid inferences that properly reflect the uncertainty due to missing values.

Initiation of ART is based on a CD4 count (threshold) and CD4 counts improve while being on ART. An interaction term was included in models to test if the effectiveness of ART was different at different CD4 levels. Wald tests were used to determine if the interaction was significant. Sub-hazard ratios for the interaction term factors were estimated by exponentiation of coefficients from the linear combinations of coefficients (Stata help file: lincom). Also, the effect of ART versus no-ART prescription on TB incidence was evaluated within each of the CD4 strata (primary and secondary models).

Kaplan Meier curves were used to characterize the failure function to first TB episode. A secondary analysis for time to first TB was prepared with a Cox proportional hazards regression model, which estimates hazard ratios and 95% confidence intervals for risk factors for incident TB.

A Poisson regression model, which is used to model count data and contingency tables, was used to evaluate the effect of ART on multiple TB episodes (counts).

For all models included in this analyses, a crude model was prepared as well as the adjusted model with and without the interaction term.

2.5.6. Sensitivity analysis

A sensitivity analysis using Cox proportional hazards model, restricted to TB information obtained solely from the HIV programme, was used to investigate the effect of ETR linkage on findings.

A summary of the planned statistical analyses is presented in Table 13

Table 13. Summary of statistical analyses for evaluating the effect of ART on TB incidence.

Investigate	Model used	Effect estimate
Baseline demographics: a) Descriptive statistics (Number (%), mean(SD), median) b) Testing for trend over time	a) Chi-squared test; Hierarchical linear regression b) Extension of Wilcoxon rank sum test	p-value
Baseline demographics	a) Descriptive statistics b) TB incidence rate c) TB incidence rate ratio	Number (%), mean(SD), median; IR, IRR
Primary analysis for effect of ART on time to first TB episode taking into account competing risk of death. (TB from HIV programme and ETR.)	a) Crude competing risks model. No imputation. b) Adjusted competing risks model. Imputed data [10 sets].	SHR
Secondary analysis for effect of ART on time to first TB episode. (TB from HIV programme and ETR.)	a) Crude Cox proportional hazards model. No imputation. b) Adjusted Cox proportional hazards model. Imputed data [10 sets]. c) Adjusted Cox proportional hazards model. No imputation.	HR
Sensitivity analysis for effect of ART on time to first TB episode but limited to TB information from HIV programme only	a) Adjusted Cox proportional hazards model. No imputation.	HR

Investigate	Model used	Effect estimate
Secondary analysis to model count data and the effect of ART on multiple TB episodes. (TB from HIV programme and ETR)	a) Crude Poisson regression model. No imputation. b) Primary analysis for effect of ART on incidence rate of multiple TB episodes: Adjusted Poisson regression model. Imputed data [25 sets].	IRR

SHR: Sub-hazard ratio; HR: Hazard ratio; IR: Incidence rate; IRR: Incidence rate ratio

2.6. The effect of ART on mortality (time to death) in co-infected HIV-TB patients (Aim 2)

TB remains a major cause of death in patients infected with HIV and accounted for almost 25% of AIDS-related mortality globally (WHO report, 2011). The effectiveness of ART on mortality was described previously for the Free State HIV programme for all patients enrolled between 2004 and 2005 (Fairall LR *et al*., 2008). The current analysis aimed at understanding the effect of ART in the co-infected HIV-TB Free State population six years after ART was first implemented in the public sector.

2.6.1. Specific aims

1. To evaluate the effect of receiving ART or initiating ART during or after TB treatment compared with not receiving ART, on survival in co-infected HIV-TB patients.
2. To identify risk factors associated with time to death in HIV patients diagnosed with TB.
3. To evaluate the effect of linkage with ETR on survival outcomes.

2.6.2. Study population

All HIV positive adults enrolled in the HIV programme between May 2004 and June 2010 who were treated for TB treatment after enrolment. Co-infected HIV-TB patients were identified from TB treatment status information collected in the HIV programme and linkage to the ETR. ART initiation could have been before, during or after TB treatment, or not at all.

Inclusion criteria:

- Patients 16 years or older of age at the start of TB treatment.

Exclusion criteria:

- HIV negative patients.
- Patients with TB prior to their enrolment in the HIV programme, in whom the date of HIV diagnosis is not known.
- Patients without a valid SA ID and one visit to the HIV programme in whom it was not possible to evaluate survival time (Pettit AC *et al*., 2011).
- Patients who died on the date they enrolled for the HIV programme and so could not contribute follow-up time to the analysis.

2.6.3. Study design

This was a retrospective cohort study comparing survival in co-infected HIV-TB patients on ART and not on ART using observational public sector HIV and TB data. Survival time was from the date when TB treatment was started until death or censoring (see Section 2.6.4).

Two models were used to evaluate survival, a marginal structural model (MSM) and Cox proportional hazards model. In contrast to standard methods, marginal structural regression models appropriately account for selection bias and adjust for time-varying covariates which are simultaneously confounders and intermediate variables (Robins JM et al., 2000; Hernan MA et al., 2000; Cole SR et al., 2003). Selection bias occurs when subjects are selected into the exposure group (ART in this case) according to factors that are part of the outcome (death in this case). CD4 count, weight and AIDS progression were part of the outcomes of treatment leading to death as well as being key variables used to select patients for treatment. Selection bias in estimation of treatment effects is also known as confounding by indication.

When CD4 counts (time-varying confounders) are themselves predicted by ART (previous exposure) they become intermediate variables on the pathway between ART exposure and death. See the directional acyclic graph (DAG) in Figure 17 displaying the confounding problem. MSM helps to solve the problem of should one adjust for the confounder or not (Steenland K, 2013).

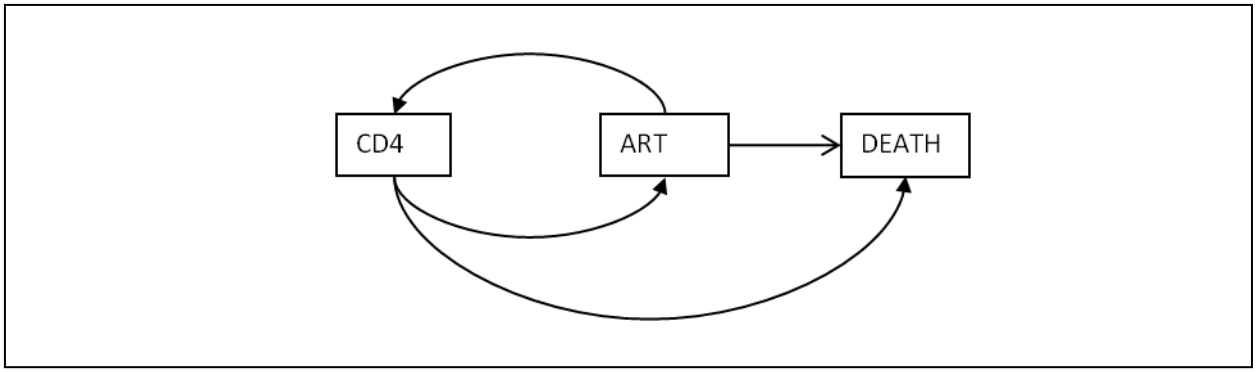

Figure 17: Directional acyclic graph to demonstrate confounding scenario.

"Marginal structural models apply inverse-probability weights based on an 'exposure' model which assesses the probability for each subject that they have received their own exposure and confounder history up to each time t, with the follow-up period divided into T (t=1 to T) categories. These weights are then used in standard regression models (eg, pooled logistic regression models across the categories of follow-up time) assessing treatment/exposure effects on disease, and their use amounts to analysing a pseudo-population where the time-varying confounding is eliminated. Such weights also enable the investigator to account for censoring which may be due to past exposure or confounder history ('informative' censoring)" (Steenland K, 2013).

For the Cox model and marginal structural model, patient follow-up time (survival time) was taken from start of TB treatment to censoring and divided in no-ART follow-up time and ART follow-up time. Three groups of patients were identified:

1. Patients who were treated for TB before ART was started:
   No-ART follow-up time was from the TB treatment start date to one day before ART was started, ART follow-up time was from the date ART was started to the censor date (see Section 2.6.4). This resulted in two records per patient.
2. Patients who were treated for TB after ART was started:
   Follow-up time was from the date when TB treatment started to the censor date (ART follow-up time only). This resulted in one record per patient.
3. Patients who were only treated for TB, and have not received ART by the end of follow-up:
   Follow-up time was from the date when TB treatment was started to the censor date (no-ART follow-up time only). This resulted in one record per patient.

Figure 18 gives a graphical depiction of the three groups of patients and how follow-up time was divided.

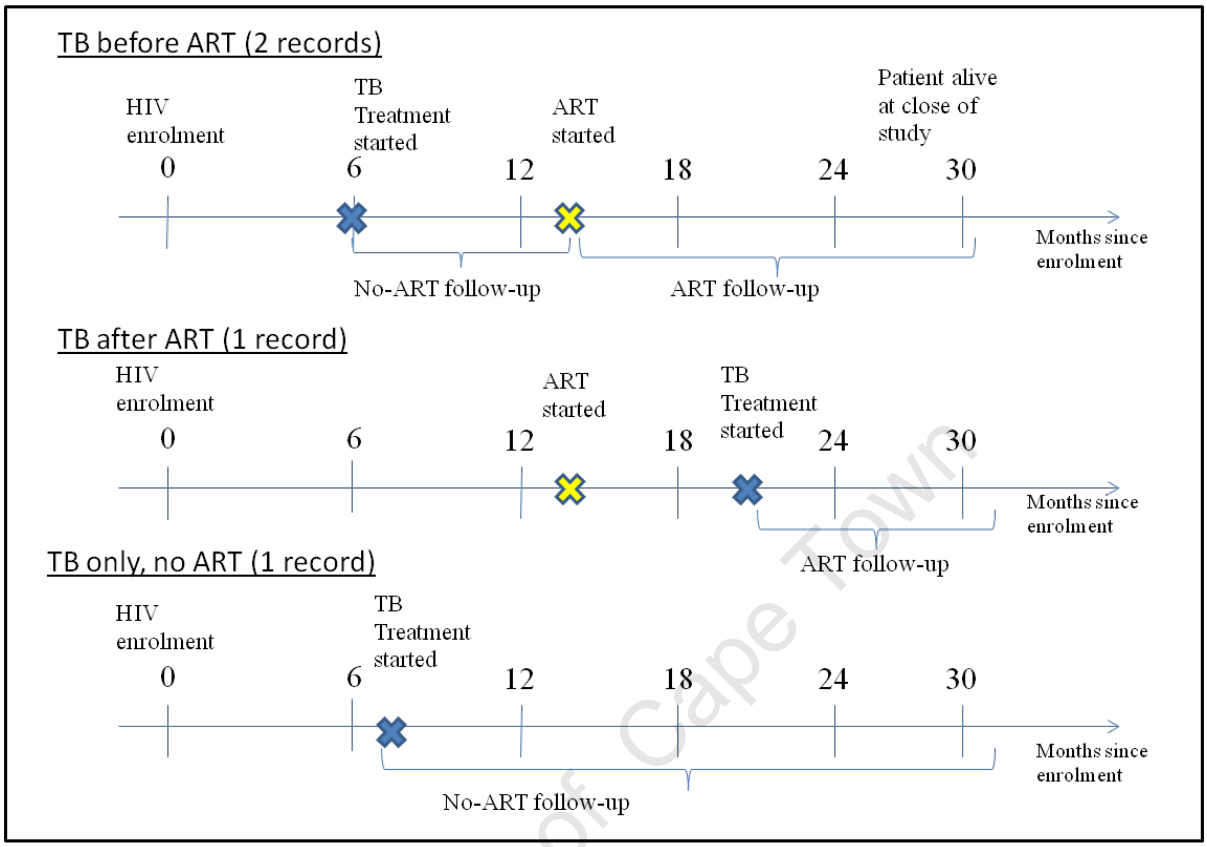

Figure 18: Time lines indicating how no-ART and ART follow-up time was divided.

For marginal structural models the dataset was transformed into a longitudinal format to create one record per month (person-month) per patient. For each patient a person-month record was created between the date when follow-up was started and the censor date. Each monthly record reported on available weight, CD4 cell count, ART and cotrimoxazole prescription for the patient. It was assumed that once a patient received ART or cotrimoxazole therapy, they remained on it. If there were more than one CD4 cell count or weight measurement per month, the first measurement during that month was used. Baseline or latest weights and CD4 values were carried forward to person-months for which they were missing. Thereafter weights and CD4 values that were still missing (because there were no previous measurements) were imputed using multiple imputation by chained equations (Royston P. 2004; Cole SR *et al*., 2003; Fairall LR *et al*., 2008).

2.6.4. Data variables

The data collection process for the HIV programme and TB programme was described in section 2.3.2 and section 2.3.4. respectively. TB patients were identified using both sources of TB information, the HIV programme and ETR. The data variables used for analyses were mostly from the HIV database after linkage to the laboratory, hospital and death population registry data.

Censor date

For patients with two records (no-ART and ART record), the no-ART record was censored one day before ART was started. The ART record was censored at one of the following endpoints/dates:

a) Death date for patients who died.
b) Patients with a valid SA ID number and not known to be dead when the last linkage with the death population register was done (2010/08/03), were assumed to be alive on this date, and so had the date of linkage (2010/08/03) as their censor date.
c) For patients without a valid South African ID number and not known to be dead, the date of their most recent visit to any health facility was used, for example hospital visit, laboratory test date, HIV programme visit or TB treatment completion date.

In patients who did not start ART by the end of follow-up, and so had only no-ART follow-up time (i.e. one record) the record was censored at one of the following dates:

a) Death date for patients who died
b) Patients with a valid SA ID number and not known to be dead when the last linkage with the death population register was done (2010/08/03), were assumed to be alive on this date, and so had the date of linkage (2010/08/03) as their censor date.
c) For patients without a valid South African ID number and not known to be dead, the date of their most recent visit to any health facility was used, for example hospital visit, laboratory test date, HIV programme visit or TB treatment completion date.

Patients who were lost to follow up to the ART programme might have had a visit at other health facilities after their last ART programme visit. Using their last health system contact date in the absence of a valid South African ID number gave them maximum follow-up time**.**

Survival time

Survival time was defined as the time in days (or months) between the date TB treatment was started and the censor date. Months were used for the marginal structural model where one record for each month of follow-up was created and for presentation purposes with the Kaplan Meier curve.

Outcome variable

The outcome was whether or not the patient had died.

A binary variable (DIED) was created to capture vital status for patients at the censor date of each no-ART and ART record. If a patient was deceased at the censor date, DIED was set to '1', otherwise DIED was set to '0'. For patients who died during the ART period and contributed follow-up time to no-ART and ART records, DIED was set to '0' for the no-ART record, while DIED was set to '1' for the ART record. Figure 19 gives a graphical representation of how follow-up time was divided between no-ART and ART and how the outcome variable was coded for each particular record.

Explanatory variables

The primary explanatory variable for mortality was whether or not the patient received ART. Other explanatory variables included in the model were follow-up time in days (survival time), sex, age, history of TB and district.

A binary variable (PREPOSTART) was created to identify no-ART and ART follow-up records. For no-ART records, PREPOSTART was set to '0', and for ART records PREPOSTART was set to '1'. Age categories were created instead of using age as a continuous variable. Age was modelled as cubic splines with knots at the 25^{th}, 50^{th}, and 75^{th} percentiles to optimise adjustment for confounding without assuming a linear relationship (Fairall LR *et al*., 2008). Age categories thus included 16-29 years, 30-35 years, 36-42 years, and 43 to 100 years.

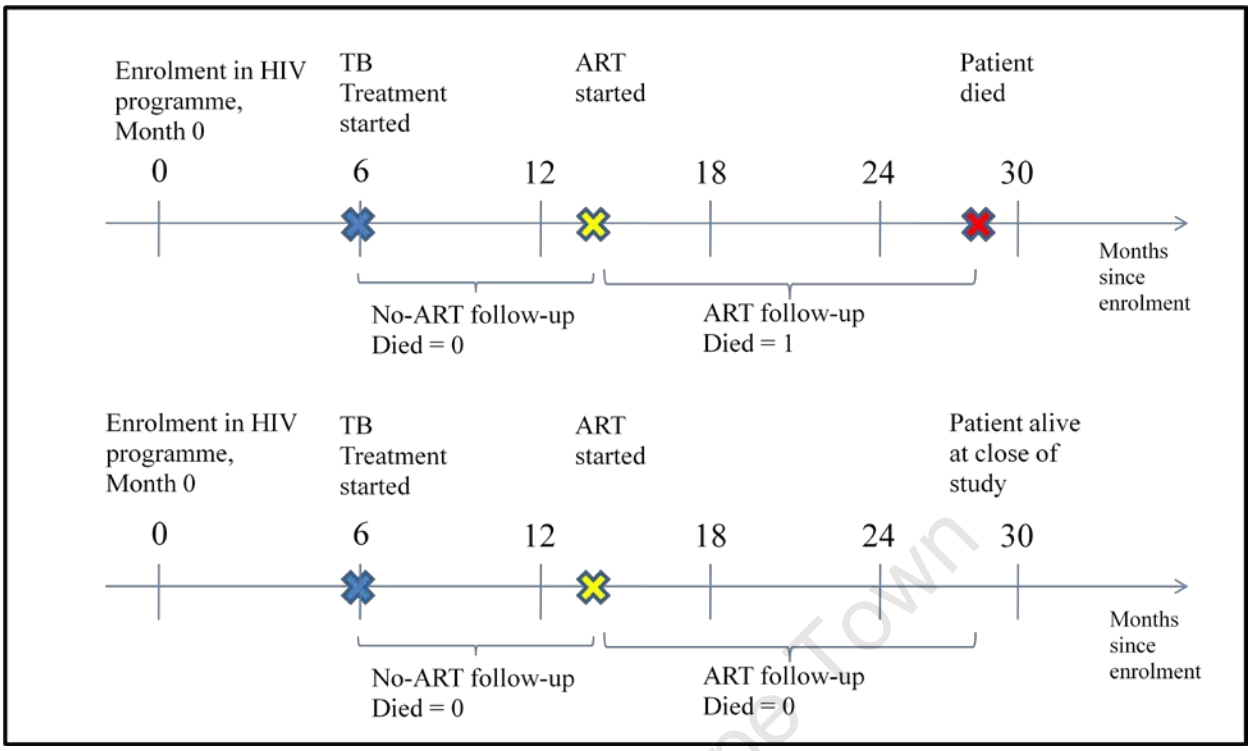

Figure 19: Graphical representation of no-ART and ART follow-up time records for time to death analysis.

CD4 cell count, weight and cotrimoxazole prescription were also included in the models. For no-ART records, a CD4 cell count and weight documented closest to the start of TB treatment were searched for, and for ART records a CD4 cell count and weight closest to the date when ART was started. The no-ART CD4 cell count and weight values included values from the period twelve months before TB treatment was started until three months after TB treatment was started. Where multiple values were available, the value closest to the TB treatment start date was used. For ART records, a narrower window period for CD4 cell counts and weights was used since CD4 cell count and weight increase rapidly after a patient has started ART. Values from the period six months before ART was started until fourteen days after ART was started were searched, and the value closest to the date when ART was started was used in the model. A six month window period was used for the ART record compared to the twelve month window period for the no-ART record. A CD4 cell count was a qualifying criteria mostly for patients to start ART therefore it was more likely to find a CD4 cell count within a six month window period close to the ART start date, whereas with TB treatment the patient could have a CD4 cell count above 200 and only be seen bi-annually for follow-up monitoring in the HIV programme at the time. Therefore a wider window period was chosen for a CD4 cell count closest to the TB start date.

Categorical variables for CD4 cell count and weight were created after the appropriate results were found. CD4 cell count categories were CD4 0-49, CD4 50-99, CD4 100-199, CD4 200-349, and CD4 350-2000 cells/µl. Weight categories were 0-45 kg, 46-52 kg, 53-59 kg, and 60-150 kg. As with age, the 25%, 50%, and 75% percentiles were used to define the weight categories.

Cotrimoxazole prescription was coded as a binary variable (COTRI). It was assumed that once cotrimoxazole was prescribed, the patient remained on it and the COTRI variable was set to '1'. If no record of cotrimoxazole prescription was found, it was assumed that the patient did not receive cotrimoxazole and the variable was set to '0'. In the case of patients with two records, if a patient was on cotrimoxazole during no-ART follow-up time, COTRI was set to '1' for both the no-ART and ART record. However, if cotrimoxazole was only prescribed during the ART follow-up time COTRI was set to '0' for the no-ART record and '1' for the ART record.

Explanatory variables (parameters) used to estimate the MSM weights:

A. MSMs of ART effectiveness

Treatment weight denominator:
Cotrimoxazole prescription at baseline and during follow-up, CD4 at baseline and during follow-up, weight at baseline and during follow-up, TB classification, history of TB, age, sex, district, TB treatment episode number, month of follow-up.

Treatment weight numerator:
Cotrimoxazole prescription at baseline, CD4 at baseline, weight at baseline, TB classification, history of TB, age, sex, district, TB treatment episode number, month of follow-up.

Censorship weight denominator:
Cotrimoxazole prescription at baseline and during follow-up, CD4 at baseline and during follow-up, weight at baseline and during follow-up, TB classification, history of TB, age, sex, district, TB treatment episode number, month of follow-up.

Censorship weight numerator:
Cotrimoxazole prescription at baseline, CD4 at baseline, weight at baseline, TB classification, history of TB, age, sex, district, TB treatment episode number, month of follow-up.

B. MSMs of Cotrimoxazole effectiveness

Treatment weight denominator:
ART prescription at baseline and during follow-up, CD4 at baseline and during follow-up, weight at baseline and during follow-up, TB classification, history of TB, age, sex, district, TB treatment episode number, month of follow-up.

Treatment weight numerator:
ART prescription at baseline, CD4 at baseline, weight at baseline, TB classification, history of TB, age, sex, district, TB treatment episode number, month of follow-up.

Censorship weight denominator:
ART prescription at baseline and during follow-up, CD4 at baseline and during follow-up, weight at baseline and during follow-up, TB classification, history of TB, age, sex, district, TB treatment episode number, month of follow-up.

Censorship weight numerator:
ART prescription at baseline, CD4 at baseline, weight at baseline, TB classification, history of TB, age, sex, district, TB treatment episode number, month of follow-up.

2.6.5. Statistical analysis

Patient characteristics at the start of episodes were tabulated and described as follows. Categorical variables were described as number (%) and continuous variables as mean (standard deviation) or median (inter quartile range). Binary outcome variables were compared with Pearson chi-squared test and continuous outcomes with hierarchical linear regression after the data were adjusted. A complex survey design was used since the data had a multi-level structure. Outcomes tend to be concentrated within clinics within districts. It was assumed that patients who started ART remained on therapy through to the end of follow-up.

Survival of co-infected HIV-TB patients was determined by the method of Kaplan-Meier, considering survival time in months (Kaplan EL *et al*. 1958). A crude Kaplan-Meier test was applied to estimate the probability of death and median time to death after a diagnosis of TB and the log-rank test was used to compare the median time to death between patients with and without ART.

The primary analysis was a marginal structural regression model (MSM) which used longitudinal data to account for selection bias, time-varying covariates and intermediate variables. Analysis was performed at person-month level, using a longitudinal data extract with multiple records per patient (see section 2.6.3). Probability weights were used to adjust for confounding by time-varying covariates which were also treatment outcomes. Two logistic regression models were used to estimate stabilised inverse probability of censoring treatment weights. Both models used censorship as outcome and included baseline covariates and months of follow-up. However, the second model had extra time-varying covariates. Stabilised inverse probability of treatment and censoring weights were produced by combining the weights. The probability weights were used in the final pooled logistic regression model with death as outcome. A Cox proportional hazards regression model was used since subjects were removed from the population at risk after the outcome had occurred (Fairall LF *et al*, 2008). A sensitivity analysis with inverse probability of treatment weights truncated at 1^{st}/99^{th} percentiles or 5^{th}/95^{th} percentiles were done to evaluate the robustness of the MSMs and the treatment weights.

Secondary analyses used Cox proportional hazard models (with and without imputation) to evaluate the independent effect of ART upon patient survival and to identify risk factors associated with survival in co-infected HIV-TB patients. To account for the non-independence of outcomes in patients who received ART and who therefore each have two records, Cox regression analyses included Huber-White robust estimation of errors, using Stata's "cluster" option. The only variables to be considered for inclusion in the final multivariable models were those that were available and considered clinically relevant. The results from the final model were expressed in terms of Hazard Ratio (HR) and 95% confidence intervals.

The statistical significance of differences in effects between subgroups were tested by adding interaction terms to the models. An interaction between CD4 cell count and ART prescription was modelled by including an interaction term in the primary and secondary models. Post-model statistics were used to determine if the interaction was significant and hazard ratios for the interaction term factors were estimated by exponentiation of coefficients from the linear combinations of coefficients (Stata help file: lincom). Also, the effect of ART versus no-ART prescription on survival was evaluated within each of the CD4 strata (primary and secondary models).

2.6.6. Sensitivity analysis

A sensitivity analysis using a Cox proportional hazards model without imputation, and restricted to TB information obtained solely from the HIV programme, was used to investigate the effect of ETR linkage on findings.

A summary of planned statistical analyses is displayed in Table 14.

Table 14. Summary of statistical analyses to evaluate the effect of ART on time to death.

Investigate	Model used	Effect estimate
Baseline demographics: a) Descriptive statistics (Number (%), mean(SD), median) b) Testing for trend over time	a) Chi-squared test; Hierarchical linear regression b) Extension of Wilcoxon rank sum test	p-value
Baseline demographics	a) Descriptive statistics	Number (%), mean(SD), median
Estimating probability of survival after TB diagnosis (start of TB treatment episode) for ART and pre-ART patients	a) Kaplan-Meier	
Primary analysis for effect of ART on time to death taking into account selection bias, time-varying covariates and intermediate variables. (TB from HIV programme and ETR.)	a) Adjusted MSM. Imputed data [25 sets].	HR
Primary analysis for effect of cotrimoxazole on time to death taking into account selection bias, time-varying covariates and intermediate variables. (TB from HIV programme and ETR.)	a) Adjusted MSM. Imputed data [25 sets].	HR
Secondary analysis for effect of ART on time to death (TB from HIV programme and ETR.)	a) Crude Cox proportional hazards model. No imputation. b) Adjusted Cox proportional hazards model. Imputed data [25 sets]. c) Adjusted Cox proportional hazards model. No imputation.	HR

Investigate	Model used	Effect estimate
Sensitivity analysis for effect of ART on time to death but limited to TB information from HIV programme only	a) Adjusted Cox proportional hazards model. No imputation.	HR

MSM: Marginal structural model; HR: Hazard ratio;

2.7. The effect of ART on TB treatment outcomes in ART-naïve co-infected HIV-TB patients (Aim 3)

The Free State Province's ETR records treatment outcomes for patients, using standardised guidelines consistent with WHO guidelines. Linkage of the ETR with the HIV programme database made it possible to add additional information, like death status, clinical and laboratory information, to perform sophisticated analyses with explanatory and confounding variables, which are mostly absent from the ETR. TB treatment outcomes and the effect of ART on those outcomes were evaluated for the period 2004 to 2008, using descriptive statistics and logistic regression models.

2.7.1. Specific aims

1. Describe and compare baseline characteristics of patients, grouped by:
    a) Year TB treatment was started.
    b) Those who did or did not receive ART during TB treatment.
    c) Those who died during the first year of TB treatment and those who survived the first year since TB treatment commenced.
2. Evaluate effect of ART, initiated during TB treatment compared with not receiving ART, on TB treatment outcomes, as defined by WHO, between 2004 and 2008.

2.7.2. Study population

The study population included all HIV positive patients with their first episode of TB registered at a public sector TB facility between January 2004 and December 2008. Patients were

considered to have TB through information obtained exclusively from the ETR since TB outcomes are studied in this analysis. The TB register extract, which was used for linkage, was received in June 2010, allowing up to 18 months follow-up time for ascertainment of TB treatment outcomes for patients who started TB treatment in December 2008.

Inclusion criteria:

- Patients 16 years or older when they started TB treatment.
- First TB treatment episode recorded on the ETR for the patient between January 2004 and December 2008.

Exclusion criteria:

- Patients starting ART before their first TB episode.
- Patients with missing TB treatment outcomes.
- Patients known to have TB from either database (ETR or HIV database) but who could not be linked.

2.7.3. Study design

This was a retrospective cohort study which analysed data for individual patients from the provincial ETR linked to the HIV programme. TB treatment outcomes were compared in patients who did not receive ART and patients who received ART during their TB episode. Patients were assumed to have stayed on ART once they have started treatment.

2.7.4. Data variables

The data collection process for the HIV programme was described in section 2.3.2, TB programme data collection in section 2.3.4 and WHO definitions of TB treatment outcomes were provided in section 2.3.3.

Outcome variables

TB treatment outcomes routinely recorded in the ETR at the end of each TB episode were the primary outcomes. TB outcome is a seven category variable (cured, completed, died, failed, transferred out, defaulted, not evaluated) and mutually exclusive, however each possible

outcome was re-coded as a binary variable for logistic regression purposes. Success was an additional binary category for patients classified as either cured or completed. Two additional death variables were added for this study. The first death variable was for all deaths that occurred during the TB episode, whether recorded by the ETR or HIV database. Non-fatal outcomes on the ETR were overwritten if linkage to HIV database identified that a patient died during the episode. The second binary death variable was for deaths that occurred during the year following commencement of TB treatment (either during TB treatment or thereafter), and was derived using information from both databases.

Explanatory variables

The primary explanatory variable was whether or not a co-infected patient started ART during TB treatment. A binary ART variable was created and set to 1 if the patient received ART at any point after TB treatment started. An additional ART categorical variable was created to capture the timing of ART. The first category included all patients who never started ART by the end of follow-up, or started ART after completion of TB treatment. The second category included patients who started ART in the first 61 days of TB treatment (early ART), and the last category included patients who started ART after 61 days of TB treatment but before TB treatment was completed (late ART).

Covariates

Other explanatory variables included in the model were sex, age, ART facility and district, anatomical classification of TB (pulmonary, extra-pulmonary or both sites), whether new or retreatment case, and TB smear results, and CD4 cell count measured closest to start of TB treatment.

A window period was used to determine baseline values; for CD4 cell counts the period twelve months before TB treatment was started and three months after TB treatment was started was searched for values, and the value closest to the TB treatment start date was used.

Multiple imputation was used to impute missing CD4 cell counts due to a high proportion of missing data even after linkage with the laboratory results (Royston P, 2005). CD4 cell count categories were generated after imputation to optimise adjustment for confounding without assuming linear relationships between these variables and the respective outcome. CD4 cell

count categories were CD4 0-24, CD4 25-49, CD4 50-99, CD4 100-199, CD4 200-349, and CD4 350-2000 cells/µl.

2.7.5. Statistical analysis

Baseline characteristics for patients were compared by the year patients started TB treatment, by those who received ART and those who did not, and by those who died and did not die during the first year of treatment, adjusting standard errors of variable estimates for clustering of estimates within clinics within districts using Stata version 11. The significance of a trend over time for baseline characteristics was evaluated using an extension of the Wilcoxon rank sum test. Categorical variables were compared with Pearson's chi-squared test and continuous variables with one-way ANOVA (analysis of variance).

Logistic regression analyses aimed to estimate the effect of ART on TB treatment outcomes, adjusting for differences in patient characteristics in the ART and no-ART groups to account for selection bias. Separate analyses were carried out for each treatment outcome. All analyses were done at the individual patient level, accounting for stratification by health district and intra-clinic correlation of outcomes using Huber-White adjustment. Results are presented with odds ratios, 95% confidence intervals and p-values.

Significant survivor bias was expected as only survivors were able to receive ART. To address this, a sub-analysis was included to evaluate the effect of ART compared with no-ART on non-fatal TB outcomes. To ignore the effect of ART timing (early or late ART), an additional set of sub-analyses was done for patients who died during the first year since TB treatment commenced, and for all the other non-fatal TB treatment outcomes. These analyses evaluated the effect of ART initiated during TB treatment (coded as a binary variable, ignoring when it was initiated) on the outcomes. A summary of the planned statistical analyses is presented in Table 15.

Table 15. Summary of statistical analyses to evaluate the effect of ART on TB treatment outcomes.

Investigate	Model used	Effect estimate
Baseline demographics: a) Descriptive statistics (Number (%), mean(SD), median) b) Testing for trend over time	a) Chi-squared test; One-way ANOVA b) Extension of Wilcoxon rank sum test	p-value
Effect of ART (early, late or no ART) on TB treatment outcomes. Each outcome coded as a binary variable.	Logistic regression	OR
Survivor bias expected therefore a sub-analysis to evaluate of effect of ART on TB treatment outcomes but limited to a sample of patients with non-fatal outcomes	Logistic regression	OR
Sub-analysis ignoring timing of ART to evaluate effect of ART on TB treatment outcomes with ART coded as a binary variable	Logistic regression	OR

OR: Odds ratio;

2.8. Ethics

Approval to complete longitudinal analyses of patients recorded in the Free State's HIV programme warehouse has previously been provided by the Research Ethics Committee of the University of Cape Town (REC/REF 082/2006), initially as an independent project and later as part of the NIH-funded International Epidemiological Databases to Evaluate AIDS (IeDEA) programme. Specific approval to complete the analyses described in this proposal was sought through and granted an addendum application to the main cohort proposal (REC/REF 082/2006).

It was not feasible to obtain consent for the use of medical records from the tens of thousands of patients included in these analyses. However, ethical principles for the use of medical records without a patient's consent (Haines A et al., 2000) were followed. The research is of significant benefit to large scale ART programmes, and the inclusion of patients' medical record data in the

analyses in no way influenced individuals' care. The data (except for TB linked data) have been used by the candidate, her supervisors and their collaborators for programme evaluation (monitoring and evaluation reports), a randomised clinical trial (Fairall LR et al., 2012) and observational evaluation of ART effectiveness in the programme cohort (Fairall LR et al., 2008) on behalf of the provincial health department.

Patient confidentiality was protected by removing all personal identifiers (name, surname, DOB, address, South African ID number) when data were shared with supervisors. Data extracts were provided in an anonymised, unlinked format and site information was coded as categorical variables. Only the candidate kept a dataset with personal identifiers which was used for the linkage process and stored the password protected back-up copies of the dataset in a secure location at UCT Lung Institute. The databases used to prepare the dataset for this project were removed from the SQL database server after completion of the analyses.

3. AN EVALUATION OF THE EFFECTIVENESS OF ANTIRETROVIRAL TREATMENT (ART) ON INCIDENCE OF TUBERCULOSIS IN HIV POSITIVE ADULTS

The Free State HIV programme database includes tens of thousands of patients followed for several years, and includes follow-up time among patients receiving and not yet receiving ART. This allows for an evaluation of the effectiveness of ART on the incidence of TB (time to first episode of TB during follow-up, number of TB episodes per 100 person years) in this high risk population, results of which are presented in this chapter.

3.1. Selection and characteristics of patients included in this analysis

A total of 97 476 patients were documented as being enrolled in the HIV programme between May 2004 and June 2010, of which 23 401 patients were excluded because:

a) they were younger than 16 years when they enrolled in the HIV programme (n=7 063), or
b) they were documented as being HIV negative (n=1 839) [1], or
c) they had no means of being followed because there was no valid South African ID number with which to link to the death population register and they only had information for one visit recorded (n=4 616), or
d) there was no visit information recorded for them on the programme (n=8 026) [2], or
e) discrepancies in the sequencing of key dates could not be resolved (n=1 021), or
f) they died on the date of enrolling in the programme and so could not contribute follow-up time to this analysis (n=836) (Figure 20).

[1] The structured clinical records implemented as part of the Free State's HIV programme made provision for the process of HIV testing to be documented. This was used, at least during the early stages of the programme, to document all testing completed during enrolment in the programme, regardless of outcome. As the programme expanded and clinicians battled to cope with large patient numbers, the practise of completing documentation for patients testing negative, was dropped.

[2] Patients were registered (capturing personal contact details) on the HIV programme database but no visit data were recorded for the patient, either because the patient never returned to the HIV service, or the visit data was never captured, or the patient was registered a second time without noticing the first registration. Often one of the registrations did not have enough personal information (identifiers) available to identify it as a duplicate registration, and therefore the first registration with missing visit information remained on the system.

Of the remaining 74 075 patients who were included in the final analyses, 26 480 (36%) had not started ART by the date on which their follow-up time was censored, and 47 595 (64%) patients had started ART.

Selection of patients for inclusion in these analyses is displayed in Figure 20, which also shows the categorisation of the follow-up records. Follow-up time for patients was measured from enrolment in the HIV programme to the first diagnosis of TB and divided into no-ART and ART follow-up time (Methods 2.7.4). No-ART follow-up was defined from the day patients enrolled in the HIV programme until one day before ART treatment was started or the censor date. Both patients who had not started ART by the end of follow-up (n=26 480) and patients who had started ART during follow-up (n=39 326), contributed time to no-ART follow-up. ART follow-up time was from the date when ART treatment was started until the censor date. Ten percent of ART follow-up records (4 572 / 43 898) was contributed by patients who started ART on the day they were documented as being enrolled in the HIV programme, therefore they only contributed ART follow-up time (Figure 20). Patients who had started TB treatment before ART was started, only contributed no-ART follow-up time to this analysis (n=3 697) since follow-up was censored at the first diagnosis of TB.

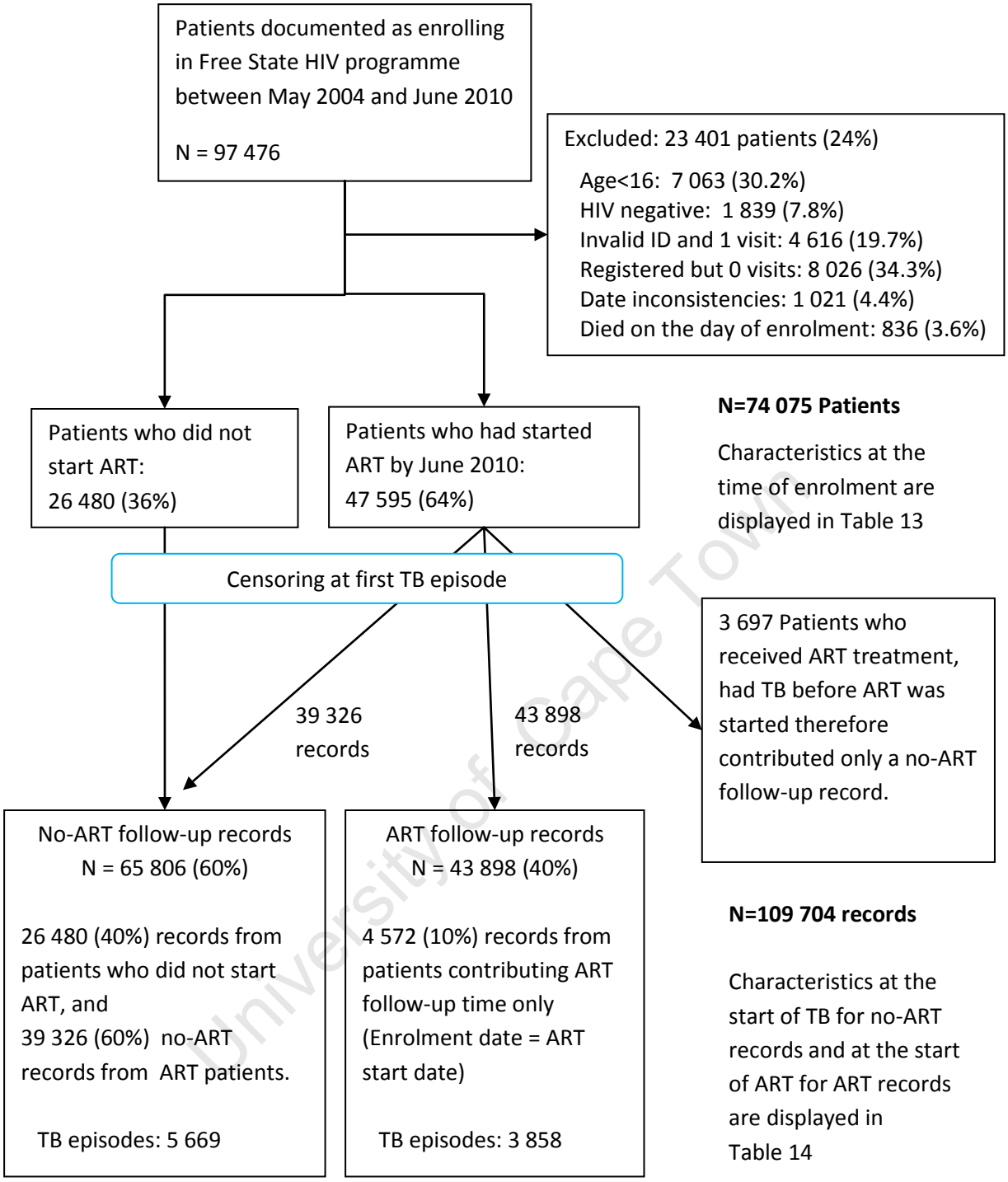

Figure 20: Selection of patients for analyses

Baseline characteristics for these patients are displayed in Table 16. The majority of patients included in this analysis were young women in their thirties (67.4% women, mean 37 years). The patients' CD4 counts at enrolment suggested they had advanced disease with a median CD4 cell count below 200 cells/µl (IQR: 66-234). Only 24.8% of patients with an available CD4 result, had a CD4 cell count above 200 cells per µl. A SA ID number was available for most of the patients (88.7%), permitting follow-up through linkage with the national death population

register. Almost a fifth (19.4%) of the patients had a history of TB treatment or were receiving TB treatment when they were registered on the HIV database. Twenty five percent (18 357 / 74 075) of the patients died, 66% (12 123 / 18 357) of them before they could be started on ART.

Table 16: Characteristics of patients included in this analysis. Data presented for the overall group, and disaggregated into those who did not receive ART by the end of follow-up, and patients who received ART during follow-up. The latter two categories are mutually exclusive.

Factors	Overall	Did not receive ART by the end of follow-up	Received ART by the end of follow-up	p value [a]
	n (%)	n (%)	n (%)	
Number of patients	74 075	26 480	47 595	
Women	49 938 (67.4)	17 121 (64.7)	32 817 (68.9)	<0.001
Age in years: Mean (SD)	37 (10)	36 (10)	37 (9)	<0.001
Mean weight at enrolment, kg. Mean (SD) [b]	58 (13) n=59 089 (79.8)	57 (14) n=18 556 (70.1)	59 (13) n=40 542 (85.2)	<0.001
Mean CD4 cell count at enrolment, cells/µl., Mean (SD) [b]	181 (171) n=60 731 (82.1)	254 (231) n=21 453 (81.0)	142 (108) n=39 278 (82.5)	<0.001
Median CD4 cell count at enrolment, cells/µl. Median (IQR) [b]	140 (66-234) n=60 731 (82.1)	192 (73-371) n=21 453 (81.0)	127 (64-188) n=39 278 (82.5)	
Number of patients per CD4 cell count category at enrolment				<0.001
CD4 missing	13 344(18.0)	5 363 (18.0)	7 981 (18.0)	
CD4 0-49	11 563(15.6)	4 583(15.4)	6 980 (15.7)	
CD4 50-99	10 420 (14.1)	3 401 (11.4)	7 019 (15.8)	
CD4 100-199	20 445 (27.6)	5 208 (17.5)	15 237 (34.4)	
CD4 200-349	10 854 (14.6)	5 029 (16.9)	5 825(13.1)	
CD4 >= 350	7 448 (10.1)	6 141(20.7)	1 307 (2.9)	
Valid SA ID number	65 728 (88.7)	23 314 (88.0)	42 414 (89.1)	0.003
History of TB at enrolment in HIV programme	14 382 (19.4)	3 142 (11.9)	11 240 (23.6)	<0.001

ART: Antiretroviral therapy; SD: Standard deviation; IQR: Inter quartile range.

a Comparing patients who did not receive ART by end of follow-up with patients who received ART for complex survey data, adjusted for district strata and clustering within subjects, using hierarchical linear regression for continuous outcomes and logistic regression for binary outcomes.

b Value from period six months before enrolment in HIV programme and six months after enrolment date but closest to the enrolment date.

Table 17 displays patient characteristics used in the analyses to adjust for confounding. For no-ART follow-up records, characteristics at the time of enrolling in the HIV programme are shown. For ART records, characteristics at the time of starting ART are shown. Categories are not mutually exclusive as ART patients may have contributed follow-up time to both. Whereas Table 16 reports on a total of 74 075 patients, Table 17 reports on all 109 704 no-ART and ART records contributed by these 74 075 patients.

Out of a total of 109 704 records, no-ART follow-up records contributed 60% (65 806 / 109 704) of the records. Overall, the same trend in data was observed as in Table 16, that is women contributed most of the records and the mean age was 36 (SD 10) at enrolment and 37 (SD 10) at the start of ART. A history of TB at enrolment was reported for almost a fifth of the records; 19.6% of no-ART follow-up records, and 24.6% of ART follow-up records reported TB before HIV programme enrolment. The majority of TB episodes recorded after enrolment occurred during no-ART follow-up period (60%, 5 669 / 9 527).

No-ART follow-up records had a total of 63 342 person years (py) and ART follow-up records a total of 79 321 person years. The overall TB incidence rate for this Free State HIV positive cohort was 6.68 per 100 py ((5 669+3 858 TB cases) / (63 342 py+79 321 py) × 100). The incidence rate of TB in patients not receiving ART by the end of follow-up was 8.95 per 100 person years (5 669 TB cases / 63 342 py × 100), while for patients receiving ART the incidence rate of TB was 4.86 per 100 person years (3 858 TB cases / 79 321 py × 100). Thus the crude TB incidence rate ratio (IRR) was 0.54 (CI: 0.52-0.56) for patients on ART compared to patients not receiving ART at the end of follow-up (Table 17).

Table 17: Characteristics of patients included in the analysis at the time of first enrolling in the HIV programme (no-ART follow-up record) and at the time of starting ART (ART follow-up record).

Factors	No-ART follow-up period n (%)	ART follow-up period n (%)	p value [a]
Number of patients	65 806	43 898	
Women	44 310 (67.3)	30 599 (69.7)	<0.001
Age in years: Mean (SD)	36 (10)	37 (10)	<0.001
Mean weight at start of record, kg. Mean (SD) [b]	58 (14) n=53 400 (81.1)	59 (13) n=33 079 (75.3)	<0.001
Mean CD4 cell count at start of record, cells/µl. Mean (SD) [b]	184 (174)	129 (90)	<0.001
Median CD4 cell count at start of record, cells/µl. Median (IQR) [b]	142 (66-239) n=56 145 (85.3)	122 (61-176) n=35 876 (81.7)	<0.001
Number of patients per CD4 cell count category at start of record [b]			<0.001
CD4 missing	9 662 (14.7)	8 022 (22.4)	
CD4 0-49	10 647 (16.2)	7 260 (20.2)	
CD4 50-99	9 576 (14.6)	7 257 (20.2)	
CD4 100-199	18 678 (28.4)	16 432 (45.8)	
CD4 200-349	10 022 (15.2)	4 353 (12.1)	
CD4 >= 350	7 221 (11.0)	574 (1.6)	
Valid SA ID number	58 488 (88.9)	39 471 (89.9)	<0.001
History of TB at enrolment	12 904 (19.6)	10 790 (24.6)	<0.001
TB episode during follow-up period	5 669 (8.6)	3 858 (8.8)	0.314
Person years (time to first TB)	63 342	79 321	<0.001

ART: Antiretroviral therapy; SD: Standard deviation; IQR: Inter quartile range.

a Comparing patients who did not receive ART by end of follow-up with patients who received ART for complex survey data, adjusted for district strata and clustering within subjects, using hierarchical linear regression for continuous outcomes and logistic regression for binary outcomes.

b No-ART: Weight or CD4 cell count from period six months before enrolment in HIV programme and six months after enrolment date but closest to the enrolment date.

ART: Weight or CD4 cell count between six months before ART was started and fourteen days after ART was started but closest to the date when ART was started.

3.2. Effect of ART on time to TB

3.2.1. Primary analysis

The primary analysis, a competing risks proportional hazards regression model, estimated associations of patient and facility characteristics with rates of TB incidence, ART and death. The model controlled for sex, age, weight, CD4 cell count, district, and history of TB. CD4 cell counts were missing for 18% of patients, and weight for 20% of patients and were assumed to be missing at random (Table 16). Ten imputed datasets, using chained equations, were used to replace each missing value with a set of plausible values that represent the uncertainty about the right value to impute (Rubin D, 1987; Royston P, 2005). For all other analyses in this thesis 25 sets of data were imputed, but due to the large size of the TB incidence dataset and computing power limitations, only 10 sets were used. The interaction between ART and CD4 cell count at enrolment and at the start of ART was also modelled by adding an interaction term.

The crude competing risks regression model (data not shown) indicated that the hazard of TB was reduced by 32% (SHR 0.68 CI: 0.65-0.71) for patients receiving ART at the time of TB diagnosis compared to patients who had not yet started ART when they were diagnosed with TB, allowing for the competing risk of death. Figure 21 displays the crude cumulative incidence of TB over time with death as competing risk for patients receiving ART or not by the end of follow-up. For on-ART patient records, 559/43 898 (1.27%) started TB treatment within 30 days of starting ART and 1 137 (2.6%) died. Within 60 days of starting ART 1 008/43 898 (2.30%) started TB treatment and 1 827 (4.2%) died.

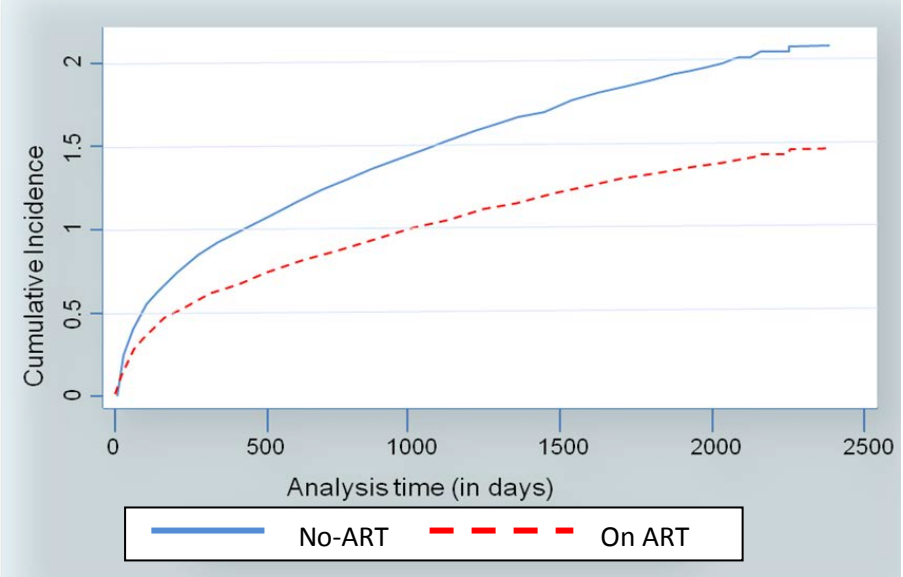

Figure 21: Cumulative incidence of TB with death as competing risk. Crude competing risks model; imputed data [10 sets].

The effect estimate obtained from an adjusted competing risks model using imputed data was very similar to the crude competing risks model (Table 18). ART reduced the hazard of TB by 36% while adjusting for the competing risk of death (SHR 0.64 CI: 0.61-0.67).

The competing risks model shown in the first column of Table 18 did not include a variable to model the interaction between receiving ART and the CD4 cell count at the start of ART, whereas the model shown in the second column included an interaction term to model the interaction between ART usage and CD4 cell count at enrolment and at the start of ART. The interaction was significant (p=0.021).

Table 18 shows that the hazard of TB in HIV infected patients was reduced by 36% (SHR 0.64 CI: 0.61-0.67) when patients received ART compared to patients not receiving ART for the competing risk model without the interaction term, while the competing risks model with the interaction term included, the hazard of TB was reduced by 38% (SHR 0.62 CI: 0.56-0.69). The model without the interaction term was selected as the primary model since the primary aim was to estimate the effectiveness of ART in all patients. For patients with CD4 cell counts between 200 and 350 cells/µl (SHR 0.82 CI: 0.76-0.89) or above 350 cells/µl (SHR 0.55 CI: 0.50-0.61), the hazard of TB was lower than for patients with a CD4<50 cells/µl. Men had an increased hazard of developing TB (SHR 1.33 CI: 1.27-1.39) as well as patients between the ages of 30 and 39 (SHR 1.22 CI: 1.01-1.50) compared to patients younger than 20 years. Increases in weight reduced the hazard of TB. Surprisingly, a history of TB at HIV programme enrolment

reduced the hazard of TB (SHR 0.86 CI: 0.82-0.91) compared to patients with no history of TB reported at episode enrolment.

Table 18: Factors associated with the effect of ART on incidence of TB in HIV positive patients using a competing risks model; imputed data [10 sets].

Factors	Competing risks regression model [a] (TB from HIV programme and ETR)			Computing risks regression model with interaction term [a] (TB from HIV programme and ETR)		
	SHR	**95% CI**	**p value**	**SHR**	**95% CI**	**p value**
ART treatment status						
No ART	1			1		
ART	0.64	0.61 - 0.67	<0.001	0.62	0.56 - 0.69	<0.001
CD4 cell count at start of follow-up record (cells/µl) [d]						
CD4 0-49	1			1		
CD4 50-99	1.17	1.10 - 1.25	<0.001	1.21	1.10 - 1.33	<0.001
CD4 100-199	1.01	0.95 - 1.07	0.73	0.99	0.91 - 1.08	0.80
CD4 200-349	0.82	0.76 - 0.89	<0.001	0.79	0.72 - 0.87	<0.001
CD4 >350	0.55	0.50 - 0.61	<0.001	0.53	0.48 - 0.60	<0.001
Interaction between CD4 and ART [b,c]						
No ART and CD4 <50				1		
ART and CD4 50-99				0.93	0.81 – 1.07	0.30
ART and CD4 100-199				1.05	0.92 – 1.19	0.47
ART and CD4 200-349				1.13	0.95 – 1.34	0.18
ART and CD4 >350				1.44	1.02 – 2.03	0.34
Sex						
Women	1			1		
Men	1.33	1.27 - 1.39	<0.001	1.33	1.27 - 1.39	<0.001
Age category at HIV programme enrolment (years)						
<20	1			1		
20-29	1.21	0.99 - 1.49	0.07	1.21	0.98 - 1.48	0.07
30-39	1.22	1.01 -1.50	0.05	1.22	1.00 - 1.50	0.06
40-49	1.14	0.93 - 1.40	0.22	1.13	0.92- 1.40	0.23
>50	0.97	0.78 - 1.20	0.76	0.96	0.78- 1.19	0.74
Weight at start of follow-up record (kg) [d]						
30-45	1			1		
46-52	0.88	0.82 - 0.95	0.001	0.88	0.82 - 0.96	0.002

Factors	Competing risks regression model [a] (TB from HIV programme and ETR)			Computing risks regression model with interaction term [a] (TB from HIV programme and ETR)		
	SHR	**95% CI**	**p value**	**SHR**	**95% CI**	**p value**
53-59	0.79	0.73 - 0.86	<0.001	0.79	0.73 - 0.86	<0.001
60-160	0.60	0.56 - 0.65	<0.001	0.60	0.56 - 0.65	<0.001
History of TB at HIV programme enrolment						
No previous TB	1			1		
Previous TB	0.86	0.82-0.91	<0.001	0.87	0.82-0.91	<0.001
District						
Fezile Dabi	1			1		
Lejweleputswa	0.86	0.80 - 0.93	<0.001	0.86	0.80-0.93	<0.001
Motheo	1.11	1.04 - 1.19	0.001	1.11	1.04 - 1.19	0.002
Thabo Mofutsanyane	0.80	0.74 - 0.85	<0.001	0.79	0.74 - 0.85	<0.001
Xhariep	2.06	1.90 - 2.23	<0.001	2.06	1.90 - 2.23	<0.001

SHR: Sub-hazard ratio; CI: Confidence interval

a Competing risks regression model also adjusting for clustering within districts and intra-patient correlation of outcomes.

b Test for equality of coefficients for interaction term, p=0.021

c Point estimates and confidence intervals were computed for linear combinations of coefficients after the estimation command. Recalculated values displayed in Table 19.

d No-ART: Weight and CD4 cell count from period six months before enrolment in HIV programme and six months after enrolment date but closest to the enrolment date.

ART: Weight and CD4 cell count between six months before ART was started and fourteen days after ART was started but closest to the date when ART was started.

Addition of ART-CD4 interaction terms significantly improved the model's fit with the data (p=0.021). That is, that the estimated effectiveness of ART was significantly different at different CD4 levels, as shown in Table 16. This shows that ART was similarly effective at preventing TB across all CD4 categories, but not when the CD4 cell count was more than 350 cells/µl (SHR 0.90 CI: 0.66-1.23).

Table 19: Effect of ART on time to TB within each CD4 stratum; Competing risks model with imputed data [10 sets].

Factor:	**Adjusted competing risks model** [a, b, c] (TB from HIV programme and ETR)		
Interaction between CD4 category (cells/µl) and ART	**SHR**	**95% CI**	**p value**
CD4 0-49 and on ART	0.62	0.56 - 0.69	<0.001
CD4 50-99 and on ART	0.58	0.53 - 0.64	<0.001
CD4 100-199 and on ART	0.65	0.61 - 0.70	<0.001
CD4 200-349 and on ART	0.70	0.62 - 0.80	<0.001
CD4 >=350 and on ART	0.90	0.66 - 1.23	0.50

SHR: Sub-hazard ratio; CI: Confidence interval

a Adjusting for sex, age, CD4, weight, history of TB, and for clustering within districts and intra-patient correlation of outcomes.

b Point estimates and confidence intervals were computed for linear combinations of coefficients after the estimation command.

c Test for equality of coefficients for interaction term, p=0.021

3.2.2. Secondary analysis: Cox Regression Models

An adjusted Kaplan Meier curve was used to estimate the probability (incidence) of active TB for patients receiving ART and patients not receiving ART by the end of follow-up (Figure 22). The log-rank test for equality of survivor functions confirmed a significant difference in the incidence of TB for patients who had not started ART by the end of follow-up compared to patients who had started ART ($p<0.001$). This was investigated further with Cox models.

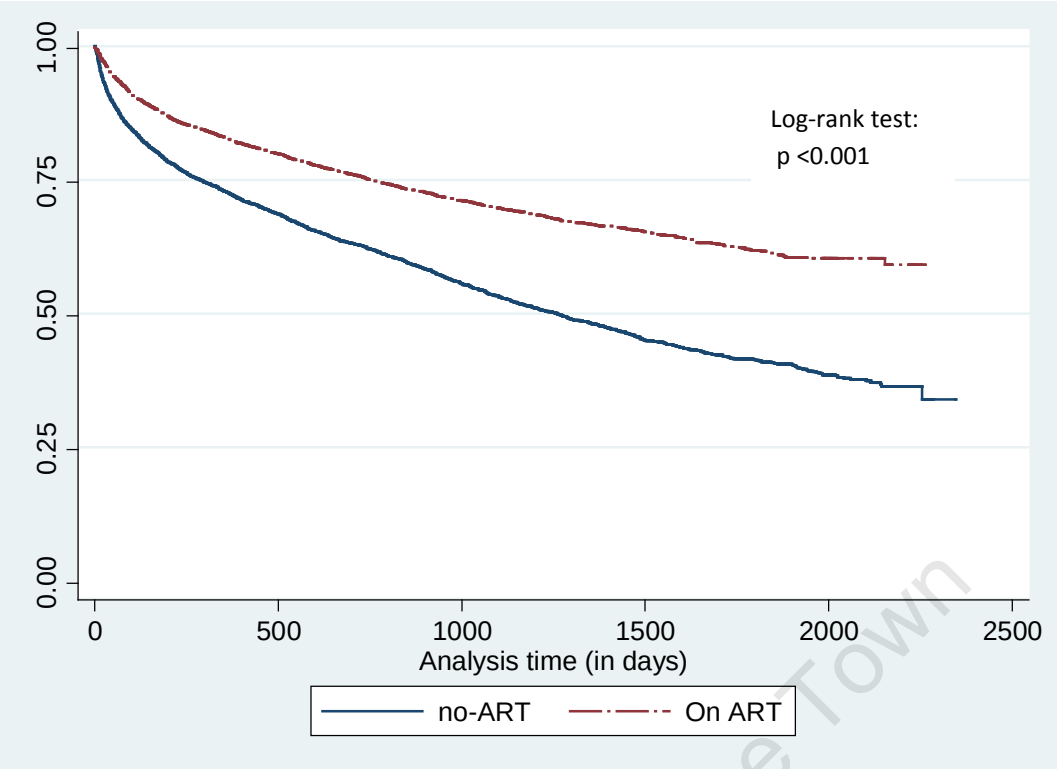

Figure 22: Adjusted Kaplan-Meier survival estimates by treatment status. For patients who received ART, time is since commencement of ART. For patients who did not receive ART by the end of follow-up, time is since enrolment in the HIV programme. Adjusted for sex, age, weight, CD4, cotrimoxazole prescription, and for clustering within clinics within districts and intra-patient correlation of outcomes.

Cox proportional hazards models were also used to assess the effectiveness of ART on the incidence of TB. The crude Cox proportional hazards model (data not shown), which only adjusted for intra-patient correlation of outcomes, showed that the hazard of TB was reduced by 39% (HR 0.61 CI: 0.59-0.64) for patients who received ART compared to patients not receiving ART by end of follow-up. Explanatory variables added to the model included sex, age, weight, CD4 cell count, history of TB at enrolment, and an interaction term modelling CD4 cell count and ART prescription.

Table 20 shows the results for two adjusted Cox models using imputed data (10 sets used for imputation), one model without the interaction between CD4 and ART and one model with the interaction modelled. The hazard of TB was reduced by 49% (HR 0.51, CI: 0.49-0.53) in the model without the interaction term, which is a larger effect than that obtained from the competing risks model (SHR 0.64, CI: 0.61-0.67, Table 18). The Cox model with the interaction term included had an even larger TB hazard reduction, 58% (HR 0.42, CI: 0.38-0.47, Table 20). For both Cox models, men were more at risk of developing TB, patients between 20 and 49 years had an increased hazard of TB compared to patients younger than 20 years, and patients

with higher body weights at enrolment were associated with a lower hazard of being diagnosed with TB.

Table 20: Factors associated with the effect of ART on incidence of TB patients using a Cox proportional hazards model; imputed data [10 sets].

Factors	Adjusted Cox proportional hazards model [a] (TB from HIV prog. and ETR)			Adjusted Cox proportional hazards model with interaction term [a] (TB from HIV prog. and ETR)		
	HR	**95% CI**	**p value**	**HR**	**95% CI**	**p value**
ART Treatment status						
No ART	1					
ART	0.51	0.49-0.53	<0.001	0.42	0.38-0.47	<0.001
Baseline CD4 at HIV programme enrolment (cells/µl) [d]						
CD4 0-49	1			1		
CD4 50-99	1.06	0.97-1.13	0.21	1.04	0.95-1.13	0.16
CD4 100-199	0.83	0.77-0.89	<0.001	0.74	0.68-0.81	<0.001
CD4 200-349	0.49	0.45-0.54	<0.001	0.51	0.46-0.57	<0.001
CD4 >350	0.27	0.24-0.30	<0.001	0.31	0.28-0.35	<0.001
Interaction between ART and CD4 [b, c]						
No ART and CD4 0-49				1		
ART and CD4 50-99				1.03	0.88-1.19	0.73
ART and CD4 100-199				1.28	1.13-1.46	<0.001
ART and CD4 200-349				1.62	1.36-1.94	<0.001
ART and CD4 >350				2.22	1.57-3.13	<0.001
Sex						
Women	1			1		
Men	1.39	1.33 - 1.45	<0.001	1.39	1.33 - 1.45	<0.001
Age at HIV programme enrolment (years)						
<20	1			1		
20-29	1.27	1.03-1.57	0.02	1.27	1.03-1.56	0.03
30-39	1.31	1.07-1.62	0.009	1.31	1.07-1.62	0.01
40-49	1.23	1.00-1.52	0.05	1.23	1.00-1.52	0.05
>50	1.08	0.87-0.34	0.48	1.08	0.87-1.34	0.49
Weight at HIV programme enrolment (kg)[d]						
30-45	1			1		
46-52	0.81	0.75-0.87	<0.001	0.81	0.75-0.87	<0.001
53-59	0.70	0.63-0.74	<0.001	0.69	0.63-0.74	<0.001

Factors	Adjusted Cox proportional hazards model [a] (TB from HIV prog. and ETR)			Adjusted Cox proportional hazards model with interaction term [a] (TB from HIV prog. and ETR)		
	HR	**95% CI**	**p value**	**HR**	**95% CI**	**p value**
60-160	0.50	0.47-0.54	<0.001	0.50	0.47-0.54	<0.001
History of TB						
No previous TB	1			1		
Previous TB	0.82	0.78-0.87	<0.001	0.83	0.78-0.87	<0.001
District						
Fezile Dabi	1			1		
Lejweleputswa	0.94	0.82-0.95	0.001	0.88	0.82-0.95	0.001
Motheo	1.01	1.05-1.20	0.001	1.11	1.04-1.19	0.002
Thabo Mofutsanyane	0.87	0.78-0.90	<0.001	0.84	0.78-0.90	<0.001
Xhariep	2.41	1.94-2.23	<0.001	2.09	1.94-2.27	<0.001

HR: Hazard ratio; CI: Confidence interval

a Cox proportional hazards model also adjusting for clustering within districts and intra-patient correlation of outcomes.

b Test for equality of coefficients for interaction term, p<0.001

c Point estimates and confidence intervals were computed for linear combinations of coefficients after the estimation command. Recalculated values displayed in Table 21.

d No-ART: Weight and CD4 cell count from period six months before enrolment in HIV programme and six months after enrolment date but closest to the enrolment date.

ART: Weight and CD4 cell count between six months before ART was started and fourteen days after ART was started but closest to the date when ART was started.

Table 21 shows the post-estimation analysis to recalculate the hazard ratios within the CD4 strata for the interaction term in the Cox model. This shows that although ART is effective in reducing the hazard of TB in all strata, as CD4 cell count increases the effectiveness of ART is less prominent, especially for CD4 above 350 cells/µl (HR 0.93 CI: 0.68-1.28).

Table 21: Effect of ART on time within each CD4 stratum; Cox model with imputed data [10 sets].

Factor:	Adjusted Cox proportional hazards model [a, b, c] (TB from HIV programme and ETR)		
Interaction between CD4 category (cells/µl) and ART	**HR**	**95% CI**	**p value**
CD4 0-49 and on ART	0.42	0.38-0.47	<0.001
CD4 50-99 and on ART	0.43	0.39-0.48	<0.001
CD4 100-199 and on ART	0.54	0.50-0.58	0.002
CD4 200-349 and on ART	0.68	0.60-0.77	0.003

Factor:	Adjusted Cox proportional hazards model [a, b, c] (TB from HIV programme and ETR)		
Interaction between CD4 category (cells/µl) and ART	**HR**	**95% CI**	**p value**
CD4 >350 and on ART	0.93	0.68-1.28	0.67

HR: Hazard ratio; CI: Confidence Interval

a Adjusting for sex, age, CD4, weight, history of TB, and for clustering within districts and intra-patient correlation of outcomes.

b Point estimates and confidence intervals were computed for linear combinations of coefficients after the estimation command.

c Test for equality of coefficients for interaction term, p<0.001

The complete case analysis of the adjusted Cox model is shown in Table 22. ART reduced the hazard of TB by 39% (HR 0.61 CI: 0.58-0.65) in the model without the interaction term, which was less effective in reducing the hazard of developing TB compared to the adjusted Cox model using imputed data (10 sets), with a 49% hazard reduction (HR 0.51 CI: 0.49-0.53, Table 20).

Unlike in previous models, the inclusion of an interaction term for CD4 and ART prescription made a substantial difference to the Cox complete case analysis model (Table 22). The effect estimate from the adjusted Cox model with the interaction term, limited to complete cases, showed that ART was more effective in preventing TB than the model without the interaction term; the hazard of TB was reduced by 53% (HR 0.47, CI: 0.41-0.52) compared to the 39% reduction in hazard from the model without the interaction term (HR 0.61 CI: 0.58-0.65). A post-analysis estimation test indicated that the interaction term included in the model was significant (p<0.001).

Table 22 also shows that patients from Xhariep district (HR 2.44 CI: 2.23-2.68), and men were more at risk of TB (HR 1.40 CI: 1.33-1.47), while patients with higher CD4 cell counts and body weights had a decreased hazard of developing TB. As with the other models, patients with a history of TB when they enrolled in the HIV programme experienced a protective effect in the incidence of subsequent TB (HR 0.81 CI: 0.76-0.86) compared to patients with no history of TB at enrolment.

Table 22: Factors associated with the effect of ART on incidence of TB patients using a Cox proportional hazards model; no imputation.

Factors	Adjusted Cox proportional hazards model [a] (TB from HIV prog. and ETR)			Adjusted Cox proportional hazards model with interaction term [a] (TB from HIV prog. and ETR)		
	HR	**95% CI**	**p value**	**HR**	**95% CI**	**p value**
ART Treatment status						
No ART	1					
ART	0.61	0.58-0.65	<0.001	0.47	0.41-0.52	<0.001
Baseline CD4 at HIV programme enrolment (cells/µl) d						
CD4 0-49	1			1		
CD4 50-99	1.06	0.97-1.13	0.21	1.02	0.92 - 1.13	0.71
CD4 100-199	0.83	0.77-0.89	<0.001	0.70	0.64 - 0.77	<0.001
CD4 200-349	0.49	0.45-0.54	<0.001	0.40	0.36 - 0.45	<0.001
CD4 >350	0.27	0.24-0.30	<0.001	0.23	0.21 - 0.25	<0.001
Interaction between ART and CD4 [b, c]						
No ART and CD4 0-49				1		
ART and CD4 50-99				1.09	0.93 - 1.27	0.23
ART and CD4 100-199				1.45	1.25 - 1.66	<0.001
ART and CD4 200-349				1.99	1.67 - 2.40	<0.001
ART and CD4 >350				3.59	2.49 – 5.11	<0.001
Sex						
Women	1			1		
Men	1.40	1.33-1.47	<0.001	1.41	1.33 - 1.48	<0.001
Age at HIV programme enrolment (years)						
18-29	1			1		
30-35	1.13	1.06-1.21	<0.001	1.14	1.06 - 1.22	<0.001
36-42	1.06	0.98-1.14	0.13	1.06	0.99 - 1.14	0.09
43-99	0.98	0.91-1.06	0.69	0.99	0.92 - 1.07	0.77
Weight at HIV programme enrolment (kg)[d]						
30-45	1			1		
46-52	0.78	0.73-0.82	<0.001	0.78	0.73 - 0.83	<0.001
53-59	0.66	0.61-0.70	<0.001	0.65	0.61 - 0.70	<0.001
60-160	0.50	0.46-0.53	<0.001	0.49	0.46 - 0.53	<0.001
History of TB						
No previous TB	1			1		
Previous TB	0.81	0.76-0.86	<0.001	0.82	0.77 - 0.87	<0.001

Factors	Adjusted Cox proportional hazards model [a] (TB from HIV prog. and ETR)			Adjusted Cox proportional hazards model with interaction term [a] (TB from HIV prog. and ETR)		
	HR	**95% CI**	**p value**	**HR**	**95% CI**	**p value**
Cotrimoxazole prescribed at any point during follow-up						
No cotrimoxazole	1			1		
Cotrimoxazole prescribed	0.49	0.46-0.52	<0.001	0.49	0.46 - 0.52	<0.001
District						
Fezile Dabi	1			1		
Lejweleputswa	0.94	0.86-1.03	0.20	0.94	0.86 - 1.03	0.16
Motheo	1.03	0.95-1.12	0.47	1.01	0.94 - 1.10	0.63
Thabo Mofutsanyane	0.88	0.82-0.96	0.003	0.87	0.80 - 0.95	0.002
Xhariep	2.44	2.23-2.68	<0.001	2.42	2.21 - 2.66	<0.001

HR: Hazard ratio; CI: Confidence Interval

a Cox proportional hazards model also adjusting for clustering within districts and intra-patient correlation of outcomes.

b Test for equality of coefficients for interaction term, p<0.001

c Point estimates and confidence intervals were computed for linear combinations of coefficients after the estimation command. Recalculated values displayed in Table 23.

d No-ART: Weight and CD4 cell count from period six months before enrolment in HIV programme and six months after enrolment date but closest to the enrolment date.

ART: Weight and CD4 cell count between six months before ART was started and fourteen days after ART was started but closest to the date when ART was started.

A post-estimation analysis was done on the Cox model to recalculate the hazard ratios for the interaction term within the different CD4 strata (Table 23). This showed that ART had an almost similar effect at the different CD4 levels, ranging between 48% (CD4 50-99, HR 0.52 CI: 0.46-0.56) and 63% (CD4 200-350, HR 0.37 CI: 0.32-0.43) effectiveness in reducing the incidence of TB.

Table 23: Effect of ART on time to TB within each CD4 stratum; Cox model with no imputation.

Factor:	Adjusted Cox proportional hazards model [a, b, c] (TB from HIV programme and ETR)		
Interaction between CD4 category (cells/µl) and ART	**HR**	**95% CI**	**p value**
CD4 0-49 and on ART	0.47	0.41-0.52	<0.001
CD4 50-99 and on ART	0.52	0.46-0.56	<0.001
CD4 100-199 and on ART	0.47	0.43-0.52	<0.001
CD4 200-349 and on ART	0.37	0.32-0.43	<0.001
CD4 >350 and on ART	0.38	0.27-0.54	<0.001

Factor:	Adjusted Cox proportional hazards model [a, b, c] (TB from HIV programme and ETR)		
Interaction between CD4 category (cells/µl) and ART	**HR**	**95% CI**	**p value**

HR: Hazard ratio; CI: Confidence Interval

a Adjusting for sex, age, CD4, weight, history of TB, and for clustering within districts and intra-patient correlation of outcomes.

b Point estimates and confidence intervals were computed for linear combinations of coefficients after the estimation command.

c Test for equality of coefficients for interaction term, $p<0.001$

A sensitivity analysis was performed, using an adjusted Cox proportional hazards model with complete case data, to evaluate the effect of ART on TB incidence (Table 24). This dataset included TB reported by the HIV programme only. A total of 10 472 subjects were identified as receiving TB treatment, 5 583 had a TB episode during the no-ART period and 4 889 patients' first episode of TB was reported while receiving ART. These categories are mutually exclusive since censoring was done at the first TB episode.

The Cox model results displayed in Table 24 shows that ART reduced the hazard of TB by 52% (HR 0.48, CI 0.45-0.51) in this sensitivity analysis - a larger effect size when compared with the equivalent model using a combination of ETR information and the HIV programme TB information , HR 0.61 CI: 0.58-0.65 (Table 22).

Table 24: Sensitivity analysis: Factors associated with the effect of ART on incidence of TB patients using a Cox proportional hazards model with TB identified from HIV programme only; no imputation.

Factors		Adjusted Cox proportional hazards model [a] (TB from ART programme only)		
		HR	**95% CI**	**p value**
ART Treatment status				
	No ART	1		
	ART	0.48	0.45-0.51	<0.001
Baseline CD4 at HIV programme enrolment (cells/µl) [b]				
	CD4 0-49	1		
	CD4 50-99	0.90	0.84-0.98	0.01
	CD4 100-199	0.71	0.66-0.75	<0.001
	CD4 200-349	0.62	0.56-0.67	<0.001
	CD4 >350	0.50	0.45-0.56	<0.001
Sex				
	Women	1		
	Men	1.43	1.36-1.51	<0.001
Age at HIV programme enrolment (years)				
	18-29	1		
	30-35	1.03	0.96-1.10	0.41
	36-42	0.96	0.89-1.03	0.24
	43-99	0.86	0.79-0.92	<0.001
Weight at HIV programme enrolment (kg) [b]				
	30-45	1		
	46-52	0.81	0.76-0.86	<0.001
	53-59	0.71	0.66-0.76	<0.001
	60-160	0.55	0.51-0.60	<0.001
History of TB				
	No previous TB	1		
	Previous TB	5.26	4.98-5.57	<0.001
Cotrimoxazole prescribed at any point during follow-up				
	No cotrimoxazole	1		
	Cotrimoxazole prescribed	1.65	1.55-1.75	<0.001
District				
	Fezile Dabi	1		
	Lejweleputswa	0.67	0.62-0.74	<0.001
	Motheo	1.28	1.18-1.39	<0.001
	Thabo Mofutsanyane	0.75	0.69-0.82	<0.001

Factors	Adjusted Cox proportional hazards model [a] (TB from ART programme only)		
	HR	**95% CI**	**p value**
Xhariep	2.16	1.97-2.36	<0.001

HR: Hazard ratio; CI: Confidence interval

a Cox proportional hazards model also adjusting for clustering within clinics within districts and intra-patient correlation of outcomes.

b No-ART: Weight and CD4 cell count from period six months before enrolment in HIV programme and six months after enrolment date but closest to the enrolment date.

ART: Weight and CD4 cell count between six months before ART was started and fourteen days after ART was started but closest to the date when ART was started.

3.3. Effect of ART on incidence rate of multiple TB episodes

All of the above analyses modelled a single episode of TB, and considered follow-up time until TB was diagnosed. However, some patients experienced multiple TB episodes after enrolment, either during the no-ART follow-up time, or during ART follow-up time, or during both periods. A Poisson regression model was used to determine the association between TB incidence rate and ART treatment, accounting for multiple TB episodes that occurred after enrolment in the HIV programme.

Table 25 summarises data for patients who contributed time to no-ART follow-up and ART follow-up for the Poisson model as well as the total number of TB episodes that occurred during these follow-up periods. The overall incidence rate (IR) of TB was 9.08/100 py [(6 471+7 346) TB cases / (65 738 py+86 404 py) × 100], which was much higher than the 6.68 / 100 py reported for first (single) episode TB (Table 17). The TB incidence rate in patients not receiving ART by the end of no-ART follow-up was 9.8 / 100 py (6 471 TB cases / 65 738 py × 100), while the incidence rate of TB during ART follow-up was 8.5/100 person years (7 346 TB cases / 86 404 py × 100). The crude incidence rate ratio (IRR) for TB was 0.85 (CI: 0.82-0.87) for patients who received ART during follow-up compared to patients who had not received ART during follow-up.

Table 25: Characteristics of patients included in the Poisson model evaluating the effect of ART on number of TB episodes at the time of first enrolling in the HIV programme (no-ART follow-up record) and at the time of starting ART (ART follow-up record). Categories are not mutually exclusive.

	No-ART follow-up period n (%)	ART follow-up period n (%)	p value [a]
Number of patients contributing a record	65 806	47 595	
Person years (TB risk free follow-up)	65 738	86 404	<0.001
History of TB	13 186 (19.6)	10 569 (22.2)	<0.001
Median CD4 cell count at start of follow-up period, cells/µl. Median (IQR) [b]	140 (66-226) n=58 318 (86.5)	126 (64-182) n=40 147 (84.4)	
Weight at start of follow-up period, kg, Mean (SD) [b]	58 (13) n=56 801 (84.2)	59 (13) n=37 703 (79.2)	<0.001
Total number of TB episodes during follow-up period	6 471	7 346	<0.001
Patients with more than one TB episode during follow-up period	1 046	1 368	<0.001

ART: Antiretroviral therapy; SD: Standard deviation; IQR: Inter quartile range.

a Comparing patients who did not receive ART by end of follow-up with patients who received ART, adjusted for district strata and clustering within subjects, using hierarchical linear regression for continuous outcomes and logistic regression for binary outcomes.

b No-ART: Value from period six months before enrolment in HIV programme and six months after enrolment date but closest to the enrolment date.

ART: Value between six months before ART was started and fourteen days after ART was started.

Table 26 shows an adjusted Poisson regression evaluating the effect of ART on TB incidence among these patients accounting for multiple episodes of TB, using 25 sets of imputed data in two models. One model included an interaction term for ART and CD4 cell count at the start of each follow-up record and one model did not include the interaction term. The incidence rate ratio (IRR) in the model without an interaction term showed that ART reduced the incidence rate of TB by 34% (IRR 0.66, CI: 0.64-0.69). The model with the interaction term included showed that ART was more effective, reducing the incidence of TB by 68% (IRR 0.32 CI: 0.30-0.35).

As before, men were at greater risk of TB (1.37 CI: 1.30-1.43), patients between the age of 30 and 43 had an increased hazard of developing TB compared to patients in their twenties, and district variability in the incidence of TB was evident.

In Table 26, the model without the interaction term showed that, compared to patients with CD4<50 cells/µl, patients with a CD4 cell count between 50 and 99 and between 200 and 350 cells/µl were at higher risk (CD4 50-99: IRR 1.10 CI: 1.02-1.18), while patients with a CD4 above 350 cells/µl had a lower rate of developing TB (IRR 0.75 CI: 0.69-0.83). The model

estimated the effect of a history of TB at enrolment into the HIV programme on subsequent episodes of TB, adjusting for all other factors, including whether or not they had started ART. Patients who had TB before or during enrolling in the HIV programme had a 47% higher risk of TB compared to patients who had no history of TB at enrolment (IRR 1.47 CI: 1.39-1.55). Between 2005 and 2009 a sequential increase in the hazard of developing TB was noted, ranging from IRR 1.13 (CI: 1.03-1.23) in 2005 to IRR 3.39 (CI: 3.09-3.71) in 2009, probably due to more intensive screening and case-finding.

Table 26: Factors associated with the effect of ART on TB using a Poisson regression model; imputed data [25 sets].

Factors	Adjusted Poisson regression model [a] (TB from HIV prog. and ETR)			Adjusted Poisson regression model with interaction term [a] (TB from HIV prog. and ETR)		
	IRR	**95% CI**	**p value**	**IRR**	**95% CI**	**p value**
ART Treatment status						
No ART	1			1		
ART	0.66	0.64-0.69	<0.001	0.32	0.30-0.35	<0.001
Baseline CD4 at HIV programme enrolment (cells/µl) [c]						
CD4 0-49	1			1		
CD4 50-99	1.10	1.02 - 1.18	0.02	1.01	0.91 - 1.10	0.942
CD4 100-199	0.97	0.91 – 1.04	0.39	0.77	0.71 - 0.84	<0.001
CD4 200-349	1.10	1.02 – 1.19	0.02	0.51	0.47 - 0.57	<0.001
CD4 >350	0.75	0.69 - 0.83	<0.001	0.28	0.25 - 0.31	<0.001
Interaction between ART and CD4 [b]						
No ART and CD4 0-49				1		
ART and CD4 50-99				1.18	1.06 - 1.33	0.003
ART and CD4 100-199				1.48	1.35 - 1.64	<0.001
ART and CD4 200-349				4.28	3.82 - 4.80	<0.001
ART and CD4 >350				13.37	11.68 - 15.31	<0.001
Sex						
Women	1			1		
Men	1.37	1.30-1.43	<0.001	1.37	1.30-1.43	<0.001
Age at HIV programme enrolment (years)						
18-29	1			1		
30-35	1.17	1.10 - 1.24	<0.001	1.16	1.09 - 1.23	<0.001
36-42	1.12	1.05 - 1.19	<0.001	1.11	1.04 - 1.18	<0.001

Factors	Adjusted Poisson regression model [a] (TB from HIV prog. and ETR)			Adjusted Poisson regression model with interaction term [a] (TB from HIV prog. and ETR)		
	IRR	**95% CI**	**p value**	**IRR**	**95% CI**	**p value**
43-99	1.02	0.96 - 1.09	0.52	1.01	0.95 - 1.07	0.745
Weight at HIV programme enrolment (kg) [c]						
30-45	1			1		
46-52	0.87	0.82 - 0.92	<0.001	0.88	0.83 - 0.93	<0.001
53-59	0.74	0.69 - 0.78	<0.001	0.74	0.70 - 0.79	<0.001
60-160	0.55	0.52 - 0.59	<0.001	0.55	0.52 - 0.59	<0.001
History of TB						
No previous TB	1			1		
Previous TB	1.47	1.39 - 1.55	<0.001	1.43	1.36 - 1.52	<0.001
Year of enrolment in HIV programme						
2004	1			1		
2005	1.13	1.03-1.23	0.008	1.14	1.04-1.24	0.004
2006	1.38	1.27-1.51	<0.001	1.37	1.26-1.49	<0.001
2007	1.93	1.77-2.10	<0.001	1.86	1.72-2.02	<0.001
2008	2.25	2.07-2.46	<0.001	2.11	1.94-2.29	<0.001
2009	3.39	3.09-3.71	<0.001	3.21	2.94-3.51	<0.001
2010	5.87	5.21-6.61	<0.001	5.33	4.74-5.99	<0.001
District						
Fezile Dabi	1			1		
Lejweleputswa	0.91	0.84 - 0.98	0.02	0.89	0.82 - 0.96	0.003
Motheo	1.18	1.10 - 1.27	<0.001	1.12	1.04 - 1.20	0.001
Thabo Mofutsanyane	0.89	0.82 - 0.96	<0.001	0.87	0.81 - 0.93	<0.001
Xhariep	2.13	1.96 - 2.30	<0.001	2.12	1.96 - 2.29	<0.001

IRR: Incidence rate ratio; CI: Confidence interval

a Cox proportional hazards models also adjusting for district, and intra-patient correlation of outcomes.

b Test for equality of coefficients for interaction term, p<0.001

c No-ART: Value from period six months before enrolment in HIV programme and six months after enrolment date but closest to the enrolment date.

ART: Value between six months before ART was started and fourteen days after ART was started.

A post-estimation analysis to recalculate the hazard ratios for the interaction term was performed on the Poisson regression with the interaction term to evaluate the effect of ART on multiple episodes of TB within each of the CD4 strata separately (Table 27). This analysis showed that ART was effective in reducing the hazard of TB in CD4 strata less than 200 cells/µl but is associated with an increased TB incidence rate if CD4 > 350 cells/µl (IRR 4.34, CI: 3.92-4.81).

Table 27: Effect of ART on incidence rate of multiple TB within each CD4 stratum; Poisson regression model with imputed data [25 sets].

Factor:	Adjusted Poisson regression model [a,b,c] (TB from HIV programme and ETR)			
Interaction between CD4 category (cells/µl) and ART	**IRR**	**95% CI**		**p value**
CD4 0-49 and on ART	0.32	0.30	0.35	<0.001
CD4 50-99 and on ART	0.39	0.36	0.41	<0.001
CD4 100-199 and on ART	0.48	0.46	0.51	<0.001
CD4 200-349 and on ART	1.39	1.29	1.50	<0.001
CD4 >350 and on ART	4.34	3.92	4.81	<0.001

a Adjusting for sex, age, CD4, weight, history of TB, and for clustering within districts and intra-patient correlation of outcomes.

b Point estimates and confidence intervals were computed for linear combinations of coefficients after the estimation command.

c Test for equality of coefficients for interaction term, p<0.001

3.4. Summary

The adjusted competing risks regression model showed that ART reduced the hazard of developing TB by 36% (SHR 0.64 CI: 0.61-0.67). The Cox models that ignored the competing risk of death over-estimated the effectiveness of ART on time to TB, producing point estimates corresponding to a reduction in hazard of TB of between 39% (Cox model using complete case data, HR 0.61 CI: 0.58-0.65) and 49% (Cox model with imputed data, HR 0.51 CI: 0.49-0.53). Similarly, ignoring TB information obtained through linkage with the ETR over-estimated the effect of ART on time to TB (HR 0.48 CI: 0.45-0.51).

Since initiation of ART is strongly associated with a CD4 cell count, heterogeneity across the CD4 strata in the ART prescription was expected and tested for with the inclusion of an interaction term between CD4 category and ART. The interaction term was significant for all the models. The competing risks model without the interaction term had a similar effect estimate as the model with it (SHR 0.64 vs. SHR 0.62 respectively), while the Cox and Poisson models with an interaction terms showed that ART was more effective than in the models without the interaction term (HR 0.42 for Cox and interaction term vs. HR 0.51 (no interaction); IRR 0.32 for Poisson and interaction term vs. IRR 0.66 (no interaction)).

The competing risks models with ART-CD4 interaction included showed that the effect of ART decreased with increasing CD4 levels, and ART had no significant effect for CD4>350

(Table 19 and Table 21). The Cox model did not show ART effectiveness to depend on CD4 (Table 23), but this model was probably biased because it did not account for the competing risk of death.

The Poisson models which evaluated the effectiveness of ART on number of TB episodes also showed a substantial effect with ART reducing the incidence rate of TB by 34% (IRR 0.66 CI: 0.64-0.69).

The analyses were limited by the proportion of patients with missing weight or CD4 values or both. This is evident in the difference between estimates obtained in Cox models using complete case data (HR 0.61 CI: 0.58-0.65) and imputed data (HR 0.51 CI: 0.49-0.53). Post estimation analyses by CD4 strata showed that ART was most effective at preventing TB in patients with lower CD4 counts and had minimal impact in patients with CD4 counts in excess of 350.

The models also highlighted other associations between risk factors and the development of TB for example men were at risk. Patients with a higher body weight and age at enrolment had a lower hazard of TB, and as CD4 cell count increased the hazard of TB decreased. A history of TB treatment at enrolment in the HIV programme was protective in all the models, except the sensitivity analysis model which used TB information reported by the HIV programme only (HR 5.26 CI: 4.98-5.57) and the Poisson regression model accounting for multiple episodes of TB (IRR 1.47 CI: 1.39-1.55). Unfortunately the TB register did not record the number of TB treatment episodes the patient had. Thus if a patient had a TB episode during the study period it could have been their second or third episode since the pre-TB information was unavailable. This limited the possibility to evaluate if a first episode of TB was associated with further episodes compared to people without prior TB.

District heterogeneity was seen in all the models, Motheo and Xhariep districts were mostly at risk while in the other districts ART reduced the hazard of TB incidence compared to Fezile Dabi district.

Overall, ART is associated with a reduced hazard of developing TB.

Table 28: Summary of primary outcome for models to evaluate the effect of ART on TB incidence. Models displayed in this table do not include the interaction term.

Model used	**Effect of ART on incidence of TB** (TB from HIV programme and ETR)		
	Effect estimate	**95% CI**	**p value**
Primary analysis for effect of ART on time to first TB episode: Adjusted competing risks model. Imputed data [10 sets].	SHR 0.64	0.61-0.67	<0.001
Crude competing risks model. No imputation.	SHR 0.68	0.65-0.71	<0.001
Adjusted Cox proportional hazards model. Imputed data [10 sets].	HR 0.51	0.49-0.53	<0.001
Crude Cox proportional hazards model. No imputation.	HR 0.61	0.59-0.64	<0.001
Adjusted Cox proportional hazards model. No imputation.	HR 0.61	0.58-0.65	<0.001
Primary analysis for effect of ART on incidence rate of multiple TB episodes: Adjusted Poisson regression model. Imputed data [25 sets].	IRR 0.66	0.64-0.69	<0.001
Crude Poisson regression model. No imputation.	IRR 0.85	0.82-0.87	<0.001

SHR: Sub-hazard ratio; HR: Hazard ratio; IRR: Incidence rate ratio

4. AN EVALUATION OF THE EFFECTIVENESS OF THE PRESCRIPTION OF ART AND COTRIMOXAZOLE ON MORTALITY (TIME TO DEATH) IN CO-INFECTED HIV-TB PATIENTS

The Free State HIV programme database includes mostly HIV positive patients. Linkage with the provincial ETR was crucial to identify the sub-population of HIV patients co-infected with TB, due to suspected under-reporting of TB in the HIV programme. The HIV programme database also documented follow-up time before and after ART was started and could therefore support (survival) analyses to evaluate the effectiveness of ART on mortality in co-infected people. This chapter reports on the effect of ART on survival in co-infected HIV-TB patients by applying different types of regression models.

4.1. Selection and characteristics of patients included in this analysis.

Between 2004 and 2010 a total of 97 476 patients were documented as being enrolled in the HIV programme, of whom 12 258 were identified as co-infected HIV-TB patients. The source of TB information was either from the HIV programme database or from the ETR. A total of 1 760 co-infected HIV-TB patients were excluded because:

a) they were younger than 16 when they enrolled in the HIV programme (n=760), or
b) they were documented as being HIV negative (n=66)[1], or
c) they had no means of being followed because there was no valid South African ID number with which to link to the population register and they only had information for one visit recorded (n=772), or
d) they died on the date of enrolling in the programme and so could not contribute follow-up time to this analysis (n=162).

[1] The structured clinical records implemented as part of the Free State's HIV programme made provision for the process of HIV testing to be documented. This was used, at least during the early stages of the programme, to document all testing completed during enrolment in the programme, regardless of outcome. As the programme expanded and clinicians battled to cope with large patient numbers, the practise of completing documentation for patients testing negative, was dropped

Of the remaining 10 498 patients who were included in the analysis, 2 679 (25.5%) patients had not started ART by the date on which their follow-up time was censored and 7 819 (74.5%) patients had.

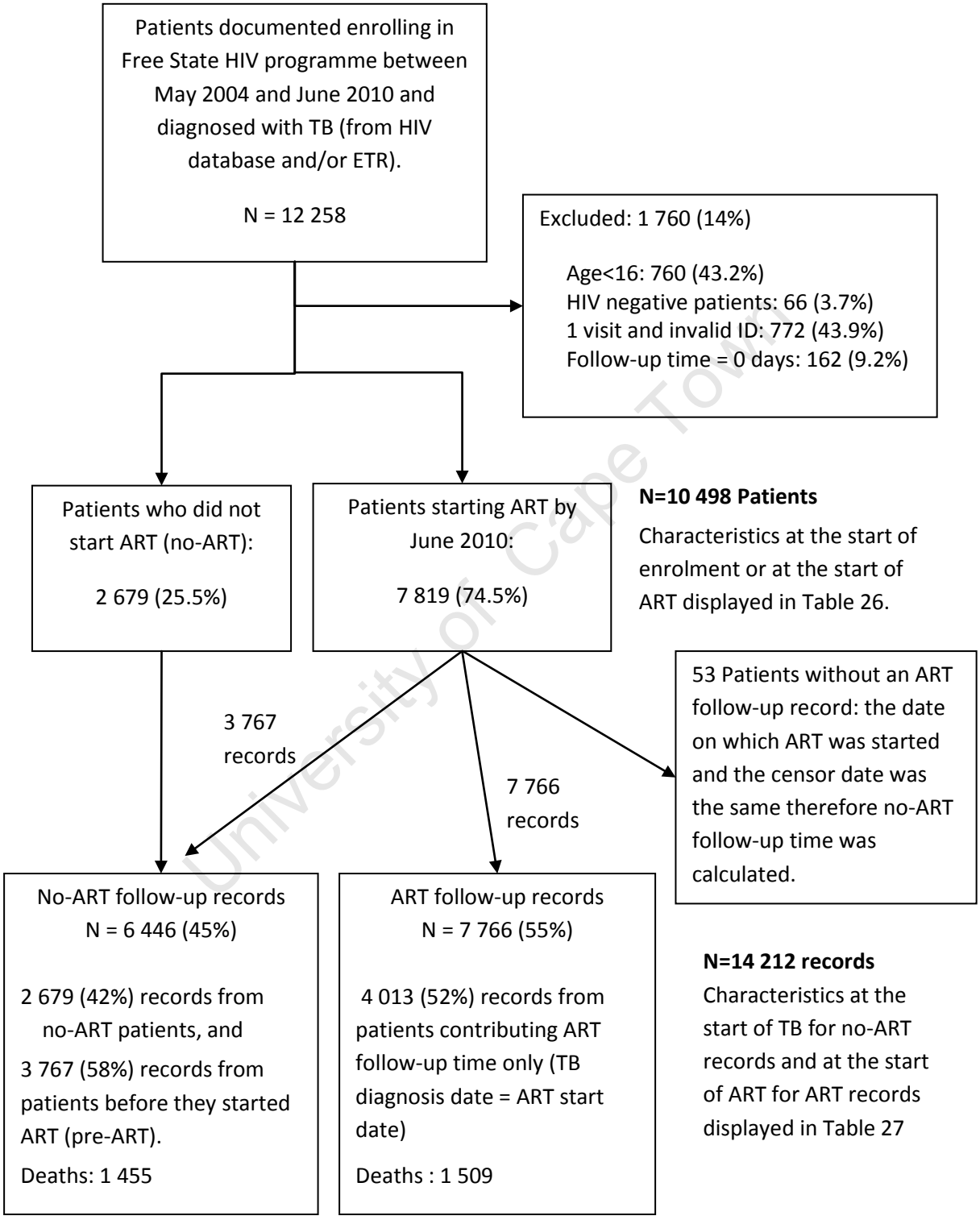

Figure 23: Enrolment of patients in analysis.

Table 29 displays the characteristics of co-infected HIV-TB patients at the time they first enrolled in the HIV programme. They were mostly women in their thirties (60.3% women, mean 36 years) and from the low median CD4 cell count of 122 cells/µl (IQR: 59-203) at enrolment, it was apparent that these patients had advanced disease. Despite this, a quarter of the patients (25.5%, 2 679 / 10 498) had not received ART by the end of follow-up. Ninety percent (89.6%) of patients had a SA ID numbers, allowing follow-up through linkage with the national death population register. Almost a fifth (16.5%) of the patients had a history of TB treatment or were receiving TB treatment when they were first registered on the HIV programme database. Fifty four percent (1 455 / 2 679) of the co-infected patients who had not received ART by the end of follow-up died, compared to 19.3% (1 509 / 7 819) of deaths reported in the patients who had received ART by the end of follow-up (Figure 23).

Table 29: Characteristics of patients included in this analysis. Data presented for the overall group, and disaggregated into those who did not receive ART by the end of follow-up, and patients who received ART during follow-up. The latter two categories are mutually exclusive.

	Overall	**Did not receive ART by end of follow-up n (%)**	**Received ART by the end of follow-up n (%)**	**p-value [a]**
Number of patients	10 498	2 679	7 819	
Women	6 331 (60.3)	1 501 (56.0)	4 830 (61.8)	<0.001
Age in years : Mean (SD)	36 (9)	36 (9)	36 (9)	<0.001
Mean weight at enrolment, kg. Mean (SD) [b]	55 (12) n=8 836 (84.2)	54 (12) n=2 068 (77.2)	55 (12) n=6 768 (86.6)	<0.001
Mean CD4 cell count at enrolment, cells/µl. Mean (SD) [b]	160 (151) n=9 026 (86.0)	222 (212) n=2 253 (84.1)	139 (116) n=6 773 (86.6)	<0.001
Median CD4 cell count at enrolment, cells/µl. Median (IQR) [b]	122 (59-203) n=9 026 (86.0)	160 (66-316) n=2 253 (84.1)	115 (58-185) n=6 773 (86.6)	
Number of patients per CD4 cell count category at enrolment [b]				<0.001
CD4 missing	1 472 (14.0)	426 (15.9)	1 046 (13.4)	
CD4 0-49	1 882 (17.9)	444 (16.6)	1 438 (18.4)	
CD4 50-99	1 848 (17.6)	356 (13.3)	1 492 (19.1)	
CD4 100-199	2 989 (28.5)	506 (18.9)	2 483 (31.8)	
CD4 200-349	1 466 (14.0)	470 (17.5)	996 (12.7)	
CD4 >= 350	841 (8.0)	477 (17.8)	364 (4.6)	
Valid South African ID number	9 407 (89.6)	2 349 (87.7)	7 058 (90.3)	<0.001
History of TB at enrolment in HIV programme	1 733 (16.5)	294 (10.9)	1 439 (18.4)	<0.001

	Overall	Did not receive ART by end of follow-up n (%)	Received ART by the end of follow-up n (%)	p-value [a]

ART: Antiretroviral therapy; SD: Standard deviation; IQR: Inter quartile range.

a Comparing patients who did not receive ART by end of follow-up with patients who received ART for complex survey data, adjusted for district strata and clustering within subjects, using hierarchical linear regression for continuous outcomes and logistic regression for binary outcomes.

b Value from period six months before enrolment in HIV programme and six months after enrolment date but closest to the enrolment date.

As described in Methods, Section 2.6.3, patient follow-up time (survival time) was described from the start of TB treatment to censoring and divided into two types of follow-up records: records during which patients did not receive ART (no-ART) and records where patients did receive ART by the end of follow-up.

No-ART follow-up was defined from the day when TB treatment was started until one day before ART treatment was started, or the censor date for patients who did not receive ART by the end of follow-up. Patients who had not started ART by the end of follow-up (n=2 679) and patients who had started ART during follow-up (n=3 767), contributed time to no-ART follow-up. ART follow-up time was from the ART treatment start date until the censor date; 4 013 patients' TB diagnosis was on the same date as their ART was started and therefore they only contributed ART follow-up time (Figure 23).

Table 30 displays patient characteristics at the start of these records, used in the analyses to adjust for confounding. For no-ART follow-up records, characteristics at the time of TB diagnosis are shown and for ART records, characteristics at the time of starting ART are shown. Categories are not mutually exclusive as ART patients may have contributed follow-up time to both, therefore Table 29 reports on a total of 10 498 patients whereas Table 30 reports on all 14 212 no-ART and ART records contributed by these 10 498 patients.

Forty percent of the records were no-ART records (6 446 / 14 212) and contributed a total of 5 218 person-years (py), while ART records contributed a total of 13 202 py. Women contributed 57.5% of the follow-up records. Patients were older (p=0.001) when they started ART (41 years, ±9) than when they started TB treatment (37 years, ±9) and the majority of patients had a CD4 cell count below 200 (65.2% for patients not receiving ART and 78.6% for patients receiving ART by the end of follow-up). The ETR identified 65% (9 304 / 14 212) of the TB episodes for this analysis (data not presented in Table).

Table 30: Characteristics of patients included in the analysis at the time of starting TB treatment (no-ART follow-up records) and at the time of starting ART (ART follow-up records).

	No-ART follow-up period n (%)	ART follow-up period n (%)	p-value [a]
Number of patients	6 446 (100)	7 766 (100)	
Person years (time to censor date)	5 218	13 202	
Women	3 706 (57.5)	4 801 (61.8)	<0.001
Age in years: Mean (SD)	37 (9)	41 (9)	0.001
Mean weight at start of TB or ART, kg. Mean (SD) [b]	53 (11) n=4 312 (66.9)	55 (12) n=5 899 (75.9)	0.001
Mean CD4 cell count at start of TB or ART, cells/µl. Mean (SD) [b]	153 (151) n=5 465 (84.8)	115 (85) n=6 745 (86.9)	0.001
Median CD4 cell count at start of TB or ART, cells/µl. Median (IQR) [b]	115 (55-194) n=5 465 (84.8)	105 (53-162) n=6 745 (86.9)	<0.001
Number of patients per CD4 cell count category at start of TB or ART			<0.001
CD4 missing	981 (15.2)	1 021 (13.1)	
CD4 0-49	1 235 (19.2)	1 558 (20.1)	
CD4 50-99	1 190 (18.5)	1 652 (21.3)	
CD4 100-199	1 773 (27.5)	2 891 (37.2)	
CD4 200-349	796 (12.3)	549 (7.1)	
CD4 >= 350	471 (7.3)	95 (1.2)	
Valid South African ID number	5 745 (89.1)	7 048 (90.8)	0.001
Cotrimoxazole prescribed during follow-up period	4 325 (67.1)	7 407 (95.4)	<0.001
History of TB at enrolment in HIV programme	683 (10.6)	1 437 (18.5)	<0.001

Abbreviations: ART- Antiretroviral therapy, SD - Standard deviation, IQR- Inter quartile range.

a Comparing patients who did not receive ART by end of follow-up with patients who received ART for complex survey data, adjusted for district strata and clustering within subjects, using hierarchical linear regression for continuous outcomes and logistic regression for binary outcomes

b No-ART: Value from period twelve months before TB diagnosis and three months after TB diagnosis but closest to the TB diagnosis date.

ART: Value between six months before ART was started and fourteen days after ART was started.

4.2. Effect of ART on time to death

The dramatic effect of ART on survival among co-infected patients is shown in Figure 24. The probability of survival after a diagnosis of TB was estimated by the Kaplan Meier method, comparing patients receiving ART and patients not receiving ART by the end of follow-up. There was a significant difference ($p<0.001$) in the median time to death between patients not receiving ART during follow-up and patients receiving ART during follow-up.

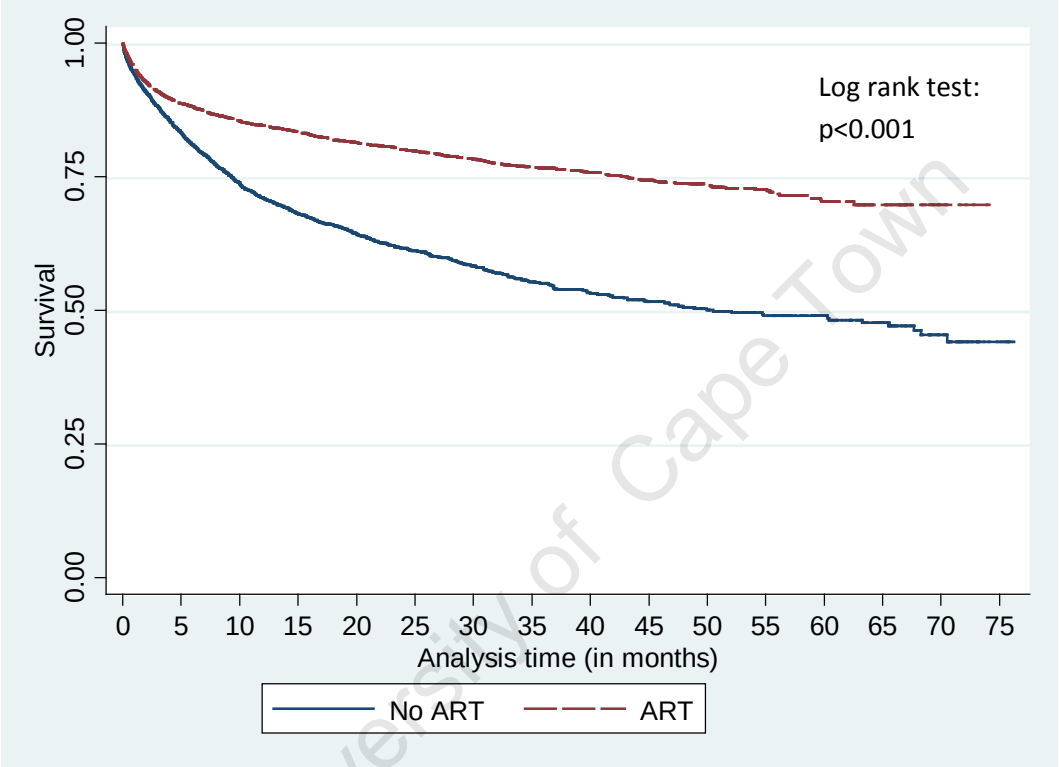

Figure 24: Crude Kaplan-Meier survival estimates by treatment status. For patients who did not receive ART by end of follow-up, time is months since start of TB treatment. For patients who received ART, time is since commencement of ART therapy.

4.2.1. Primary analysis

Marginal structural models were used to evaluate the effect of ART and cotrimoxazole therapy on death in these co-infected patients. Multiple imputation was used to impute missing CD4 cell counts and weights due to almost 20% missing CD4 cell counts and 16% missing weights overall (Table 29). It was decided to use multiple imputation since results from complete case analysis could be biased and not generaliseble to the population.

Table 31 shows results of the marginal structural models (MSM). ART reduced the hazard of death in co-infected patients by 47% (HR 0.53, CI: 0.47-0.61) compared with patients who had

not received ART by the end of follow-up. Men were more at risk of death compared to women (HR 1.29, CI: 1.15-1.45). As CD4 cell count and weight increased, the hazard of death decreased and for each extra month of follow-up since commencement of ART, the hazard of death decreased by 8% (HR 0.92 CI: 0.88-0.95) (Table 31). There was no interaction between ART and CD4 category (p=0.61) at the start of TB treatment, therefore the interaction term was not included in the model. The pooled, un-weighted model showed that ART was effective at reducing the hazard of death by 37% (HR 0.63 CI:0.57-0.69) but less so compared to the weighted marginal structural model (HR 0.53, CI:0.47-0.61). The age effect was large compared to the other covariates since <20 years was used as the reference group, which was a small group and fewer deaths occurred in this group. The hazard ratios for the other covariates in the un-weighted model displayed the same trends as the MSM. The model evaluating the impact of cotrimoxazole therapy on death showed no effect in this analysis (HR 1.02, CI: 0.71-1.46).

Table 31: Factors associated with death using marginal structural regression models (MSM); imputed data [25 sets].

Factors	Primary MSM model for effect of ART on death [a]			Pooled unweighted logistic regression model for effect of ART on death			MSM model for effect of cotrimoxazole on death [b]		
	HR	95% CI	Wald p value	HR	95% CI	Wald p value	HR	95% CI	Wald p value
ART Treatment status									
No ART	**1**			**1**			1		
ART	**0.53**	**0.47 - 0.61**	**<0.001**	**0.63**	**0.57 - 0.69**	**<0.001**	0.63	0.56 - 0.71	<0.001
Cotrimoxazole prescribed during follow-up									
No cotrimoxazole	1			1			**1**		
Cotrimoxazole prescribed	0.92	0.83 - 1.02	0.13	1.07	0.94 - 1.23	0.32	**1.02**	**0.71 - 1.46**	**0.92**
Sex									
Women	1			1			1		
Men	1.29	1.15 - 1.45	<0.001	1.30	1.20 - 142	<0.001	1.50	1.27 - 1.77	<0.001
Age at start of TB (years)									
<20	1			1					
20-29	1.63	0.41 - 6.47	0.48	5.87	2.66 - 12.96	<0.001	5.87	2.40 - 14.34	<0.001
30-39	1.44	0.36 - 5.71	0.60	5.86	2.66 - 12.87	<0.001	4.56	1.84 - 11.34	0.001
40-49	1.47	0.38 - 5.77	0.58	5.96	2.71 - 13.10	<0.001	5.37	2.20 - 13.11	<0.001
50-59	1.82	0.46 - 7.14	0.39	7.30	3.31 - 16.08	<0.001	6.50	2.65 - 15.95	<0.001
CD4 cell count at start of TB or ART (cells/µl) [c]									
CD4 0-49	1			1			1		
CD4 50-99	0.66	0.55 - 0.77	<0.001	0.54	0.46 - 0.65	<0.001	0.77	0.62 - 0.96	0.02
CD4 100-199	0.48	0.40 - 0.57	<0.001	0.33	0.28 - 0.39	<0.001	0.66	0.51 - 0.86	0.002
CD4 200-349	0.42	0.31 - 0.56	<0.001	0.19	0.16 - 0.24	<0.001	0.45	0.34 - 0.59	<0.001
CD4 >350	0.26	0.19 - 0.36	<0.001	0.14	0.11 - 0.18	<0.001	0.39	0.30 - 0.52	<0.001

Factors	Primary MSM model for effect of ART on death [a]			Pooled unweighted logistic regression model for effect of ART on death			MSM model for effect of cotrimoxazole on death [b]		
	HR	95% CI	Wald p value	HR	95% CI	Wald p value	HR	95% CI	Wald p value
TB episode during follow-up									
One TB episode	1			1			1		
Two TB episodes	0.92	0.73- 1.17	0.51	0.96	0.77 - 1.20	0.745	0.91	0.72 - 6.46	0.44
Three TB episodes	1.20	0.80 - 1.81	0.37	1.21	0.82 - 1.78	0.327	2.64	1.07 - 14.34	0.034
Weight at start of TB or ART (kg) [c]									
30-45	1			1			1		
46-52	0.72	0.85 - 0.85	<0.001	0.44	0.38 - 0.51	<0.001	0.87	0.68 - 1.11	0.25
53-59	0.51	0.40 - 0.64	<0.001	0.27	0.22 - 0.33	<0.001	0.75	0.57 - 0.99	0.05
60-160	0.46	0.37 - 0.59	<0.001	0.17	0.13 - 0.21	<0.001	0.58	0.44 - 0.78	<0.001
History of TB at start of TB									
No previous TB	1			1			1		
Previous TB	1.08	0.88 - 1.32	0.48	1.06	0.87 - 1.29	0.52	1.18	0.96 - 1.46	0.12
Follow-up time									
No follow-up time	1			1			1		
Extra month	0.92	0.88 - 0.95	<0.001	0.92	0.89 - 0.95	<0.001	0.95	0.94 - 0.97	<0.001
TB Classification									
Both	1			1			1		
Pulmonary	0.63	0.31 - 1.30	0.21	1.02	0.77 - 1.34	0.89	1.06	0.72 - 1.53	0.77
Extra-pulmonary	0.53	0.26 - 1.11	0.10	0.89	0.69 - 1.16	0.39	1.02	0.71 - 1.46	0.93
No classification	0.50	0.24 - 1.00	0.05	0.80	0.61 - 1.03	0.09	0.68	0.44 - 1.05	0.08

Factors		Primary MSM model for effect of ART on death [a]			Pooled unweighted logistic regression model for effect of ART on death			MSM model for effect of cotrimoxazole on death [b]		
		HR	95% CI	Wald p value	HR	95% CI	Wald p value	HR	95% CI	Wald p value
District										
	Fezile Dabi	1			1			1		
	Lejweleputswa	1.23	0.96 - 1.58	0.09	1.08	0.93 - 1.26	0.27	0.60	0.35 - 1.03	0.07
	Motheo	1.06	0.81 - 1.39	0.65	1.03	0.90 - 1.18	0.65	0.85	0.61 - 1.17	0.33
	Thabo Mofutsanyane	1.43	1.12 - 1.83	0.01	1.21	1.05 - 1.39	0.01	0.87	0.65 - 1.16	0.36
	Xhariep	1.09	0.85 - 1.41	0.49	0.93	0.80 - 1.08	0.36	0.70	0.52 - 0.96	0.03

MSM: Marginal structural model; HR: Hazard ratio; CI: Confidence interval

a weighted by inverse probability of receiving ART

b weighted by inverse probability of receiving cotrimoxazole therapy

c TB: CD4 cell count and weight from period twelve months before TB diagnosis and three months after TB diagnosis but closest to the TB diagnosis date.
ART: CD4 cell count and weight between six months before ART was started and fourteen days after ART was started.

To estimate the effect of ART within each CD4 stratum, an additional MSM model was performed in which an interaction term was included, modelling ART prescription and CD4 cell count at the start of TB treatment, stratified by CD4 categories. Table 32 shows the recalculated hazard ratios and confidence intervals for the interaction term from this model, which are the point estimates and confidence intervals computed for linear combinations of coefficients. This shows that the effect of ART was similar in reducing the hazard of death in the CD4 categories less than 350 cells/µl. The interaction term was not significant (p=0.061).

Table 32: Effect of ART on death within each CD4 category at the start of TB treatment

Factor	**Adjusted MSM model** [a, b, c] (TB from HIV programme and ETR)		
	HR	**95% CI**	**p value**
CD4 0-49 and on ART	0.53	0.47 - 0.61	<0.001
CD4 50-99 and on ART	0.49	0.38 - 0.62	<0.001
CD4 100-199 and on ART	0.61	0.47 - 0.80	<0.001
CD4 200-349 and on ART	0.50	0.31 - 0.79	0.004
CD4 >= 350 and on ART	0.66	0.42 - 1.06	0.09

MSM: Marginal structural model; HR: Hazard ratio; CI: Confidence interval

a MSM adjusting for ART, CD4, CD4×ART interaction, sex, age, weight, history of TB, and for clustering within districts and intra-patient correlation of outcomes.

b Point estimates and confidence intervals were computed for linear combinations of coefficients.

c Test for equality of coefficients for interaction term, p=0.61

Summary statistics of the distribution of stabilised weight for treatment and censorship showed that the MSM dataset of 60 331 484 observations had a mean (standard deviation) of 1.12 (3.56), median (inter quartile range) of 0.93 (0.80-1.12), and a variance of 12.70.

MSMs might be biased by extreme inverse probability weights. Therefore sensitivity analyses were done truncating inverse probability weights at the 1st and 99th percentiles, and the 5th and 95th percentiles (Table 33). This did not affect the estimated effect of ART, but it did affect the estimated effect of cotrimoxazole compared to the primary analyses shown in Table 33. That is, truncation of inverse probability weights at different percentiles resulted in cotrimoxazole significantly reducing mortality.

Table 33: Sensitivity analysis with inverse probability of treatment weights truncated at 1st/99th percentiles or 5th/95th percentiles; imputed data [25 sets].

		Adjusted MSM model [a] Truncated at 1st/99th percentile			Adjusted MSM model [a] Truncated at 5th/95th percentile		
Reference category	**Explanatory variable**	**HR**	**95% CI**	**p-value**	**HR**	**95% CI**	**p value**
Model to evaluate effect of ART on death							
No ART	ART	0.53	0.47 - 0.58	<0.001	0.53	0.48 - 0.59	<0.001
Model to evaluate effect of cotrimoxazole on death							
Cotrimoxazole not prescribed	Cotrimoxazole prescribed	0.83	0.75 - 0.92	<0.001	0.81	0.74 - 0.90	<0.001

a Also adjusted for sex, age, baseline CD4, baseline weight, number of TB episodes, history of TB, follow-up time, TB classification, district and intra-patient correlation of outcomes.

4.2.2. Secondary analysis: Cox regression.

Cox proportional hazards models were used for secondary analyses. A crude Cox model indicated that the hazard of death was reduced by 49% (HR 0.51, CI: 0.48-0.55) in co-infected patients (data not shown in table). Due to a high proportion of patients with a missing CD4 cell count or weight or both values, multiple imputation was used to assign these values. An interaction term was included to model the effect of ART and CD4 cell count at the start of TB treatment. The interaction term was significant (p=0.0031).

Table 34 shows the results for adjusted Cox models with and without an interaction term, using 25 sets of imputed data. The Cox model without the interaction term demonstrated that the hazard of death was reduced by 46% (HR 0.54, CI: 0.49-0.60) for patients receiving ART compared to patients not receiving ART by the end of follow-up, and by 51% (HR 0.49, CI:0.42-0.57) in the model with the interaction term. These effect estimates were very similar to the 47% reduction in hazard of death demonstrated by the marginal structural model (HR 0.53, CI: 0.47-0.61).

Men (HR 1.25, CI: 1.15-1.35), patients older than 20 years (HR>7.09), and patients who had a history of TB at the time they started TB treatment (HR 1.11, CI: 1.01-1.24) had a higher risk of dying. Patients with increasing CD4 cell counts and weights had a lower risk of dying compared to patients with a CD4 cell count below 50 cells/µl or a weight of 30-45kg.

Table 34: Cox proportional hazards model to evaluate the effect of ART on death; imputed data [25 sets].

Factors	Adjusted Cox model [a] (TB from HIV programme and ETR)			Adjusted Cox model with interaction term [a] (TB from HIV programme and ETR)		
	HR	**95% CI**	**p value**	**HR**	**95% CI**	**p value**
ART Treatment status						
No ART	1			1		
ART	0.54	0.49 - 0.60	<0.001	0.49	0.42 - 0.57	<0.001
CD4 cell count at start of TB or ART (cells/µl) [d]						
CD4 0-49	1			1		
CD4 50-99	0.68	0.61 - 0.76	<0.001	0.70	0.60 - 0.82	<0.001
CD4 100-199	0.52	0.47 - 0.58	<0.001	0.47	0.40 - 0.54	<0.001
CD4 200-349	0.37	0.31 - 0.43	<0.001	0.31	0.25 - 0.37	<0.001
CD4 >350	0.29	0.23 - 0.36	<0.001	0.26	0.21 - 0.33	<0.001
Interaction between CD4 category (cells/µl) and ART [b, c]						
CD4<50 and no ART				1		
CD4 50-99 and on ART				0.96	0.78 - 1.19	0.71
CD4 100-199 and on ART				1.21	0.99 - 1.48	0.05
CD4 200-349 and on ART				1.72	1.25 - 2.34	0.001
CD4 >350 and on ART				1.49	0.78 - 2.83	0.225
Cotrimoxazole prescribed ever						
No cotri prescribed	1			1		
Cotri prescribed	0.56	0.50 - 0.63	<0.001	0.58	0.50 - 0.62	<0.001
Sex						
Women	1			1		
Men	1.25	1.15 - 1.35	<0.001	1.25	1.16 - 1.35	<0.001
Age at start of TB (years)						
< 20				1		
20-29	7.09	2.28 - 22.01	0.001	7.07	2.28 - 21.91	0.001
30-39	7.09	2.29 - 21.90	0.001	7.04	2.28 - 21.73	0.001
40-49	8.02	2.60 - 24.83	<0.001	8.01	2.59 - 24.71	<0.001
50-100	9.90	3.18 - 30.52	<0.001	9.83	3.18 - 30.37	<0.001

Factors	Adjusted Cox model [a] (TB from HIV programme and ETR)			Adjusted Cox model with interaction term [a] (TB from HIV programme and ETR)		
	HR	**95% CI**	**p value**	**HR**	**95% CI**	**p value**
Weight category at start of TB or ART (kg) [d]						
30-45	1			1		
46-52	0.72	0.64 - 0.80	<0.001	0.72	0.64 - 0.81	<0.001
53-59	0.60	0.53 - 0.69	<0.001	0.60	0.53 - 0.69	<0.001
60-160	0.50	0.44 - 0.57	<0.001	0.50	0.43 - 0.57	<0.001
History of TB at start of TB						
No previous TB	1			1		
Previous TB	1.11	1.01 - 1.24	0.04	1.12	1.01 - 1.24	0.03
District						
Fezile Dabi	1			1		
Lejweleputswa	0.73	0.36 - 1.47	0.38	0.71	0.35 - 1.45	0.03
Motheo	0.52	0.16 - 1.71	0.28	0.51	0.15 - 1.68	0.32
Thabo Mofutsanyane	0.77	0.31 - 1.92	0.58	0.76	0.30 - 1.88	0.27
Xhariep	0.46	0.21 - 1.03	0.06	0.45	0.20 - 1.01	0.05

HR: Hazard ratio; CI: Confidence interval

a Cox proportional hazards models also adjusting for district, and intra-patient correlation of outcomes.

b Test for equality of coefficients for interaction term, p=0.0031

c Point estimates and confidence intervals were computed for linear combinations of coefficients after the estimation command. Recalculated values displayed in Table 19.

d TB: CD4 cell count and weight from period twelve months before TB diagnosis and three months after TB diagnosis but closest to the TB diagnosis date.

ART: CD4 cell count and weight between six months before ART was started and fourteen days after ART was started.

Adding CD4-ART interaction terms significantly improved the model's fit with the data. (p=0.0031). That is, the estimated effectiveness of ART was different between CD4 strata (Table 35). This shows that ART was most effective in preventing death if baseline CD4 was less than 200 cells/µl, but was less effective with a baseline CD4 of 200 cells/µl or more.

Table 35: Effect of ART on time to death within each CD4 stratum; Imputed data [25 sets].

Factor:	Adjusted Cox model [a] (TB from HIV programme and ETR)		
Interaction between CD4 category (cells/µl) and ART	**HR**	**95% CI**	**p value**
CD4 0-49 and on ART	0.49	0.42 - 0.57	<0.001
CD4 50-99 and on ART	0.47	0.40 - 0.56	<0.001
CD4 100-199 and on ART	0.59	0.51 - 0.69	<0.001
CD4 200-349 and on ART	0.84	0.64 - 1.11	0.211
CD4 >= 350 and on ART	0.73	0.19 - 1.36	0.319

HR: Hazard ratio; CI: Confidence interval

a Adjusting for ART, CD4, CD4×ART interactions, sex, age, weight, history of TB, cotrimoxazole prescription, and for clustering within districts and intra-patient correlation of outcomes. p values for all interaction terms=0.0031

A second set of adjusted Cox proportional hazards models, with and without the interaction term, was used to evaluate the effect of ART restricted to complete case data (Table 36). The results from this complete case analysis were similar to those obtained from both the marginal structural model (Table 31) and imputed Cox models (Table 34). The complete case analysis without the interaction term (Table 36) illustrated that the hazard of death in co-infected patients was 47% less (HR 0.53, CI: 0.47-0.60) for patients who received ART, the marginal structural model showed a 47% reduction in hazard of death in co-infected HIV-TB patients (HR 0.53, CI: 0.47-0.61, Table 31), and the imputed Cox model (without the interaction term) showed a 46% reduction in hazard of death (HR 0.54, CI 0.49-0.60, Table 34). The complete case Cox model with the interaction term included, showed a 53% reduction in hazard of death (HR 0.47, CI: 0.39-0.56, Table 36).

The Cox model showed (Table 36) that men and older patients were more at risk of dying (Men: HR 1.37, CI:1.24-1.51; HR 1.18 (CI: 1.03-1.36) for patients 45-99 years old compared to 16-32 year old patients). Better clinical prognostic factors such as a higher CD4 cell count and weight were associated with better survival (CD4 more than 350 cells/µl: HR 0.28, CI: 0.21-0.36; weight between 60-160 kg: HR 0.50 CI: 0.44-0.58). A history of TB was negatively associated with hazard of death, increasing the hazard of death by 14% (HR 1.14 CI: 1.01-1.29).

Table 36: Factors associated with death in co-infected HIV-TB patients using a Cox proportional hazards model; no imputation.

Factors	Adjusted model [a] (TB from HIV programme and ETR)			Adjusted model with interaction term included [a] (TB from HIV programme and ETR)		
	HR	**95% CI**	**p value**	**HR**	**95% CI**	**p value**
ART Treatment status						
No ART	1			1		
ART	0.53	0.47 - 0.60	<0.001	0.47	0.39 - 0.56	<0.001
Interaction between CD4 category (cells/µl) and ART [b, c]						
CD4<50 and no ART				1		
CD4 50-99 and on ART				0.89	0.70 - 1.14	0.35
CD4 100-199 and on ART				1.39	1.09 - 1.76	0.006
CD4 200-349 and on ART				2.04	1.41 - 2.95	<0.001
CD4 >350 and on ART				1.76	0.85 - 3.65	0.13
CD4 cell count at start of TB or ART (cells/µl) [d]						
CD4 0-49	1			1		
CD4 50-99	0.65	0.57 - 0.73	<0.001	0.69	0.58 - 0.83	<0.001
CD4 100-199	0.51	0.45 - 0.57	<0.001	0.41	0.34 - 0.49	<0.001
CD4 200-349	0.35	0.29 - 0.43	<0.001	0.26	0.21 - 0.33	<0.001
CD4 >350	0.28	0.21 - 0.36	<0.001	0.24	0.18 - 0.32	<0.001
Cotrimoxazole prescribed ever						
No cotrimoxazole prescribed	1			1		
Cotrimoxazole prescribed	0.65	0.55 - 0.75	<0.001	0.64	0.55 - 0.75	<0.001
Sex						
Women	1			1		
Men	1.37	1.24 - 1.51	<0.001	1.37	1.25 - 1.51	<0.001
Age at start of TB or ART (years)						
16-32	1			1		
33-38	1.04	0.91 - 1.20	0.54	1.04	0.91 -1.20	0.56
39-44	1.07	0.93 -1.23	0.37	1.07	0.93 -1.24	0.32
45-99	1.18	1.03 -1.36	0.02	1.19	1.04 -1.37	0.01

Factors	Adjusted model [a] (TB from HIV programme and ETR)			Adjusted model with interaction term included [a] (TB from HIV programme and ETR)		
	HR	**95% CI**	**p value**	**HR**	**95% CI**	**p value**
Weight at start of TB or ART (kg)						
30-45	1			1		
46-52	0.68	0.60 – 0.77	<0.001	0.68	0.6 - 0.77	<0.001
53-59	0.60	0.52 – 0.68	<0.001	0.60	0.53 - 0.68	<0.001
60-160	0.50	0.44 – 0.58	<0.001	0.50	0.44 - 0.58	<0.001
History of TB at start of TB						
No previous TB	1			1		
Previous TB	1.14	1.01 - 1.29	0.03	1.15	1.02 - 1.30	0.03
District						
Fezile Dabi	1			1		
Lejweleputswa	0.71	0.34 - 1.47	0.35	0.70	0.34 - 1.46	0.35
Motheo	0.0	0.0 - 0.0	<0.001	0.0	0.0 - 0.0	<0.001
Thabo Mofutsanyane	0.95	0.35 - 2.62	0.93	0.90	0.33 - 2.49	0.84
Xhariep	0.46	0.19 - 1.13	0.09	0.46	0.19 - 1.10	0.08

a Cox proportional hazards models also adjusting for clinic and intra-patient correlation of outcomes.

b Test for equality of coefficients for interaction term, p<0.001.

c Point estimates and confidence intervals were computed for linear combinations of coefficients after the estimation command. Recalculated values displayed in Table 37.

d TB: CD4 cell count and weight from period twelve months before TB diagnosis and three months after TB diagnosis but closest to the TB diagnosis date.

ART: CD4 cell count and weight between six months before ART was started and fourteen days after ART was started.

The post-estimation analysis restricted to complete case analysis to recalculate the hazard ratios for the interaction term per CD4 stratum (Table 37), showed that ART was effective in preventing death if baseline CD4 was less than 200 cells/µl or more.

Table 37: Effect of ART on death within each CD4 stratum at the start of TB treatment; no imputation.

	Adjusted Cox model [a, b] (TB from HIV programme and ETR)		
Factor: Interaction between CD4 category (cells/µl) and ART	**HR**	**95% CI**	**p value**
CD4 0-49 and on ART	0.47	0.39 – 0.56	<0.001

	Adjusted Cox model [a, b] (TB from HIV programme and ETR)		
Factor: Interaction between CD4 category (cells/µl) and ART	**HR**	**95% CI**	**p value**
CD4 50-99 and on ART	0.41	0.34 – 0.51	<0.001
CD4 100-199 and on ART	0.65	0.54 – 0.78	<0.001
CD4 200-349 and on ART	0.95	0.68 – 1.33	0.767
CD4 >=350 and on ART	0.82	0.40 – 1.67	0.582

HR: Hazard ratio; CI: Confidence interval

a Adjusting for ART, CD4, CD4×ART interaction, sex, age, weight, history of TB, cotrimoxazole prescription, and for clustering within clinic within districts and intra-patient correlation of outcomes.

b Point estimates and confidence intervals were computed for linear combinations of coefficients.

Post estimation graphs were carried out to test the validity of the Cox model. Since the lines were not crossing in the complimentary log-log plot (Figure 25), Cox proportional hazards model were appropriate to use.

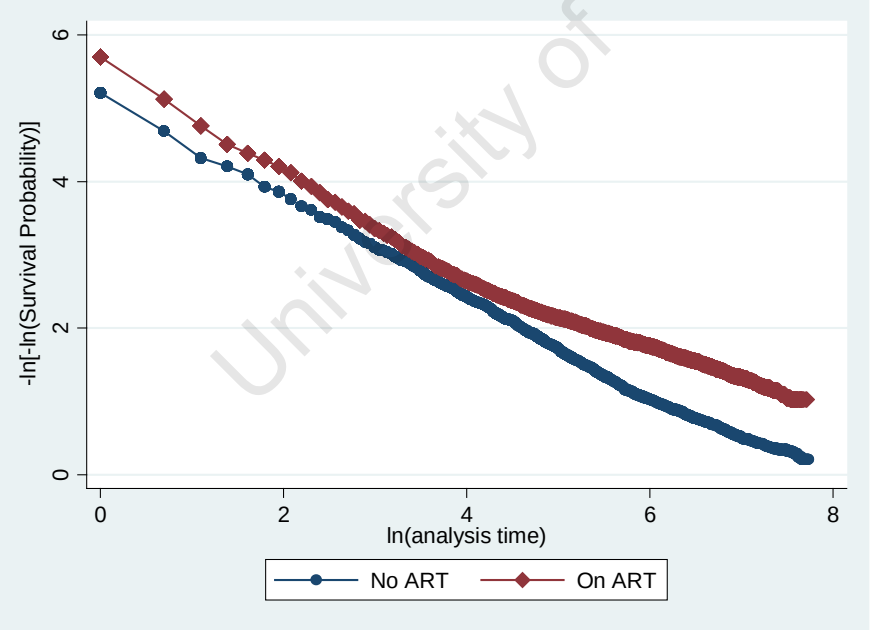

Figure 25: Complimentary log-log plot.

A Cox model using complete case data was used in a sensitivity analysis to evaluate the impact of excluding TB information from the ETR, obtained through linkage (Table 38). The

sensitivity analysis showed that ignoring TB information obtained from the ETR would have over-estimated the effectiveness of ART on time to death. The hazard ratio from the sensitivity analysis was 0.39 (CI: 0.29-0.54) compared to HR 0.47 CI: 0.39-0.56 (Table 38) in the equivalent model which included ETR info. The same trends were seen for age, weight, CD4 cell count and district in the sensitivity model as with the adjusted Cox model (Table 36).

Table 38: Sensitivity analysis: Factors associated with the effect of ART on death using a Cox proportional hazards model with TB information restricted to HIV programme; no imputations.

Factors	**Sensitivity analysis model** [a] (TB from HIV programme database only)		
	HR	**95% CI**	**p value**
ART Treatment status			
No ART	1		
ART	0.39	0.29 - 0.54	<0.001
CD4 cell count at start of TB or ART (cells/µl)			
CD4 0-49	1		
CD4 50-99	0.63	0.46 - 0.87	0.01
CD4 100-199	0.47	0.34 - 0.64	<0.001
CD4 200-349	0.27	0.18 - 0.4	<0.001
CD4 >350	0.29	0.18 - 0.46	<0.001
Cotrimoxazole prescribed ever			
No cotrimoxazole prescribed	1		
Cotrimoxazole prescribed	0.73	0.56 - 0.96	0.02
Interaction between CD4 category (cells/µl) and ART [b]			
CD4<50 and no ART	1		
CD4 50-99 and on ART	1.06	0.7 - 1.61	0.78
CD4 100-199 and on ART	1.23	0.82 - 1.84	0.32
CD4 200-349 and on ART	2.79	1.49 - 5.22	<0.001
CD4 >350 and on ART	1.55	0.45 - 5.36	0.48
Cotrimoxazole prescribed ever			
No cotrimoxazole prescribed	1		
Cotrimoxazole prescribed	0.73	0.56 - 0.96	0.02
Sex			
Women	1		

Factors	Sensitivity analysis model [a] (TB from HIV programme database only)		
	HR	**95% CI**	**p value**
Men	1.38	1.16 - 1.63	<0.001
Age at start of TB or ART (years):			
16-32	1		
33-38	1.21	0.95 - 1.55	0.12
39-44	1.27	1.00 - 1.63	0.05
45-99	1.25	0.98 - 1.6	0.07
CD4 cell count at start of TB or ART (cells/µl)			
CD4 0-49	1		
CD4 50-99	0.63	0.46 - 0.87	0.01
CD4 100-199	0.47	0.34 - 0.64	<0.001
CD4 200-349	0.27	0.18 - 0.4	<0.001
CD4 >350	0.29	0.18 - 0.46	<0.001
Weight at start of TB or ART (kg):			
30-45	1		
46-52	0.66	0.53 - 0.81	<0.001
53-59	0.52	0.41 - 0.65	<0.001
60-160	0.38	0.3 - 0.49	<0.001
History of TB at the start of TB			
No previous TB	1		
Previous TB	0.98	0.78 - 1.23	0.86
District:			
Fezile Dabi	1		
Lejweleputswa	0.56	0.21 - 1.52	0.25
Motheo	0.0	0.0 - 0.0	<0.001
Thabo Mofutsanyane	0.67	0.12 - 3.84	0.66
Xhariep	0.39	0.12 - 1.27	0.12

HR: Hazard ratio; CI: Confidence interval

a Cox proportional hazards models also adjusting for clinic, and intra-patient correlation of outcomes.

b Test for equality of coefficients for interaction term, p<0.001

c No-ART: CD4 cell count and weight from period twelve months before TB diagnosis and three months after TB diagnosis but closest to the TB diagnosis date.

ART: CD4 cell count and weight between six months before ART was started and fourteen days after ART was started.

4.3. Summary of results

Most of the co-infected HIV-TB patients received ART (75%). This co-infected population had an overall median CD4 cell count of 122 cells/µl and 64% of them had a CD4 cell count below 200. More than fifty percent (54.3% (1 455 / 2 679) of the patients who had not received ART by the end of follow-up died, while 19.3% (1 509 / 7 819) of patients who had received ART by the end of follow-up died.

The marginal structural model adjusting for baseline characteristics as well as time-varying covariates showed that ART reduced the hazard of death by 47% (HR 0.53 CI: 0.47-0.61). There was no association between cotrimoxazole and death, probably because of the co-linearity between ART and cotrimoxazole therapy (HR 1.02 CI: 0.71-1.46). The Cox models, adjusting for baseline characteristics only, showed similar effectiveness of ART, producing point estimates corresponding to a reduction in hazard of death of 46% (Cox model with imputed data, HR 0.54 CI: 0.49-0.60), and 47% (Cox model using complete case data, HR 0.53 CI: 0.47-0.60). The two Cox models with an interaction term found similar results overall compared to the models without an interaction term and showed that ART was effective in reducing the hazard of death between 51% and 53%. Using TB information from the HIV programme database (ignoring TB information from ETR linkage) over-estimated the effect of ART on death, a 61% reduction in the hazard of death was found (HR 0.39 CI: 0.29-0.54). The shortage of repeated (and missing) CD4 measurements over time, because most patients were tested every 6-12 months, could be a reason why the results from the MSM and Cox model were very similar. The analyses were not limited by the proportion of patients with missing weight or CD4 values or both. All of the models found that ART reduced the hazard of death between 46% and 49% (MSM: HR 0.53 $p<0.001$, Cox with imputed data: HR 0.54 $p<0.001$, Cox with complete case data HR 0.53 $p<0.001$).

The models also highlighted other associations between risk factors and the risk of death for example men and patients who had TB before enrolling in the HIV programme were at greater risk, and as patients' age increased, their hazard of death increased. The hazard of death decreased for patients with a higher body weight and CD4 cell count at enrolment. Post

estimation analyses by CD4 strata also showed that ART was less effective at preventing death in patients with CD4 counts above 200 cells/µl.

Two district homogeneity trends were seen. In the marginal structural model all the districts were at risk compared to Fezile Dabi district, whereas in the other models all the districts had a lower hazard of mortality compared to Fezile Dabi.

Overall, ART was effective in preventing death in co-infected patients, especially in those with advanced immune-suppression.

Table 39: Estimates effectiveness of ART from different regression models.

Model used	**Effect of ART on survival** (TB from HIV programme and ETR)		
	Effect estimate	**95% CI**	**p value**
Primary analysis: Marginal structural model. Imputed data [25 sets].	HR 0.53	0.47 - 0.61	<0.001
Crude Cox proportional hazards model. No imputation.	HR 0.51	0.48 - 0.55	<0.001
Adjusted Cox proportional hazards model. Imputed data [25 sets].	HR 0.54	0.49 - 0.60	<0.001
Adjusted Cox proportional hazards model. No imputation	HR 0.53	0.47 - 0.60	<0.001
HR: Hazard ratio; CI: Confidence interval			

5. EFFECT OF ART ON TB TREATMENT OUTCOMES

HIV and TB programmes are managed separately due to verticalisation in most health care settings, therefore large sample size surveillance data at the intersection of these programmes are scarce, and currently not available in South Africa. Linkage of the Free State public sector provincial databases made it possible to identify these inter-programme co-infected HIV-TB patients, and to evaluate the effect of ART on their TB treatment outcomes. Linkage with the population registry provided crucial vital status information for patients, providing an opportunity to monitor patients after they have completed TB treatment. HIV database linkages contributed additional co-variates for inclusion in logistic regression models as potential confounders and risk factors associated with TB treatment outcomes.

5.1. Cohort and TB treatment outcomes

In total, 9 276 patients presenting with a first episode of TB between 2004 and 2008, and linked to the HIV programme database, were included in analyses. HIV related information in the ETR was limited and greatly improved with linkage to the HIV programme database. For example, HIV status was recorded for only 23.5% (2 184 / 9 276) of included patients in the ETR, CD4 counts for 14% (1 272 / 9 276) of patients, and death for 8.3% of patients (771 / 9 276) (data not shown in table). Subsequent to linkage with the HIV programme database, which included linkages with the laboratory databases and death population register, HIV status was available for 100% of patients, CD4 counts for 49.4% and death information for 29% of patients (including deaths during and after TB treatment). Patients who had died during TB treatment and had been registered as dead on the ETR increased from 8.3% to 9.8% after the linkage.

Before linkage 11 896 patients on the HIV programme database were identified as being on TB treatment, 73% of them were linked to the ETR (Figure 13). A manual search was conducted, searching for the remainder of unlinked TB patients on the ETR to identify patients and possible reasons why matches could not be identified with prior linking rules. This helped to improve algorithms for linkage.

Table 40 shows baseline characteristics of the 9 276 co-infected HIV-TB patients enrolled in the TB programme between 2004 and 2008. New enrolments of co-infected HIV-TB patients increased substantially from 1 133 patients during 2004 to 2 833 in 2008. Overall, slightly more women (56%) enrolled and the mean age for patients was 36 years. Enrolment age increased slightly during the study period, increasing from 35 to 37 years ($p<0.001$ testing for trend). SA ID numbers were mostly collected by the HIV programme and ascertainment of ID numbers ranged between 84% and 89% during the study period. A significant decrease in ascertainment of ID numbers over time was present ($p<0.001$ testing for trend).

Overall, 75% of patients presented with pulmonary TB, and 22% with extra-pulmonary TB. Patients were more likely to be smear positive (51.7%) and least likely to have no evidence of a smear being taken (17%). New TB cases were registered for 84% of patients and 16% retreatment cases overall. Testing for trend over time shows a significant increase in new TB cases, no smear results, smear negative results and all three types of TB (pulmonary TB, extra-pulmonary TB or both types), while a significant decrease is seen in retreatment cases and smear positive results ($p<0.001$ testing for trend).

Smear positive results decreased from 59% in 2004 to 49% in 2008, new TB cases increased from 80% in 2004 to 86% in 2008, and the proportion of patients with a missing CD4 cell count declined from 80% in 2004 to 41% in 2008 (Table 40). This co-infected population is overall a very immuno-compromised group with a median CD4 cell count as low as 105 cells/µl, IQR 49-184 cells/µl. Of the 4 581 patients with an available CD4 cell count close to the TB treatment start date, 78.5% (3 597) had a CD4 cell count of 200 or less and were therefore eligible for ART at the time of starting TB treatment according to national guidelines at the time (National Antiretroviral Treatment Guidelines. April 2004). Yet, only 41% (1 484) patients eligible for ART on the basis of their CD4 count started ART during their TB treatment episode.

In the combined group (Table 40), 21.5% ((1 534+467) / 9 276) of patients started ART during TB treatment. Of the group who did not start ART during their TB treatment episode, 50% (3 640 / 7 275) had not started ART by June 2010 and 50% (3 635 / 7 275) started ART after their TB episode. From 2006 onwards there was an increase in the number of patients starting

ART early during TB treatment, most likely reflecting policy changes in terms of concomitant treatment of TB and HIV (PALSA PLUS 2006 edition). Testing for trend shows a significant increase in the number of patients starting ART during the first 61 days since TB treatment commenced ($p=0.004$), and also for patients starting ART after two months of TB treatment but before completion of the TB treatment episode ($p<0.001$). From 2006 to 2008, the percentage of co-infected patients starting ART during the first 14 days of TB treatment increased from 7.8% (4 / 51) to 17.6% (42 / 239); similarly patients starting ART between 15 and 28 days of TB treatment increased from 17.7% (6 / 51) to 27.2% (65 / 239) over the same period (data not shown).

Table 40: Baseline characteristics of co-infected HIV-TB patients according to the year in which TB treatment was started

	2004	2005	2006	2007	2008	Combined	
	n (%)	n (%)	n (%)	n (%)	n (%)	N (%)	p value[b]
Number of patients	1 133 (12.2)	1 406 (15.2)	1 773 (19.1)	2 131 (23)	2 833 (30.5)	9 276 (100)	
Women	642 (56.6)	779 (55.4)	1 039 (58.6)	1 210 (56.7)	1 565 (55.2)	5 235 (56.4)	0.477
Age in years: Mean (SD)	35 (8.4)	36 (8.9)	36 (8.8)	37 (9.2)	37 (9.2)	36 (9)	<0.001
Valid South African ID number	1 007 (88.9)	1 249 (88.8)	1 569 (88.5)	1 806 (84.5)	2 387 (84.2)	8 018 (86.4)	<0.001
CD4 cell count at time of starting TB treatment, cells/µl (median, IQR) [a]	102 (52-195)	121 (55-208)	120 (55-213)	97 (46-177)	103 (48-175)	105 (49-184)	
CD4 cell count at time of starting TB treatment per category, cells/µl (median, IQR)							<0.001
Missing CD4	902 (79.6)	938 (66.7)	948 (53.5)	796 (37.4)	1 211 (41.3)	4 695 (50.6)	
0-24	25 (2.2)	56 (4)	105 (5.9)	168 (7.9)	225 (7.7)	579 (6.2)	
25-49	29 (2.6)	49 (3.5)	78 (4.4)	195 (9.2)	229 (7.8)	580 (6.3)	
50-99	59 (5.2)	92 (6.5)	157 (8.9)	327 (15.3)	386 (13.2)	1 021 (11)	
100-199	63 (5.6)	148 (10.5)	260 (14.7)	380 (17.8)	566 (19.3)	1 417 (15.3)	
200-349	33 (2.9)	76 (5.4)	145 (8.2)	155 (7.3)	218 (7.4)	627 (6.8)	
350-2000	22 (1.9)	47 (3.3)	80 (4.5)	110 (5.2)	98 (3.3)	357 (3.8)	
District of enrolment							<0.001
Fezile Dabi District	62 (5.5)	142 (10.1)	191 (10.8)	213 (10)	374 (13.2)	982 (10.6)	
Lejweleputswa District	249 (22)	260 (18.5)	338 (19.1)	451 (21.2)	671 (23.7)	1 969 (21.2)	
Motheo District	409 (36.1)	468 (33.3)	613 (34.6)	645 (30.3)	719 (25.4)	2 854 (30.8)	
Thabo Mofutsanyane	292 (25.8)	373 (26.5)	431 (24.3)	554 (26)	755 (26.7)	2 405 (25.9)	
Xhariep District	121 (10.7)	163 (11.6)	200 (11.3)	268 (12.6)	314 (11.1)	1 066 (11.5)	
Type of TB							0.08
Pulmonary TB	846 (74.7)	1 079 (76.7)	1 287 (72.6)	1 566 (73.5)	2201 (77.7)	6 979 (75.3)	
Extra-pulmonary TB	259 (22.9)	284 (20.2)	23.01 (25.5)	495 (23.2)	578 (20.4)	2 024 (21.8)	
Both types	28 (2.5)	43 (3.1)	78 (4.4)	70 (3.3)	53 (1.9)	272 (2.9)	

	2004	2005	2006	2007	2008	Combined	
	n (%)	n (%)	n (%)	n (%)	n (%)	N (%)	p value[b]
New or Retreatment case							<0.001
New case	909 (80.2)	1 148 (81.7)	1 469 (82.9)	1 827 (85.7)	2 433 (85.9)	7 786 (83.9)	
Retreatment case	224 (19.8)	258 (18.4)	304 (17.2)	304 (14.3)	400 (14.1)	1 490 (16.1)	
Sputum smear results							<0001
Positive smear results	648 (57.2)	833 (59.3)	938 (52.9)	997 (46.8)	1 376 (48.6)	4 792 (51.7)	
Negative smear results	316 (27.9)	385 (27.4)	546 (30.8)	733 (34.4)	896 (31.6)	2 876 (31)	
No evidence of smear	169 (14.9)	188 (13.4)	289 (16.3)	401 (18.8)	561 (19.8)	1 608 (17.3)	
ART category							<0.001
Patients never started ART or started ART after completion of TB treatment episode	1 098 (96.9)	1 295 (92.1)	1 446 (81.6)	1 405 (65.9)	2 031 (71.7)	7 275 (78.4)	
Starting ART during the first 61 of TB treatment	3 (0.3)	10 (0.7)	51 (2.9)	164 (7.7)	239 (8.4)	467 (5)	
Starting ART after 61 days of TB treatment but before TB treatment episode is completed	32 (2.8)	101 (7.2)	276 (15.6)	562 (26.4)	563 (19.9)	1 534 (16.5)	

SD: Standard deviation; CI: Confidence interval

a CD4 test date in the period 12 months prior and three months after TB diagnosis, but closest to the TB diagnosis date if multiple CD4 test dates are available.

b Testing for trend over time: binary and continuous variables were checked using an extension of the Wilcoxon rank sum test, ordered categorical variables were checked with Spearman's rank correlation and nominal categorical variables were checked with multinomial logit models.

Table 41 shows baseline characteristics of patients who did and did not receive ART during their TB treatment. Only 5% (467 / 9 276) of patients received ART during the first 61 days of TB treatment; a further 16.5% (1 534 / 9 276) after two months of TB treatment but before TB treatment was completed, and the majority (78.4%, 7 275 / 9 276) received ART either after the TB treatment episode (50%, 3 635 of 7 275) or had yet to start ART by the end of follow-up in June 2010 (50%, 3 640 of 7 275). The average duration of a TB episode is six months, two months of intensive therapy (consisting of four drugs) and a continuation phase of four months consisting of a two drug regimen. Follow-up time for patients registered in the ETR is limited to the duration of the TB treatment episode, whereas linkage with the HIV database allows for longer follow-up.

A gradient in median CD4 cell count is seen across the three ART categories (Table 41). Patients starting ART during the first 61 days of TB treatment had the lowest median CD4 count of 73 cells/µl (IQR: 32-138 cells/µl), consistent with severe immuno-suppression and in part explaining why they received ART earlier than their counterparts. The highest median CD4 cell count (119 cells/µL IQR: 55-222 cells/µl) was found in the group of patients who did not start ART during their TB treatment or did not start ART at all. This also demonstrates the importance of adjusting for CD4 counts in the analysis of outcomes.

For all ART categories the predominant type of TB was pulmonary, accounting for between 71% (331 / 466) and 76% (5 533 / 7 275) of patients. Patients presenting with both pulmonary and extra-pulmonary TB accounted for between 2.8% and 3.5%. The percentage of new TB cases reported for each of the ART categories was similar between the groups (84% - 85.4%). Smear negative results were more common in patients starting ART during the first 61 days of TB treatment (205 / 467, 43.8%) compared to the other groups; 29.1% (2 117 / 7 275) of patients who did not receive ART during TB treatment, and 36.1% (554 / 1 534) of patients who received ART after 61 days of TB treatment but during TB treatment, had smear negative results.

Table 41: **Baseline characteristics for patients on ART during TB treatment or no ART. ART category divided in patients who started ART in the first 61 days of TB treatment (early ART) or thereafter (late ART).**

	Did not start ART by end of follow-up or started ART after completion of TB	Started ART in the first 61 days of TB treatment (early ART)	Started ART after 61 days of TB treatment but before TB treatment episode is completed (late ART)	p value [a]
	n (%)	**n (%)**	**n (%)**	
No. of patients (N=9276)	7 275 (78.5)	467 (5.0)	1 534 (16.5)	
Women (%) (N=9276)	4 091 (56.2)	270 (57.8)	874 (57.0)	0.717
Age in years (mean, SD)	36 (9.0)	36 (8.6)	37 (8.7)	0.008
Median CD4 cell count closest to starting TB treatment, cells/µl (median, IQR)	119 (55-222)	73 (32-138)	92 (44-155)	<0.001
CD4 cell count closest to starting TB treatment per category, cells/µl (median, IQR) [b]				<0.001
Missing CD4	4 349 (59.8)	60 (12.8)	286 (18.6)	
CD4 0-24	327 (4.5)	78 (16.7)	174 (11.3	
CD4 25-49	329 (4.5)	66 (14.1)	185 (12.1)	
CD4 50-99	609 (8.4)	108 (23.1)	304 (19.8)	
CD4 100-199	823 (11.3)	121 (25.9)	473 (30.8)	
CD4 200-349	510 (7.0)	28 (6.0)	89 (5.8)	
CD4 350-2000	328 (4.5)	6 (1.3)	23 (1.5)	
Type of TB				0.009
Pulmonary TB	5 533 (76.0)	331 (71.0)	1 115 (72.7)	
Extra-pulmonary TB	1 542 (21.2)	117 (25.1)	365 (23.8)	
Both types of TB	200 (2.8)	18 (3.9)	54 (3.5)	
New or Retreatment case				0.662
New case	6 100 (83.9)	399 (85.4)	1 287 (83.9)	
Retreatment case	1 175 (16.1)	68 (14.6)	247 (16.1)	
Sputum smear results				<0.001
Smear positive	3 862 (53.1)	188 (40.3)	742 (48.4)	
Smear negative	2 117 (29.1)	205 (43.8)	554 (36.1)	
No evidence of smear	1 296 (17.8)	74 (15.9)	238 (15.5)	
Died during the year after TB treatment started	1 350 (18.6)	70 (15.0)	174 (11.3)	<0.001

a Comparing ART categories (no ART during TB, ART during first two months of TB treatment, ART after two months) with Pearson's chi-squared test for categorical variables and one-way ANOVA for continuous variables.

b CD4 test date in the period 12 months prior and three months after TB diagnosis but closest to the TB diagnosis date if multiple CD4 test dates are available.

Table 42 shows a summary of TB treatment outcomes recorded in the TB register between 2004 and 2008 for co-infected HIV-TB patients. Treatment success (cured or completed treatment outcome) accounted for the largest proportion of outcomes and declined significantly (p<0.001 testing for trend) from 79.7% (903 / 1 133) in 2004 to 70.9% (2 008 / 2 833) in 2008, with an overall success rate of 73.9% (6 855 / 9 276). The proportion of patients cured during the study period also declined over time (p<0.001 testing for trend), from 40% (452 / 9 276) in 2004 to 34.6% (979 / 9 276) in 2008, with an overall cured outcome of 36.4% (3 373 / 9 276).

Mortality increased over time, with 4.8% (54 / 1 133) deaths recorded in 2004, which increased to 9.6% (273 / 2 833) in 2008, overall mortality of 8.3% (771 / 9 276; p<0.001 testing for trend). TB treatment completion stayed the same over the study period and accounted for 37.5% (3 482 / 9 276) of TB treatment outcomes overall. Other TB treatment outcomes which changed minimally over time include defaulted, failed, not evaluated, and transferred out.

Table 42: TB register outcomes grouped by the year in which TB treatment was started for co-infected HIV-TB patients

	2004	2005	2006	2007	2008	Combined	
	n (%)	n (%)	n (%)	n (%)	n (%)	N	p-value [a]
Number of patients starting TB treatment	1 133 (12.2)	1 406 (15.2)	1 773 (19.1)	2 131 (23)	2 833 (30.5)	9 276 (100)	
TB outcome recorded on ETR register							
Completed	451 (39.8)	527 (37.5)	656 (37)	819 (38.4)	1029 (36.3)	3 482 (37.5)	0.166
Cured	452 (39.9)	589 (41.9)	670 (37.8)	683 (32.1)	979 (34.6)	3 373 (36.4)	<0.001
Success [b]	903 (79.7)	1 116 (79.4)	1 326 (74.8)	1 502 (70.5)	2 008 (70.9)	6 855 (73.9)	<0.001
Defaulted	62 (5.5)	59 (4.2)	58 (3.3)	80 (3.8)	89 (3.1)	348 (3.8)	0.956
Died [c]	54 (4.8)	104 (7.4)	138 (7.8)	202 (9.5)	273 (9.6)	771 (8.3)	<0.001
Failed	17 (1.5)	25 (1.8)	22 (1.2)	37 (1.7)	30 (1.1)	131 (1.4)	0.281
Moved	29 (2.6)	23 (1.6)	114 (6.4)	153 (7.2)	219 (7.7)	538 (5.8)	<0.001
Not Evaluated	3 (0.3)	4 (0.3)	3 (0.2)	4 (0.2)	12 (0.4)	26 (0.3)	0.431
Transferred out	65 (5.7)	75 (5.3)	112 (6.3)	153 (7.2)	202 (7.1)	607 (6.5)	0.077

a Testing for trend over time: binary variables were checked using an extension of the Wilcoxon rank sum test.

b Success is defined as cure or completion as TB outcome.

c Deaths as recorded by the TB register only, no deaths added from HIV programme database.

Table 43 displays summary information for deaths recorded during the first year of TB treatment, grouped by the year when TB treatment commenced and the period when death occurred. ETR recorded deaths during TB treatment, while the group of patients with deaths recorded during the first year since TB treatment commenced, includes deaths during TB as well as deaths recorded after TB treatment was completed, but within one year. Linkage with the HIV programme and death population register provided vital status information for TB patients, specifically after completion of their TB episode. This made it possible to study deaths during the first year after TB treatment commenced.

Additional death information after linkage with the HIV programme database increased the proportion of patients who died during TB treatment by a small number (1.5% additional deaths). Over the study period, deaths occurring during the first year of TB treatment almost doubled, increasing from 10% (115 / 1 133) for patients who had started TB treatment in 2004 and are known to have died, to 17.2% (1 594 / 2 833) in 2008. Importantly, the data show that overall nearly the same number of patients who died during their TB treatment (9.8%, 906 / 9 276), died in the six months immediately after their TB treatment was completed (7.4%, (1 594 – 906) / 9 276).

Table 43: Summary of death information: the time period when death occurred and the source of death information, grouped by the year in which TB treatment was started.

	2004	2005	2006	2007	2008	Combined	
	n (%)	n (%)	n (%)	n (%)	n (%)	N (%)	p value [a]
Number of patients starting TB treatment	1 133 (12.2)	1 406 (15.2)	1 773 (19.1)	2 131 (23.0)	2 833 (30.5)	9 276 (100)	
Number of patients with death registered as their outcome on the TB register	54 (4.8)	104 (7.4)	138 (7.8)	202 (9.5)	273 (9.6)	771 (8.3)	<0.001
Number of patients who died during TB treatment (TB register supplemented by information from the HIV programme database)	62 (5.5)	129 (9.2)	162 (9.1)	238 (11.2)	315 (11.1)	906 (9.8)	<0.001
Number of patients known to have died during the first year after TB treatment was started [b] (information drawn from both TB register and HIV programme database)	115 (10.0)	225 (16.1)	297 (16.7)	395 (18.8)	562 (19.8)	1 594 (17.2)	<0.001

a Testing for trend over time: binary variables were checked using an extension of the Wilcoxon rank sum test

b Includes patients who died during TB and after completion of TB treatment but within one year of TB start date

Table 44 displays baseline characteristics for co-infected patients who survived the first year since TB treatment commenced and those who did not, divided into categories of when death occurred (during TB or after TB). A large proportion of co-infected patients who died during the first year since commencement of TB treatment were not receiving ART (85%, (765+584) / (906+688)). Overall patients presented with a low CD4 cell count; the median CD4 cell count was 74 cells/µl (IQR: 29-140 cells/µl) for patients who died during the first year of TB treatment, compared to a median CD4 cell count of 121 cells/µl (IQR: 58-198 cells/µl) for patients who survived the first year of TB treatment. Overall, women accounted for 46.8% ((414+332) / (906+688)) of the deaths recorded during the first year of TB treatment.

Of patients who died during TB treatment, 6.1% (55 / 906) of deaths occurred in the group who started ART within the first two months of TB treatment, and 9.5% (86 / 906) of deaths occurred in the group who started ART after two months of TB treatment (Table 44). Interestingly, in the group who died after TB treatment completion but within one year since TB treatment commenced, only 2.2% (15 / 688) of patients died who received ART during the first two months of TB treatment, whereas 12.9% (89 / 688) of patients died who started ART after two months of TB treatment but during the TB episode. A large proportion of patients who died during the six months after completing TB treatment (62%, 430 / 688), had a successful TB outcome registered. Migration seemed to be more prevalent among those who died after TB treatment (data not shown), with higher rates of transfer (12%, 82 / 688) and moved (17%, 113 / 668) reported, compared with the overall population which had a transfer rate of 6.5% and moved of 5.8% (Table 42).

Table 44: Baseline characteristics for co-infected HIV-TB patients with no deaths recorded during first year of TB treatment (survivors), deaths during TB treatment and deaths after TB treatment but within one year since TB treatment commenced.

	Survived first year since TB treatment commenced	Died during TB treatment [a]	Died after TB treatment completion but within a year of starting TB treatment [b]	Combined (All three groups)	p value [c]
	n (%)	n (%)	n (%)	N (%)	
Number of patients	7 682 (82.8)	906 (9.8)	688 (7.4)	9 276 (100)	
Women	3 295 (42.9)	414 (45.7)	332 (48.3)	4 041 (43.6)	0.717
TB duration (median, IQR), days	186 (181-215)	92 (43-149)	183 (88-193)	184 (172-210)	<0.001
Median CD4 cell count closest to starting TB treatment, cells/µl (median, IQR)	121 (58-198)	64 (28-111)	74 (29-140)	105 (49-184)	<0.001
CD4 cell count closest to starting TB treatment per category, cells/µl					0.010
0-24	340 (9.6)	151 (23.1)	88 (22.6)	579 (12.6)	
25-49	416 (11.8)	112 (17.2)	52 (13.4)	580 (12.7)	
50-99	737 (20.8)	190 (29.1)	94 (24.2)	1 021 (22.3)	
100-199	1 178 (33.3)	140 (21.4)	99 (25.5)	1 417 (30.9)	
200-349	548 (15.5)	41 (6.3)	38 (9.8)	627 (13.7)	
350->	320 (9)	19 (2.9)	18 (4.6)	357 (7.8)	
Type of TB					0.601
Pulmonary	5 795 (75.4)	666 (73.6)	518 (75.3)	6 979 (75.3)	
Extra-pulomonary	1 662 (21.6)	215 (23.8)	147 (21.4)	2 024 (21.8)	
Both types of TB	225 (2.9)	24 (2.7)	23 (3.3)	272 (2.9)	
New or Retreatment case					0.001
New case	6 433 (83.7)	742 (81.9)	611 (88.8)	7 786 (83.9)	
Retreat case	1 249 (16.3)	164 (18.1)	77 (11.2)	1 490 (16.1)	
Sputum smear results					<0.001
Smear positive	4 103 (53.4)	357 (39.4)	332 (48.3)	4 792 (51.7)	
Smear negative	2 346 (30.5)	326 (36)	204 (29.7)	2 876 (31)	
No evidence of smear	1 233 (16.1)	223 (24.6)	152 (22.1)	1 608 (17.3)	

	Survived first year since TB treatment commenced	Died during TB treatment [a]	Died after TB treatment completion but within a year of starting TB treatment [b]	Combined (All three groups)	p value [c]
	n (%)	n (%)	n (%)	N (%)	
ART category					<0.001
Patients never started ART or started ART after completion of TB treatment episode	5 926 (77.1)	765 (84.4)	584 (84.9)	7 275 (78.4)	
Starting ART during the first 61 days of TB treatment	397 (5.2)	55 (6.1)	15 (2.2)	467 (5)	
Starting ART after 61 days of TB treatment but before TB treatment episode is completed	1 359	86 (9.5)	89 (12.9)	1 534 (16.5)	

a Death information from ETR and HIV database

b Death information from HIV database

c Comparing three death categories (survived first year after TB treatment started, died during TB treatment, died after TB but within one year of TB diagnosis) with Pearson's chi-squared test for categorical variables and one-way ANOVA for continuous variables.

5.2. Effect of ART on TB treatment outcomes in co-infected patients

Table 45 shows the effect of ART initiated either during the first two months of TB treatment or later (>2 months) during TB treatment, compared to ART not being provided during TB treatment, on death (during and after TB treatment but within one year of TB commencement) and other non-fatal TB treatment outcomes. The adjusted models show a significant reduction in death when ART is administered during TB treatment. In a model considering deaths reported by ETR alone, ART initiated within the first two months of TB treatment was associated with reducing the hazard of death by 50% (OR 0.50 CI: 0.35-0.73), while ART initiated after two months of TB treatment, the hazard of dying during TB treatment was reduced by 76% (OR 0.24 CI: 0.18-0.32). The adjusted risk of dying during the first year of TB treatment was decreased by 66% (OR 0.34 CI: 0.26-0.44) when ART was administered during the first two months of TB treatment, and by 71% (OR 0.29 CI: 0.23-0.37) when administered after two months of TB treatment.

It was shown in Table **41** that patients who started ART during the first two months of TB treatment had a lower median CD4 cell count (73 cells/µl IQR: 32-138 cells/µl) compared to patients who started ART after two months of TB treatment (92 cells/µl IQR: 44-155 cells/µl). The overall low median CD4 cell count together with the gradient in odds ratios for early and late ART from the logistic regression models (Table 45), suggest that starting ART later during the course of TB treatment is more effective than starting ART early, which is consistent with survivor bias expected of this observational cohort.

Table 45 further shows that ART initiated during (after two months of) TB treatment improved cured (OR 1.68 CI: 1.42-2.0), completed (OR 3.32 CI: 2.78-3.95), and successful (OR 4.09 CI: 3.28-5.11) TB treatment outcomes. Undesirable outcomes like defaulted (OR 0.60 CI: 0.44-0.82) and failed (OR 0.73 CI: 0.41-1.30) are less likely to occur with concomitant ART and TB treatment.

Table 45: Odds ratios of each TB treatment outcome comparing patients who did not receive ART during their TB treatment, with those who received ART early (first 2 months of TB treatment) or later (between 2 months and end of TB treatment): logistic regression models

			CRUDE [a]		ADJUSTED [b]	
Outcome variable	**Explanatory variable**	**Reference category: No ART during first year of TB treatment**	**OR**	**95% CI**	**OR**	**95% CI**
Number of patients with death registered as their outcome on the TB register [c]	Early ART [f]	1	1.20	0.81-1.76	0.50	0.35-0.73
	Late ART [g]	1	0.45	0.35-0.58	0.24	0.18-0.32
Number of patients who died during TB treatment (TB register supplemented by information from the HIV programme database) [d]	Early ART	1	1.17	0.82-1.68	0.49	0.35-0.68
	Late ART	1	0.52	0.43-0.63	0.28	0.22-0.35
Number of patients known to have died during the first year after TB treatment was started (information drawn from both TB register and HIV programme database) [e]	Early ART	1	0.77	0.57-1.05	0.34	0.26-0.44
	Late ART	1	0.55	0.44-0.68	0.29	0.23-0.37
Cured	Early ART	1	0.64	0.46-0.89	1.06	0.77-1.46
Cured	Late ART	1	1.12	0.96-1.31	1.68	1.42-2.0
Completed	Early ART	1	1.61	1.33-1.95	1.80	1.40-2.30
Completed	Late ART	1	1.81	1.58-2.07	3.32	2.78-3.95

Success	Early ART	1	1.08	0.84-1.41	1.84	1.41-2.41
Success	Late ART	1	2.71	2.23-3.30	4.09	3.28-5.11
Moved	Early ART	1	0.62	0.35-1.08	0.40	0.24-0.69
Moved	Late ART	1	0.06	0.02-0.14	0.04	0.02-0.10
Failed	Early ART	1	0.67	0.27-1.64	0.96	0.36-2.59
Failed	Late ART	1	0.65	0.38-1.11	0.73	0.41-1.30
Defaulted	Early ART	1	0.98	0.72-1.33	1.09	0.80-1.50
Defaulted	Late ART	1	0.56	0.42-0.75	0.60	0.44-0.82

OR: Odds ratio; CI: Confidence interval

a Crude odds ratio adjusted for district and intra-cluster correlation of outcomes

b Also adjusted for district, year when TB treatment started, sex, age, baseline CD4, timing of CD4, anatomical site of TB, sputum smear results, new or retreatment case of TB and intra-clinic correlation

c Death during TB episode; Death source: TB register

d Death during TB episode; Death source: TB register and HIV database

e Death during first year on TB treatment; Death source: TB register and HIV database

f Early ART – ART started during first two months of TB treatment

g Late ART – ART started after the first two months of TB treatment but before TB episode was completed

Table 46 and Table 47 display odds ratios from adjusted logistic regression models on the effect of ART for patients who died during the first year since commencement of TB treatment (includes patient who died during TB treatment and those who died after completion of TB treatment, Table 46), and patients who died during TB treatment (excludes patients who died after TB treatment completed, Table 47). Factors associated with a greater odds of death are patients older than 50 years (OR 1.39 CI: 1.14-1.69 (Table 46); OR 1.46 CI: 1.18-1.8 (Table 47)), and patients with a CD4 cell count less than 100 cells/µl, but particularly severely immuno-compromised patients with a CD4 cell count less than 25 cells/µl (OR 3.64 CI: 2.92-4.55 (Table 46); OR 3.20 CI: 2.61-3.93 (Table 47)).

Retreatment TB cases had a 40% higher odds of death during TB treatment compared to patients diagnosed with a new TB episode, OR 1.40 CI: 1.2-1.62 (Table 47), however this hazard declines after completion of TB treatment (OR 0.88 CI: 0.99-1.12, Table 46). A missing CD4 cell count or a CD4 taken more than 30 days after diagnosis of TB treatment was not associated with an increased odds of death, however patients with no smear result had an increased odds of death (OR 1.31 CI: 1.11-1.54 (Table 46); OR 1.25 CI: 1.00-1.57 (Table 47)). Patients with smear positive sputum results (OR 0.74 CI: 0.62-0.88, Table 46) or extra-pulmonary TB (OR 0.82 CI: 0.69-0.97, Table 46) were associated with better survival.

Table 46: Multivariable model for the effect of ART and baseline covariates on patients who died during the first year of TB treatment, includes patients who died during TB treatment and patients who died after completion of TB treatment.

Reference category	Explanatory variable	OR	P>|t|	95% CI	
Fezile Dabi District	Lejweleputswa District	0.93	0.677	0.66	1.31
	Motheo District	1.24	0.220	0.88	1.75
	Thabo Mofutsanyane	1.20	0.258	0.87	1.66
	Xhariep District	1.15	0.402	0.83	1.59
Year when TB treatment was started: 2006	2004	0.58	<0.001	0.43	0.77
	2005	0.95	0.545	0.80	1.12
	2007	1.05	0.643	0.84	1.32
	2008	1.07	0.501	0.88	1.29
Women	Men	1.10	0.134	0.97	1.24
Age category: 30-39 years	<20	0.67	0.181	0.37	1.20
	20-29	0.96	0.618	0.84	1.11
	40-49	1.05	0.529	0.91	1.21
	>50	1.39	0.001	1.14	1.69
CD4 cell count closest to starting TB treatment: 100-199 cells/µl	CD4 0-24	3.20	<0.001	2.61	3.93
	CD4 25-49	1.81	<0.001	1.50	2.19
	CD4 50-99	1.66	<0.001	1.42	1.94
	CD4 200-349	0.67	0.002	0.52	0.86
	CD4 350-2000	0.58	0.001	0.42	0.80
CD4 test date during the period one year preceding TB diagnosis and 30 days there after	CD4 test date after 30 days of TB treatment	0.70	0.003	0.56	0.89
	CD4 missing	0.36	<0.001	0.29	0.46
Anatomical classification of TB: Pulmonary	Extra-pulmonary	0.82	0.020	0.69	0.97
	Both types of TB	0.97	0.886	0.62	1.50
TB Smear results: Smear negative	Smear positive	0.74	0.001	0.62	0.88
	No smear results	1.31	0.001	1.11	1.54
Treatment case: New case	Retreatment case	0.99	0.909	0.88	1.12
ART use: Did not receive ART during TB treatment episode	Started ART during first two months of TB treatment	0.34	<0.001	0.26	0.44
	Started ART between two months and end of TB treatment	0.29	<0.001	0.23	0.37
No valid South African ID number	Valid South African ID number	1.63	0.013	1.11	2.41

Reference category	Explanatory variable	OR	P>|t|	95% CI	

OR: Odds ratio; CI: Confidence interval

Died during first year of TB treatment; Death source: TB Register and HIV database

Also adjusted for intra-clinic correlation

Table 47: Multivariable model for the effect of ART and baseline covariates on patients who died during TB treatment. [a]

Reference category	Explanatory variable	OR	P>|t|	95% CI	
Fezile Dabi District	Lejweleputswa	0.88	0.467	0.63	1.23
	Motheo	1.27	0.098	0.96	1.69
	Thabo Mofutsanyane	1.50	0.003	1.15	1.97
	Xhariep	1.54	<0.001	1.24	1.91
Year when TB treatment was started: 2006	2004	0.64	0.005	0.46	0.87
	2005	1.01	0.947	0.77	1.33
	2007	1.08	0.582	0.82	1.41
	2008	0.97	0.804	0.78	1.21
Women	Men	1.01	0.845	0.88	1.16
Age category: 30-39 years	<20	0.87	0.693	0.44	1.73
	20-29	0.95	0.575	0.78	1.15
	40-49	1.13	0.145	0.96	1.34
	>50	1.46	<0.001	1.18	1.80
CD4 cell count closest to starting TB treatment: 100-199 cells/µl	CD4 0-24	3.64	<0.001	2.92	4.55
	CD4 25-49	2.25	<0.001	1.73	2.94
	CD4 50-99	1.89	<0.001	1.54	2.31
	CD4 200-349	0.55	0.002	0.37	0.80
	CD4 350-2000	0.40	<0.001	0.26	0.61
CD4 test date during the period one year preceding TB diagnosis and 30 days there after	CD4 test date after 30 days of TB treatment	0.64	<0.001	0.51	0.80
	CD4 missing	0.29	<0.001	0.23	0.36
Anatomical classification of TB: Pulmonary	Extra-pulmonary	0.89	0.306	0.72	1.11
	Both types of TB	0.88	0.652	0.51	1.52
Smear results: Smear negative	Smear positive	0.66	<0.001	0.55	0.80
	No smear results	1.25	0.049	1.00	1.57
Treatment case: New case	Retreatment case	1.40	<0.001	1.20	1.62

Reference category	Explanatory variable	OR	P>|t|	95% CI	
ART use: **Did not receive ART during TB treatment episode**	Started ART during first two months of TB treatment	0.49	<0.001	0.35	0.68
	Started ART between two months and end of TB treatment	0.28	<0.001	0.22	0.35
No valid South African ID number	Valid South African ID number	0.78	0.168	0.55	1.11

OR: Odds ratio; CI: Confidence interval

a Died during TB treatment; Death source: TB Register and HIV database

b Also adjusted for intra-clinic correlation

Table 48 displays a summary of adjusted logistic regression models for the effect of ART on non-fatal TB treatment outcomes. Patients who received ART after two months of TB treatment (late ART), had a 1.7 times better odds of being cured (OR 1.67 CI: 1.42-2.0), a 3.3 times better odds of completing TB treatment (OR 3.32 CI: 2.78-3.95), and a 4 times better odds of a successful treatment outcome (OR 4.09 CI: 3.28-5.11) compared to not being on ART during TB treatment. The first two years (2004 and 2005) were associated with better TB treatment outcomes for cured, completed and success compared to 2006, while 2007 and 2008 were associated with poorer outcomes. Although not significant, men were more at risk of failing (OR 1.36 CI: 0.96-1.92) and defaulting (OR 1.15 CI: 0.98-1.34) compared to women.

Patients older than fifty years were associated with an 18% odds of not being cured (OR 0.82 CI: 0.68-0.99) and the odds of defaulting almost doubled (OR 1.73 CI: 0.99-3.01) for patients younger than 20 years old compared to patients between 30 and 39 years old (Table 48). Immuno-compromised patients (CD4 cell count less than 100 cells/µl) had a lower odds of being cured, completing TB treatment or a successful TB treatment outcome (ORs between 0.45 and 0.84), while patients with a CD4 cell count above 200 were associated with better cured, completed and successful TB outcomes compared to patients with a CD4 cell count between 100-199 cells/µl (ORs between 1.12 and 1.57). Patients who failed treatment were four times more likely to fail if they were smear positive compared to smear negative patients (OR 3.95 CI: 1.62-9.63).

Subgroup analyses were done to evaluate the effect of ART on TB outcomes (cured, completed, successful, failed) for patients who survived TB treatment (excluding all deceased patients), and

found a similar gradient in odds ratios for starting ART early and starting ART after two months of TB treatment. Logistic regression models for subgroup analyses are provided in Appendix A.

Table 49 shows a summary of logistic regression models evaluating the effect of ART on TB treatment outcomes, when the two ART categories were combined (ART initiated within two months of TB treatment and ART initiated after two months of TB treatment). The trends are very similar as discussed for Table 46 and Table 48. Overall, the odds of death during the first year since commencement of TB treatment was reduced by 70% (OR 0.30 CI: 0.25-0.38), and ART was associated with improved favourable outcomes: cured (OR 1.52 CI: 1.27-1.83), completed (OR 2.84 CI: 2.37-3.42), success (OR 3.25 CI: 2.62-4.06).

Table 48: A summary (odds ratio and confidence interval) of multivariable regression models for the effect of ART and baseline covariates on other TB treatment outcomes. Individual detailed tables are provided for TB outcomes listed in this table. See Table 50 to Table 55.

		Cured	Completed	Success	Failed	Defaulted	Moved
Reference category	**Explanatory variable**	**OR (CI)**	**OR (CI)**	**OR (CI)**	**OR (CI)**	**OR (CI)**	**OR (CI)**
Fezile Dabi District	Lejweleputswa	0.88 (0.69-1.14)	1.19 (0.76-1.86)	1.34 (0.95-1.90)	0.42 (0.16-1.09)	0.83 (0.53-1.30)	0.80 (0.41-1.54)
	Motheo	0.50 (0.39-0.64)	1.57 (1.03-2.39)	1.41 (0.99-2.02)	0.79 (0.35-1.79)	0.80 (0.51-1.23)	0.40 (0.24-0.67)
	Thabo Mofutsanyane	0.95 (0.75-1.19)	1.46 (0.93-2.28)	1.64 (1.15-2.33)	0.66 (0.28-1.53)	0.63 (0.43-0.93)	0.25 (0.14-0.44)
	Xhariep	0.40 (0.28-0.58)	1.41 (0.89-2.24)	1.45 (0.99-2.12)	0.63 (0.26-1.53)	1.02 (0.63-1.65)	0.05 (0.02-0.13)
Year when TB treatment was started: 2006	2004	1.05 (0.88-1.26)	1.75 (1.32-2.32)	1.41 (1.13-1.75)	1.05 (0.55-2.03)	1.15 (0.81-1.61)	0.35 (0.21-0.58)
	2005	1.11 (0.96-1.28)	1.54 (1.18-2.01)	1.33 (1.08-1.62)	1.24 (0.71-2.18)	0.94 (0.76-1.17)	0.24 (0.16-0.35)
	2007	0.74 (0.61-0.88)	0.81 (0.65-1.01)	0.80 (0.67-0.95)	1.76 (0.98-3.14)	1.19 (1.0-1.43)	1.20 (0.86-1.65)
	2008	0.79 (0.66-0.94)	0.85 (0.69-1.05)	0.92 (0.77-1.10)	1.02 (0.61-1.70)	1.10 (0.92-1.31)	1.09 (0.75-1.58)
Women	Men	0.96 (0.86-1.07)	1.12 (1.0-1.25)	0.95 (0.84-1.06)	1.36 (0.96-1.92)	1.15 (0.98-1.34)	0.88 (0.72-1.07)
Age category: 30-39 years	<20	1.12 (0.68-1.87)	0.82 (0.44-1.55)	0.80 (0.47-1.36)	1.19 (0.25-5.61)	1.73 (0.99-3.01)	0.89 (0.38-2.11)
	20-29	0.99 (0.87-1.14)	0.95 (0.84-1.07)	0.93 (0.82-1.04)	0.89 (0.61-1.29)	1.08 (0.90-1.30)	1.34 (1.07-1.68)
	40-49	1.02 (0.90-1.15)	0.91 (0.78-1.06)	1.03 (0.94-1.14)	0.78 (0.50-1.23)	0.84 (0.72-0.99)	1.05 (0.88-1.26)
	>50	0.82 (0.68-0.99)	0.96 (0.79-1.17)	0.89 (0.76-1.05)	1.01 (0.54-1.87)	0.84 (0.67-1.06)	1.07 (0.76-1.52)
CD4 cell count closest to starting TB treatment: 100-199 cells/µl	CD4 0-24	0.65 (0.52-0.81)	0.52 (0.39-0.70)	0.45 (0.36-0.56)	1.44 (0.58-3.58)	1.04 (0.73-1.46)	1.28 (0.87-1.88)
	CD4 25-49	0.79 (0.65-0.96)	0.75 (0.59-0.95)	0.65 (0.55-0.78)	1.16 (0.54-2.48)	1.03 (0.77-1.36)	1.10 (0.76-1.60)
	CD4 50-99	0.84 (0.71-0.99)	0.84 (0.67-1.04)	0.75 (0.63-0.89)	1.20 (0.59-2.44)	0.90 (0.70-1.16)	1.13 (0.81-1.57)
	CD4 200-349	1.12 (0.91-1.37)	1.33 (1.05-1.69)	1.33 (1.06-1.69)	1.10 (0.50-2.41)	0.88 (0.64-1.22)	0.91 (0.62-1.33)
	CD4 350-2000	1.29 (1.04-1.59)	1.51 (1.11-2.06)	1.57 (1.20-2.07)	1.36 (0.65-2.85)	0.81 (0.54-1.21)	0.71 (0.41-1.24)
CD4 test date during the period one year preceding TB diagnosis and 30 days there after	CD4 test date after 30 days of TB treatment	1.39 (1.21-1.60)	0.85 (0.67-1.08)	1.03 (0.83-1.29)	1.53 (0.7-3.38)	1.67 (1.2-2.31)	0.75 (0.47-1.20)
	CD4 missing	1.67 (1.42-1.96)	1.41 (1.21-1.64)	1.60 (1.32-1.96)	1.48 (1.0-2.18)	1.44 (1.18-1.76)	0.76 (0.57-1.02)

		Cured	Completed	Success	Failed	Defaulted	Moved
Reference category	**Explanatory variable**	**OR (CI)**	**OR (CI)**	**OR (CI)**	**OR (CI)**	**OR (CI)**	**OR (CI)**
Anatomical classification of TB: Pulmonary	Extra-pulmonary	0.01 (0-0.01)	1.27 (1.09-1.49)	1.20 (1.03-1.40)	0.26 (0.07-0.99)	0.93 (0.73-1.18)	0.88 (0.70-1.09)
	Both TB types	0.65 (0.47-0.91)	1.41 (1.07-1.86)	1.12 (0.88-1.42)	0.93 (0.36-2.44)	0.94 (0.68-1.3)	1.16 (0.57-2.38)
Smear results: Smear negative	Smear positive		0.02 (0.02-0.03)	1.19 (1.04-1.37)	3.95 (1.62-9.63)	0.90 (0.71-1.13)	0.87 (0.65-1.17)
	No smear results		0.43 (0.35-0.53)	0.44 (0.35-0.54)	0.83 (0.31-2.23)	1.38 (1.11-1.71)	4.80 (3.16-7.30)
Treatment case: New case	Retreatment case	0.64 (0.56-0.72)	0.99 (0.83-1.19)	0.79 (0.70-0.89)	1.43 (0.86-2.37)	1.16 (0.99-1.35)	0.74 (0.56-0.98)
ART use: Did not receive ART during TB treatment episode	Started ART during first two months of TB treatment	1.06 (0.77-1.46)	1.80 (1.40-2.30)	1.84 (1.41-2.41)	0.96 (0.36-2.59)	1.09 (0.80-1.50)	0.40 (0.24-0.69)
	Started ART between two months and end of TB treatment	1.68 (1.42-2.0)	3.32 (2.78-3.95)	4.09 (3.28-5.11)	0.73 (0.41-1.30)	0.60 (0.44-0.82)	0.04 (0.02-0.10)
No valid South African ID number	Valid SA ID number	1.15 (1.01-1.31)	1.16 (0.98-1.37)	1.25 (1.08-1.43)	0.96 (0.55-1.67)	0.72 (0.58-0.91)	1.19 (0.84-1.69)

OR: odds ratio; CI: Confidence interval

Also adjusted for intra-clinic correlation

Table 49: A summary (odds ratio and confidence interval) of multivariable regression models for the effect of ART and baseline covariates on other TB treatment outcomes. ART coded as a binary variable, ART received during TB treatment or not.

		Died [a]	Cured	Completed	Success	Failed	Defaulted	Moved
Reference category	**Explanatory variable**	**OR (CI)**	**OR (CI)**	**OR (CI)**	**OR (CI)**	**OR (CI)**	**OR (CI)**	**OR (CI)**
Fezile Dabi District	Lejweleputswa	0.93 (0.67-1.31)	0.88 (0.69-1.15)	1.18 (0.76-1.85)	1.33 (0.95-1.9)	0.42 (0.17-1.10)	0.83 (0.54-1.31)	0.81 (0.42-1.56)
	Motheo	1.24 (0.88-1.75)	0.51 (0.40-0.65)	1.58 (1.04-2.41)	1.42 (1.0-2.03)	0.79 (0.35-1.79)	0.79 (0.52-1.23)	0.40 (0.24-0.68)
	Thabo Mofutsanyane	1.20 (0.88-1.67)	0.95 (0.76-1.19)	1.45 (0.93-2.29)	1.63 (1.15-2.33)	0.66 (0.29-1.54)	0.63 (0.44-0.93)	0.25 (0.14-0.45)
	Xhariep	1.15 (0.83-1.60)	0.40 (0.28-0.59)	1.40 (0.89-2.24)	1.45 (1.0-2.12)	0.63 (0.27-1.53)	1.02 (0.63-1.65)	0.05 (0.02-0.14)
Year when TB treatment was started: 2006	2004	0.58 (0.44-0.77)	1.05 (0.88-1.27)	1.73 (1.31-2.3)	1.40 (1.13-1.75)	1.06 (0.55-2.04)	1.15 (0.82-1.63)	0.35 (0.22-0.58)
	2005	0.95 (0.81-1.13)	1.11 (0.96-1.28)	1.53 (1.18-2.01)	1.33 (1.09-1.63)	1.24 (0.71-2.19)	0.94 (0.76-1.18)	0.24 (0.17-0.36)
	2007	1.05 (0.85-1.32)	0.73 (0.62-0.88)	0.81 (0.65-1.01)	0.80 (0.67-0.96)	1.76 (0.99-3.15)	1.2 (1.01-1.44)	1.19 (0.87-1.66)
	2008	1.07 (0.89-1.30)	0.78 (0.66-0.93)	0.84 (0.69-1.04)	0.91 (0.76-1.09)	1.03 (0.62-1.71)	1.11 (0.94-1.33)	1.10 (0.76-1.61)
Women	Men	1.10 (0.98-1.24)	0.96 (0.87-1.07)	1.12 (1.01-1.25)	0.95 (0.85-1.07)	1.36 (0.97-1.92)	1.15 (0.98-1.35)	0.88 (0.72-1.07)
Age category: 30-39 years	<20	0.67 (0.38-1.21)	1.12 (0.68-1.85)	0.82 (0.44-1.55)	0.79 (0.47-1.35)	1.20 (0.26-5.65)	1.74 (1.01-3.02)	0.88 (0.38-2.10)
	20-29	0.97 (0.84-1.12)	0.99 (0.87-1.14)	0.95 (0.85-1.07)	0.92 (0.83-1.04)	0.89 (0.62-1.30)	1.09 (0.91-1.31)	1.35 (1.09-1.70)
	40-49	1.05 (0.91-1.21)	1.03 (0.91-1.16)	0.91 (0.79-1.07)	1.04 (0.95-1.15)	0.78 (0.50-1.23)	0.84 (0.72-0.99)	1.04 (0.88-1.25)
	>50	1.38 (1.14-1.69)	0.82 (0.69-1.0)	0.97 (0.80-1.19)	0.90 (0.77-1.06)	1.01 (0.55-1.88)	0.84 (0.67-1.06)	1.07 (0.77-1.52)

		Died [a]	Cured	Completed	Success	Failed	Defaulted	Moved
Reference category	**Explanatory variable**	**OR (CI)**	**OR (CI)**	**OR (CI)**	**OR (CI)**	**OR (CI)**	**OR (CI)**	**OR (CI)**
CD4 cell count closest to starting TB treatment: 100-199 cells/µl	CD4 0-24	3.21 (2.62-3.95)	0.64 (0.52-0.80)	0.51 (0.39-0.69)	0.44 (0.36-0.55)	1.44 (0.58-3.60)	1.05 (0.75-1.49)	1.31 (0.90-1.91)
	CD4 25-49	1.82 (1.51-2.20)	0.78 (0.65-0.96)	0.75 (0.60-0.95)	0.65 (0.55-0.78)	1.16 (0.55-2.49)	1.03 (0.78-1.38)	1.10 (0.76-1.61)
	CD4 50-99	1.66 (1.43-1.95)	0.83 (0.71-0.99)	0.83 (0.67-1.04)	0.74 (0.63-0.89)	1.20 (0.60-2.44)	0.91 (0.71-1.17)	1.13 (0.82-1.59)
	CD4 200-349	0.67 (0.53-0.86)	1.11 (0.91-1.37)	1.33 (1.05-1.69)	1.33 (1.06-1.68)	1.10 (0.50-2.42)	0.89 (0.65-1.22)	0.91 (0.63-1.33)
	CD4 350-2000	0.58 (0.43-0.80)	1.28 (1.04-1.59)	1.50 (1.11-2.05)	1.57 (1.20-2.06)	1.36 (0.65-2.85)	0.81 (0.55-1.22)	0.71 (0.41-1.24)
CD4 test date during the period one year preceding TB diagnosis and 30 days there after	CD4 test date after 30 days of TB treatment	0.70 (0.56-0.89)	1.45 (1.25-1.69)	0.89 (0.71-1.13)	1.09 (0.88-1.37)	1.50 (0.7-3.22)	1.59 (1.15-2.2)	0.71 (0.44-1.14)
	CD4 missing	0.36 (0.29-0.46)	1.69 (1.44-1.99)	1.42 (1.23-1.66)	1.62 (1.34-1.98)	1.46 (1.01-2.15)	1.42 (1.17-1.73)	0.75 (0.57-1.01)
Anatomical classification of TB: Pulmonary	Extra-pulmonary	0.82 (0.69-0.97)	0.01 (0.01-0.02)	1.28 (1.09-1.50)	1.20 (1.04-1.41)	0.26 (0.07-1.0)	0.93 (0.74-1.18)	0.87 (0.7-1.09)
	Both TB types	0.97 (0.63-1.51)	0.65 (0.47-0.91)	1.4 0 (1.07-1.84)	1.11 (0.88-1.41)	0.94 (0.36-2.44)	0.95 (0.69-1.31)	1.18 (0.59-2.42)
Smear results: Smear negative	Smear positive	0.74 (0.63-0.88)		0.02 (0.02-0.03)	1.20 (1.05-1.39)	3.94 (1.62-9.65)	0.89 (0.71-1.12)	0.86 (0.64-1.16)
	No smear results	1.31 (1.12-1.55)		0.43 (0.35-0.54)	0.44 (0.36-0.55)	0.83 (0.32-2.23)	1.37 (1.12-1.70)	4.75 (3.10-7.28)
Treatment case: New case	Retreatment case	0.99 (0.89-1.12)	0.64 (0.57-0.73)	0.99 (0.83-1.2)	0.79 (0.71-0.89)	1.43 (0.87-2.36)	1.15 (0.99-1.36)	0.74 (0.57-0.99)
ART use: Did not receive ART during TB treatment episode	Started ART during TB treatment	0.30 (0.25-0.38)	1.52 (1.27-1.83)	2.84 (2.37-3.42)	3.25 (2.62-4.06)	0.77 (0.44-1.37)	0.71 (0.55-0.92)	0.13 (0.09-0.21)

		Died [a]	Cured	Completed	Success	Failed	Defaulted	Moved
Reference category	**Explanatory variable**	**OR (CI)**	**OR (CI)**	**OR (CI)**	**OR (CI)**	**OR (CI)**	**OR (CI)**	**OR (CI)**
No valid South African ID number	Valid SA ID number	1.63 (1.11-2.41)	1.15 (1.02-1.32)	1.17 (0.99-1.39)	1.26 (1.10-1.45)	0.96 (0.55-1.67)	0.72 (0.58-0.90)	1.17 (0.83-1.67)

OR: odds ratio; CI: Confidence interval

a Died – patients who died within one year of TB treatment diagnosis, either during TB treatment or thereafter.

Also adjusted for intra-clinic correlation

5.3. Summary of results

A large proportion of co-infected patients were eligible for ART at time of starting TB treatment, yet only 1 484 / 4 581 (41%) received ART during treatment, and for most it was after two months of TB treatment, despite guideline changes in 2006 recommending earlier ART (PALSA PLUS 2006 edition). Patients who did not receive any ART during their TB treatment had the highest percentage of deaths during the first year 1 350 / 7 275 (18.6%), followed by patients who started ART early during their treatment 70 / 467 (15%) and lastly patients who started ART later during their TB treatment 174 / 1 534 (11.3%). Surprisingly, for patients not receiving ART during TB treatment the percentage of deaths during the first year increased from 10.4% (114 / 1 098 patients registered) in 2004 to 23.4 % (476 / 2 031 patients registered) in 2008, despite better access to ART (Uebel KE *et al*. 2010), change in policy to start ART earlier (NDOH NSP 2007-2011), and a shift towards integrated TB and HIV services (Data by year not shown).

Almost as many patients who died during TB treatment (9.8%) died in the six month period immediately after TB treatment was completed (7.4%). A large proportion of these patients who died soon after TB treatment was successfully completed, had not yet started ART. For them migration may in part explain failure to start ART as high rates of moved and transferred out outcomes were reported.

Treatment success in co-infected HIV TB patients worsened during the study period (80% in 2004, 71% in 2008) and was lower than the WHO proposed target of 85% (WHO guidelines, 2011). Deaths in this co-infected group increased from 4.8% in 2004 to 9.6% in 2008. The proportion of TB patients co-infected with HIV increased from 12.2% to 30.5% over the study period and could possibly have an impact on the quality of service facilities can provide due to the patient load (Table 42). Even though this dataset is a sub-sample of the patients in the ETR, the total number of records in the ETR increased from 8.2% (16 395 / 199 998) in 2003 to 16.1% (32 299 / 199 998) in 2008, confirming the increasing caseload. TB remains a challenging public health problem although more healthcare resources have been made available (Uebel KE *et al*., 2010).

In summary, ART was extremely effective in reducing the odds of death (either during TB treatment or thereafter but within one year of TB diagnosis), and improving favourable TB outcomes (cured, completed, success). Results from this observational cohort of co-infected

patients suggest that starting ART after two months of TB treatment is more effective than starting ART early during TB treatment but this is likely due to survivor bias and a small sample size for patients starting ART early (5%, 467 / 9 276). Severely immuno-compromised patients were probably prioritised to start ART early and more likely to die during treatment, and 'healthier' patients were more likely to start ART later during TB treatment with a better chance of survival, contributing to immortal bias (Sterne JA *et al.*, 2009; Levesque LE *et al.*, 2010).

Patients registered as retreatment cases were more at risk of unfavourable outcomes, they had a higher odds of death, failing or defaulting and were less likely to be cured, to complete treatment or to have a successful TB treatment outcomes. Overall being male and older than 50 years were identified as a risk factor for most of the TB treatment outcomes.

Table 50: Multivariable regression model for the effect of ART and baseline covariates on cured.

Reference category	Explanatory variable	OR	P>|t|	95% Confidence Interval	
Fezile Dabi District	Lejweleputswa	0.88	0.347	0.69	1.14
	Motheo	0.50	<0.001	0.39	0.64
	Thabo Mofutsanyane	0.95	0.634	0.75	1.19
	Xhariep	0.40	<0.001	0.28	0.58
Year when TB treatment was started: 2006	2004	1.05	0.585	0.88	1.26
	2005	1.11	0.165	0.96	1.28
	2007	0.74	0.001	0.61	0.88
	2008	0.79	0.008	0.66	0.94
Women	Men	0.96	0.47	0.86	1.07
Age category: 30-39 years	<20	1.12	0.648	0.68	1.87
	20-29	0.99	0.934	0.87	1.14
	40-49	1.02	0.737	0.90	1.15
	>50	0.82	0.038	0.68	0.99
CD4 cell count closest to starting TB treatment: 100-199 cells/µl	CD4 0-24	0.65	<0.001	0.52	0.81
	CD4 25-49	0.79	0.018	0.65	0.96
	CD4 50-99	0.84	0.043	0.71	0.99
	CD4 200-349	1.12	0.295	0.91	1.37
	CD4 350-2000	1.29	0.02	1.04	1.59
CD4 test date during the period one year preceding TB diagnosis and 30 days there after	CD4 test date after 30 days of TB treatment	1.39	<0.001	1.21	1.60
	CD4 missing	1.67	<0.001	1.42	1.96
Anatomical classification of TB: Pulmonary	Extra-pulmonary	0.01	<0.001	0.00	0.01
	Both TB types	0.65	0.012	0.47	0.91
Treatment case: New case	Retreatment case	0.64	<0.001	0.56	0.72
ART use: Did not receive ART during TB treatment episode	Started ART during first two months of TB treatment	1.06	0.74	0.77	1.46
	Started ART between two months and end of TB treatment	1.68	<0.001	1.42	2.00
No valid South African ID number	Valid South African ID number	1.15	0.036	1.01	1.31
Also adjusted for intra-clinic correlation					

Table 51: Multivariable regression model for the effect of ART and baseline covariates on completed TB outcome.

Reference category	Explanatory variable	OR	P>|t|	95% Confidence Interval	
Fezile Dabi District	Lejweleputswa	1.19	0.454	0.76	1.86
	Motheo	1.57	0.036	1.03	2.39
	Thabo Mofutsanyane	1.46	0.097	0.93	2.28
	Xhariep	1.41	0.147	0.89	2.24
Year when TB treatment was started: 2006	2004	1.75	<0.001	1.32	2.32
	2005	1.54	0.001	1.18	2.01
	2007	0.81	0.06	0.65	1.01
	2008	0.85	0.133	0.69	1.05
Women	Men	1.12	0.043	1.00	1.25
Age category: 30-39 years	<20	0.82	0.542	0.44	1.55
	20-29	0.95	0.39	0.84	1.07
	40-49	0.91	0.205	0.78	1.06
	>50	0.96	0.687	0.79	1.17
CD4 cell count closest to starting TB treatment: 100-199 cells/µl	CD4 0-24	0.52	<0.001	0.39	0.70
	CD4 25-49	0.75	0.018	0.59	0.95
	CD4 50-99	0.84	0.111	0.67	1.04
	CD4 200-349	1.33	0.019	1.05	1.69
	CD4 350-2000	1.51	0.009	1.11	2.06
CD4 test date during the period one year preceding TB diagnosis and 30 days there after	CD4 test date after 30 days of TB treatment	0.85	0.177	0.67	1.08
	CD4 missing	1.41	<0.001	1.21	1.64
Anatomical classification of TB: Pulmonary	Extra-pulmonary	1.27	0.003	1.09	1.49
	Both TB types	1.41	0.016	1.07	1.86
Smear results: Smear negative	Smear positive	0.02	<0.001	0.02	0.03
	No smear results	0.43	<0.001	0.35	0.53
Treatment case: New case	Retreatment case	0.99	0.93	0.83	1.19
ART use: Did not receive ART during TB treatment episode	Started ART during first two months of TB treatment	1.80	<0.001	1.40	2.30
	Started ART between two months and end of TB treatment	3.32	<0.001	2.78	3.95
No valid South African ID number	Valid South African ID number	1.16	0.085	0.98	1.37
Also adjusted for intra-clinic correlation					

Table 52: Multivariable regression model for the effect of ART and baseline covariates on successful TB outcome. Cured and completed are combined for successful treatment outcomes.

Reference category	Explanatory variable	OR	P>|t|	95% Confidence Interval	
Fezile Dabi District	Lejweleputswa	1.34	0.099	0.95	1.90
	Motheo	1.41	0.057	0.99	2.02
	Thabo Mofutsanyane	1.64	0.006	1.15	2.33
	Xhariep	1.45	0.054	0.99	2.12
Year when TB treatment was started: 2006	2004	1.41	0.002	1.13	1.75
	2005	1.33	0.006	1.08	1.62
	2007	0.80	0.013	0.67	0.95
	2008	0.92	0.361	0.77	1.10
Women	Men	0.95	0.359	0.84	1.06
Age category: 30-39 years	<20	0.80	0.409	0.47	1.36
	20-29	0.93	0.212	0.82	1.04
	40-49	1.03	0.534	0.94	1.14
	>50	0.89	0.159	0.76	1.05
CD4 cell count closest to starting TB treatment: 100-199 cells/µl	CD4 0-24	0.45	<0.001	0.36	0.56
	CD4 25-49	0.65	<0.001	0.55	0.78
	CD4 50-99	0.75	0.001	0.63	0.89
	CD4 200-349	1.33	0.016	1.06	1.69
	CD4 350-2000	1.57	0.001	1.20	2.07
CD4 test date during the period one year preceding TB diagnosis and 30 days there after	CD4 test date after 30 days of TB treatment	1.03	0.79	0.83	1.29
	CD4 missing	1.60	<0.001	1.32	1.96
Anatomical classification of TB: Pulmonary	Extra-pulmonary	1.20	0.02	1.03	1.40
	Both TB types	1.12	0.35	0.88	1.42
Smear results: Smear negative	Smear positive	1.19	0.014	1.04	1.37
	No smear results	0.44	<0.001	0.35	0.54
Treatment case: New case	Retreatment case	0.79	<0.001	0.70	0.89
ART use: Did not receive ART during TB treatment episode	Started ART during first two months of TB treatment	1.84	<0.001	1.41	2.41
	Started ART between two months and end of TB treatment	4.09	<0.001	3.28	5.11
No valid South African ID number	Valid South African ID number	1.25	0.002	1.08	1.43
Also adjusted for intra-clinic correlation					

Table 53: Multivariable regression model for the effect of ART and baseline covariates on failed TB outcome.

Reference category	Explanatory variable	OR	P>|t|	95% Confidence Interval	
Fezile Dabi District	Lejweleputswa	0.42	0.076	0.16	1.09
	Motheo	0.79	0.573	0.35	1.79
	Thabo Mofutsanyane	0.66	0.332	0.28	1.53
	Xhariep	0.63	0.31	0.26	1.53
Year when TB treatment was started: 2006	2004	1.05	0.879	0.55	2.03
	2005	1.24	0.454	0.71	2.18
	2007	1.76	0.057	0.98	3.14
	2008	1.02	0.934	0.61	1.70
Women	Men	1.36	0.082	0.96	1.92
Age category: 30-39 years	<20	1.19	0.824	0.25	5.61
	20-29	0.89	0.534	0.61	1.29
	40-49	0.78	0.286	0.50	1.23
	>50	1.01	0.978	0.54	1.87
CD4 cell count closest to starting TB treatment: 100-199 cells/µl	CD4 0-24	1.44	0.431	0.58	3.58
	CD4 25-49	1.16	0.697	0.54	2.48
	CD4 50-99	1.20	0.613	0.59	2.44
	CD4 200-349	1.10	0.82	0.50	2.41
	CD4 350-2000	1.36	0.419	0.65	2.85
CD4 test date during the period one year preceding TB diagnosis and 30 days there after	CD4 test date after 30 days of TB treatment	1.53	0.288	0.70	3.38
	CD4 missing	1.48	0.051	1.00	2.18
Anatomical classification of TB: Pulmonary	Extra-pulmonary	0.26	0.048	0.07	0.99
	Both TB types	0.93	0.887	0.36	2.44
Smear results: Smear negative	Smear positive	3.95	0.002	1.62	9.63
	No smear results	0.83	0.714	0.31	2.23
Treatment case: New case	Retreatment case	1.43	0.168	0.86	2.37
ART use: Did not receive ART during TB treatment episode	Started ART during first two months of TB treatment	0.96	0.937	0.36	2.59
	Started ART between two months and end of TB treatment	0.73	0.285	0.41	1.30
No valid South African ID number	Valid South African ID number	0.96	0.874	0.55	1.67
Also adjusted for intra-clinic correlation					

Table 54: Multivariable regression model for the effect of ART and baseline covariates on defaulted TB outcome.

Reference category	Explanatory variable	OR	P>|t|	95% Confidence Interval	
Fezile Dabi District	Lejweleputswa	0.83	0.419	0.53	1.30
	Motheo	0.80	0.309	0.51	1.23
	Thabo Mofutsanyane	0.63	0.018	0.43	0.93
	Xhariep	1.02	0.942	0.63	1.65
Year when TB treatment was started: 2006	2004	1.15	0.437	0.81	1.61
	2005	0.94	0.594	0.76	1.17
	2007	1.19	0.049	1.00	1.43
	2008	1.10	0.285	0.92	1.31
Women	Men	1.15	0.094	0.98	1.34
Age category: 30-39 years	<20	1.73	0.054	0.99	3.01
	20-29	1.08	0.405	0.90	1.30
	40-49	0.84	0.039	0.72	0.99
	>50	0.84	0.137	0.67	1.06
CD4 cell count closest to starting TB treatment: 100-199 cells/µl	CD4 0-24	1.04	0.838	0.73	1.46
	CD4 25-49	1.03	0.853	0.77	1.36
	CD4 50-99	0.90	0.424	0.70	1.16
	CD4 200-349	0.88	0.444	0.64	1.22
	CD4 350-2000	0.81	0.295	0.54	1.21
CD4 test date during the period one year preceding TB diagnosis and 30 days there after	CD4 test date after 30 days of TB treatment	1.67	0.002	1.20	2.31
	CD4 missing	1.44	<0.001	1.18	1.76
Anatomical classification of TB: Pulmonary	Extra-pulmonary	0.93	0.539	0.73	1.18
	Both TB types	0.94	0.715	0.68	1.30
Smear results: Smear negative	Smear positive	0.90	0.345	0.71	1.13
	No smear results	1.38	0.003	1.11	1.71
Treatment case: New case	Retreatment case	1.16	0.073	0.99	1.35
ART use: Did not receive ART during TB treatment episode	Started ART during first two months of TB treatment	1.09	0.583	0.80	1.50
	Started ART between two months and end of TB treatment	0.60	0.001	0.44	0.82
No valid South African ID number	Valid South African ID number	0.72	0.005	0.58	0.91
Also adjusted for intra-clinic correlation					

Table 55: Multivariable regression model for the effect of ART and baseline covariates on moved TB outcome.

Reference category	Explanatory variable	OR	P>|t|	95% Confidence Interval	
Fezile Dabi District	Lejweleputswa	0.80	0.502	0.41	1.54
	Motheo	0.40	0.001	0.24	0.67
	Thabo Mofutsanyane	0.25	<0.001	0.14	0.44
	Xhariep	0.05	<0.001	0.02	0.13
Year when TB treatment was started: 2006	2004	0.35	<0.001	0.21	0.58
	2005	0.24	<0.001	0.16	0.35
	2007	1.20	0.279	0.86	1.65
	2008	1.09	0.664	0.75	1.58
Women	Men	0.88	0.199	0.72	1.07
Age category: 30-39 years	<20	0.89	0.792	0.38	2.11
	20-29	1.34	0.01	1.07	1.68
	40-49	1.05	0.587	0.88	1.26
	>50	1.07	0.69	0.76	1.52
CD4 cell count closest to starting TB treatment: 100-199 cells/µl	CD4 0-24	1.28	0.213	0.87	1.88
	CD4 25-49	1.10	0.619	0.76	1.60
	CD4 50-99	1.13	0.482	0.81	1.57
	CD4 200-349	0.91	0.618	0.62	1.33
	CD4 350-2000	0.71	0.228	0.41	1.24
CD4 test date during the period one year preceding TB diagnosis and 30 days there after	CD4 test date after 30 days of TB treatment	0.75	0.228	0.47	1.20
	CD4 missing	0.76	0.065	0.57	1.02
Anatomical classification of TB: Pulmonary	Extra-pulmonary	0.88	0.241	0.70	1.09
	Both TB types	1.16	0.684	0.57	2.38
Smear results: Smear negative	Smear positive	0.87	0.365	0.65	1.17
	No smear results	4.80	<0.001	3.16	7.30
Treatment case: New case	Retreatment case	0.74	0.035	0.56	0.98
ART use: Did not receive ART during TB treatment episode	Started ART during first two months of TB treatment	0.40	0.001	0.24	0.69
	Started ART between two months and end of TB treatment	0.04	<0.001	0.02	0.10
No valid South African ID number	Valid South African ID number	1.19	0.323	0.84	1.69
Also adjusted for intra-clinic correlation					

6. DISCUSSION

6.1. Main findings

This thesis reports on a series of cohort studies carried out among patients infected with HIV and receiving public sector HIV and TB healthcare in the Free State province of South Africa between 2004 and 2010. Each study was aimed at estimating the effectiveness of ART by comparing health outcomes between patients receiving ART and those not receiving ART, while adjusting for other prognostic variables. For patients who started ART after the start of cohort follow-up, this included comparing outcomes before and after ART was initiated, except for the study of TB treatment outcomes in which patients who ever received ART during treatment were compared with those who never did. All of these studies showed that ART, provided at large scale, up to six years after public sector ART was implemented, was effective at preventing TB in patients with HIV, and at preventing death in co-infected HIV-TB patients.

The main findings are summarised in Table 56.

Table 56: Summary of the best estimates of effects of ART in each study

Study population	Aim - Estimated effect of ART on:	Statistical model	Primary result
All HIV positive patients; TB data sources – ETR and HIV programme	Time to first occurrence of TB	Competing risks regression model	SHR 0.64, CI: 0.61-0.67
	TB incidence rate, accounting for multiple episodes of TB	Poisson regression model	IRR 0.66, CI: 0.64-0.69
HIV-TB co-infected patients; TB data sources – ETR and HIV programme	Time to death	Marginal structural model	HR 0.53, CI:0.47-0.61
		Cox proportional hazard model	HR 0.47, CI:0.39-0.56
HIV-TB co-infected patients; TB source – ETR only	TB outcomes: death during first year of TB diagnosis	Logistic regression model	Treatment success: OR 3.25 CI: 2.61-4.05 Died during TB treatment: OR 0.33 CI: 0.26-0.42 Died during first year: OR 0.30 CI: 0.24-0.38

ETR: electronic TB register; SHR: sub-hazard ratio; IRR: Incidence rate ratio; HR: Hazard ratio; OR: odds ratio

They show that, in people with HIV, ART was associated with a 36% lower hazard of a first TB episode and a 34% lower TB incidence rate. In co-infected HIV-TB patients ART was

associated with roughly halving of the hazard of death. Among patients on the ETR, receipt of ART was associated with a 67-70% reduction in mortality between 2004 and 2008.

6.2. Effect of ART on incidence of TB

In this resource-limited, public health programme setting in the Free State province, the overall incidence of TB in this HIV positive population was 6.68 / 100 person years (py), with an incidence rate of 8.95 / 100 py for patients who did not receive ART and 4.86 / 100 py for patients who received ART. These rates are similar to a study compiling data from eight Sub-Saharan African HIV programmes with 10.5 / 100 py during no-ART period and 5.4 / 100 py during ART (Nicholas S *et al*., 2011) but slightly higher than reported from other studies in Africa. Van Rie and colleagues (2009) reported a TB incidence rate of 4.2 / 100 py for TB patients on ART at a clinic in Thembalethu, while Hermans and colleagues (2009) found a 2.65 / 100 py rate for patients on ART at a clinic in Uganda.

Multivariable adjustment for the most important prognostic features showed that ART was independently associated with reducing the hazard of developing TB in this HIV positive population. Although the Cox regression model indicated that ART was more effective (53% risk reduction) than the competing risks regression (CRR) model (36% risk reduction), the CRR model is probably the more accurate model since it accounted for death as a competing risk. The Cox regression model made no use of the pertinent information that death was a reason for censorship of follow-up. When censorship of follow-up is due to a competing risk – such as when death makes incident TB impossible – Cox regression estimates of the effects of covariates - such as ART - on survival may be biased. These CRR results are similar to the results of our earlier study of the Free State HIV programme cohort between 2004 and 2005, using a marginal structural model which appropriately adjusted for time-varying covariates (HR for time to TB: 0.61 CI: 0.46-0.81) (Fairall LR *et al*., 2008). Taken together, these results suggest that other studies of TB incidence estimates using Cox regression may have overestimated ART effectiveness.

The most recent systematic review and meta-analysis of the effect of ART on TB incidence in a range of 0 to >350 cells/µl CD4 categories (Suthar AB *et al*., 2012b), reported an overall effect estimate of 65% hazard reduction (HR 0.35 CI: 0.28-0.44), which is higher than this study's competing risk and Cox models' effect estimates (36% and 53% respectively). Other studies have also reported a greater impact of ART on TB incidence but differences in study

designs and statistical methods could partly explain the differences. A single South African facility participating in a randomised controlled trial before ART was made publically available was used to estimated the effect of ART with a Poisson regression model, resulting in an adjusted rate ratio of 0.19 CI: 0.09-0.38 (Badri M *et al*., 2002). Golub and co-workers (2007) reported on a multi facility study done in Brazil, which had a low overall TB incidence rate of 2.28 / 100 py and a TB incidence rate of 1.90 / 100 py for patients on ART. They reported an adjusted RH (relative hazard) of 0.41 ($p<0.001$) for patients receiving ART vs. no-ART, but stated that an underreporting of TB (due to death) could have biased their results. Underreporting of TB can occur when patients appear to be lost to follow-up on an HIV treatment programme, while attending a TB treatment programme (and defaulted/transferred/died) and not being counted as a TB case when data from a non-integrated HIV setting is reported on.

The present study provided original estimates of the effectiveness of ART in preventing TB in different strata, defined according to CD4 level at the start of follow up. It showed that ART effectiveness was similar and significant among patients with CD4<350 cells/µl at baseline, but had no significant effect among patients with CD4>350 cells/µl at baseline. Although the review by Suthar and colleagues suggested that ART was less effective with increasing baseline CD4 levels, this was not significantly different (Suthar AB *et al*., 2012b). In contrast, a recent cohort study by the HIV-Causal Collaboration (2012) from facilities in America and Europe, showed that ART was not effective in patients with baseline CD4<50 cells/µl at baseline, but was highly effective at higher levels, including CD4>350 cells/µl.

The effect of ART on multiple TB episodes was estimated with a Poisson regression model and showed that ART was less effective in preventing multiple TB episodes (IRR 0.66, CI: 0.64-0.69) than in preventing the first episode, in the total population. However it was highly effective in preventing multiple episodes in patients with CD4<200 cells/µl (IRR 0.32-0.48). The surprising finding that, in patients with CD4>200 cells/µl, the IRRs for ART were 1.39 and 4.34, suggest that these patients might have had indications for ART other than depressed CD4 counts, which were associated with higher TB risks.

This study also showed that a history of TB significantly increased the hazard of TB; this is keeping with previous studies reporting the influence of previous TB (Golub JE *et al*., 2007, Moore D *et al*., 2007, Van Rie A *et al*., 2011, Nicholas S *et al*., 2011).

In summary, this study provides evidence that ART reduces TB infection in people with HIV. As people infected with TB are liable to infect others in the population, ART is therefore an important part of the overall TB control strategy in South Africa, directly of benefit to people with HIV.

6.3. Effect of ART on mortality

In this public health programme, with constrained resources, the hazard of death was halved for HIV positive patients receiving ART compared to those not receiving ART. The survival benefit persisted for patients with very low CD4 cell counts. Most patients had advanced immune-suppression, which is consistent with other studies done in Sub-Saharan Africa (Lawn SD *et al*., 2006; Moore D *et al*., 2007; Westreich D *et al*., 2009; Worodia W *et al*., 2012). The primary analysis used marginal structural models to optimise adjustment for time-varying as well as baseline values of CD4 cell count, weight, and cotrimoxazole prescription, using inverse probability of treatment weighting. A comparison of weighted and un-weighted models that accounted for time varying covariates (Table 31) suggests that un-weighted models underestimated the effect of ART on death, that is, 37% hazard reduction with un-weighted model vs. 47% hazard reduction with MSM. Although MSM and Cox models had similar results there was one slight discrepancy: with MSM the effectiveness of ART was not significantly different at different baseline CD4 levels, whereas with Cox models ART had no significant effect at CD4 levels above 200 cells/µl. This discrepancy may be due to MSM being able to adjust for time-varying CD4 levels as well as for baseline CD4 levels.

These results, which appropriately controlled for confounding, found ART to be much less effective in preventing deaths than did two studies from Thailand (Sanguangwongese N *et al*., 2007; Akksilp S *et al*., 2007). The Thailand studies used public health system surveillance data to compare survival between co-infected HIV-TB patients receiving ART or not. After controlling for confounders (CD4, cotrimoxazole treatment, type and location of TB, type of facility), patients treated with ART had one sixth the hazard of death compared to TB patients not receiving ART (Sanguangwongese N *et al*., 2007; Akksilp S *et al*., 2007). However, these studies' patients were only followed for the duration of TB treatment and not thereafter like this study and 41% of patients had missing CD4 results (Akksilp S *et al*., 2007).

Another study from Thailand (Manosuthi W *et al*., 2006), which followed patients after TB treatment commenced, showed that the hazard of death in TB patients was 20 times higher for patients who did not receive ART than for patients who received ART. Fifty percent of patients not receiving ART died within one year of TB diagnosis. A major difference between that study and the study presented in this thesis was that their patients were categorised in either an ART group or no-ART group, whereas with this analyses follow-up time was divided into no-ART and ART follow-up, appropriately accounting for time at risk of death with or without ART. Also, in the latter Thai study CD4 counts were missing for more than 50% of the patients and were therefore not adjusted for in primary analysis. A secondary analysis limited to patients with available CD4 counts showed that CD4 was not identified as a risk factor for mortality, which is surprising (Manosuthi W *et al*., 2006). These Thai results are likely to be biased by systematic differences between patients who did and did not receive ART. Other possible reasons for these different effect measures could also be due to differences in host susceptibility, pathogen virulence, or intensity of TB transmission (Sanguangwongese N *et al*., 2007).

There is little other direct evidence of the effectiveness ART on mortality in HIV-TB patients. Most of the other studies reporting on associations between ART and mortality in co-infected patients, were randomised controlled trials investigating the optimal timing of ART initiation, that is, comparing mortality in co-infected patients starting ART either earlier or later during TB treatment (Abdool Karim SA *et al*., 2010; Velasco M *et al*., 2009). Showing that initiating ART earlier in TB treatment reduced mortality, these trials provide the most rigorous evidence available that ART is effective. However, they are unable to estimate the absolute effectiveness of ART compared to no ART.

In the primary MSM of the effect of cotrimoxazole on mortality, it was surprising that cotrimoxazole appeared to have no effect at all. However, in the MSMs in which the inverse probability of treatment weights were truncated, cotrimoxazole was weakly associated with reduced mortality (HRs 0.81 and 0.83). The latter estimates are more likely to be correct, because MSMs are recognised as being at risk of bias by extremely large weights resulting from extremely low cumulative probabilities of treatment in a small number of observations, in which case truncated weights are recommended (Cole & Hernan, 2008). In the Cox models, cotrimoxazole was associated with greater effects on mortality (HRs 0.56 - 0.65) than in the MSMs. However, the Cox models of cotrimoxazole effectiveness did not account

for the precise date on which cotrimoxazole was started; they only accounted for whether a patient received cotrimoxazole at all when on ART or not on ART. In contrast, all the analyses of ART effectiveness, and the MSM of cotrimoxazole effectiveness did account for the precise date on which cotrimoxazole started and so are more likely to be accurate.

6.4. Effect of ART on TB treatment outcomes

The third study was of 9 276 co-infected HIV-TB patients whose TB treatment outcomes were recorded in the provincial ETR, as required by the South African national TB programme and WHO. In addition to death during treatment, the most important treatment outcomes were "cure" (conversion from positive to negative sputum smear) and treatment completion. As many HIV- infected TB patients are sputum negative at the start of TB treatment, and so thus cannot be "cured" as defined, the composite outcome of cure or completion – defined as "success" – is thus important. In this HIV positive study population 73.9% of patients had a successful treatment outcome, and 8.3% died during treatment. These results are similar to results from European cohorts (Faustini A *et al*., 2005; Falcon D *et al*., 2005), but mortality was lower than in two other studies from Taiwan and Italy, which reported death rates of 21.6% (Yen YF *et al*., 2012) and 14.6% (Girardi E *et al*., 2012) respectively. For HIV-negative patients mortality for TB patients is usually in the order of 3.6% (Mugusi FM *et al*., 2009). Cohort and research setting differences are most probably the main reason for variations seen in TB outcomes, such as the proportion of intravenous drug users, migrants, elderly patients, which differed between the cohorts.

In this study logistic regression models were used to estimate the effect of ART on TB treatment outcomes while controlling for confounders. Unlike the studies with time to TB and time to death as outcomes, these analyses did not account for how long it took to attain each outcome, but only which outcomes were recorded at the end of treatment, or whether death occurred within a year of starting TB treatment. The study showed that ART was significantly and independently associated with lower odds of unfavourable outcomes (death, treatment failure, default) and higher odds of favourable outcomes (cure, completion and success). This is in line with other studies which also reported that the use of ART was associated with a reduction in the risk of death (Girardi E *et al*., 2012; Kwan C *et al*., 2011; Boulle A *et al*., 2010).

Internationally there is limited evidence of the effect of ART on TB treatment outcomes other than death. Studies by Yen and colleagues (2012) and by others (Obuku EA *et al*., 2012; Muniyande M *et al*., 2007) reported strong associations between unemployment and lower education levels with unfavourable TB treatment outcomes. Although not recorded in our data, these factors could also be of importance and applicable to the Free State population which has an unemployment rate of 38.9% (South African Government information). Studies have shown that patients with better socioeconomic status are likely to receive additional diagnostic procedures and superior treatment (Younis MZ *et al*., 2005) as well as patients with a better education (Berisha M *et al*., 2009; Lu SH *et al*., 2009).

In this study population only 20.5% of patients started ART during their TB treatment episode, that is, 5.0% during initiation phase (early ART) and 16.5% during continuation phase (late ART). In these analyses, ART started after two months of TB treatment appeared to be more effective than ART started earlier, which is the opposite of what randomised trials have shown (Naidoo K *et al*., 2013; Velasco M *et al*., 2012). The reason for this discrepancy in the present study could be because the comparison of patients who started ART early or late is subject to "immortal time bias" (Sterne JA *et al*., 2009; Levesque LE *et al*., 2010). That is, TB patients would only be able to start ART later if they had not already died earlier, and those who did not die earlier would tend to be healthier at the start TB treatment than those who died earlier. The adjustment for baseline prognostic variables would have gone some way toward avoiding this bias, but would be unlikely to avoid it altogether.

An important finding for this study population was the high mortality (7.4% of patients who died during the first year, 118 / 1 594) between completion of TB treatment and a year after TB treatment started, even in people who successfully complete TB treatment. This is consistent with previously published studies but mainly in cohorts not receiving ART (Mugusi FM *et al*., 2009; Harries AD *et al*., 2001). These deaths after successful TB completion could be caused by HIV-related causes other than TB. The presence of TB in HIV positive patients is known to accelerate the progression of HIV disease due to an activation of the immune system, which subsequently enhances viral replication. Patients with low CD4 counts and increased viral loads are more susceptible to opportunistic infections which may cause death (Freudenberg N *et al*., 2006; Mugusi FM *et al*., 2009; Geldmacher C *et al*., 2010). High rates of transfer and movement elsewhere were seen for

this sub-population who died after treatment compared to the overall population, which suggests that they moved back to their home towns to be taken care of by family.

In terms of other risk factors, this study's results showed that advanced age (Farah MG *et al*., 2005; Falzon D *et al*., 2005; Yen FE *et al*., 2010), history of TB, positive smears (Tessema B *et al*., 2009), lower weight and CD4 cell counts were risk factors for unfavourable outcomes of TB treatment. Male sex was associated with unfavourable TB outcomes which confirmed the results of another report (Falzon D *et al*., 2005), but contradicts the results of other studies (Pablos-Mendez A *et al*., 1996; Silva DR *et al*., 2010).

In summary, this study shows that receipt of ART during TB treatment greatly improves all treatment outcomes.

6.5. Strengths and limitations of this research

Compared with previous studies from dedicated non-government organisations and research-supported ART services in South Africa (Bekker LG *et al*., 2006; Coetzee D *et al*., 2004) and elsewhere (Tessema B *et al*., 2009; Lee JJ *et al*., 2009; Silva DR *et al*., 2010), this study investigated a large heterogeneous province-wide population of HIV positive and co-infected HIV-TB patients, representing diverse geographical and clinical settings and exposed to all inefficiencies of real-world healthcare services (Sanguanwongse N *et al*., 2007). It is thus unusually representative of the impact of ART applied across a health system, making the research findings generalisable to a wide variety of settings (D'Agostino RB *et al*., 2007). The studies also showed variation in care in a national programme which is typically a function of when patients present themselves at clinics. Ingle SM *et al*., reported in 2010b that facility-level characteristics rather than patient-level characteristics were associated with the odds of starting ART treatment across facilities in the Free State Province.

Ideally, evidence of ART effectiveness would come from randomised controlled trials comparing ART with placebo. However these would not be ethically possible because of the known effectiveness of ART, based on the long history of randomised trials comparing single ART drugs with placebo, dual ART with single ART and triple ART with dual ART, as well as clinical and public health experience of ART. The fact that so many patients had to wait so long to initiate ART, while unfortunate, provided powerful comparative evidence of risks associated with not receiving ART.

The present studies aimed to make optimal use of all available data and to avoid the biases, which are a continual problem with observational studies of treatment effectiveness. The main methodological problem is that people receiving ART tend to be systematically different from those not receiving ART, both because of the indications for initiating treatment and because of the effects of ART after treatment. The fact that, for most patients, the most important prognostic factors were recorded at their time of registration with the HIV programme and periodically thereafter made it possible to adjust statistically for these differences.

The relative completeness of personal identifiers such as first name, surname, address, sex, and DOB on both datasets supported a robust data linkage procedure between the two public sector programmes without a unique patient identifier in the majority of TB records. The opportunity to combine patient information from five large scale databases (ART, TB, laboratory, hospital, death population register) made it possible to establish an integrated research platform to address several research questions about this cohort.

The strict and standardised systems for ART and TB data collection helped in decreasing recording errors and information bias, as patient data was captured from clinical records and registers completed and kept on site by the ART and TB facilities. Although underreporting of HIV information on the TB register, and of TB status information on the HIV programme was seen, the linked dataset provided a more accurate status of co-morbidity. Also, due to the high percentage of available South African ID numbers for the HIV programme patients, active follow-up of survival through data linkage with the death population register was possible and helped to avoid survivor bias.

Patients co-infected with TB were identified from both the HIV programme and provincial TB register to overcome non-reported TB episodes for HIV patients who defaulted from the HIV programme while attending a TB facility, or were not screened for TB during a HIV visit, or a visit form was not captured on the HIV system (missing information), or were not scheduled for a HIV visit during the time of TB due to a high CD4 cell count, which required less regular visits to the HIV programme.

The studies of effect of ART on TB incidence, and on time to death in co-infected patients used relatively sophisticated cohort study designs, compared to most previous African studies, to optimally use data on person-time at risk and on time-varying treatment and prognostic

variables. This made it possible to include the experience of the majority of patients before and after they started ART, as well as that of the minority of patients who never received ART or who were already receiving it at the time of enrolment with the HIV programme. The large sample sizes provided ample statistical power to adjust a wide range of prognostic variables and to provide precise effectiveness estimates.

A key limitation of these studies, other than their observational design, was missing data about prognostic variables, especially CD4 counts. Between 13% and 50% CD4 counts were missing, depending on the time point for CD4 cell counts used for various analyses. CD4 cell counts recorded close to enrolment were missing for about 13%-18% of patients, and CD4 cell counts closest to the date when TB treatment was started were missing for 50% of patients. Body weight was another important prognostic variable, which was missing for 16%-22% of patients. In order not to exclude them from the primary analyses, for patients who had one or more missing CD4 or weight measurements, multiple imputation of missing CD4 counts and weights was used. It is reassuring that the sensitivity analyses, that is, complete case analyses excluding patients with any missing CD4 or weight data, showed very similar results to the primary analyses with imputed data.

Also, according to national policy at the time, patients with CD4>200 should not have started ART unless they had stage 4 AIDS. As we did not have other information on whether patients had stage 4 AIDS, CD4>200 was thus simultaneously a good and a bad prognostic indicator. Therefore the estimated risks and modification of ART effectiveness associated with CD4>200 should be interpreted with caution.

Another limitation was uncertainty about whether those patients assumed to be receiving ART were in fact receiving it. In the absence of accurate information about treatment adherence or default, all analyses assumed that patients who received ART continued to do so throughout follow-up. It is likely that some patients assumed to be receiving ART were no longer doing so. The same applies to estimates of cotrimoxazole effectiveness. However, if misclassification was non-differential then the final outcome would be to underestimate ART's effectiveness. It is also likely that for some patients the dates of initiation of ART and cotrimoxazole and of TB diagnosis could have been inaccurately recorded, such errors could have biased effectiveness estimates in either direction, but without information about such errors it is difficult to predict the direction of bias.

The exact date of TB diagnosis was not available on either the ETR or HIV programme database therefore the assumption was made that diagnosis and treatment was on the same date. This would overestimate time to TB disease, but it would not bias ART effectiveness estimates unless the delay from TB disease to treatment was different in people on ART than in people not on ART. If the delay was shorter in people on ART than in people not on ART, then ART effectiveness would be underestimated, and vice versa.

Incomplete and incorrect entries in either of the registers (ART, TB or death) would have limited the generalisability of the results. The completeness of the TB register has always been questioned (Van Hest NAH *et al*., 2006; Melosini L *et al*., 2012) and there is potential for misclassification of patients as being TB free when there is no record of the patient in the ETR, which was shown when patients themselves reported to be on TB treatment at an ART programme visit but they could not be linked to a patient record on the ETR. Also, the dataset had a small proportion of patients with CD4 cell counts above 350 cells/µl and the relevant outcomes (death, TB), which limits the ability to look at the effect of ART in less immuno-suppressed patient groups and could explain differences noted from North American and European cohorts on the benefits of ART in patients with high CD4 cell counts.

The deterministic data matching process was complex and to improve data linkage accuracy the number of rules can increase dramatically. This data linkage process could be difficult to reproduce on different datasets and each dataset might need different rules to suit the data depending on the format of the data variables. This limits reproducibility but the process developed and described in this thesis sets a platform on how to develop rules to link datasets. Probabilistic matching is a more automated process but requires human input to calibrate matching rules and weights (probabilities) for each dataset, which is time consuming, needs sophisticated software and computers with extensive processing power. A sample run was done with a server in Bern, Switzerland and probabilistic General Record Linkage Software (GRLS) but due to the large datasets and limited computing power of the servers the linkage process was aborted.

In summary, these studies made exceptionally thorough efforts to avoid confounding, selection biases and information biases, but it remains possible that unmeasured or hidden factors could have influenced treatment outcomes and thus biased the effectiveness estimates. However, these limitations do not change the main findings: ART was highly effective but not as effective as it could be.

6.6. Recommendations for policy and research

This thesis suggests that ART is effective at preventing death in co-infected patients even when provided at scale within a province-wide resource-constrained setting, supporting the ongoing scale-up of treatment programmes. These findings also support recent changes in ART treatment polices which recognise the risk imposed by co-infection, and recommend ART for those HIV-infected with TB irrespective of the CD4 count (Circular H116/2012 and H3/2013 for Western Cape). There is not conclusive evidence that initiating ART in patients with a CD4 cell count above 350 cells/µl will benefit patients (Geffen N, 2011; Suthar AB *et al*., 2012b). A study done by South African Centre for Epidemiological Modelling and Analysis (SACEMA) predicts that the best threshold for starting ART is 350 cells/µl not to increase the population incidence of TB (Williams BG, 2013).

Although the optimal timing of ART initiation in co-infected patients was beyond the scope of this thesis, findings are consistent with current recommendations to start ART early in co-infected HIV-TB patients, especially in those who are severely immuno-compromised (Naido K *et al*., 2013; Circular H116/2012 and H3/2013 for Western Cape).

Mortality in co-infected patients remains unacceptably high and this thesis and suggests that ART may be less effective at reducing deaths compared with other reports (Sanguangwongese N et al., 2007; Akksilp S et al., 2007; Manosuthi W et al., 2006). Furthermore, linking the ETR and national death population registers showed that a sizeable proportion of co-infected patients die within the six months immediately after they have completed TB treatment, many of them despite successful TB treatment outcomes. Currently there is no consensus regarding the reasons for this high mortality. Some studies have shown that TB was not an independent predictor of mortality in co-infected patients, which instead was explained by low baseline CD4 cell counts (Lawn SD *et al*., 2009; Escombe AR *et al*. 2008; Zacharian R *et al*., 2006; Westreich D *et al*., 2009). Research has also shown that TB activates the immune system which in turn leads to an increase in the replication of the HIV virus, placing the patient at risk of other opportunistic infections and death. A recent analysis from the IeDEA collaboration in Southern Africa showed that TB confers increased risk for cryptococcal disease and pnemocystis pneumonia (Fenner L *et al*., 2013). Further research is urgently needed to elucidate the reasons for this high mortality, so that strategies can be devised to tackle this problem. In the interim, it is important that care of co-infected patients include adequate screening for other concurrent opportunistic infections. One important

development in this regard has been the introduction of routine laboratory-based screening for cryptococcal disease in all patients with CD4 counts below 100 (Jarvis JN *et al*., 2012).

The linkage of data from the ART and TB programmes showed that patients treated in the TB programme were not routinely enrolled in the ART programme, despite advanced levels of immuno-suppression and often with fatal consequences. Some eight years into the rollout of ART in South Africa's public sector, integration of these programmes at clinic level remains suboptimal, despite high levels of co-infection. In part this can be attributed to the restriction of ART services to selected accredited sites, in contrast to the vastly more decentralised model adopted by the TB programme where almost every fixed primary care facility operates a TB service. A study from Cape Town showed that patients attending an integrated service had a 60% greater chance of starting and reduced waiting times for ART initiation (Kerschberger B *et al*., 2012). The integration agenda must be actively pursued to ensure that TB patients are actively screened for HIV, eligibility for ART assessed and treatment started without delay. Hopefully this will become easier as South Africa works towards making ART available in all its primary care facilities (NDOH National Strategic Plan 2012-2016).

Vulnerable populations identified from this thesis include men and older patients, in keeping with findings from other cohorts (Cornell M *et al*, 2012a; Cornell M *et al*, 2012b; Grieg J *et al*, 2012). Special care and strategies need to be developed to target these patient groups to improve their care-seeking and retention in care.

This thesis showed that ART substantially reduces the risk of TB in HIV-infected adults and should form a key part of strategies to control TB in high HIV burden settings, alongside WHO's collaborative activities for HIV-TB co-infections (Intensified case-finding, Izoniazid preventative therapy, Infection control (WHO 2012). Unfortunately izoniazid prophylaxis was only rarely prescribed to patients in the Free State cohort, preventing any assessment of its effect on TB incidence. Reports that IPT has now been rolled to out to 373 000 South Africans provides some reassurance that this effective intervention is finally being adopted on a large scale (WHO Global TB report, 2012). But this thesis confirmed the work of others which shows that the risk of TB on ART remains higher than background rates (Lawn SD *et al*., 2010). Research on the additive effects of ART/IPT and optimal duration of IPT for the prevention of TB is urgently needed and results from the two randomised trials currently underway in South Africa and Côte d'Ivoire are awaited (Gideon HP *et al*., 2009; Clinical trial ANRS 12136 TEMPRANO).

These results provide further justification for the strengthening and expansion of ART in South Africa. However, they also show that in this province-wide programme, ART was not as effective as it could be, compared with the experience of ART in high income countries, and of smaller scale better resourced ART programmes in South Africa. Within this province, the big differences in treatment outcomes between different districts and facilities show wide variations in the quality of care, raising the question of what the reasons are for those differences, and of how the worst performing districts and facilities should be helped to obtain the best possible outcomes. Future research should be targeted at finding the reasons why the province- and country-wide public sector ART programmes in South Africa are not as effective as they could be, how they could be strengthened, and at evaluating the effectiveness and cost-effectiveness of new interventions aimed at improving ART delivery. Cohort, operational research and qualitative studies are needed to understand the reasons for effective and ineffective ART delivery, and randomised trials, quasi-experimental and cohort studies are needed to evaluate innovative ways of strengthening ART. Efforts to link clinical, laboratory, mortality and health system data, as in this study, would greatly enhance the ability to carry out such studies, especially on such a large scale. At the same time smaller studies in sentinel sites which are able to ensure complete high quality data are also needed.

In summary, this thesis demonstrated substantial and durable clinical benefits for co-infected HIV-TB patients who were able to access public sector treatment interventions during a 7-year period in the Free State Province. It advanced our current understanding by demonstrating that ART in HIV positive and co-infected HIV-TB patients was strongly associated with better TB treatment outcomes, reducing the presence of TB and improving survival. It is strongly recommended that health services should be integrated such that patients are diagnosed earlier with HIV/TB to prevent impairment of the immune system which leads to poorer outcomes and to implement co-morbidity screening for this vulnerable population.

REFFERENCES

Harvard referencing format

Abdool Karim SS, Naidoo K, Grobler A, Padayatchi N, Baxter C, Gray A, Gengiah T, Nair G, Bamber S, Singh A, Khan M, Pienaar J, El-Sadr W, Friedland G, Abdool Karim Q. 2010. Timing of initiation of antiretroviral drugs during tuberculosis therapy. *N Engl J Med*, 362:697-706

Abdool Karim SS, Naidoo K, Grobler A, Baxter C, Gray A, Gengiah T, Gengiah S, Naidoo A, Jithoo N, Nair G, El-Sadr WM, Friedland G, Abdool Karim Q. 2011. Integration of antiretroviral therapy with Tuberculosis treatment. *N Engl J Med*, 365:1492-1501

Akksilp S, Karnkawinpong O,Wattanaamornkiat W, Viriyakitja D, Monkongdee P, Sitti W, Rienthong D, Siraprapasiri R, Wells CD, Tappero JW, Varma JK. 2007. Antiretroviral Therapy during Tuberculosis Treatment and Marked Reduction in Death Rate of HIV-Infected Patients, Thailand. *Emerg. Infect. Dis*, 13(7):1001-1007

Akolo C, Adetifa I, Shepperd S, Volmink J. 2010. Treatment of latent tuberculosis infection in HIV infected persons. *Cochrane Database Syst Rev*, 1.

Andersen PK, Geskus RB, de Witte T, Putter H. 2012. Competing risks in epidemiology: possibilities and pitfalls. *Int J Epidemiol*, 41(3):861-70

Andia I, Biraro S, Taseera K, Muzoora C, Asiimwe C, Musinguzi N et al. 2010. Outcomes of early versus late initiation of antiretroviral therapy in HIV-positive patients co-infected with tuberculosis in South-western Uganda. International AIDS Conference July 18-23 2010, Vienna, Austria.

Ansari NA, Kombe AH, Kenyon TA, Hone NM, Tappero JW, Nyirenda ST, Binkin NJ, Lucas SB. 2002. Pathology and causes of death in a group of 128 predominantly HIV-positive patients in Botswana, 1997–1998. *Int J Tuberc Lung Dis*, 6:55-63

Badri M, Wilson D, Wood R. 2002. Effect of highly active antiretroviral therapy on incidence of tuberculosis in South Africa: a cohort study. Lancet, 359:2059-2064

Bassett IV, Wang B, Chetty S, Mazibuko M, Bearnot B, Giddy J, Lu Z, Losina E, Walensky RP, Freedberg KA. 2009. Loss to care and death before antiretroviral therapy in Durban, South Africa. *J Acquir Immune Defic Syndr*, 51:135-139

Bekker LG, Myer L, Orrell C, Lawn S, Wood R. 2006. Rapid scale-up of a community based HIV treatment service: programme performance over 3 consecutive years in Guguletu, South Africa. *SAMJ*, 96(4):315-320

Berisha M, Zheki V, Zadzhmi D, Gashi S, Hokha R, Begoli I. 2009. Level of knowledge regarding tuberculosis and stigma among patients suffering from tuberculosis. *Georgian Med News* , 166:89-93

Blanc FX, Sok T, Laureillard D, Borand L, Rekacewicz C, Nerrienet E, Madec Y, Marcy O, Chan S, Prak N, Kim C, Lak KK, Hak C, Dim B, Sin CI, Sun S, Guillard B, Sar B, Vong S, Fernandez M, Fox L, Delfraissy JF, Goldfeld AE. 2011. Early (2 weeks) vs. late (8 weeks) initiation of highly active antiretroviral treatment (HAART) significantly enhance survival of severely immunosuppressed HIV-infected adults with the newly diagnosed tuberculosis: results of the CAMELIA clinical trial. *BMC Proc*, 5(1):O11

Bong CN, Chen SC, Jong YJ, et al. 2007. Outcomes of HIV-infected children with tuberculosis who are started on antiretroviral therapy in Malawi. *Int J Tuberc Lung Dis*, 11:534-538

Bonnet MM, Pinoges LL, Varaine FF, Oberhauser BB, O'Brien DD, Kebede YY, Hewison CC, Zachariah RR, Ferradini LL. 2006. Tuberculosis after HAART initiation in HIV-positive patients from five countries with a high tuberculosis burden. *AIDS*, 20(9):1275-9

Boulle A, Clayden P, Cohen K, Cohen T, Conradie F, Dong K, Geffen N, Grimwood A, Hurtado R, Kenyon C, Lawn S, Maartens G, Meintjes G, Mendelson M, Murray M, Rangaka M, Spencer D, Taljaard J, Variava E, Venter WDF, Wilson D. 2010. Prolonged deferral of antiretroviral therapy in the SAPIT trial: Did we need a clinical trial to tell us that this would increase mortality? *SAMJ*, 100(9): Sep 2010

Bradley CJ, Penberthy L, Devers KJ, Holden DJ. 2010. Health Services Research and Data Linkages: Issues, Methods, and Directions for the Future. *Health Services Research*, 45:5 (Part II)

Brinkhof MW, Egger M, Boulle A, May M, Hosseinipour M, Sprinz E, Braitstein P, Dabis F, Reiss P, Bangsberg DR, Rickenbach M, Miro JM, Myer L, Mocroft A, Nash D, Keiser O, Pascoe M, van der Borght S, Schechter M. Antiretroviral Therapy in Low-Income Countries Collaboration of the International epidemiological Databases to Evaluate AIDS IeDEA); ART Cohort Collaboration. 2007. Tuberculosis after initiation of antiretroviral therapy in low-income and high-income countries. *Clin Infect Dis*, 45:1518-1521

Cain KP, Kanara N, Laserson KF, Vannarith C, Sameourn K, Samnang K, Qualls ML, Wells CD, Varma JK. 2007. The epidemiology of HIV-associated tuberculosis in rural Cambodia. *Int J Tuberc Lung Dis*, 1(1008-1013):

Charalambous S, Grant AD, Innes C, Hoffmann CJ, Dowdeswell R, Pienaar J, Fielding KL, Churchyard GJ. 2010. Association of isoniazid preventive therapy with lower early mortality in individuals on antiretroviral therapy in a workplace programme. *AIDS*, 24(Suppl 5) S5-S13

Coetzee D, Hildebrand K, Boulle A, Maartens G, Louis F, Labatala V, Reuter H, Ntwana N, Goemaere E. 2004. Outcomes after two years of providing antiretroviral treatment in Khayelitsha, South Africa. *AIDS*, 18(6):887-895

Cohen MS, Chen YQ, McCauley M *et al*., 2011. *N Engl J Med* 365(6): 493-505

Cohen T, Murray M, Wallengren K, Alvarez GG, Samuel EY, Wilson D. 2010. The prevalence and drug sensitivity of tuberculosis among patients dying in hospital in KwaZulu-Natal, South Africa: a postmortem study. PLoS Med, 7:e1000296

Cohen K, Meintjes G. 2010. Management of individuals requiring antiretroviral therapy and TB treatment. Curr Opin HIV *AIDS*, 5:61-69

Cole SR, Hernan M. 2008. Constructing inverse probability weights for marginal structural models. *Am J Epidemiol*, 168:656-664

Colebunders RL, Braun MM, Nzila N, Dikilu K, Muepu K, Ryder R. 1989. HIV infection in patients with tuberculosis in Kinshasa, Zaire. *American Review of Respiratory Diseases*, 139(5):1082-5

Cornell M, Schomaker M, Garone DB, Giddy J, Hoffmann CJ, Lessells R, Maskew M, Prozesky H, Wood R, Johnson LF, Egger M, Boulle A, Myer L. (for International Epidemiologic Databases to Evaluate AIDS Southern Africa (IeDEA-SA) Collaboration). 2012. Gender Differences in Survival among Adult Patients Starting Antiretroviral Therapy in South Africa: A Multicentre Cohort Study. *PLoS Med*, 9(9):e1001304

Cornell M, McIntyre J, Myer L. 2012. Men and antiretroviral therapy in Africa: our blind spot. *Trop Med Int Health*, 16(7):828-829

D'Agostino RB Jr, D'Agostino RB Sr. 2007. Estimating treatment effects using observational data. *JAMA*, 297:314-316

De Cock KM, Chaisson RE. 1999. Will DOTS do it? A reappraisal of tuberculosis control in countries with high rates of HIV infection. *Int J Tuberc Lung Dis*, 3:457-465

De Cock KM, Soro B, Coulibaly IM, Lucas SB. 1992. Tuberculosis and HIV in sub-Saharan Africa. *Journal of the American Medical Association*; 268(12):1581-1587

Dean GL, Edwards SG, Ives NJ, Matthews G, Fox EF, Navaratne L, Fisher M, Taylor GP, Miller R, Taylor CB, de Ruiter A, Pozniak AL. 2002. Treatment of tuberculosis in HIV-infected persons in the era of highly active antiretroviral therapy. *AIDS*, 16:75-83

Degu WA, Lindquist L, Aderaye G, Aklillu E, Wold H, Ali GY, Worku A, Sonnerborg A, Makonnen E. 2012. Randomized clinical trial to determine efficacy and safety of ART 1 week after TB therapy in patients with CD4 counts <200 cells/mL [abstract]. Paper #144.

Del Amo J, Malin AS, Pozniak A, De Cock KM. 1999. Does tuberculosis accelerate the progression of HIV disease? Evidence from basic science and epidemiology *AIDS*, 13(10):1151-1158

Dembélé M, Saleri N, Carvalho AC, Saouadogo T, Hien AD, Zabsonre I, Koala ST, Simporé J, Matteelli A. 2010. Incidence of tuberculosis after HAART initiation in a cohort of HIV-positive patients in Burkina Faso. *Int J Tuberc Lung Dis*, 14(3):318-23

Dhasmana DJ, Dheda K, Ravn P, Wilkinson RJ, Meintjes G. 2008. Immune reconstitution inflammatory syndrome in HIV-infected patients receiving antiretroviral therapy: pathogenesis, clinical manifestations and management [review]. *Drugs*, 68(2):191-208

Dheda K, Lampe FC, Johnson MA, Lipman MC. 2004. Outcome of HIV-associated tuberculosis in the era of highly active antiretroviral therapy. J Infect Dis, 190:1670-1676

Durovni B, Cavalcante SC, Saraceni V, Vellozo V, Israel G, King BS, Cohn S,Efron A, Pacheco AG, Moulton LH, Chaisson RE, Golube GE. 2010. The implementation of isoniazid preventive therapy in HIV clinics: the experience from the TB/HIV in Rio (THRio) Study. *AIDS*. 24(Suppl 5):S49–S56.

El-Sony AI, Khamis AH, Enarson DA, Baraka O, Mustafa SA, Bjune G. 2002. Treatment results of DOTS in 1797 Sudanese tuberculosis patients with or without HIV coinfection. *Int J Tuberc Lung Dis*, 6(12):1058-1066

Eshun-Wilson I, Taljaard JJ, Nachega JB. 2012. Sub-optimal CD4 T-lymphocyte responses among HIV infected patients who develop TB during the first year of ART. *AIDS Clin Res*, 3(135):

Etard JF, Ndiaye I, Thierry-Mieg M, Guèye NF, Guèye PM, Lanièce I, Dieng AB, Diouf A, Laurent C, Mboup S, Sow PS, Delaporte E. 2006. Mortality and causes of death in adults receiving highly active antiretroviral therapy in Senegal: a 7-year cohort study. *AIDS*, 20(1181-1189):

Fairall LR, Bachmann MO, Louwagie GMC, van Vuuren C, Chikobvu P, Steyn D, Staniland GH, Timmerman V, Msimanga M, Seebregts CJ, Boulle A, Nhiwatiwa R, Bateman ED, Zwarenstein MF, Chapman RD. 2008. Effectiveness of antiretroviral treatment in the South African public-sector programme: cohort study. *Arch Intern Med*, 168(1):86-93

Fairall LR, Bachmann MO, Lombard C, Timmerman V, Uebel K, Zwarenstein M, Boulle A, Georgeu D, Colvin CJ, Lewin S, Faris G, Cornick R, Draper B, Tshabalala M, Kotze E, Van Vuuren C, Steyn D, Chapman R, Bateman E. 2012. Task shifting of antiretroviral treatment from doctors to primary-care nurses in South Africa (STRETCH): a pragmatic, parallel, cluster-randomised trial. *Lancet*, 380(9845):889–898

Falzon D, Le Strat Y, Belghiti F, and Infuso A. 2005. Exploring the determinants of treatment success for tuberculosis cases in Europe. *Int J Tuberc Lung Dis*, 9(11):1224-1229

Farah MG, Tverdal A, Steen TW, Heldal E, Brantsaeter AB, Bjune G. 2005. Treatment outcome of new culture positive pulmonary tuberculosis in Norway. *BMC Public Health*, 5:14

Fayad U, Uthurusamy R. 2002. Evolving data mining into solutions for insight. *Communications of the association of computing machinery*, 45(8):28-31

Faustini A, Hall AJ, Perucci CA. 2005. Tuberculosis treatment outcomes in Europe: a systematic review. *European Respiratory Journal*, 26(3):503-510

Fellegi IF, Sunter AB. 1969. A theory for record linkage. *J Am Stat Assoc*, 64:1183-1210

Fenner L, Reid SE, Fox MP, Garone D, Wellington M, Prozesky H, Zwahlen M, Schomaker M, Wandeler G, Kancheya N, Boulle A, Wood R, Henostroza G, Egger M (IeDEA Southern Africa) 2013. Tuberculosis and the risk of opportunistic infections and cancers in HIV-infected patients starting ART in Southern Africa. *Trop Med Int Health*, 18(2):194-198

Fine JP, Gray RJ. 1999. A proportional hazards model for the subdistribution of a competing risk. *J Am Stat Assoc*; 94:496–509

Franke MF, Robins JM, Mugabo J, Kaigamba F, Cain LE, Fleming JG, Murray MB. Effectiveness of Early Antiretroviral Therapy Initiation to Improve Survival among HIV-Infected Adults with Tuberculosis: A Retrospective Cohort Study. 2011. *PLoS Med*. 8(5): e1001029.

Freudenberg N, Fahs M, Galea S, Greenberg A. 2006. The impact of New York City's 1975 fiscal crisis on the tuberculosis, HIV, and homicide syndemic. *Am J Public Health*, 96:424-434

Geffen N. 2011. When to start antiretroviral therapy in adults: the resulfs of HPTN 052 move us closer to a 'Test-and-Treat' policy. *S Afr J of HIV Med*, 12(3)

Geldmacher C, Ngwenyama N, Schuetz A, Petrovas C, Reither K, Heeregrave EJ, Casazza JP, Ambrozak DR, Louder M, Ampofo W, Pollakis G, Hill B, Sanga E, Saathoff E, Maboko L, Roederer M, Paxton WA, Hoelscher M, Koup RA. 2010. Preferential infection and depletion of Mycobacterium tuberculosis-specific CD4 T cells after HIV-1 infection. *J Exp Med*, 207(13):2869-2881

Gideon HP, Du Toit E, Maartens G, van Cutsem G, Wilkinson KA, Wilkonson RJ, Rangakal MX. 2009. Evaluation of IGRA for detection of prevalent tuberculosis (TB) amongst asymptomatic HIV-1 infected adults on combined antiretroviral treatment (ART) being screened for a TB prevention study in Khayelitsha, South Africa. Proceedings of the 5th International AIDS Society Conference on HIV Pathogenesis, Treatment and Prevention, Cape Town, South Africa; July 19–22, 2009. Abstract TUPEB154.107

Girardi E, Palmieri F, Angeletti C, Vanacore P, Matteelli A, Gori A, Carbonara C, and Ippolito G. 2012. Impact of Previous ART and of ART Initiation on Outcome of HIV-Associated Tuberculosis. *Clin Dev Immunol*, ID 931325.

Goluba JE, Saracenib V, Cavalcanteb VC, Pachecob AG, Moultona LH, Kinga BS, Efrona A, Moorea RD, Chaissona RE, and Durovni B. 2007. The impact of antiretroviral therapy and isoniazid preventive therapy on tuberculosis incidence in HIV-infected patients in Rio de Janeiro, Brazil. *AIDS*, 21:1441-1448

Grieg J, Casas EC, O'Brien DP, Mills EJ, Ford N. 2012. Association between older age and adverse outcomes on antiretroviral therapy: a cohort analysis of programme data from nine countries. *AIDS*, Suppl 1:31-37

Grimwade K, Sturm AW, Nunn AJ, Mbatha D, Zungu D, Gilks CF. 2005. Effectiveness of cotrimoxazole prophylaxis on mortality in adults with tuberculosis in rural South Africa. AIDS. 28;19(2):163-8.

Grimwade K and Swingler GH. Cotrimoxazole prophylaxis for opportunistic infections in adults with HIV. 2009. Editorial Group: *Cochrane HIV/AIDS Group* (http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD003108/full)

Gupta A, Wood R, Kaplan R, Bekker L-G, Lawn SD. 2012. Tuberculosis Incidence Rates during 8 Years of Follow-Up of an Antiretroviral Treatment Cohort in South Africa: Comparison with Rates in the Community. *PLoS ONE*, 7(3):e34156

Haar CH, Cobelens FG, Kalisvaart NA, van Gerven PJ, van der Have JJ. 2007. HIV-related mortality among tuberculosis patients in The Netherlands, 1993–2001. *Int J Tuberc Lung Dis*, 11:1038-1041

Haines A, Ashcroft R, Coggon D, Coulter A, Doyal L, Gadd E, Gillis C, Pfeffer N, Wadsworth M, Haines A, Ashcroft R, Coggon D, Coulter A, Doyal L, Gadd E, Gillis C, Pfeffer N, Wadsworth M, Walker P, Wand M. 2000. Personal information in medical research. http://www.mrc.ac.uk London: MRC

Harries AD, Hargreaves NJ, Kemp J, Jindani A, Enarson DA, Maher D, Salaniponi FM. 2001. Deaths from tuberculosis in sub-Saharan African countries with a high prevalence of HIV-1. *Lancet*, 357(9267):1519-1523

Havlir DV, Kendall MA, Ive P, Kumwenda J, Swindells S, Qasba SS, Luetkemeyer AF, Hogg E, Rooney J, Wu X, Hosseinipour MC, Lalloo U, Veloso VG, Some FF, Kumarasamy N, Padayatchi N, Santos BR, Reid S, Hakim J, Mohapi L, Mugyenyi P, Sanchez J, Lama JR, Pape JW, Sattler FR, Asmelash A, Moko E, Sawe F, Andersen J, Sanne I. For the AIDS Clinical Trials Group Study A5221. 2011. Timing of Antiretroviral Therapy for HIV-1 Infection and Tuberculosis. *N Engl J Med*, 365(16):1482-1491

Hermans SM, Kiragga AN, Schaefer P, Kambugu A, Hoepelman AI, Manabe YC. 2010. Incident tuberculosis during antiretroviral therapy contributes to suboptimal immune reconstitution in a large urban HIV clinic in sub-Saharan Africa. *PLoS ONE*, 5(5):e10527

Hermans SM, Castelnuovo B, Katabira C, Mbidde P, Lange JM, Hoepelman AI, Coutinho A, Manabe YC. 2012. Integration of HIV and TB services results in improved TB treatment outcomes and earlier prioritized ART initiation in a large urban HIV clinic in Uganda. *J Acquir Immune Defic Syndr*. 60(2):e29-35.

Hernan MA, Brumbacm B, Robins JM. 2000. Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. Epidemiology 11(5): 561-570

Heunis C. 2011. Accuracy of Tuberculosis Routine Data and Nurses' Views of the TB-HIV Information System in the Free State, South Africa. *Journal of the association of nurses in AIDS care,* 22 (1): 67-73.

HIV Causal collaboration 2012. Impact of Antiretroviral Therapy on Tuberculosis Incidence Among HIV-Positive Patients in High-Income Countries. *Clin Infect Dis*, 54(9):1364–1372

Hser Y, Evans E. 2008. Cross-system data linkage for treatment outcome evaluation: Lessons learned from the California Treatment Outcome project. *Evaluation and program planning*, 31:125-135

Ingle SM, May M, Uebel K, Timmerman V, Kotze E, Bachmann M, Sterne JAC, Egger M, Fairall L, for IeDEA-Southern Africa. 2010. Outcomes in patients waiting for antiretroviral treatment in the Free State Province, South Africa: prospective linkage study. *AIDS*, 24:2717-2725

Ingle SM, May M, Uebel K, Timmerman V, Kotze E, Bachmann M, Sterne JAC, Egger M, Fairall L, for IeDEA-Southern Africa. 2010b. Differences in access and patient outcomes across antiretroviral treatment clinics in the Free State Province, South Africa: a prospective cohort study. *S Afr Med J*, 100:675-681b

Jaro MA. 1989. Advances in record-linkage methodology as applied to matching the 1983 Census of Tampa, Florida. *JASA*, 89:414-420.

Jaro MA. Probabilistic linkage of large public health data files. 1995. *Statistics in Medicine*, 14:491-498.

Johnson WD, Diaz RM, Flanders WD, Goodman M, Hill AN, Holtgrave D, Malow R, McClellan WM. 2008. Behavioural interventions to reduce risk for sexual transmission of HIV among men who have sex with men. *Cochrane Database Syst Rev,* 16(3):CD001230.

Joubert G. and Ehrlich R. 2007. Epidemiology: A research manual for South Africa, 2nd edition. Published by Oxford University Press South Africa.

Jarvis JN, Govender N, Chiller T, Park BJ, Longley N, Meintjes G, Bekker LG, Wood R, Lawn SD, Harrison TS. 2012. Cryptococcal antigen screening and preemptive therapy in patients initiating antiretroviral therapy in resource-limited settings: a proposed algorithm for clinical implementation. *J Int Assoc Physicians AIDS Care (Chic)*, 11(6):374-379

Jayasundera CI, Attapattu M, Kumarasinghe MP. 1993. Atypical presentations of pulmonary tuberculosis diagnosed by fibreoptic bronchoscopy. *Postgrad Med J*, 69(814):621-623

Jones JL, Hanson DL, Dworkin MS, DeCock KM. 2000. HIV-associated tuberculosis in the era of highly active antiretroviral therapy. The Adult/Adolescent Spectrum of HIV Disease Group. *Int J Tuberc Lung Dis*, 4(11):1026-1031

Kanyerere HS, Mpunga J, Tweya H, Edginton M, Harries AD, Hinderaker SG, Chimbwandira F, Gonani A, Mbendera K. 2012. Timing of antiretroviral therapy and effects on tuberculosis treatment outcomes in HIV-co-infected patients in Malawi. *Public Health Action*, 2(4):174-177

Kaplan EL, Meier P. 1958. Nonparametric estimation from incomplete observations. *J Amer Statt Assoc*, 53:457–481

Kerkhoff AD, Kranzer K, Samandari T, Nakiyingi-Miiro J, Whalen CC, Harries AD, Lawn SD. 2012. Systematic review of TST responses in people living with HIV in under-resourced settings: implications for isoniazid preventive therapy. *PLoS ONE*, 7(11):e49928

Kerschberger B, Hilderbrand K, Boulle AM, Coetzee D, Goemaere E, De Azevedo V, Van Cutsem G. 2012. The Effect of Complete Integration of HIV and TB Services on Time to Initiation of Antiretroviral Therapy: A Before-After Study. *PLoS ONE*, 7(10):e76988

Kotze E, McDonald T. 2010. Using record linkage to conform the patient dimension of an antiretroviral therapy data warehouse. *Proceedings of the IASTED African Conference Health informatics*.

Kwan C, Ernst JD 2011. HIV and tuberculosis: a deadly human syndemic. *Clin Microbiol Rev*, 24(2):351-376

Kwara A, Ramachandran G, Swaminathan S. 2010. Dose adjustment of the nonnucleoside reverse transcriptase inhibitors during concurrent rifampicin-containing tuberculosis therapy: one size does not fit all. *Expert Opin Drug Metab Toxicol*, 6:55-68

Lawn SD, Badri M, Wood R 2005. Tuberculosis among HIV-infected patients receiving HAART: long term incidence and risk factors in a South African cohort. *AIDS*, 19:2109-2116

Lawn SD, Bekker LG, Miller RF. 2005. Immune reconstitution disease associated with mycobacterial infections in HIV-infected individuals receiving antiretrovirals. *Lancet Infect Dis*, 5(6):361-73

Lawn SD, Myer L, Orrell C, Bekker LG, Wood R. 2005. Early mortality among adults accessing a community-based antiretroviral service in South Africa: implications for programme design. *AIDS*, 19:2141-2148

Lawn SD,Myer L, Bekker LG. 2006. Burden of tuberculosis in an antiretroviral treatment programme in sub-Saharan Africa: impact on treatment outcomes and implications for tuberculosis control. *AIDS*, 20:1605-1612

Lawn SD, Myer L, Bekker LG, Wood R 2007. Tuberculosis-associated immune reconstitution disease: incidence, risk factors and impact in an antiretroviral treatment service in South Africa. *AIDS*, 21:335-341

Lawn SD, Wood R. 2007. Optimum time to initiate antiretroviral therapy in patients with HIV-associated tuberculosis: there may be more than one right answer. *J Acquir Immune Defic Syndr*, 46:121-123

Lawn SD, Edwards DJ, Wood R. 2007. Concurrent drug therapy for tuberculosis and HIV infection in resource-limited settings: present status and future prospects. *Future HIV Therapy*, 1(4):387-398

Lawn SD, Wilkinson RJ, Lipman MC, Wood R. 2008. Immune reconstitution and "unmasking" of tuberculosis during antiretroviral therapy. *Am J Respir Crit Care Med*, 177(7):680-685

Lawn SL, Myer L, Edwards D, Bekker L-G, Wood R 2009. Short-term and long-term risk of tuberculosis associated with CD4 cell recovery during antiretroviral therapy in South Africa. *AIDS*, 23:1715-1717

Lawn SD, Churchyard GJ. 2009. Epidemiology of HIV-associated tuberculosis. *Curr Opin HIV AIDS*, 4:325–333

Lawn SD, Myer L, Edwards D, Bekker LG, Wood R. 2009. Short-term and long-term risk of tuberculosis associated with CD4 cell recovery during antiretroviral therapy in South Africa. *AIDS*, 23(13):1717–1725

Lawn SD, Wood R, De Cock KM, Kranzer K, Lewis JJ, Churchyard GJ. 2010. Antiretrovirals and isoniazid preventive therapy in the prevention of HIV-associated tuberculosis in settings with limited health-care resources. *Lancet Infect Dis*, 10:489-498

Lawn, Torok E, Wood R. 2011. Optimum time to start antiretroviral therapy during HIV-associated opportunistic infections. *Curr Opin Infect Dis*, 24:34-42

Lawn SD, Campbell L, Kaplan R, Little F, Morrow C, Wood R; IeDEA-Southern Africa. 2011. Delays in starting antiretroviral therapy in patients with HIV-associated tuberculosis accessing non-integrated clinical services in a South African township. *BMC Infect Dis*, 11:258

Levesque LE, Hanley JA, Kezuoh, Suissa S 2010. Problem of immortal time bias in cohort studies: example using statins for preventing progression of diabetes. *BMJ*, 340:b5087

López-Gatell H, Cole SR, Margolick JB, Witt MD, Martinson J, Phair JP, Jacobson LP; Multicenter AIDS Cohort Study. 2008. Effect of tuberculosis on the survival of HIV-infected men in a country with low tuberculosis incidence. *AIDS*, 22:1869-1873

Louwagie GM, Bachmann MO, Meyer K, Booysen Fle R, Fairall LR, Heunis C. 2007. Highly active antiretroviral treatment and health related quality of life in South African adults with human immunodeficiency virus infection: A cross-sectional analytical study. *BMC Public Health*. 14(7):244.

Lu SH, Tian BC, Kang XP, Zhang W, Meng XP, Zhang JB, Lo SK. 2009. Public awareness of tuberculosis in China: a national survey of 69 253 subjects. *Int J Tuberc Lung Dis*, 13:1493-1499

Lucas SB, Hounnou A, Peacock C, Beaumel A, Djomand G, N'Gbichi JM, Yeboue K, Hondé M, Diomande M, Giordano C, et al 1993. The mortality and pathology of HIV infection in a west African city. *AIDS*, 7(1569-1579):

Mahajan V, Verma SK. 2008. HIV-Tuberculosis Co-Infection. *Internet J of Pulmonary Medicine* 10(1).

Makombe SD, Harries AD, Yu JK, Hochgesang M, Mhango E, Weigel R, Pasulani O, Fitzgerald M, Schouten EJ, Libamba E. 2007. Outcomes of tuberculosis patients who start antiretroviral therapy under routine programme conditions in Malawi. *Int J Tuberc Lung Dis*, 11:412-416

Manosuthi W, Tantanathip P, Chimsuntorn S, Eampokarap B, Thongyen S, Nilkamhang S, Sungkanuparph S. 2010. Treatment outcomes of patients co-infected with HIV and tuberculosis who received a nevirapine-based antiretroviral regmen: a four-year prospective study. *Int J Inf Dis*, 14:e1013-e1017

Mayosi BM, Lawn JE, van Niekerk A, Bradshaw D, Abdool Karim SS, Coovadia HM; Lancet South Africa team. 2012. Health in South Africa: changes and challenges since 2009. *Lancet*, 380(9858):2029-2043

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

Meintjes G, Lawn SD, Scano F, Maartens G, French MA, Worodria W, Elliott JH, Murdoch D, Wilkinson RJ, Seyler C, John L, van der Loeff MS, Reiss P, Lynen L, Janoff EN, Gilks C, Colebunders R. 2008. International Network for the Study of HIV-associated IRIS. Tuberculosis-associated immune reconstitution inflammatory syndrome: case definitions for use in resource-limited settings. *Lancet Infect Dis*, 8(8):516-523

Melosini L, Vetrano U, Dente FL, Cristofano M, Giraldi M, Gabbrielli L, Novelli F, Aquilini F, Rindi L, Menichetti F, Freer G, Paggiaro PL. 2012. Evaluation of underreporting tuberculosis in Central Italy by means of record linkage. *BMC Public Health* 12:472

Mohan A, Sharma SK. 2008. Fibreoptic bronchoscopy in the diagnosis of sputum smear-negative pulmonary tuberculosis: current status. *Indian J Chest Dis Allied Sci*, 50(1):67-78

Moolphate S, Aung MN, Nampaisan O, Nedsuwan S, Kantipong P, Suriyon N, Hansudewechakul C, Yanai H, Yamada N, Ishikawa N. 2011. Time of highest tuberculosis death risk and associated factors: an observation of 12 years in Northern Thailand. *International Journal of General Medicine*, 4:181-190

Moore D, Liechty C, Ekwaru P, Were W, Mwima G, Solberg P, Rutherford G, Mermin J. 2007. Prevalence, incidence and mortality associated with tuberculosis in HIV-infected patients initiating antiretroviral therapy in rural Uganda. *AIDS*, 21(6):713-719

Mugusi FM, Mehta S, Villamor E, Urassa W, Saathoff E, Bosch RJ, Fawzi WW. 2009. Factors associated with mortality in HIV-infected and uninfected patients with pulmonary tuberculosis. *BMC Public Health*, 9:409

Muniyandi M, Ramachandran R, Gopi PG, Chandrasekaran V, Subramani R, Sadacharam K, Kumaran P, Santha T, Wares F, Narayanan PR. 2007. The prevalence of tuberculosis in different economic strata: a community survey from South India. *Int J Tuberc Lung Dis*, 11(9):1042-1045

Naidoo K, Baxter C, Abdool Karim SS. 2013. When to start antiretroviral therapy during tuberculosis treatment? *Curr Opin Infect Dis*, 26:35-42

Newcombe HB, Kennedy JM. 1962. Record linkage. *Communication of the ACM*, 5: 563-566

Nicholas S Sabapathy K, Ferreyra C, Varaine F, Pujades-Rodr´ıguez M. (for the AIDS Working Group of Me´decins Sans Frontie`res). 2011. Incidence of Tuberculosis in HIV-Infected Patients Before and After Starting Combined Antiretroviral Therapy in 8 Sub-Saharan African HIV Programs. *J Acquir Immune Defic Syndr*, 57:311-318

Niemi M, Backman JT, Fromm MF, Neuvonen PJ, Kivisto KT. 2003. Pharmacokinetic interactions with rifampicin: clinical relevance. *Clin Pharmacokinet*, 42:819-850

Obuku EA, Meynell C, Kiboss-Kyeyune J, Blankley S, Atuhairwe C, Nabankema E, Lab M, Jeffrey N, Ndungutse D. 2012. Socio-demographic determinants and prevalence of Tuberculosis knowledge in three slum populations of Uganda. *BMC Public Health*, 12:536

Pablos-Mendez A, Sterling TR, Frieden TR. 1996. The relationship between delayed or incomplete treatment and all-cause mortality in patients with tuberculosis. *JAMA*, 276:1223-1228

Parker JD, Woodruff TJ, Akinbami LJ, Kravets N. Linkage of the US national health interview survey to air monitoring data: An evaluation of different strategies. 2008. *Environmental Research*, 106 :384-392

Perriens JH, St Louis ME, Mukadi YB, Brown C, Prignot J, Pouthier F, Portaels F, Willame JC, Mandala JK, Kaboto M, Ryder RW, Roscigno G, Piot P. 1995. Pulmonary tuberculosis in HIV-infected patients in Zaire. A controlled trial of treatment for either 6 or 12 months. *New England Journal of Medicine*; 332(12):779-84

Pettit AC, Jenkins CA, Stinnette SE, Rebeiro PF, Blackwell RB, Raffanti SP, Shepherd BE, Sterling TR. 2011. Tuberculosis Risk Before and After Highly Active Antiretroviral Therapy Initiation: Does HAART Increase the Short-Term TB Risk in a Low Incidence TB Setting? *J Acquir Immune Defic Syndr*, 57:305-310

Qauad MG, Zhang H. 2009. Accuracy of public health data linkages. *Matern Child Health J*, 13:531-538

Rajasekaran S, Raja K, Jeyaseelan L, Vijilat S, Priya K, Mohan K, Parvez A, Mahilmaran A, Chandrasekar C. 2009. Post-HAART tuberculosis in adults and adolescents with HIV in India: incidence, clinical and immunological profile. *Indian J Tuberc*, 56(2):69-76

Rana FS, Hawken MP, Mwachari C, Bhatt SM, Abdullah F, Ng'ang'a LW, Power C, Githui WA, Porter JD, Lucas SB. 2000. Autopsy study of HIV-1-positive and HIV-1-negative adult medical patients in Nairobi, Kenya. *J Acquir Immune Defic Syndr*, 24:23-29

Robins JM, Hernan MA, Brumback B. 2000. Marginal structural models and causal inference in epidemiology. *Epidemiology* 11: 550-560

Rogot F, Feinleib M, Ockay KA, Schwartz SH, Bilgrad R, Patterson JE. 1983. The feasibility of linking samples to the National Death Index for epidemiologic studies: a progress report. *American journal of public health*, 74: 1256

Roos LL, Wajda A, Nicol JP. 1986. The art and science of record linkage: Methods that work with few identifiers. *Comput Biol Med*, Vol 16, No 1: 45-57

Royston P. 2004. Multiple imputation of missing values. *Stata J*. 4(3):227-241.

Rubin D. 1987. Multiple imputation for nonresponse in surveys. New York: Wiley.

Sanguanwongse N, Cain KP, Suriya P, Nateniyom S, Yamada N, Wattanaamornkiat K, Sumnapan S, Sattayawuthipong W, Kaewsa-ard S, Ingkaseth S, Varma JK. 2008. Antiretroviral Therapy for HIV-Infected Tuberculosis Patients Saves Lives but Needs to Be Used More Frequently in Thailand. *J Acquir Immune Defic Syndr*, 48:181-189

Saraceni V, King BS, Cavalcante SC, Golub JE, Lauria LM, Moulton LH, Chaisson RE, Durovni B. 2008. Tuberculosis as primary cause of death among AIDS cases in Rio de Janeiro, Brazil. *Int J Tuberc Lung Dis*, 12:769-772

Sarder N, Hossain S, Huda M, Eunus A, Rahman M. 2006. Consequence on Treatment of TB Patients Affected by HIV/AIDS: A Conceptual Research. *Am J Infect Dis*, 2(4):210-218

Schneider H, Blaaw D, Gilson L, Chabikuli N, Goudge J. 2006. Health systems and access to antiretroviral drugs for HIV in Southern Africa: service delivery and human resource challenges. *Reproductive Health Matters* 14(27):12-23.

Shaweno D, Worku A. 2012. Tuberculosis treatment survival of HIV positive TB patients on directly observed treatment short-course in Southern Ethiopia: A retrospective cohort study. *BMC Res Notes*, 5(1):682

Silva DR, Menegotto DM, Schulz LF, Gazzana MB, Dalcin Pde T. 2010. Factors associated with mortality in hospitalized patients with newly diagnosed tuberculosis. *Lung*, 188:33-41

StataCorp. 2009. Stata Statistical Software: Release 11. College Station, TX: StataCorp LP.

Statistics South Africa. Mortality and Causes of Death in South Africa, 1997-2003: Findings From Death Notification. Pretoria: Statistics South Africa; 2005.

Steenland K. 2013. Marginal structural models to control for time-varying confounding in occupational and environmental epidemiology. *Occup Environ Med* Published. Online First: 9 July 2013, doi:10.1136/oemed-2013-101629

Sterne JA, May M, Costagliola D, de Wolf F, Phillips AN, Harris R, Funk MJ, Geskus RB, Gill J, Dabis F, Miró JM, Justice AC, Ledergerber B, Fätkenheuer G, Hogg RS, Monforte AD, Saag M, Smith C, Staszewski S, Egger M, Cole SR. (When To Start Consortium). 2009. Timing of initiation of antiretroviral therapy in AIDS-free HIV-1-infected patients: a collaborative analysis of 18 HIV cohort studies. *Lancet*, 373(9672):1352-1362 April 18

Steyn F, Van Rensburg D, Engelbrecht M. 2006. Human resource for ART in the Free State public sector: recording achievements, identifying challenges. *Acta Academica* Supplementum:94-139

Stiell A, Forster AJ, Stiell IG, van Walraven C. 2003. Prevalence of information gaps in the emergency department and the effect on patient outcomes. *CMAJ* 169(10):1023-8

Straetemans M, Bierrenbach AL, Nagelkerke N, Glaziou P, van der Werf MJ. 2010. The effect of tuberculosis on mortality in HIV positive people: a meta-analysis. *PLoS One*, 5(12):e15241.

Styblo K. 1985. The relationship between the risk of tuberculosis infection and the risk of developing infectious tuberculosis. *Bull Int Union Tuberc Lung Dis*, 60:117-119

Sungkanuparph S, Eampokalap B, Chottanapund S, Thongyen S, Manosuthi W. 2006. Impact of drug-resistant tuberculosis on the survival of HIV-infected patients. *Int J Tuberc Lung Dis*, 11(3):325-330

Suthar AB, Granich R, Mermin J, Van Rie A. 2012a. Effect of cotrimoxazole on mortality in HIV-infected adults on antiretroviral therapy: a systematic review and meta-analysis. *Bull World Health Organ*, 90(2):128C-138C

Suthar AB, Lawn SD, del Amo J, Getahun H, Dye C, Sculier D, Sterling TR, Chaisson RE, Williams BG, Harries AD, Granich RM. 2012b. Antiretroviral Therapy for Prevention of Tuberculosis in Adults with HIV: A Systematic Review and Meta-Analysis. *PLoS Med*, 9(7):e1001270

Tanser F, Bärnighausen T, Grapsa E, Zaidi J, Newell ML. 2013. High coverage of ART associated with Decline in risk of HIV acquisition in rural KwaZulu-Natal. *Science*, 339(6122): 966-971

Tessema B, Muche A, Bekele A, Reissig D, Emmrich F, Sack U. 2009. Treatment outcome of tuberculosis patients at Gondar University Teaching Hospital, Northwest Ethiopia. A five–year retrospective study. *BMC Public Health*, 9:371

Tromp M, Raveli AC, BOnsel GJ, Hasman A, Reitsma JB. 2011. Results from simulated data sets: probabilistic record linkage outperforms deterministic record linkage. *J Clin Epidemiol.* 64(5):565-72

Thornton SN, Hood SK. 2005. Reducing duplicate patient creation using a probabilistic matching algorithm in an open access community data sharing environment. *AMIA Annu Symp Proc,* 1135-1136

Uebel KE, Timmerman V, Ingle SM, van Rensburg DHC, Mollentze WF. 2010. Towards universal ARV access: Achievements and challenges in Free State Province, South Africa. *S Afr Med J* , 100:589-593

Van Hest NAH, Smit F, Baars HWM, De Vries G, De Haas PEW, Westenend PJ, Nagelkerke NHD, Richardus JH. 2006. Completeness of notification of tuberculosis in The Netherlands: how reliable is record-linkage and capture-recapture analysis? *Epidemio Infect* 135, 1021-1029

Van Rie A, Westreich D, Sanne I. 2011. Tuberculosis in patients receiving antiretroviral treatment: incidence, risk factors and prevention strategies. *J Acquir Immune Defic Syndr,* 56(4):349-355

Velasco M, Castilla V, Sanz J, Gaspar G, Condes E, Barros C, Cervero M, Torres R, Guijarro C. (COMESEM Cohort). 2009. Effect of simultaneous use of highly active antiretroviral therapy on survival of HIV patients with Tuberculosis. *J Acquir Immune Defic Syndr,* 50:148-152

Violari A, Paed FC, Cotton MF, Gibb DM, Babiker AG, Steyn J, Madhi SA, Paed FC, Jean-Philippe P, McIntyre JA, (F.R.C.O.G. for the CHER Study Team). 2008. Early antiretroviral therapy and mortality among HIV-infected infants. *N Engl J Med,* 359:2233-2244

Westreich D, MacPhail P, Van Rie A, Malope-Kgokong B, Ive P, Rubel D, Boulme R, Erond J, Sanne I. 2009. Effect of pulmonary tuberculosis on mortality in patients receiving HAART. *AIDS,* 23:707-715

Williams BG, Granich R, De Cock KM, Glaziou P, Sharma A, Dye C. 2010. Antiretroviral therapy for tuberculosis control in nine African countries. *PNAS,* 107 (45): 19485-19489

Williams BG. 2013. Could ART increase the population level incidence of TB? South African Centre for Epidemiological Modelling and Analysis (SACEMA). http://arxiv.org/abs/1302.0503.

Winkler WE. Overview of record linkage and current research directions. Research report series, Statistics 2006

Wong EB, Omar T, Setlhako G, Osih R, Murdoch D, Martinson N, Feldman C, Bangsberg D, Venter WDF. 2010. Causes of death in ART-treated adults: a post-mortem study from Johannesburg. Abstracts of the XVIII International AIDS Conference. International AIDS Society. Vienna, Austria. Abstract #WEPE0154.

Worodria W, Massinga-Loembe M, Mazakpwe D, Luzinda K, Menten J, Van Leth F, Mayanja-Kizza H, Kestens L, Mugerwa RD, Reiss P, Colebunders R. (TB-IRIS Study Group) 2011. Incidence and predictors of mortality and the effect of Tuberculosis immune reconstitution inflammatory syndrome in a cohort of TB/HIV patients commencing antiretoviral therapy. *J Acquir Immune Defic Syndr*, 58:32-37

Yen YF, Muh-Yong Yen MY, Shih HC, Deng CY. 2012. Risk factors for unfavorable outcome of pulmonary tuberculosis in adults in Taipei, Taiwan. *Trans R Soc Trop Med Hyg*, 106:303-308

Yotebieng M, Van Rie A, Moultrie H, Cole SR, Adimora A, Behets F. 2010. Effect on mortality and virological response of delaying antiretroviral therapy initiation in children receiving tuberculosis treatment. *AIDS*, 24(1341-1349):

Younis MZ, Rivers PA, Fottler MD. 2005. The impact of HMO and hospital competition on hospital costs. *J Health Care Finance*, 31:60-74

Zachariah R, Fitzgerald M, Massaquoi M, Pasulani O, Arnould L, Makombe S, Harries AD. 2006. Risk factors for high early mortality in patients on antiretroviral treatment in a rural district of Malawi. *AIDS*, 20:2355-2360

Zachariah R, Fitzgerald M, Massaquoi M, Acabu A, Chilomo D, Salaniponi FM, Harries AD. 2007. Does antiretroviral treatment reduce case fatality among HIV-positive patients with tuberculosis in Malawi? *Int J Tuberc Lung Dis*, 11(8):848-853

Zingmond DS, Ye Z, Ettner SL, Liu H. 2004. Linking hospital discharge and death records accuracy and sources of bias. *Journal of Clinical Epidemiology*, 57:21-29

REFERENCES: INTERNET LINKS

WORLD HEALTH ORGANISATION (WHO) : Tested on 10 February 2013

WHO Interim policy 2004	http://www.who.int/hiv/pub/tb/tbhiv/en/index.html
WHO Global TB Report, 2009	http://www.who.int/tb/publications/global_report/2009/key_points/en/index.html
WHO Global TB report, 2011	http://www.who.int/entity/tb/publications/globalreport/2011/gtbr11_full.pdf
WHO Global TB Report, 2012	http://www.who.int/tb/publications/global_report/en/
WHO Millennium Development Goals, 2012	http://www.who.int/topics/millennium_development_goals/en/
WHO TB and HIV	http://www.who.int/hiv/topics/tb/en/
WHO TB fact sheet, 2012	http://www.who.int/mediacentre/factsheets/fs104/en/index.html
WHO TB South Africa 2012	http://www.who.int/tb/country/data/profiles/en/index.html
WHO for IPT, 2011	http://whqlibdoc.who.int/publications/2011/9789241500708_eng.pdf
WHO preventative therapy policy, 1998	http://whqlibdoc.who.int/hq/1998/WHO_TB_98.255.pdf
WHO Global Health Observatory	http://www.who.int/gho/en/
WHO Health Strategy:	http://www.who.int/hiv/aboutdept/strategyconsultation/en/index.html

SOUTH AFRICAN LINKS : Tested on 10 February 2013

South African government information:	http://www.info.gov.za/aboutsa/provinces.htm#freestate
ART Treatment Guidelines 2004	http://www.kznhealth.gov.za/arv/arv5.pdf
National Strategic plan, 2000-2005	http://www.info.gov.za/otherdocs/2000/aidsplan2000.pdf
National Strategic plan, 2007-2012	http://www.doh.gov.za/docs/policy/2007/part1.pdf
National Strategic plan, 2012-2016	http://www.doh.gov.za/docs/stratdocs/2012/NSPsum.pdf
Antenatal sentinel survey report, 2012	http://www.doh.gov.za/docs/reports/2012/Antenatal_Sentinel_survey_Report2012_final.pdf

OTHERS : Tested on 10 February 2013

Irin News, Report 94365	http://www.avert.org/aidssouthafrica.htm
Stats SA July 2007	http://www.statssa.gov.za/publications/P0302/P03022007.pdf
World Bank Report 2012	http://www.tradingeconomics.com/south-africa/incidence-of-tuberculosis-per-100-000-people-wb-data.html

UNAIDS LINKS : Tested on 10 February 2013

UNAIDS HIV epidemic report, 2010	http://www.unaids.org/globalreport/Global_report.htm
UNAIDS, Aids at 30, 2011	http://www.unaids.org/unaids_resources/aidsat30/aids-at-30.pdf
UNAIDS global report 2012	http://www.unaids.org/en/resources/campaigns/20121120_globalreport2012/globalreport/
UNGASS Global AIDS progress response, 2012	http://www.unaids.org/en/regionscountries/countries/southafrica/
UN Millennium goals	http://www.un.org/millenniumgoals/bkgd.shtml

APPENDIX A:
LOGISTIC REGRESSTION MODELS TO EVALUATE THE EFFECT OF ART IN SURVIVORS

Table 1: Multivariable regression model for the effect of ART and baseline covariates on cured TB outcome in patients who did not die during TB treatment.

Reference category	Explanatory variable	OR	P>|t|	95% Confidence Interval	
Fezile Dabi District	Lejweleputswa	0.85	0.226	0.66	1.10
	Motheo	0.50	0.000	0.39	0.64
	Thabo Mofutsanyane	0.99	0.953	0.79	1.25
	Xhariep	0.41	0.000	0.27	0.61
Year when TB treatment was started: 2006	2004	1.00	0.980	0.81	1.22
	2005	1.14	0.110	0.97	1.33
	2007	0.74	0.003	0.61	0.90
	2008	0.80	0.013	0.66	0.95
Women	Men	0.96	0.524	0.86	1.08
Age category: 30-39 years	<20	1.13	0.653	0.67	1.90
	20-29	0.99	0.896	0.87	1.13
	40-49	1.03	0.627	0.91	1.17
	>50	0.86	0.099	0.72	1.03
CD4 cell count closest to starting TB treatment: 100-199 cells/µl	CD4 0-24	0.82	0.085	0.65	1.03
	CD4 25-49	0.88	0.232	0.72	1.08
	CD4 50-99	0.92	0.345	0.77	1.10
	CD4 200-349	1.04	0.719	0.84	1.28
	CD4 350-2000	1.18	0.129	0.95	1.46
CD4 test date during the period one year preceding TB diagnosis and 30 days there after	CD4 test date after 30 days of TB treatment	1.28	0.002	1.10	1.50
	CD4 missing	1.37	0.000	1.16	1.61
Anatomical classification of TB: Pulmonary	Extra-pulmonary	0.00	0.000	0.00	0.01
	Both TB types	0.64	0.010	0.45	0.90
Treatment case: New case	Retreatment case	0.66	0.000	0.58	0.75
ART use: Did not receive ART during TB treatment episode	Started ART during first two months of TB treatment	0.94	0.729	0.68	1.31
	Started ART between two months and end of TB treatment	1.39	0.000	1.17	1.65
No valid South African ID number	Valid South African ID number	1.09	0.153	0.97	1.23

Death source: Deaths registered on TB register and HIV programme database

Table 2: Multivariable regression model for the effect of ART and baseline covariates on completed TB outcome in patients who did not die during TB treatment.

Reference category	Explanatory variable	OR	P>|t|	95% Confidence Interval	
Fezile Dabi District	Lejweleputswa	1.15	0.584	0.70	1.91
	Motheo	1.93	0.008	1.19	3.13
	Thabo Mofutsanyane	1.92	0.007	1.19	3.10
	Xhariep	1.76	0.031	1.05	2.94
Year when TB treatment was started: 2006	2004	1.60	0.002	1.19	2.13
	2005	1.62	0.002	1.20	2.19
	2007	0.75	0.014	0.60	0.94
	2008	0.77	0.022	0.61	0.96
Women	Men	1.19	0.009	1.04	1.35
Age category: 30-39 years	<20	0.74	0.437	0.35	1.57
	20-29	0.91	0.162	0.80	1.04
	40-49	0.96	0.690	0.80	1.16
	>50	1.09	0.493	0.86	1.38
CD4 cell count closest to starting TB treatment: 100-199 cells/µl	CD4 0-24	0.82	0.262	0.57	1.17
	CD4 25-49	0.94	0.656	0.70	1.26
	CD4 50-99	0.99	0.963	0.77	1.29
	CD4 200-349	1.18	0.240	0.89	1.56
	CD4 350-2000	1.36	0.081	0.96	1.93
CD4 test date during the period one year preceding TB diagnosis and 30 days there after	CD4 test date after 30 days of TB treatment	0.72	0.009	0.56	0.92
	CD4 missing	0.99	0.928	0.85	1.16
Anatomical classification of TB: Pulmonary	Extra-pulmonary	1.30	0.008	1.07	1.57
	Both TB types	1.27	0.198	0.88	1.84
Smear results: Smear negative	Smear positive	0.01	0.000	0.01	0.02
	No smear results	0.34	0.000	0.26	0.45
Treatment case: New case	Retreatment case	1.10	0.385	0.88	1.38
ART use: Did not receive ART during TB treatment episode	Started ART during first two months of TB treatment	1.56	0.004	1.15	2.10
	Started ART between two months and end of TB treatment	2.74	0.000	2.24	3.34
No valid South African ID number	Valid South African ID number	1.15	0.197	0.93	1.44

Death source: Deaths registered on TB register and HIV programme database

Table 3: Multivariable regression model for the effect of ART and baseline covariates on successful TB outcome in patients who did not die during TB treatment.

Reference category	Explanatory variable	OR	P>|t|	95% Confidence Interval	
Fezile Dabi District	Lejweleputswa	1.38	0.100	0.94	2.02
	Motheo	1.71	0.007	1.16	2.52
	Thabo Mofutsanyane	2.31	0.000	1.56	3.41
	Xhariep	1.97	0.005	1.22	3.16
Year when TB treatment was started: 2006	2004	1.30	0.065	0.98	1.72
	2005	1.49	0.001	1.17	1.90
	2007	0.76	0.006	0.62	0.92
	2008	0.88	0.222	0.72	1.08
Women	Men	0.94	0.419	0.82	1.09
Age category: 30-39 years	<20	0.69	0.209	0.38	1.23
	20-29	0.88	0.039	0.78	0.99
	40-49	1.10	0.094	0.98	1.24
	>50	1.03	0.781	0.84	1.26
CD4 cell count closest to starting TB treatment: 100-199 cells/µl	CD4 0-24	0.66	0.005	0.50	0.88
	CD4 25-49	0.81	0.089	0.64	1.03
	CD4 50-99	0.91	0.372	0.73	1.13
	CD4 200-349	1.16	0.309	0.87	1.53
	CD4 350-2000	1.34	0.057	0.99	1.82
CD4 test date during the period one year preceding TB diagnosis and 30 days there after	CD4 test date after 30 days of TB treatment	0.79	0.072	0.61	1.02
	CD4 missing	1.03	0.805	0.82	1.30
Anatomical classification of TB: Pulmonary	Extra-pulmonary	1.24	0.020	1.03	1.49
	Both TB types	1.10	0.544	0.81	1.50
Smear results: Smear negative	Smear positive	1.03	0.752	0.88	1.20
	No smear results	0.35	0.000	0.27	0.46
Treatment case: New case	Retreatment case	0.85	0.032	0.73	0.99
ART use: Did not receive ART during TB treatment episode	Started ART during first two months of TB treatment	1.59	0.006	1.14	2.21
	Started ART between two months and end of TB treatment	3.86	0.000	2.89	5.14
No valid South African ID number	Valid South African ID number	1.19	0.025	1.02	1.39

Death source: Deaths registered on TB register and HIV programme database

Table 4: Multivariable regression model for the effect of ART and baseline covariates on failed TB outcome in patients who did not die during TB treatment.

Reference category	Explanatory variable	OR	P>|t|	95% Confidence Interval	
Fezile Dabi District	Lejweleputswa	0.42	0.068	0.16	1.07
	Motheo	0.80	0.587	0.36	1.79
	Thabo Mofutsanyane	0.67	0.343	0.29	1.54
	Xhariep	0.65	0.337	0.27	1.57
Year when TB treatment was started: 2006	2004	1.04	0.895	0.55	2.00
	2005	1.27	0.402	0.73	2.23
	2007	1.79	0.044	1.02	3.17
	2008	1.03	0.908	0.62	1.71
Women	Men	1.35	0.091	0.95	1.90
Age category: 30-39 years	<20	1.17	0.847	0.25	5.52
	20-29	0.87	0.475	0.60	1.27
	40-49	0.78	0.278	0.49	1.22
	>50	1.04	0.891	0.56	1.95
CD4 cell count closest to starting TB treatment: 100-199 cells/µl	CD4 0-24	1.66	0.269	0.67	4.10
	CD4 25-49	1.25	0.564	0.58	2.70
	CD4 50-99	1.27	0.501	0.63	2.59
	CD4 200-349	1.06	0.886	0.48	2.33
	CD4 350-2000	1.30	0.489	0.62	2.71
CD4 test date during the period one year preceding TB diagnosis and 30 days there after	CD4 test date after 30 days of TB treatment	1.40	0.395	0.64	3.06
	CD4 missing	1.28	0.221	0.86	1.92
Anatomical classification of TB: Pulmonary	Extra-pulmonary	0.26	0.047	0.07	0.98
	Both TB types	0.93	0.891	0.36	2.45
Smear results: Smear negative	Smear positive	3.80	0.003	1.56	9.26
	No smear results	0.88	0.794	0.33	2.35
Treatment case: New case	Retreatment case	1.48	0.134	0.89	2.46
ART use: Did not receive ART during TB treatment episode	Started ART during first two months of TB treatment	0.89	0.819	0.32	2.46
	Started ART between two months and end of TB treatment	0.65	0.156	0.36	1.18
No valid South African ID number	Valid South African ID number	0.93	0.790	0.53	1.63

Death source: Deaths registered on TB register and HIV programme database

APPENDIX B:
HIV PROGRAMME CLINICAL VISIT FORMS