

https://dspace.org/

Markers of adherence among HIV-positive adults on antiretroviral therapy at Themba Lethu Clinic

Nnambalirwa, Maria Tegulifa

2015-05-05
http://hdl.handle.net/10539/17621

Downloaded from a DSpace repository. https://dspace.org/

MARKERS OF ADHERENCE AMONG HIV-POSITIVE ADULTS ON ANTIRETROVIRAL THERAPY AT THEMBA LETHU CLINIC



RESEARCH REPORT SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE IN EPIDEMIOLOGY AND BIOSTATISTICS

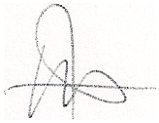
BY

MARIA NNAMBALIRWA

September, 2014

DECLARATION:

I, Maria Tegulifa Nnambalirwa (student number: 608668), declare that this research report is my own work. It is being submitted for the degree of Master of Science in the field of Epidemiology and Biostatistics at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at this or any other university.



M.T Nnambalirwa (Dr)

……19th…… day of …….September.…….2014

DEDICATION:

This work is dedicated to my parents for their unwavering support, encouragement and prayers during the research process. I would also like to dedicate it to my siblings, who have kept me in good spirits during this arduous journey.

ABSTRACT

Introduction: The prevalence of the Human Immunodeficiency Virus (HIV) in South Africa was 17.8% among 15 to 49 year olds in 2010. Antiretroviral therapy (ART) has thus played a crucial role in mitigating the impact of the HIV epidemic. Themba Lethu Clinic is one of the largest single clinics providing ART in South Africa. One of the challenges of ART provision is ensuring adherence to taking the medication. To date there has been no clear consensus on the ideal way to measure adherence in resource limited settings (RLS). Viral load is perhaps the best and most reliable indicator of poor adherence but is expensive and not easily accessible or available in many RLS. Surrogate markers such as mean cell volume (MCV), CD4 cell count, self-reported adherence and missed visits have been shown to be useful to measure adherence but their reliability remains unclear. The aim of the study was to identify other markers that can be used to measure adherence using viral load as the gold standard.

Materials and methods: The study was a retrospective analysis of HIV-positive ART-naïve adults (≥ 18 years) initiating standard first-line ART at the Themba Lethu Clinic in Johannesburg, South Africa between April 2004 and January 2012. The association between the last self-reported adherence, change in MCV calculated from baseline to 6 months, change in CD4 count calculated from baseline to 6 months (≥ or < the expected increase of 50 cells/mm^3 at 6 months) and missed visits (defined as a scheduled appointment that had been missed by ≥ 7 days but not by more than 3 months) and poor adherence (defined as a viral load ≥ 400copies/ml after 6 months on ART) was tested using Poisson regression models with robust error variance to estimate incidence rate ratio (IRR) and 95% confidence interval (CI). The IRR was used to approximate the relative risk (RR) of poor adherence. Interacting variables were stratified by each other, to create a new variable. The diagnostic accuracy of each identified marker of adherence was also tested using sensitivity, specificity, positive predictive values and negative predictive values.

Results: 7160 patients were eligible for the study and of these 63.2% were female. The median age was 36.7 years. The median CD4 count was 101 cells/mm^3 at baseline and 18.9% of the patients had poor adherence at 6 months. Variables associated with poor adherence at 6 months were change in CD4 count stratified by change in MCV at 6 months (change in CD4 count ≥ expected and change in MCV ≥ 14.5fL; adjusted relative risk (aRR) 1, change in CD4 count ≥ expected and change in MCV < 14.5fL; aRR 3.11 95% CI 2.41 – 4.02, change in CD4 < expected and change in MCV ≥ 14.5fL; aRR 1.23 95% CI 0.76 – 2.00 and change

in CD4 count < expected and change in MCV < 14.5fL; aRR 6.98 95% CI 5.35 – 9.09), CD4 count at baseline (> 200 cells/mm^3; aRR 1, 101 – 200 cells/mm^3; aRR 1.05 95% CI 0.80 – 1.38, 51 – 100 cells/mm^3; aRR 1.08 95% CI 0.80 – 1.47 and ≤ 50cells/mm^3; aRR 1.34 95% CI 1.02 – 1.76) , WHO stage at baseline (stage I; aRR 1, stage II; aRR 1.16 95% CI 0.90 – 1.48, stage III; aRR 1.27 95% CI 1.04 – 1.55 and stage IV; aRR 1.44 95% CI 1.12 – 1.84) and MCV at baseline (< 80fL; aRR 1, 80 – 100fL; aRR 1.33 95% CI 1.01 – 1.75 and > 100fL aRR 0.98 95% CI 0.62 – 1.55). Sensitivity and specificity of the change in CD4 stratified by change in MCV at 6 months to predict poor adherence were 86.5% and 37.3% respectively for all eligible patients. For patients on AZT-based regimens the variables associated with poor adherence at 6 months were change in CD4 count at 6 months (≥ expected; aRR 1 and < expected; aRR 7.66 95% CI 0.98 – 59.91) and pregnancy during the first 6 months on ART (Never pregnant; aRR 1 and pregnant during follow up; aRR 9.11 95% CI 2.17 – 38.25). Sensitivity and specificity of the change in CD4 count at 6 months to predict poor adherence were 64.7% and 75.2% respectively for all eligible patients on AZT-based regimens. Sensitivity and specificity of pregnancy during the first 6 months on ART to predict poor adherence were 20% and 97.6% respectively for all eligible patients on AZT-based regimens.

Discussion: Change in CD4 count stratified by change in MCV at 6 months was an expected marker of adherence as CD4 count is expected to rise in adherent patients on ART and since most patients (62.9%) were on d4T or AZT-based regimens. Pregnancy during the first 6 months on ART appeared as a marker of adherence for patients on AZT-based regimens before multiple imputation possibly due to missing data hence results for this variable should be interpreted with caution. Contrary to previous studies, self-reported adherence was not associated with poor adherence at 6 months before multiple imputation. This could have been due to the fact that that > 50% of patients had missing data for this variable. The variable is also vulnerable to recall and reporting bias so even after multiple imputation, the area under the receiver operating characteristic (ROC) curve remained < 0.55. The number of missed medical visits and regimen change were also markers of adherence in a few of the models after multiple imputation and require further investigation. In conclusion, the markers of adherence to ART are change in CD4 count stratified by change in MCV at 6 months and pregnancy during the first 6 months on ART for patients on AZT-based regimens. These could help health workers identify poor adherence in the absence of viral load testing and target patients for adherence interventions to prevent virological failure.

ACKNOWLEDGEMENTS:

I thank all my lecturers, who grounded me in Epidemiology and Biostatistics. A special word of gratitude goes to my supervisors, Professor Peter Nyasulu, Mrs Lynne McNamara and Dr Denise Evans for tirelessly guiding me towards the finishing line. I am also indebted to the Clinical HIV Research Unit, Right to Care and Helen Joseph Hospital for providing the data, without which this study would not have been possible.

TABLE OF CONTENTS:

DECLARATION.....i
DEDICATION.....ii
ABSTRACT.....iii
ACKNOWLEDGEMENTS.....v
TABLE OF CONTENTS.....vi
LIST OF ABBREVIATIONS.....ix
LIST OF FIGURES.....xi
LIST OF TABLES.....xii
LIST OF APPENDICES.....xiv
CHAPTER ONE: INTRODUCTION.....1
1.1 Background.....1
1.1.1 The global and South African Situations.....1
1.1.2 Themba Lethu Clinic.....1
1.1.3 The South African National ART Guidelines.....2
1.2 Statement of the problem.....2
1.3 Justification for the study.....3
1.3.1 Definition of a marker.....3
1.3.2 Surrogate markers.....4
1.3.3 Non-surrogate markers for measuring adherence.....4
1.4 Literature review.....5
1.4.1 Adherence.....5
1.4.2 Surrogate methods for measuring adherence.....6
1.4.3 Factors that influence adherence.....7
1.5 Study aim and objectives.....8
1.5.1 Study question.....8
1.5.2 Aim.....8
1.5.3 Objectives.....9
CHAPTER TWO: MATERIALS AND METHODS.....10
2.1 Data source.....10
2.2 Study design.....10
2.3 Patient selection and eligibility criteria.....11
2.3.1 Inclusion criteria.....11
2.3.2 Exclusion criteria.....11

2.3.3 Rationale for eligibility criteria....11
2.4 Measurement....12
2.4.1 Working definitions....12
2.4.2 Outcome and explanatory variables and their definition/categories....12
2.5 Data management....14
2.5.1 Data cleaning and coding....14
2.5.2 Statistical methods....15
Descriptive statistics....15
Inferential statistics....15
Univariate and multivariate regression analyses....16
Sensitivity analyses....17
Subgroup analysis....17
Other analyses....18
Multiple imputation....18
2.6 Ethical considerations....18
CHAPTER THREE: RESULTS....19
3.1 Selection process for eligible patients....19
3.2 Statistical analyses....21
3.2.1 Descriptive analysis....21
3.2.2 Poisson regression with robust error variance....30
3.3 Measures of validity....38
3.3.1 Sensitivity, specificity, PPV and NPV....38
3.3.2 Receiver operating characteristic (ROC) curves....40
3.4 Sensitivity analyses....44
3.4.1 Descriptive analysis....44
3.4.2 Poisson regression with robust error variance for the sensitivity analysis....61
3.5 Poisson regression with robust error variance, using a viral load done 90 – 270 days after ART initiation....69
CHAPTER FOUR: DISCUSSION....77
4.1 Markers of adherence....77
4.2 Other variables....81
4.3 Strengths and limitations....84
4.4 Generalizability....86
4.5 Conclusion and recommendations....86

4.6 Dissemination.....87
REFERENCES.....88
APPENDICES.....95

LIST OF ABBREVIATIONS:

aRR : adjusted relative risk

AIC : Akaike's Bayesian information criteria

AIDS : Acquired Immune Deficiency Syndrome

ART : Antiretroviral Therapy

ARV : Antiretroviral

ATV : Atazanavir

AZT : Zidovudine

BIC : Schwarz's Bayesian information criteria

BMI : Body Mass Index

CCMT : Comprehensive Care Management and Treatment

CD4 cell : Cluster of differentiation 4 helper T lymphocyte cell

CI : Confidence interval

Csv : Comma separated values

DART : Development of Antiretroviral Therapy in Africa

ddI : Didanosine

d4T : Stavudine

EFV : Efavirenz

FBC : Full blood count

FMI test : Equal fractions missing information test

FTC : Emtricitabine

HIV : Human Immunodeficiency Virus

ICD10 : International Statistical Classification of Diseases (10^{th} revision)

IQR : Interquartile range

IRR : Incidence rate ratio

3TC : Lamivudine

LPV : Lopinavir

LTFU : Loss to follow up

MCV : Mean cell volume of red blood cells

MEM : Microelectronic monitoring

NPV : Negative predictive value

NNRTI : Non-nucleoside reverse transcriptase inhibitor

NRTI : Nucleoside reverse transcriptase inhibitor

NVP : Nevirapine

PI	:	Protease inhibitor
PPV	:	Positive predictive value
r	:	Ritonavir
RLS	:	Resource limited setting
ROC	:	Receiver operating characteristic
RR	:	Relative risk
TB	:	Tuberculosis
TDF	:	Tenofovir
VAS	:	Visual Analogue Scale
WHO	:	World Health Organization

LIST OF FIGURES:

Figure 1: Patient flow chart for the selection of patients eligible for the study.....19

Figure 2a: Receiver operating characteristic (ROC) curves for different markers of adherence for eligible patients with a documented viral load result.....40

Figure 2b: Receiver operating characteristic (ROC) curves for different markers of adherence for eligible patients with a documented viral load result after multiple imputation.....41

Figure 2c: Receiver operating characteristic (ROC) curves for different markers of adherence for all eligible patients on AZT-based regimens with a documented viral load result.....42

Figure 2d: Receiver operating characteristic (ROC) curves for different markers of adherence for all eligible patients on AZT-based regimens with a documented viral load result after multiple imputation.....43

LIST OF TABLES:

Table 1a: Baseline demographic characteristics and their associations with poor adherence for all eligible patients and eligible patients on AZT-based regimens with a documented viral load result.....21

Table 1b: Baseline clinical characteristics and their associations with poor adherence for all eligible patients and eligible patients on AZT-based regimens with a documented viral load result.....23

Table 1c: 6 month clinical characteristics and their associations with poor adherence for all eligible patients and eligible patients on AZT-based regimens with a documented viral load result.....25

Table 2: Table showing the points at which the ART adherence was reported and the points at which the MCV, CD4 count and viral load at 6 months were done for eligible patients with a documented viral load result.....29

Table 3a: Univariate and multivariate Poisson regression models with robust error variance for all eligible patients and eligible patients on AZT-based regimens with a documented viral load result.....30

Table 3b: Univariate and multivariate Poisson regression models with robust error variance for all eligible patients and eligible patients on AZT-based regimens with a documented viral load result after multiple imputation.....34

Table 4a: Sensitivity, specificity, PPV and NPV of markers to identify poor adherence, using viral load as the gold standard, for all patients with a documented viral load result.....38

Table 4b: Sensitivity, specificity, PPV and NPV of each category of change in CD4 count stratified by change in MCV at 6 months for all patients with a documented viral load result.....39

Table 5a: Baseline demographic characteristics and their associations with poor adherence for all patients including those excluded due to missing viral load results assuming that those with missing results had a viral load ≥ 400 copies/ml.....44

Table 5b: Baseline clinical characteristics and their associations with poor adherence for all patients including those excluded due to missing viral load results assuming that those with missing results had a viral load ≥ 400 copies/ml.....46

Table 5c: 6 month clinical characteristics and their associations with poor adherence for all patients including those excluded due to missing viral load results assuming that those with missing results had a viral load ≥ 400 copies/ml.....48

Table 5d: Baseline demographic characteristics and their associations with poor adherence for all patients including those excluded due to missing viral load results assuming that those with missing results had a viral load < 400 copies/ml.....52
Table 5e: Baseline clinical characteristics and their associations with poor adherence for all patients including those excluded due to missing viral load results assuming that those with missing results had a viral load < 400 copies/ml.....54
Table 5f: 6 month clinical characteristics and their associations with poor adherence for all patients including those excluded due to missing viral load results assuming that those with missing results had a viral load < 400 copies/ml.....56
Table 6: Table showing the points at which the ART adherence was reported and the points at which the MCV, CD4 count and viral load at 6 months were done for the sensitivity analyses.....60
Table 7a: Univariate and multivariate Poisson regression models with robust error variance assuming that those with missing viral load results had a viral load ≥ 400 copies/ml...61
Table 7b: Univariate and multivariate Poisson regression models with robust error variance assuming that those with missing viral load results had a viral load < 400 copies/ml...65
Table 8a: Univariate and multivariate Poisson regression models with robust error variance for all eligible patients and eligible patients on AZT-based regimens with a documented viral load done 90 to 270 after ART initiation…….....69
Table 8b: Univariate and multivariate Poisson regression models with robust error variance for all eligible patients and eligible patients on AZT-based regimens with a documented viral load (done 90 to 270 days after ART initiation) after multiple imputation.....73

LIST OF APPENDICES:

APPENDIX 1: Summary of different types of adherence measurement methods.....95

APPENDIX 2: Adherence by self-report at Themba Lethu Clinic.....101

APPENDIX 3A: Tables showing formulae for calculating sensitivity, specificity, PPV and NPV of change in MCV at 6 months and change in CD4 count at 6 months, for all patients with a documented viral load result.....102

APPENDIX 3B: Table showing formulae for calculating sensitivity, specificity, PPV and NPV of each category of change in CD4 count stratified by change in MCV at 6 months for all patients with a documented viral load result.....103

APPENDIX 3C: Table showing formulae for calculating overall sensitivity, specificity, PPV and NPV of change in CD4 count stratified by change in MCV at 6 months for all patients with a documented viral load result.....104

APPENDIX 3D: Table showing formulae for calculating sensitivity, specificity, PPV and NPV of pregnancy during the first 6 months on ART for patients on AZT with a documented viral load result.....105

APPENDIX 4A: Baseline demographic characteristics after multiple imputation for all eligible patients and eligible patients on AZT-based regimens with a documented viral load result.....106

APPENDIX 4B: Baseline clinical characteristics after multiple imputation for all eligible patients and eligible patients on AZT-based regimens with a documented viral load result.....107

APPENDIX 4C: 6 month clinical characteristics after multiple imputation for all eligible patients and eligible patients on AZT-based regimens with a documented viral load result.....108

APPENDIX 5A: Baseline demographic characteristics and their associations with poor adherence for all eligible patients and eligible patients on AZT-based regimens with a documented viral load done 90 – 270 days after ART initiation.....111

APPENDIX 5B: Baseline clinical characteristics and their associations with poor adherence for all eligible patients and eligible patients on AZT-based regimens with a documented viral load done 90 – 270 days after ART initiation.....112

APPENDIX 5C: 6 month clinical characteristics and their associations with poor adherence for all eligible patients and eligible patients on AZT-based regimens with a documented viral load done 90 – 270 days after ART initiation.....114

APPENDIX 6A: Baseline demographic characteristics after multiple imputation for all eligible patients and eligible patients on AZT-based regimens with a documented viral load done 90 – 270 days after ART initiation……118
APPENDIX 6B: Baseline clinical characteristics after multiple imputation for all eligible patients and eligible patients on AZT-based regimens with a documented viral load done 90 – 270 days after ART initiation…….....119
APPENDIX 6C: 6 month clinical characteristics after multiple imputation for all eligible patients and eligible patients on AZT-based regimens with a documented viral load done 90 – 270 days after ART initiation…….120
APPENDIX 7A: Receiver operating characteristic (ROC) curves for different markers of adherence for eligible patients with a documented viral load done 90 to 270 days after ART initiation…….....123
APPENDIX 7B: Receiver operating characteristic (ROC) curves for different markers of adherence for all eligible patients on AZT-based regimens with a documented viral load done 90 to 270 days after ART initiation……124
APPENDIX 8A: Univariate and multivariate Poisson regression models with robust error variance for all eligible patients and eligible patients on AZT-based regimens, with a documented viral load result, excluding patients with a history of prior ART use such as NVP prophylaxis and excluding patients switched to second-line or a non-standard ART regimen before 6 months on treatment.....125
APPENDIX 8B: Univariate and multivariate Poisson regression models with robust error variance after multiple imputation, for all eligible patients and eligible patients on AZT-based regimens, with a documented viral load result, excluding patients with a history of prior ART use such as NVP prophylaxis and excluding patients switched to second-line or a non-standard ART regimen before 6 months on treatment.....129
APPENDIX 9A: ROC curves for different markers of adherence for eligible patients with a documented viral load, excluding patients with a history of prior ART use such as NVP prophylaxis and excluding patients switched to second-line or a non-standard ART regimen before 6 months on treatment……133
APPENDIX 9B: ROC curves for different markers of adherence for eligible patients on AZT-based regimens with a documented viral load, excluding patients with a history of prior ART use such as NVP prophylaxis and excluding patients switched to second-line or a non-standard ART regimen before 6 months on treatment……...134
APPENDIX 10: Permission letter to use TherapyEdge-HIV^{TM}.....135

APPENDIX 11: Ethics Clearance Certificate.....136
APPENDIX 12: Right to Care Confidentiality Agreement.....137

CHAPTER ONE

INTRODUCTION:

In this chapter, the extent of the Human Immunodeficiency Virus (HIV) and Acquired Immune Deficiency Syndrome (AIDS) globally, regionally and in South Africa is presented. I also introduce key elements from the guidelines for antiretroviral therapy (ART) provision in South Africa. Good adherence to ART is defined and the consequences of poor adherence are outlined. The definition of a marker of adherence is also given. Adherence assessment, as carried out at Themba Lethu Clinic, is outlined and the challenges of measuring adherence using different methods/markers are explained. The available literature on adherence and its surrogate and non-surrogate markers is also reviewed. From here the case for the need of ideal markers of adherence at Themba Lethu Clinic is made. The chapter ends with the aims and objectives of the study.

1.1 Background:

1.1.1 The global and South African situations:

HIV/AIDS is a global pandemic, the impact of which has been mitigated by the provision of ART. The world had an estimated 34 million people living with HIV in 2010[1]. The majority of these (68%) were from sub-Saharan Africa[2]. Provision of ART is thus a major part of health service delivery on the subcontinent. In South Africa, about 5.6 million people were living with HIV in 2010, with a prevalence of 17.8% among 15 to 49 year olds[3, 4]. Approximately 1.5 million of those were in need of ART[5]. This translates to an unmet need for ART of about 45%[6].

1.1.2 Themba Lethu Clinic:

Themba Lethu is one clinic providing ART in South Africa. It is an urban clinic, located at the Helen Joseph Hospital in Johannesburg. Johannesburg is located in Gauteng province, which has the second largest number of patients on ART in South Africa[7]. The clinic is an accredited public sector Comprehensive Care, Management and Treatment (CCMT) site for HIV positive patients. It was established in 2004 and is one of the largest single clinics providing HIV treatment in South Africa[8]. By October 2011, 21 101 patients had ever received ART at Themba Lethu Clinic. ART is initiated according to the South African National Guidelines[7].

1.1.3 The South African National ART Guidelines:

In 2004, adults were initiated on ART at a CD4 count of < 200 cells/mm^3 regardless of World Health Organization (WHO) stage or WHO stage IV irrespective of CD4 cell (cluster of differentiation number 4 helper T lymphocytes) count. The first-line regimen was stavudine and lamivudine combined with either efavirenz or nevirapine (d4T/3TC/EFV or d4T/3TC/NVP). The second-line regimen was zidovudine and didanosine with ritonavir-boosted lopinavir (AZT/ddI/LPV/r) [(9)].

In 2010, the guidelines changed. ART was to be initiated at a CD4 count of ≤ 350 in pregnant women and tuberculosis (TB) co-infected patients. In patients with drug-resistant TB, ART was to be initiated as soon as possible regardless of CD4 count or WHO stage. In addition, the new first-line regimens became tenofovir and lamivudine or emtricitabine combined with either efavirenz or nevirapine (TDF/3TC/EFV, TDF/FTC/EFV, TDF/3TC/NVP or TDF/FTC/NVP). Where TDF was contraindicated, it was to be replaced with zidovudine (AZT) in the regimens (AZT/3TC/NVP or AZT/3TC/EFV). The new second-line regimen became AZT/3TC/LPV/r. For those previously on an AZT or a d4T-based first-line regimen, the new second-line regimen became TDF/3TC/LPV/r or TDF/FTC/LPV/r[(10)]. In September 2011, the National Guidelines were updated again to make HIV-positive adults with a CD4 ≤ 350 eligible for ART[(7)].

1.2 Statement of the Problem:

One of the challenges of ART provision is ensuring adherence to the drugs. Unfortunately, there has been no clear consensus on the ideal way to monitor adherence in resource-limited settings. This means that poor adherence may go undetected.

Patients at Themba Lethu Clinic have adherence measured by calculating the number of missed visits from visit schedules and by self-reporting[(7)]. There is no standardized method of measuring self-reported adherence and various studies have assessed it in different ways [(11-13)]. At Themba Lethu Clinic, self-reported treatment adherence is assessed at routine clinic visits and is based on the patient's assessment of pills taken in the last week out of those expected to have been taken. Self-reported adherence in the electronic patient management system (TherapyEdge-HIV™) at the site is categorised as > 90% (all pills), > 60-90% (most pills), > 30-60% (about half the pills), 10-30% (a few of the pills) and < 10% (none of the pills). Another proxy measure, missed visits (missed medical visit or missed ARV drug

pickup by at least 7 days), is also used to measure adherence[14]. Patients come for ARV pickups monthly for the first 6–12 months on treatment and, once stable, every 2 months thereafter. Patients also have a full blood count (FBC) from which mean cell volume of red blood cells (MCV) is obtained, CD4 count and viral load in accordance with the National Guidelines[7, 9, 10, 15]. Before 2010, the MCV was measured at baseline then at 1, 2, 3 and 6 months on ART then 6 monthly for patients on AZT-based regimens. The 2010 National ART Guidelines then stipulated that MCV was to be measured only at baseline, 1, 2, 3 and 6 months on ART for patients on AZT-based regimens. The latest (2013) National ART Guidelines, however, stipulate that an FBC, from which MCV is obtained be done only at 3 and 6 months for patients on AZT-based regimens [15]. For patients taking an AZT-based regimen, after 1 month of treatment, the MCV rises above the normal range (80 – 100fL) [16, 17]. Before 2010, CD4 count and viral load were both measured at baseline then 6 monthly thereafter. Between 2010 and 2013, the CD4 count was measured at baseline then at 6 months, 1 year and yearly thereafter once on ART. Based on the latest (2013) National Guidelines, CD4 count is now measured at baseline and one year only, meaning that subsequent monitoring of a patient's response to ART is done by viral load[15]. The viral load is measured at 6 months, 1 year and yearly thereafter. The MCV measured as part of an FBC costs R124.60 per test while a CD4 count test costs R88.20 per test.

Of these methods, viral load is the most reliable marker of adherence. It is the gold standard measure of adherence [14, 18, 19]. Studies have shown that measuring viral load at 6 months is a reliable indicator of adherence or effectiveness of treatment [20, 21]. A viral load of < 400 copies/ml indicates good adherence while a viral load of ≥ 400 copies/ml indicates poor adherence. Viral load is an expensive test (R606.40 per test). It, however, serves as the benchmark against which the efficacy of other routinely collected and more affordable markers in this study will be assessed as measures of adherence. Routinely collected markers found to be dependable measures of adherence will then be used to assess adherence at each visit.

1.3 Justification for the study:

1.3.1 Definition of a marker:

For the purposes of this study, a marker was defined as a variable that gives an indication of whether adherence is good or poor. Markers may be surrogate or non-surrogate.

1.3.2 Surrogate markers:

The surrogate clinical markers of adherence are viral load, CD4 cell count, MCV for patients on AZT, serum lactate for patients on d4T and serum lipid levels for patients on protease inhibitors (PIs)[12, 17, 22, 23]. Serum lactate and lipid levels are not measured routinely at Themba Lethu Clinic[7]. Since all the markers have disadvantages, using a combination of markers could give a more reliable measure of adherence. This study thus aims to identify a group of already routinely performed, affordable tests that can be used as markers to assess adherence levels, using viral load as the gold standard. Having these markers could help clinic staff improve adherence thereby improving the health of the individual patient by lowering viral load. A patient with an undetectable viral load is less likely to transmit the virus and, in case of transmission, is less likely to transmit drug-resistant viral strains.

1.3.3 Non-surrogate markers for measuring adherence:

Measuring drug levels in serum or hair could be considered the ideal way to test for adherence[24, 25]. This is because these are non-surrogate clinical measures. In addition, measuring drug levels in serum or hair may help in the detection or prevention of drug toxicity, which can lead to poor adherence[11]. Unfortunately, serum drug levels, unlike drug levels in hair, have the disadvantage of detecting only recent adherence[11, 24]. They are thus vulnerable to white coat adherence[11]. This is full adherence for a short interval prior to a clinic visit, making serum drug monitoring unreliable[26]. Another disadvantage of therapeutic drug monitoring using serum drug levels is the fact that it cannot routinely measure nucleoside reverse transcriptase inhibitor (NRTI) levels because active moieties are intracellular. In other words *in vivo* correlates are with intracellular NRTI concentrations rather than plasma NRTI concentrations[11, 16]. An additional disadvantage is the substantial intra-patient variation in drug concentrations [11, 16]. Furthermore drug levels in serum and hair may be low for reasons other than poor adherence e.g diet, drug interactions[11]. Measuring serum or hair drug levels is also expensive (approximately R385 for a 4 ARV drug panel) and is not routinely performed at any HIV treatment site. Directly observed therapy (DOT) is another non-surrogate measure for enhancing and monitoring adherence[27, 28]. This, however, requires the guaranteed presence of a constant treatment supporter who is willing and able to directly observe medication being taken correctly and may, thus, may not be practical.

Considering the large number of patients on ART, any marker of adherence should not only already be routinely collected and affordable but should also be obtainable without causing

lengthy waiting times. The latter cannot be said about pill count, which is discussed in more detail below.

1.4 Literature review:

1.4.1 Adherence:

Good adherence means taking at least 95% of the correct medication, in the correct quantities, at the correct time and following any associated recommendations such as those pertaining to food[13, 29-33]. EFV, for example, should be taken on an empty stomach or with a low fat meal[16]. The widely accepted definition of using 95% was derived from a prospective observational study by Paterson and colleagues (2000), of patients on unboosted protease inhibitors (PIs) in the USA[34]. Other studies have, however, shown that less than 95% or at least 80% adherence may be sufficient to maintain virological suppression[21, 35, 36]. For this reason, some studies on adherence have used definitions of taking at least 80% of prescribed medication as good adherence[37, 38], whilst others have defined good adherence as taking at least 90% of prescribed medication[39, 40].

Despite the varying definitions of adherence, there is consensus that the highest possible level of adherence needs to be achieved to delay the development of drug resistant viral strains, and thus preventing clinical deterioration. Furthermore, patients who are adherent are less likely to transmit the virus to others. The public health implications of transmission of drug resistant strains include the need for newer, more expensive drugs. Other studies have thus gone on to distinguish between complete (100%) and incomplete adherence (< 100%) [41, 42].

A cross-sectional survey of adult adherence in countries on different continents found that "overall, 57% of patients reported a 30-day recall of 100% adherence (Latin America: 89%; Africa: 73% vs North America: 45% and Asia Pacific: 47%) [38]." In the same study, 77% of Africans reported adherence of over 80%. An earlier meta-analysis had similar findings with 73% of patients in Sub-Saharan Africa found to be adherent as compared to 55% in North America[43]. Other cross-sectional studies done on levels of adherence in adults in developed and developing countries have found levels of over 70% per patient[13, 31].

1.4.2 Surrogate methods for measuring adherence:

A variety of surrogate methods are used to assess adherence around the world[43]. Adherence measurements using micro-electronic monitoring (MEM) are said to correlate more closely with viral load[11]. MEM is, however, considered too costly for developing countries (R1148 to R1283 per cap). In addition, it is vulnerable to technological malfunction[11]. Furthermore, it tells you the time the pill bottle was opened but not whether medication was actually taken or whether food restrictions were adhered to. MEM may also underestimate adherence rates among patients who remove an additional dose from their pill bottle with the intention of taking it later, making the use of pill boxes problematic[11, 33].

Pill count is prone to error as the patient may remove medication and throw it away so this measure may be more accurate if the pill count is unannounced[36, 44]. It is also time-consuming and does not tell you whether the medication was taken at the correct time with the appropriate dietary requirements. Self-reporting, on the other hand, is subject to errors from patient recall or dishonesty[11, 12]. This introduces bias and creates misclassification of adherence levels. The visual analogue scale (VAS) is also used for self-reported adherence measurement. It relies on patients estimating their adherence by placing a mark on the scale. Despite being less-time consuming, it is subject to the same errors of patient recall and dishonesty found with ordinary self-reporting. In confirmation of the limitations of both methods, one prospective cohort study by Kunutsor and co-workers in Uganda (2010), found no difference in adherence levels by self-reporting and by pill count[45]. A meta-analysis by Nieuwkerk and colleagues (2005), however, determined that "self-reported adherence measures can distinguish between clinically meaningful patterns of medication-taking behaviour[19]." In other words, self-reported adherence measures were found to distinguish between poor and good adherence as defined by detectable and undetectable viral loads respectively.

Appointment keeping may also be used to estimate adherence[14]. On-time drug pick up does not, however, guarantee adherence[46]. In addition, an adherent patient who misses an appointment due to having enough medication left over cannot be classified as non-adherent.

Surrogate clinical methods, on the other hand, may be specific for the type of drug the patient is on. In addition, these methods detect long-term adherence but are poor at detecting recent adherence levels. An example is MCV for patients on AZT[12, 17, 47, 48]. This is because, in

virtually all patients on AZT, after 4 weeks of AZT use, the MCV rises above the normal range (80-100fL)[16, 17]. Mugisha and colleagues (2012) reported that 12 weekly MCV measurements may be useful in monitoring adherence to AZT-containing regimens in low-income countries, where HIV RNA may be unavailable[18]. Some studies have also shown elevation of MCV in patients on d4T [12, 47, 48]. Studies have found that patients with virological failure have significantly smaller mean increases in MCV[12, 49]. An elevated MCV may, however, also be due to other factors such as vitamin B12 and/or folic acid deficiency, hepatic disease and alcoholism. The most common cause of HIV treatment failure is poor adherence. CD4 cell count may thus also be used as a surrogate clinical method to monitor adherence. Caution is needed in interpreting results, however, as it is a poor predictor of treatment failure[50-53]. Viral load remains the more reliable indicator of treatment failure and, hence, a more suitable proxy against which to gauge the reliability of other potential markers of adherence[14, 54]. Furthermore, there is no consensus on the expected rate of CD4 count rise for patients on ART[16, 55-57]. Interestingly, another prospective study by Oyugi and co-workers in Uganda (2004), found no difference in adherence levels measured by clinical (viral load) or non-clinical methods (self-report, visual analogue scale, MEM and pill count) [54]. Surrogate and non-surrogate methods for measuring adherence are summarized in Appendix 1.

1.4.3 Factors that influence adherence:

When looking at ways to measure adherence, it is also important to consider the factors that influence adherence. This is because they may confound any relationship between markers and adherence levels. These factors are similar across various studies in both developed and developing countries. Younger age, for example, has been associated with poorer adherence[33, 45, 58-60]. Gender has also been cited by some researchers. For instance, Godin and colleagues in a longitudinal study in the USA (2005), found male gender to be a predictor of good adherence[61]. Maqutu and co-workers in a retrospective review of patient records in South Africa (2011), also found that males have better adherence than females but discovered that after the age of 50, females have better adherence levels than males[32].

Poor adherence may also result from inadequate adherence education, stigma and non-disclosure[61-63]. In addition, Muyingo and co-workers in the Development of Antiretroviral Therapy in Africa (DART) trial in Uganda and Zimbabwe (2008), found that complete (100%) adherence was independently associated with having only one sexual partner[41].

Adverse effects, complexity and pill burden of ARV regimen and co-morbidities have also been associated with poor adherence[41, 44, 58, 63, 64]. Comorbidities compound the problem of pill burden and may also impair the ability to adhere due to aggravating severe ill health thus affecting memory or strength to take medication. In particular, neurocognitive impairment/mental ill health, including that resulting from substance abuse, are known to compromise adherence[33, 58, 63, 65]. Furthermore, forgetfulness in the absence of severe ill health or impaired neurocognitive function, being away from home and leaving medications at home, have all been cited as contributors to poor adherence[21, 41, 44, 62, 63].

Low CD4 count at ART initiation is another factor that has been linked to poor adherence[41]. Low education levels and lack of support from a partner have also been associated with poor adherence[21, 58]. Cultural or religious beliefs contrary to adherence, lack of faith in treatment, poor patient-provider relationship and inability to get time off work are other factors associated with poor adherence.[10, 16, 61, 63, 65, 66]. Political unrest, natural disaster and drug shortages could also result in poor adherence.

In essence, surrogate measures have been used to assess adherence, particularly in resource-limited settings since they are cheaper. There is consensus on viral load as the gold standard proxy marker of adherence but no consensus on the ranking of other surrogate measures. It is this lack of consensus, together with the high cost of viral load, MEM and non-surrogate measures that has prompted a study on suitable markers of adherence at Themba Lethu Clinic, while taking into consideration the factors that influence adherence.

1.5 Study aims and objectives:

1.5.1 Study question:

What are the markers of adherence to treatment among HIV positive adults on ART at Themba Lethu Clinic?

1.5.2 Aim:

To identify variables that can be used routinely to assess adherence to ART at Themba Lethu Clinic in Johannesburg, South Africa, using viral load as the gold standard.

1.5.3 Objectives:

1. To describe the demographic and clinical characteristics of patients over the age of 18, initiated on ART between 1st April 2004 and 1st January 2012 at Themba Lethu Clinic.
2. To identify any differences in adherence levels of the study subjects by demographic and clinical characteristics, using viral load at 6 months as the gold standard to define adherence or effectiveness of treatment.
3. To identify affordable and routinely collected variables or markers of adherence, using viral load at 6 months as the gold standard to define adherence or effectiveness of treatment.
4. To find the sensitivity, specificity, positive predictive values, negative predictive values and receiver operating characteristics of the identified markers of adherence.
5. To identify affordable and routinely collected variables or markers of adherence, using viral load at 6 months as the gold standard to define adherence or effectiveness of treatment, in a subgroup of patients with a documented viral load at 6 months but excluding patients with a documented history of prior ART use such as NVP prophylaxis and patients changed to a non-standard ART regimen or second-line ART before 6 months on treatment.

CHAPTER TWO

MATERIALS AND METHODS:

This chapter describes materials and methods that were used and includes study design, study population, inclusion and exclusion criteria and source of data. The outcome variable, explanatory variables and all other covariates are outlined. In addition, data processing and analysis procedures are described.

2.1 Data source:

Data were obtained from the electronic management system called TherapyEdge-HIV™. This system contains data on the Themba Lethu clinical cohort, which is made up of HIV positive patients on ART at Themba Lethu Clinic. Data have been entered into this system since the inception of the cohort in 2004[7]. The data are collected routinely for monitoring of, evaluation of and reporting on, the progress of the cohort. The progress is measured using records in the database on patient vitals (height, weight, Body Mass Index (BMI) and blood pressure), appointment keeping, ART regimen, the use of cotrimoxazole prophylaxis, other drugs prescribed, self-reported adherence, HIV-related opportunistic infections such as TB, non-HIV-related medical and surgical conditions and results of laboratory tests such as CD4 count, viral load, full blood count, liver function tests, renal function tests and treatment outcomes (down-referral, transfer out, loss to follow-up or death). The database is the ideal source of information as it captures demographic and treatment-related information on all patients in the cohort. These data are also ideal to answer the research question as Themba Lethu is one of the largest single clinics providing ART in South Africa[8]. Prior to 2007, data were entered from patient files into the database. Currently data are entered into the system as the patient is being seen. Laboratory results are downloaded from the National Health Laboratory Service (NHLS) database directly into TherapyEdge-HIV™.

2.2 Study design:

This was a retrospective cohort design using secondary data taken from the TherapyEdge-HIV™ database. The design was suitable for comparing baseline measures with outcome measures and exploring the effects of several demographic and clinical level exposures on adherence.

2. 3 Patient selection and eligibility criteria:

Being a retrospective study using secondary data, the analysis included all patients in the TherapyEdge-HIV^{TM} database who met the eligibility criteria.

2.3.1 Inclusion criteria:

The following inclusion criteria were adhered to:

- ART-naïve patients aged 18 years and above initiated on a standard first-line ART regimen between 1st April 2004 and 1st January 2012 at Themba Lethu Clinic. The standard first-line regimens considered were either of the following: d4T/3TC/EFV, d4T/3TC/NVP, TDF/3TC/NVP, TDF/FTC/NVP, TDF/3TC/EFV, TDF/FTC/EFV, AZT/3TC/EFV or AZT/3TC/NVP[9, 10]. Patients included had been on ART for at least 6 months.

2.3.2 Exclusion criteria:

-Patients under 18 years of age.

-Patients initiated on ART before 1st April 2004 or after 1st January 2012.

-Patients initiated on ART outside Themba Lethu Clinic.

-Patients with a history of prior ART use as evidenced by a baseline viral load of < 400 copies/ml or in the absence of a baseline viral load, a baseline CD4 count of > 1000 cells/mm^3 with a baseline WHO stage that was I, II or missing.

-Patients initiated on a non-standard first-line ART regimen.

-Loss to follow up (LTFU) within the first 6 months on ART i.e being more than 3 months (> 90 days) late for a scheduled visit[67].

-Transfer out or death within the first 6 months on ART.

2.3.3 Rationale for eligibility criteria:

The eligibility criteria focussed on HIV-positive ART-naïve adults so that we obtain a homogenous study sample. These individuals were initiated on the same standard ART regimens, formulations and doses at initiation. The purpose was to minimize the confounding effect of different physiological and treatment-related factors. On the other hand, patients with a history of ART use may have residual resistance to ART and may exhibit a dissimilar and poor virological response despite good adherence.

2. 4 Measurement:

2.4.1 Working definitions:

(a) Three month period: This was defined as 70 to 110 days after ART initiation to cater for people who may come earlier or later for their three month visit.

(b) Six month period: The value for a variable at 6 months was defined as the first available value 150 to 300 days after initiation. According to the National Guidelines, the first CD4 count and first viral load should be done at 6 months. In some patients, however, it was found to have been done a bit earlier or later, hence the use of 150 to 300 days. Likewise, for some patients MCV and CD4 count at 6 months were found to have been done a bit earlier or later, hence the use of 150 to 300 days.

(c) *The* ***end point:*** was defined as the date of doing the first viral load, 150 to 300 days after ART initiation. For the sensitivity analysis, however, where the end-point viral load was missing, the end-point was set at 183 days after ART initiation i.e 30.5 days × 6 as 6 months is usually a period of about 183 days.

2.4.2 Outcome and explanatory variables and their definitions/categories:

Outcome variable:

Adherence at 6 months measured by viral load at 6 months (gold standard): The first viral load done 150 to 300 days after ART initiation. This could be undetectable (< 400 copies/ml, a proxy for good adherence) or detectable (≥ 400 copies/ml, a proxy for poor adherence).

Explanatory variables:

Demographic characteristics

1. Gender: male or female.
2. Age at ART initiation: ≤ median age or > median age.
3. Education level at ART initiation: ≥ secondary school, primary school or < primary school
4. Professional status at ART initiation: Employed or unemployed.

Clinical characteristics:

1. Self-reported adherence: This is the patient's assessment of pills taken in the last week computed to a percentage. This variable was assessed as the last self-reported adherence levels on, or after, the 3 month visit (70 days post-initiation) and not more than 3 months (110 days) before the date of the viral load but at least 28 days before the viral load was done. The range was selected to capture adherence for as many patients as possible in light of the

noted inconsistency in adherence measurement. The range was also chosen as a reasonable period within which documented adherence may have an impact on the viral load at the end of the study period. Self-reported adherence in the electronic patient management system (TherapyEdge-HIV^{TM}) at the site is categorized as > 90% (all pills), > 60 – 90% (most pills), > 30 –60% (about half the pills), 10 – 30% (a few of the pills) and < 10% (none of the pills) (Appendix 2). Since there were very few people in the latter 3 categories, the variable was categorized as > 90%, > 60 – 90% and ≤ 60%.

2. Number of missed visits during the study period: A missed visit was defined as a scheduled appointment that had been missed by 7 or more days[(14)] but less than 91 days as missing by over 90 days is LTFU[(67)]. The variable was categorized as missed visits, missed ARV visits and missed medical visits. Each of these categories was further broken down into 0, 1 or ≥ 2 missed visits. An ARV visit is scheduled for refilling of ARV medication and involves the taking of weight and/or height as well as adherence assessment and TB screening. A medical visit involves all the components of an ARV visit, more comprehensive clinical assessment and scheduled laboratory tests.

3. BMI (Body Mass Index) at initiation: This is calculated as weight in kg ÷ (height in metres)^2. The variable was thus categorized according to the WHO classification as ≥ 30kg/m^2 (obese), 25 – < 30kg/m^2 (overweight), 18.5 – < 25 kg/m^2 (normal) and < 18.5kg/m^2 (underweight).

4. BMI (Body Mass Index) at 6 months: The variable was also categorized according to the WHO classification as ≥ 30kg/m^2 (obese), 25 – < 30kg/m^2 (overweight), 18.5 – < 25 kg/m^2 (normal) and < 18.5kg/m^2 (underweight).

5. Baseline haemoglobin level. This was thus categorized as ≤ 8g/dl and > 8g/dl.

6. Haemoglobin level at 6 months. This was also categorized as ≤ 8g/dl and > 8g/dl.

7. MCV at baseline: mean cell volume at baseline. The normal range is 80 – 100fL. The variable was thus categorized as < 80fL (low), 80 – 100fL (normal) and > 100fL (high).

8. MCV at 6 months: mean cell volume at 6 months. The variable was also categorized based on the normal range as < 80fL (low), 80 – 100fL (normal) and > 100fL (high).

9. Change in MCV at 6 months: The difference between the MCV at baseline and the MCV at 6 months i.e MCV at 6 months ─ MCV at baseline. Kamya and colleagues (2007) demonstrated that adults on ART with virological suppression, had significantly higher mean increases (14.5 ± 0.4 fL) than those without virological suppression (5.6 ± 2.4fL) [(49)]. The variable change in MCV at 6 months was thus categorized as ≥ 14.5fL and < 14.5fL.

10. CD4 cell count at baseline: The normal range is 500 – 1500 cells/mm^3. The variable was categorized as > 200 cells/mm^3, 101 – 200 cells/mm^3, 51 – 100 cells/mm^3 and ≤ 50 cells/mm^3. These are categories that have been used in previous studies [14, 68].

11. Change in CD4 count at 6 months in cells/mm^3: The difference between the CD4 count at 6 months (the first CD4 count done 150 to 300 days after initiation) and the CD4 count at baseline i.e CD4 count at 6 months – the CD4 count at baseline. CD4 count is expected to rise by 100 cells/mm^3/year or 50 cells/mm^3 in 6 months [14, 16, 55]. The variable was thus categorized as follows:

- ≥ expected (≥ 100 cells/mm^3/year × duration of treatment in years).
- < expected (< 100 cells/mm^3/year × duration of treatment in years).

12. WHO stage at baseline categorized as WHO stage I – IV.

13. Regimen change:

- No: Remained on the same first-line regimen prescribed at initiation during the study period.
- Yes: Regimen changed during the study period.

14. TB: Whether or not the patient was on TB treatment during the first 6 months on ART:

- Never: Never on TB treatment during the study period.
- At baseline: On TB treatment at baseline.
- During follow up: Started TB treatment during follow-up.

15. Pregnancy: Whether or not the patient was pregnant during the first 6 months on ART:

- Never: Never pregnant during the study period.
- At baseline: Pregnant at baseline.
- During follow up: Pregnant during follow-up.

16. Adverse events or side effects: Whether or not the patient experienced at least one of the following during the first 6 months on treatment as a result of ART: peripheral neuropathy, lipodystrophy, lactic acidosis, pancreatitis, anaemia (haemoglobin ≤ 8g/dl), NVP rash, rash due to EFV, renal toxicity, hepatotoxicity, neuropsychiatric disturbances.

2. 5 Data management:

2. 5 .1 Data cleaning and coding:

Data were exported from TherapyEdge-HIV™ as raw csv files then imported into SAS, for cleaning. Analytic datasets were made; some were wide files (single observation per patient) while others were long files (multiple observations per patient). The relevant data were then

exported from SAS into STATA Release 12.1 (StataCorp, College station, Texas) for analysis.

These data were contained in 7 datasets containing information on:

- Socio-demographic variables, baseline CD4 count, baseline viral load and baseline WHO stage, date of ART initiation and first ART regimen.
- Treatment outcomes (down-referral, transfer out, loss to follow-up or death).
- ARV drug use including begin and end dates and doses.
- Appointment keeping (scheduled and visit dates).
- Self-reported adherence
- Patient vitals including height, weight, Body Mass Index (BMI) and blood pressure.
- TB treatment start and stop dates.
- Medical conditions classified according to ICD10 (International Classification of Diseases: 10^{th} revision) and drugs that caused the condition, where applicable.
- Lab tests done, lab test dates and lab results.

The data underwent further cleaning and patients that did not meet the eligibility criteria were dropped from the analysis. Inconsistent and extreme values were also corrected where information was available. Values that were erroneous were set to missing. Thereafter datasets were merged, creating a single wide file for data analysis.

2.5.2 Statistical methods:

Descriptive Statistics:

The demographic and clinical characteristics (at baseline and 6 months) were described using frequencies with percentages for categorical variables and medians with interquartile ranges (IQRs) for continuous variables that were not normally distributed. Means and standard deviations were used for continuous variables that were normally distributed.

Inferential Statistics:

Inferential statistics were used to identify any differences in adherence levels of the study subjects by demographic and clinical characteristics, using viral load at 6 months to define adherence or effectiveness of treatment.

Hypothesis*:*

There is an association between demographic and clinical characteristics and adherence level measured using viral load at 6 months.

(a) For all patients, comparison of those with detectable and undetectable viral load at 6 months was done for all the variables. The chi-square test was used for categorical variables and the t-test was used for continuous variables and p-values were obtained. Non-parametric tests (Wilcoxon rank sum test or median test) were used for continuous variables that were not normally distributed. A p-value of ≤ 0.05 was considered significant.

(b) For patients on AZT, comparison of those with detectable and undetectable viral load at 6 months was done for all the variables. The chi-square test and Fisher's exact test were used for categorical variables and the t-test was used for continuous variables and p-values were obtained. Non-parametric tests (Wilcoxon rank sum test or median test) were used for continuous variables that were not normally distributed. A p-value of ≤ 0.05 was considered significant.

Univariate and multivariate regression analyses:

Log-binomial regression using RRs was replaced with Poisson regression with robust error variance due to the frequent occurrence of failure to converge. Poisson regression with robust error variance was thus used to identify affordable and routinely collected variables or markers of adherence using viral load at 6 months to define adherence or effectiveness of treatment. The IRRs were used to approximate the RRs of poor adherence.

Univariate analysis:

Association of each individual variable to viral load at 6 months was assessed. Incidence rate ratios (IRRs) were obtained and a p-value of ≤ 0.05 was considered significant.

Multivariate analysis:

Multivariate Poisson regression analysis with robust error variance was done first for all eligible patients then for all eligible patients on AZT-based regimens. Variables with a $p < 0.2$ on univariate analysis were selected for the construction of the multivariate regression models. A $p < 0.2$ was used as variables that are associated with the viral load at 6 months may appear to have no significant association at the 5% level on univariate analysis due to

confounding factors. In a model with other variables, however, some variables with a $p > 0.05$ and < 0.2 on univariate analysis may be found to be significantly associated with the viral load at 6 months ($p \leq 0.05$) on multivariate regression analysis. The variables (with $p < 0.2$ on univariate analysis) were added into the multivariate model one after the other. With each variable added, Akaike's and Schwarz's Bayesian information criteria (AIC and BIC) were assessed to find out which variable had the biggest effect on the AIC and BIC statistics. Variables that did not improve the model were left out except for CD4 count at baseline, gender and age at initiation. Interaction terms were then used to identify any significant interactions for inclusion in the model. A variable that had a significant interaction with another was stratified by the other variable, to create a new variable that was put into the model in place of the two interacting variables. Model fit was assessed using the link test (specification link test for single equation models) and $p \leq 0.05$ was considered significant. The diagnostic accuracy of each identified marker of adherence was tested using sensitivity, specificity, positive predictive values (PPVs) and negative predictive values (NPVs). Receiver operating characteristics (ROCs) of each identified marker of adherence were also investigated.

Sensitivity analyses:

Descriptive and inferential analyses including Poisson regression (with robust error variance) were done by assuming that all patients with a missing viral load result at the end-point had a viral load measurement of < 400 copies/ml. This was done first for all patients then for all patients on AZT-based regimens. The analyses were repeated assuming that all patients with a missing viral load had a viral load measurement of ≥ 400 copies/ml.

Subgroup analysis:

A subgroup analysis was then done for patients with a viral load at 6 months, excluding patients with a documented history of prior ART use such as NVP prophylaxis and patients changed to a non-standard ART regimen or second-line ART before 6 months on treatment. Poisson regression with robust error variance was done first for all eligible patients then for all eligible patients on AZT-based regimens.

Other analyses:

Since the viral load may be done as early as 4 months after ART initiation at Themba Lethu Clinic [7], analyses were repeated for all patients with a documented viral load done 90 – 270 after ART initiation.

Multiple imputation:

Multiple imputation was done using multivariate normal regression. It was done for all patients with a documented viral load result, using explanatory variables with missing values. New models were fitted first for all patients then for all patients on AZT-based regimens, using the data with imputed values. The equal fractions missing information test (FMI test) was used to assess model fit and whether additional variables improved the model. Variables that did not improve the model were left out except for CD4 count at baseline, gender and age at initiation.

2.6 Ethical considerations:

TherapyEdge-HIV™ data are covered by Ethical Approval Clearance Certificate Number M110I40. Permission to use this data was obtained from the Clinical HIV Research Unit (CHRU) of the University of the Witwatersrand (Appendix 10). The study was approved by the University of the Witwatersrand Human Research and Ethics Committee; Clearance Certificate number M120858 (Appendix 11). The Medical Superintendent of Helen Joseph Hospital authorised use of the data after receiving the Ethical Clearance Certificate and a signed confidentiality agreement (Appendix 12). All data provided had no personal identifiers.

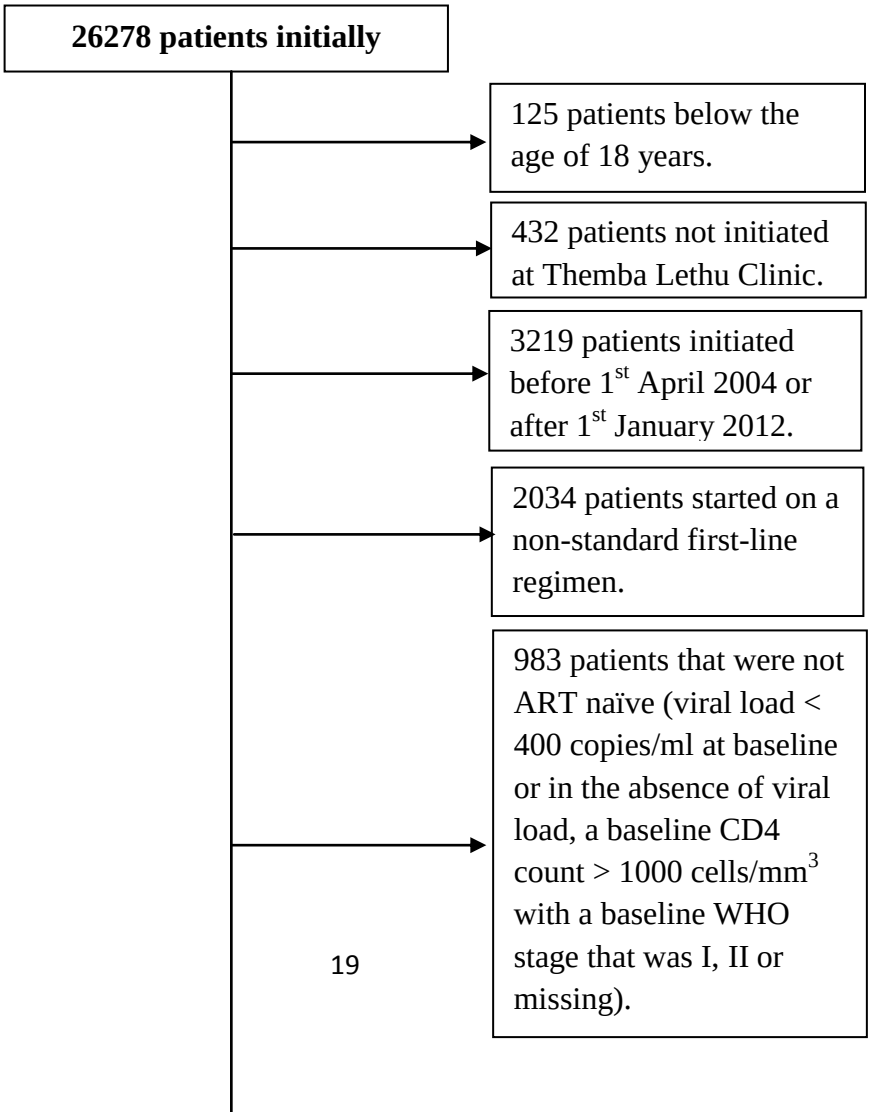
CHAPTER THREE

RESULTS:

In this chapter, the results of data analysis are presented. This begins with a step by step breakdown of the initial number of patients and how they were reduced to produce the final cohort for data analysis. The rest of the data are presented in tables. Firstly, baseline characteristics are presented. This is followed by results depicting the association of the variables to the viral load at 6 months in univariate and multivariate Poisson regression models with robust error variance. Results after multiple imputation are also presented. Sensitivity, specificity, PPV, NPV and ROC curves of identified markers of adherence are also presented. The effect of patients excluded due to missing viral load at the end-point is assessed through sensitivity analyses.

3.1 SELECTION PROCESS FOR ELIGIBLE PATIENTS:

This section shows the process by which patients were dropped from the study based on the exclusion criteria.



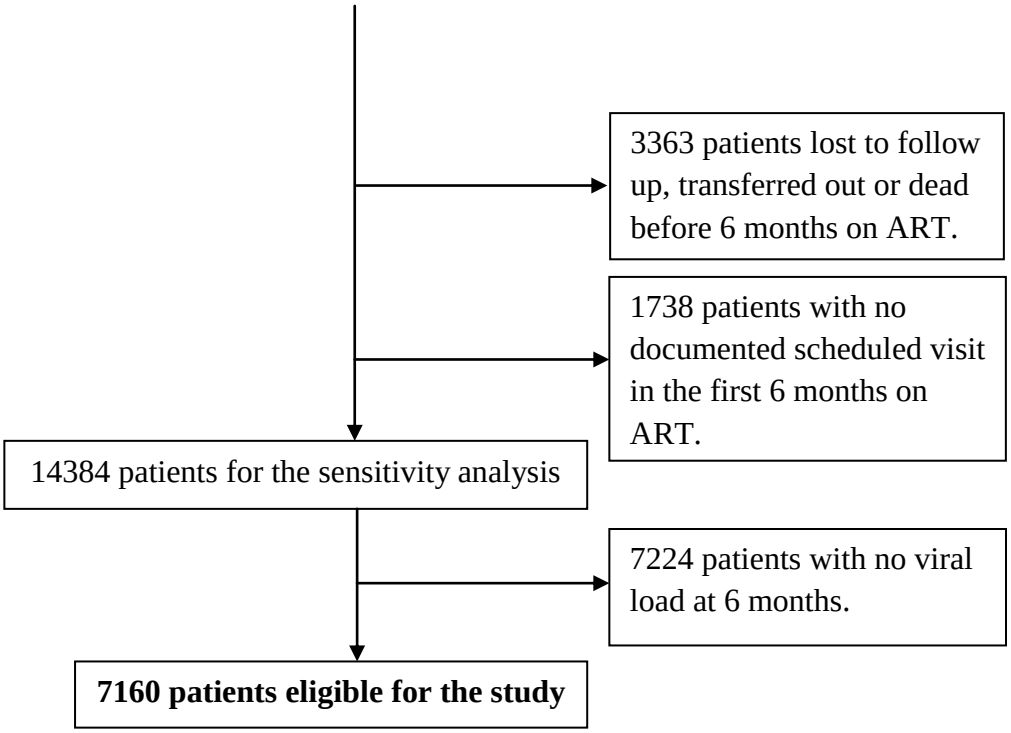


Figure 1: Patient flow chart for selection of patients eligible for the study

Patients who were not eligible for the study are shown in the boxes on the right. These patients were dropped from the dataset of 26278 patients, leaving 14384 patients, for whom sensitivity analyses were done to minimize selection bias. From the 14384 patients, a further 7224 patients with no documented viral load at 6 months, were dropped, leaving 7160 patients for the main analysis of the study. The baseline and 6 month characteristics as well as the poisson regression results were then compared to the results of the main analysis. To minimise selection bias, multiple imputation was also done, for exposure variables with missing values. The baseline and 6 month characteristics as well as the poisson regression results of this data were also compared to the results of the main analysis.

3.2: STATISTICAL ANALYSIS

3.2.1: Descriptive analysis:

In this section the baseline and 6 month characteristics of patients in the study with a documented viral load result at 6 months are presented.

Table 1a: Baseline demographic characteristics and their associations with poor adherence for all eligible patients and eligible patients on AZT-based regimens with a documented viral load result

Characteristics	**All eligible patients**				**All eligible patients on AZT-based regimens***			
	Total N = 7160	**Viral load < 400 copies/ml N = 5806**	**Viral load ≥ 400 copies/ml N = 1354**	**P-value**	**Total N = 177**	**Viral load < 400 copies/ml N = 143**	**Viral load ≥ 400 copies/ml N = 34**	**P-value**
	n (%)	n (%)	n (%)		n (%)	n (%)	n (%)	
Gender								
Females	4523 (63.2%)	3716 (64%)	807 (59.6%)	0.002	97 (54.8%)	82 (57.3%)	15 (44.1%)	0.164
Males	2637 (36.8%)	2090 (36%)	547 (40.4%)		80 (45.2%)	61 (42.7%)	19 (55.9%)	
Age at initiation in years								
Median (IQR) / Mean ± standard deviation (SD)	36.7 (31.4 – 43.4)	36.8 (31.6 – 43.5)	36.3 (30.7 – 42.9)	0.0204[þ]	41.2 ± 9.1	41.8 ± 8.9	38.4 ± 9.6	0.0242[ŧ]
Age at initiation in years								
≤ median (36.7)[α]	3603 (50.3%)	2901 (50.0%)	702 (51.8%)	0.213	60 (33.9%)	43 (30.1%)	17 (50%)	0.027
> median (36.7)	3557 (49.7%)	2905 (50.0%)	652 (48.2%)		117 (66.1%)	100 (69.9%)	17 (50%)	
Education level								
≥ Secondary school	4482 (62.6%)	3573 (61.5%)	909 (67.1%)	0.159	94 (53.1%)	76 (53.1%)	18 (52.9%)	0.840[ð]
Primary school	910 (12.7%)	744 (12.8%)	166 (12.3%)		24 (13.6%)	18 (12.6%)	6 (17.6%)	
< Primary school	220 (3.1%)	184 (3.2%)	36 (2.7%)		7 (4.0%)	6 (4.2%)	1 (2.9%)	
Missing	1548 (21.6%)	1305 (22.5%)	243 (17.9%)		52 (29.3%)	43 (30.1%)	9 (26.6%)	
Professional status								
Employed	3685 (51.5%)	3020 (52.0%)	665 (49.1%)	0.078	89 (50.3%)	71 (49.7%)	18 (52.9%)	0.665
Unemployed	3362 (47.0%)	2700 (46.5%)	662 (48.9%)		85 (48.0%)	70 (49.0%)	15 (44.1%)	

Missing	113 (1.5%)	86 (1.5%)	27 (2.0%)		3 (1.7%)	2 (1.3%)	1 (3.0%)	

* On AZT-based regimens throughout the study period. α The median age at ART initiation for all eligible patients with a documented viral load result. Þ Wilcoxon rank sum test. ŧ t-test. ð Fisher's exact test. AZT: Zidovudine. IQR: Interquartile range. SD: Standard deviation.

The table above shows the baseline demographic characteristics of all eligible patients and eligible patients on AZT-based regimens. The characteristics were described using frequencies with percentages and medians with interquartile ranges or means with standard deviations. Of the 7160 patients eligible for the study 18.9% had poor adherence while out of the 177 patients on AZT-based regimens, 19.2% had poor adherence. Almost two thirds of all eligible patients were female. This is not surprising as females are more vulnerable to HIV infection than men[(69)]. Their median age at initiation was 36.7 years (IQR: 31.4 – 43.4) but the overall age range at initiation was 18.1 – 74.4 years. The table also shows p-values for the association between different variables and poor adherence at 6 months measured by viral load at 6 months. Except where otherwise stated, the chi-square test was used.

Table 1b: Baseline clinical characteristics and their associations with poor adherence for all eligible patients and eligible patients on AZT-based regimens with a documented viral load result

Characteristics	All eligible patients				All eligible patients on AZT-based regimens*			
	Total N = 7160	**Viral load < 400 copies/ml N = 5806**	**Viral load ≥ 400 copies/ml N = 1354**	**P-value**	**Total N = 177**	**Viral load < 400 copies/ml N = 143**	**Viral load ≥ 400 copies/ml N = 34**	**P-value**
	n (%)	n (%)	n(%)		n (%)	n (%)	n (%)	
BMI at initiation in kg/m^2								
Median (IQR)	22 (19.1 – 25)	22 (19.7 – 25)	22 (19 – 25)	0.260 [M]	22.7 (20.0 – 27)	22.3 (20.0 – 29)	23 (20 – 24.0)	0.5455 [þ]
BMI at initiation in kg/m^2								
< 18.5	989 (13.8%)	782 (13.5%)	207 (15.3%)	0.184	19 (10.7%)	17 (11.9%)	2 (5.9%)	0.109[ð]
18.5 – < 25	3330 (46.5%)	2719 (46.8%)	611 (45.1%)		68 (38.4%)	51 (35.7%)	17 (50%)	
25 – < 30	1176 (16.4%)	966 (16.6%)	210 (15.5%)		20 (11.3%)	16 (11.2%)	4 (11.8%)	
≥ 30	558 (7.8%)	462 (8.0%)	96 (7.1%)		24 (13.6%)	23 (16.1%)	1 (2.9%)	
Missing	1107 (15.5%)	877 (15.1%)	230 (17.0%)		46 (26.0%)	36 (25.1%)	10 (29.4%)	
Baseline haemoglobin in g/dl								
Median (IQR) / Mean ± SD	11.7 (10.1 – 13.1)	11.7 (10.2 – 13.1)	11.5 (10.1 – 13)	0.0759[þ]	12.8 ± 1.8	12.6 ± 1.8	13.3 ± 1.4	0.0838[ŧ]
Baseline haemoglobin								
≤ 8 g/dl	336 (4.7%)	273 (4.7%)	63 (4.7%)	0.951	1 (0.6%)	1 (0.7%)	0 (0%)	1.000[ð]
> 8g/dl	6022 (84.1%)	4901 (84.4%)	1121 (82.8%)		136 (76.8%)	112 (78.3%)	24 (70.6%)	
Missing	802 (11.2%)	632 (10.9%)	170 (12.5%)		40 (22.6%)	30 (21.0%)	10 (29.4%)	
MCV at baseline in fL								
Median (IQR)	88.4 (84 – 92.4)	88.4 (83.9 – 92.3)	88.8 (84.3 – 92.6)	0.0267[þ]	91.4 (86.9 – 101.9)	91.9 (87.4 – 102.2)	89.5 (82.9 – 97.4)	0.1856[þ]
MCV at baseline								
< 80fL	648 (9.1%)	557 (9.6%)	91 (6.7%)	0.079	13 (7.3%)	9 (6.3%)	4 (11.8%)	0.374[ð]
80 – 100fL	4533 (63.3%)	3749 (64.6%)	784 (57.9%)		74 (41.8%)	60 (41.9%)	14 (41.2%)	
> 100fL	246 (3.4%)	199 (3.4%)	47 (3.5%)		38 (21.5%)	33 (23.1%)	5 (14.7%)	
Missing	1733 (24.2%)	1301 (22.4%)	432 (31.9%)		52 (29.4%)	41 (28.7%)	11 (32.3%)	
Baseline CD4 count in cells/mm^3								
Median (IQR)	101 (39 – 175)	104 (40 – 176)	90 (34 – 170)	0.0011[þ]	156 (74 – 244)	154 (74 – 243)	169 (61 – 249)	0.9246[þ]

Baseline CD4 count in cells/mm^3								
> 200	1001 (14.0%)	826 (14.2%)	175 (12.9%)	0.007	53 (29.9%)	41 (28.7%)	12 (35.3%)	0.917
101 – 200	2100 (29.3%)	1749 (30.1%)	351 (25.9%)		36 (20.3%)	30 (21.0%)	6 (17.6%)	
51 – 100	1207 (16.9%)	976 (16.8%)	231 (17.1%)		23 (13.0 %)	19 (13.3%)	4 (11.8%)	
≤ 50	1864 (26.0%)	1476 (25.4%)	388 (28.7%)		25 (14.1%)	20 (14.0%)	5 (14.7%)	
Missing	988 (13.8%)	779 (13.5%)	209 (15.4%)		40 (22.7%)	33 (23.0%)	7 (20.6%)	
WHO stage at baseline								
I	2161 (30.2%)	1826 (31.5%)	335 (24.7%)	0.003	75 (42.4%)	61 (42.7%)	14 (41.2%)	0.390[ð]
II	1135 (15.9%)	920 (15.8%)	215 (15.9%)		15 (8.5%)	13 (9.1%)	2 (5.9%)	
III	1744 (24.4%)	1409 (24.3%)	335 (24.7%)		25 (14.1%)	19 (13.3%)	6 (17.6%)	
IV	716 (10%)	572 (9.9%)	144 (10.6%)		11 (6.2%)	11 (7.7%)	0 (0%)	
Missing	1404 (19.5%)	1079 (18.5%)	325 (24.1%)		51 (28.8%)	39 (27.2%)	12 (35.3%)	

[*] On AZT-based regimens throughout the study period. [M] Median test. [Þ] Wilcoxon rank sum test. [ŧ] t-test. [ð] Fisher's exact test. AZT: Zidovudine. BMI: Body Mass Index. IQR: Interquartile range. MCV: Mean cell volume. SD: Standard deviation. WHO: World Health Organization.

The table above shows the baseline clinical characteristics of all eligible patients and eligible patients on AZT-based regimens. The characteristics were described using frequencies with percentages and medians with interquartile ranges or means with standard deviations. The majority of patients had WHO stage I (30.2%) or WHO stage III (24.4%) at initiation. A median CD4 count of 101 cells/mm^3 (IQR: 39 – 175) was obtained for all eligible patients. This is a reflection on the South African National ART Guidelines criteria for ART initiation. At baseline, 63.3% of all patients had a normal MCV. A median BMI of 22 (IQR: 19.1 – 25) kg/m^2 was found for all eligible patients. The median haemoglobin level of all eligible patients was 11.7 g/dl (IQR: 10.1 – 13.1) at the time of initiation. The table also shows p-values for the association between different variables and poor adherence at 6 months measured by viral load at 6 months. Except where otherwise stated, the chi-square test was used.

Table 1c: 6 month clinical characteristics and their associations with poor adherence for all eligible patients and eligible patients on AZT-based regimens with a documented viral load result

Characteristics	All eligible patients				All eligible patients on AZT-based regimens*			
	Total N = 7160	**Viral load < 400 copies/ml N = 5806**	**Viral load ≥ 400 copies/ml N = 1354**	**P-value**	**Total N = 177**	**Viral load < 400 copies/ml N = 143**	**Viral load ≥ 400 copies/ml N = 34**	**P- value**
	n (%)	n (%)	n (%)		n (%)	n (%)	n (%)	
Self-reported adherence[δ]				0.201				
> 90%	2718 (38.0%)	2106 (36.3%)	612 (45.2%)		45 (25.4%)	38 (26.6%)	7 (20.6%)	0.355[δ]
> 60 – 90%	300 (4.2%)	227 (3.9%)	73 (5.4%)		6 (3.4%)	5 (3.5%)	1 (2.9%)	
≤ 60%	71 (1%)	49 (0.8%)	22 (1.6%)		2 (1.1%)	1 (0.7%)	1 (2.9%)	
Missing	4071 (56.8%)	3424 (59.0%)	647 (47.8%)		124 (70.1%)	99 (69.2%)	25 (73.6%)	
Number of missed visits								
0	4617 (64.5%)	3731 (64.3%)	886 (65.4%)	0.633	107 (60.5%)	88 (61.5%)	19 (55.9%)	0.066
1	1804 (25.2%)	1475 (25.4%)	329 (24.3%)		44 (24.9%)	39 (27.3%)	5 (14.7%)	
≥ 2	623 (8.7%)	509 (8.8%)	114 (8.5%)		16 (9.0%)	10 (7.0%)	6 (17.6%)	
Missing	116 (1.6%)	91 (1.5%)	25 (1.8%)		10 (5.6%)	6 (4.2%)	4 (11.8%)	
Number of missed ARV visits								
0	5238 (73.2%)	4257 (73.3%)	981 (72.5%)	0.194	120 (67.8%)	99 (69.2%)	21 (61.8%)	0.942[δ]
1	1497 (20.9%)	1219 (21.0%)	278 (20.5%)		37 (21.0%)	30 (21.0%)	7 (20.6%)	
≥ 2	302 (4.2%)	233 (4.0%)	69 (5.1%)		10 (5.6%)	8 (5.6%)	2 (5.8%)	
Missing	123 (1.7%)	97 (1.7%)	26 (1.9%)		10 (5.6%)	6 (4.2%)	4 (11.8%)	
Number of missed medical visits								
0:	5353 (74.8%)	4360 (75.1%)	993 (73.3%)	0.160	125 (70.6%)	106 (74.1%)	19 (55.9%)	0.044[δ]
1:	1411 (19.7%)	1139 (19.6%)	272 (20.1%)		37 (21.0%)	29 (20.3%)	8 (23.5%)	
≥ 2:	273 (3.8%)	210 (3.6%)	63 (4.7%)		5 (2.8%)	2 (1.4%)	3 (8.8%)	
Missing:	123 (1.7%)	97 (1.7%)	26 (1.9%)		10 (5.6%)	6 (4.2%)	4 (11.8%)	
BMI at 6 months in kg/m^2								
Median (IQR)	23.3 (21 – 27)	23.7 (21 – 27)	23 (20.7 – 26)	0.0001[þ]	23.8 (21.6 – 27.7)	24 (21.7 – 28.2)	23 (20.7 – 25)	0.2055[þ]

BMI at 6 months in kg/m^2								
< 18.5	463 (6.5%)	351 (6.0%)	112 (8.3%)	0.008	12 (6.8%)	10 (7.0%)	2 (5.9%)	0.476^{ð}
18.5 – < 25	3344 (46.7%)	2683 (46.2%)	662 (48.9%)		72 (40.7%)	55 (38.5%)	17 (50%)	
25 – < 30	1684 (23.5%)	1375 (23.7%)	306 (22.6%)		39 (22.0%)	33 (23.1%)	6 (17.6%)	
≥ 30	835 (11.7%)	695 (11.9%)	142 (10.5%)		22 (12.4%)	20 (14.0%)	2 (5.9%)	
Missing	834 (11.6%)	702 (12.2%)	132 (9.7%)		32 (18.1%)	25 (17.4%)	7 (20.6%)	
Haemoglobin at 6 months in g/dl								
Median (IQR)	13.2 (12.1 – 14.5)	13.3 (12.1 – 14.5)	13.1 (12 – 14.3)	0.0131^{þ}	13.3 (12.3 – 14.6)	13.1 (12.1 – 14.6)	13.7 (12.5 – 14.7)	0.6436^{þ}
Haemoglobin at 6 months								
≤ 8 g/dl	50 (0.7%)	37 (0.6%)	13 (1.0%)	0.101	1 (0.6%)	0 (0%)	1 (2.9%)	0.149^{ð}
> 8g/dl	5190 (72.5%)	4298 (74.0%)	892 (65.9%)		120 (67.8%)	103 (72.0%)	17 (50%)	
Missing	1920 (26.8%)	1471 (25.4%)	449 (33.2%)		56 (31.6%)	40 (28.0%)	16 (47.1%)	
MCV at 6 months in fL								
Median (IQR)	99.6 (93 – 106.6)	100.6 (94 – 107.5)	95.2 (90.4 – 101.2)	0.0000 ^{þ}	107.6 (102.4 – 113.5)	109.2 (103.3 – 114.2)	100.2 (92.3 – 108.3)	0.0005^{þ}
MCV at 6 months								
< 80fL	218 (3.0%)	177 (3.0%)	41 (3.0%)	0.000	2 (1.1%)	1 (0.7%)	1 (2.9%)	0.000^{ð}
80 – 100fL	2896 (40.4%)	2203 (37.9%)	693 (51.2%)		23 (13.0%)	12 (8.4%)	11 (32.4%)	
> 100fL	2918 (40.8%)	2624 (45.2%)	294 (21.7%)		125 (70.6%)	113 (79.0%)	12 (35.3%)	
Missing	1128 (15.8%)	802 (13.9%)	326 (24.1%)		27 (15.3%)	17 (11.9%)	10 (29.4%)	
Change in MCV at 6 months in fL								
Median (IQR)	12.5 (5.8 – 18.5)	13.6 (6.8 – 19)	6.8 (2.6 – 11.8)	0.0000^{þ}	15.1 (4.2 – 20.7)	16.1 (5.9 – 22)	7.4 (3.2 – 16.9)	0.0336^{þ}
Change in MCV at 6 months								
≥ 14.5fL	1902 (26.5%)	1785 (30.7%)	117 (8.6%)	0.000	53 (29.9%)	49 (34.3%)	4 (11.8%)	0.055^{ð}
< 14.5fL	2626 (36.7%)	2073 (35.7%)	553 (40.8%)		52 (29.4%)	41 (28.7%)	11 (32.4%)	
Missing	2632 (36.8%)	1948 (33.6%)	684 (50.6%)		72 (40.7%)	53 (37.0%)	19 (55.8%)	
Change in CD4 count at 6 months in cells/mm^3								
Median (IQR)	129 (69 – 205)	135 (76 – 208)	97 (33 – 180)	0.0000 ^{þ}	95.5 (45 – 140)	102 (54 – 145)	52 (-9 – 61)	0.0044^{þ}

Change in CD4 count at 6 months in cells/mm^3								
≥ expected	4348 (60.7%)	3738 (64.4%)	610 (45.1%)	0.000	85 (48%)	79 (55.2%)	6 (17.6%)	0.001
< expected	1156 (16.1%)	856 (14.7%)	300 (22.2%)		37 (20.9%)	26 (18.2%)	11 (32.4%)	
Missing	1656 (23.2%)	1212 (20.9%)	444 (32.7%)		55 (31.1%)	38 (26.6%)	17 (50%)	
Change in CD4 count at 6 months stratified by change in MCV at 6 months								
-Change in CD4 count ≥ expected and change in MCV ≥ 14.5fL:	1412 (19.7%)	1331 (22.9%)	81 (6.0%)	0.000	37 (20.9%)	35 (24.5%)	2 (5.9%)	0.022$^{\delta}$
-Change in CD4 count ≥ expected and change in MCV < 14.5fL:	1863 (26.0%)	1553 (26.7%)	310 (22.9%)		29 (16.4%)	26 (18.2%)	3 (8.8%)	
-Change in CD4 count < expected and change in MCV ≥ 14.5fL:	358 (5%)	337 (5.8%)	21 (1.6%)		11 (6.2%)	9 (6.3%)	2 (5.9%)	
-Change in CD4 count < expected and change in MCV < 14.5fL:	537 (7.5%)	349 (6.0%)	188 (13.9%)		20 (11.3%)	13 (9.1%)	7 (20.5%)	
Missing:	2990 (41.8%)	2236 (38.5%)	754 (55.6%)		80 (45.2%)	60 (41.9%)	20 (58.9%)	
Regimen change								
No	6445 (90.0%)	5210 (89.7%)	1235 (91.2%)	0.103	171 (96.6%)	140 (97.9%)	31 (91.2%)	0.086$^{\delta}$
Yes	715 (10.0%)	596 (10.3%)	119 (8.8%)		6 (3.4%)	3 (2.1%)	3 (8.8%)	
TB$^{\alpha}$: On TB treatment during first 6 months on ART:								
Never	5355 (74.8%)	4345 (74.8%)	1010 (74.6%)	0.764	150 (84.7%)	120 (83.9%)	30 (88.2%)	1.000$^{\delta}$
At baseline	1493 (20.9%)	1204 (20.7%)	289 (21.3%)		24 (13.6%)	20 (14.0%)	4 (11.8%)	
During follow-up	312 (4.3%)	257 (4.5%)	55 (4.1%)		3 (1.7%)	3 (2.1%)	0 (0%)	
Pregnancy$^{\alpha}$: Pregnancy during first 6 months on ART:								
Never	4294 (60.0%)	3534 (60.9%)	760 (56.1%)	0.445	92 (52.0%)	80 (55.9%)	12 (35.3%)	0.025$^{\delta}$
At baseline	82 (1.1%)	67 (1.2%)	15 (1.1%)		0 (0%)	0 (0%)	0 (0%)	
During follow-up	147 (2.1%)	115 (1.9%)	32 (2.4%)		5 (2.8%)	2 (1.4%)	3 (8.8%)	
Missing$^{\gamma}$	2637 (36.8%)	2090 (36.0%)	547 (40.4%)		80 (45.2%)	61 (42.7%)	19 (55.9%)	

Adverse events or side effects								
No	6955 (97.1%)	5618 (96.8%)	1337 (98.7%)	0.000	176 (99.4%)	142 (99.3%)	34 (100%)	1.000[ð]
Yes	205 (2.9%)	188 (3.2%)	17 (1.3%)		1 (0.6%)	1 (0.7%)	0 (0%)	

[*] On AZT-based regimens throughout the study period. [þ] Wilcoxon rank sum test. [ð] Fisher's exact test. [δ] The last self-reported adherence levels on, or after, the 3 month visit (70 days post-initiation) and not more than 3 months (110 days) before the date of the viral load but at least 28 days before the viral load was done. [α] Variable has baseline and 6 month components. [γ] Male patients. ART: Antiretroviral Therapy. ARV: Antiretroviral. AZT: Zidovudine. BMI: Body Mass Index. IQR: Interquartile range. MCV: Mean cell volume. TB: Tuberculosis.

The table above shows the clinical characteristics of all eligible patients and eligible patients on AZT-based regimens at 6 months. The characteristics were described using frequencies with percentages and medians with interquartile ranges. There were a large number of missing values. Over half of the patients did not have a documented self-reported adherence. Missing values for a missed medical visit, missed ARV drug pickup or missing either a medical or ARV visit were the result of the patient having only one documented scheduled visit during the study period (excluding the ART initiation visit) i.e there was no possibility of the patient falling into the group with ≥ 2 missed visits. Over 60% of the patients did not miss a single visit. 40.4 % of all eligible patients had a normal MCV at 6 months and 40.8% of all eligible patients had a high MCV at 6 months. A median change in CD4 count of 129 cells/mm^3 (IQR: 69 – 205) was obtained for all eligible patients at 6 months. 5.1% of the females were pregnant during the study period. About a quarter of the patients were on TB treatment during the study period but adverse events or side effects were rare. 10% of patients had a regimen change during the first 6 months on ART. The table also shows p-values for the association between different variables and poor adherence at 6 months measured by viral load at 6 months. Except where otherwise stated, the chi-square test was used.

Table 2: Table showing the point at which the ART adherence was reported and the points at which the MCV, CD4 count and viral load at 6 months were done for eligible patients with a documented viral load result

Variable	All eligible patients	Eligible patients on AZT-based regimens*
	Median(IQR)	Median (IQR)
Self-reported adherence	63 days (IQR: 56 – 84) before the viral load was done.	56 days (IQR: 41 – 84) before the viral load was done.
MCV at 6 months	202 days (IQR: 170 – 268) after ART initiation.	203.5 days (IQR: 175 – 261) after ART initiation.
CD4 count at 6 months	204 days (IQR: 171 – 272) after ART initiation.	210.5 days (IQR: 176 – 265.5) after ART initiation.
Viral load at 6 months	203 days (IQR: 173 – 267) after ART initiation.	205 days (IQR: 176 – 261) after ART initiation.

* On AZT-based regimens throughout the study period. AZT: Zidovudine. IQR: Interquartile range. MCV: Mean cell volume.

The table above shows that despite the value for a variable at 6 months being defined as the first available value between 150 and 300 days after ART initiation, the MCV, CD4 count and viral load at 6 months for all eligible patients were done at approximately the same time as the same tests for all eligible patients on AZT-based regimens. Self-reported adherence, assessed as the last self-reported adherence levels on, or after, the 3 month visit (70 days post-initiation) and not more than 3 months (110 days) before the date of the viral load but at least 28 days before the viral load was done, had differences in the medians between all eligible patients and eligible patients on AZT-based regimens.

3.2.2 Poisson regression with robust error variance:

In this section, Poisson regression models (with robust error variance) for patients with a documented viral load result at 6 months are presented.

Table 3a: Univariate and multivariate Poisson regression models with robust error variance for all eligible patients and eligible patients on AZT-based regimens with a documented viral load result

Variable	All eligible patients				All eligible patients on AZT-based regimens*			
	Univariate Analyses		Multivariate Analysis		Univariate Analyses		Multivariate Analysis	
	Crude IRR (95% CI)	P - value	Adjusted IRR (95% CI)	P- value	Crude IRR (95% CI)	P - value	Adjusted IRR (95% CI)	P- value
Gender								
Females	1	-	1	-	1	-	-	-
Males	1.16 (1.05 – 1.28)	0.002	1.09 (0.92 – 1.28)	0.320	1.54 (0.83 – 2.83)	0.168	-	-
Age at initiation								
≤ 36.7 years	1	-	1	-	1	-	1	-
> 36.7 years	0.94 (0.85 – 1.04)	0.213	0.94 (0.80 – 1.11)	0.463	0.51 (0.28 – 0.93)	0.028	1.18 (0.31 – 4.53)	0.808
Education level								
≥ Secondary school	1	-	-	-	1	-	-	-
Primary school	0.90 (0.77 – 1.04)	0.164	-	-	1.31 (0.58 – 2.94)	0.519	-	-
< Primary school	0.81 (0.60 – 1.09)	0.167	-	-	0.75 (0.12 – 4.84)	0.759	-	-
Professional status								
Employed	1	-	-	-	1	-	-	-
Unemployed	1.09 (0.99 – 1.20)	0.078	-	-	0.87 (0.47 – 1.62)	0.666	-	-
BMI at initiation in kg/m^2								
< 18.5	1	-	-	-	1	-	-	-
18.5 – < 25	0.88 (0.76 – 1.01)	0.067	-	-	2.38 (0.60 – 9.43)	0.219	-	-
25 – < 30	0.85 (0.72 – 1.01)	0.071	-	-	1.9 (0.39 – 9.25)	0.427	-	-
≥ 30	0.82 (0.66 – 1.02)	0.079	-	-	0.40 (0.04 – 4.08)	0.436	-	-
Baseline haemoglobin								
≤ 8 g/dl	1	-	-	-	1	-	-	-
> 8g/dl	0.99 (0.79 – 1.25)	0.951	-	-	-‡	-	-	-

MCV at baseline								
< 80fL	1	-	1	-	1	-	-	-
80 – 100fL	1.23 (1.01 – 1.51)	0.042	1.33 (1.01 – 1.75)	0.044	0.61 (0.24 – 1.58)	0.314	-	-
> 100fl	1.36 (0.99 – 1.87)	0.059	0.98 (0.62 – 1.55)	0.937	0.43 (0.13 – 1.36)	0.151	-	-
Baseline CD4 count in cells/mm^3								
> 200	1	-	1	-	1	-	1	-
101 – 200	0.96 (0.81 – 1.13)	0.594	1.05 (0.80 – 1.38)	0.716	0.74 (0.30 – 1.79)	0.498	0.82 (0.07 – 9.24)	0.870
51 – 100	1.09 (0.92 – 1.31)	0.318	1.08 (0.80 – 1.47)	0.612	0.77 (0.28 – 2.14)	0.614	3.11 (0.47 – 20.67)	0.240
≤ 50	1.19 (1.01 – 1.40)	0.034	1.34 (1.02 – 1.76)	0.035	0.88 (0.35 – 2.24)	0.794	2.10 (0.12 – 37.31)	0.612
WHO stage at baseline								
I	1	-	1	-	1	-	-	-
II	1.22 (1.05 – 1.43)	0.012	1.16 (0.90 – 1.48)	0.248	0.71 (0.18 – 2.84)	0.653	-	-
III	1.24 (1.08 – 1.42)	0.002	1.27 (1.04 – 1.55)	0.021	1.29 (0.55 – 3.00)	0.560	-	-
IV	1.30 (1.09 – 1.55)	0.004	1.44 (1.12 – 1.84)	0.004	-‡	-	-	-
Self-reported adherence								
> 90%	1	-	-	-	1	-	-	-
> 60 – 90%	1.08 (0.87 – 1.34)	0.472	-	-	1.07 (0.16 – 7.40)	0.944	-	-
≤ 60%	1.38 (0.97 – 1.96)	0.077	-	-	3.21 (0.68 – 15.28)	0.142	-	-
Number of missed visits								
0	1	-	-	-	1	-	-	-
1	0.95 (0.85 – 1.07)	0.382	-	-	0.64 (0.25 – 1.61)	0.343	-	-
≥ 2	0.95 (0.80 – 1.14)	0.597	-	-	2.11 (0.99 – 4.49)	0.052	-	-
Number of missed ARV visits								
0	1	-	-	-	1	-	-	-
1	0.99 (0.88 – 1.12)	0.890	-	-	1.08 (0.50 – 2.34)	0.844	-	-
≥ 2	1.22 (0.98 – 1.51)	0.070	-	-	1.14 (0.31 – 4.21)	0.841	-	-
Number of missed medical visits								
0	1	-	-	-	1	-	-	-
1	1.04 (0.92 – 1.17)	0.532	-	-	1.42 (0.68 – 2.99)	0.352	-	-
≥ 2	1.24 (0.99 – 1.56)	0.056	-	-	3.95 (1.72 – 9.05)	0.001	-	-

BMI at 6 months in kg/m^2								
< 18.5	1	-	-	-	1	-	-	-
18.5 – < 25	0.82 (0.69 – 0.97)	0.025	-	-	1.42 (0.37 – 5.39)	0.609	-	-
25 – < 30	0.75 (0.62 – 0.91)	0.003	-	-	0.92 (0.21 – 4.01)	0.915	-	-
≥ 30	0.70 (0.56 – 0.87)	0.002	-	-	0.55 (0.49 – 3.42)	0.518	-	-
Haemoglobin at 6 months								
≤ 8 g/dl	1	-	-	-	1	-	-	-
> 8g/dl	0.66 (0.41 – 1.06)	0.085	-	-	0.14 (0.09 – 0.22)	0.000	-	-
MCV at 6 months								
< 80fL	1	-	-	-	1	-	-	-
80 – 100fL	1.27 (0.96 – 1.69)	0.096	-	-	0.96 (0.22 – 4.10)	0.952	-	-
> 100fL	0.54 (0.40 – 0.72)	0.000	-	-	0.19 (0.04 – 0.85)	0.030	-	-
Change in MCV at 6 months								
≥ 14.5 fL	1	-	-	-	1	-	-	-
< 14.5 fL	3.42 (2.83 – 4.14)	0.000	-	-	2.80 (0.95 – 8.28)	0.062	-	-
Change in CD4 count at 6 months in cells/mm^3								
≥ expected	1	-	-	-	1	-	1	-
< expected	1.85 (1.64 – 2.09)	0.000	-	-	4.21 (1.68 – 10.57)	0.002	7.66 (0.98 – 59.91)	0.052
Change in CD4 count at 6 months stratified by change in MCV at 6 months								
-Change in CD4 count ≥ expected and change in MCV ≥ 14.5fL:	1	-	1	-	1	-	-	-
-Change in CD4 count ≥ expected and change in MCV < 14.5fL:	2.90 (2.29 – 3.66)	0.000	3.11 (2.41 – 4.02)	0.000	1.91 (0.34 – 10.80)	0.462	-	-
-Change in CD4 count < expected and change in MCV ≥ 14.5fL:	1.02 (0.64 – 1.63)	0.925	1.23 (0.76 – 2.00)	0.394	3.36 (0.53 – 21.40)	0.199	-	-
-Change in CD4 count < expected and change in MCV < 14.5fL:	6.10 (4.80 – 7.76)	0.000	6.98 (5.35 – 9.09)	0.000	6.48 (1.47 – 28.50)	0.013	-	-
Regimen change								
No	1	-	-	-	1	-	-	-
Yes	0.87 (0.73 – 1.03)	0.107	-	-	2.76 (1.16 – 6.54)	0.021	-	-

TB: On TB treatment during first 6 months on ART:								
Never	1	-	-	-	1	-	-	-
At baseline	1.03 (0.91 – 1.15)	0.665	-	-	0.83 (0.32 – 2.16)	0.708	-	-
During follow-up	0.93 (0.73 – 1.20)	0.591	-	-	-‡	-	-	-
Pregnancy: Pregnancy during first 6 months on ART:								
Never	1	-	-	-	1	-	1	-
At baseline	1.03 (0.65 – 1.64)	0.889	-	-	-‡	-	-‡	-
During follow-up	1.23 (0.90 – 1.68)	0.195	-	-	4.6 (1.88 – 11.24)	0.001	9.11 (2.17 – 38.25)	0.003
Adverse events or side effects								
No	1	-	-	-	1	-	-	-
Yes	0.43 (0.27 – 0.68)	0.000	-	-	-‡	-	-	-

* On AZT-based regimens throughout the study period. ‡ Cell has no observations. ART: Antiretroviral Therapy. ARV: Antiretroviral. AZT: Zidovudine. BMI: Body Mass Index. CI: Confidence interval. IRR: Incidence rate ratio. MCV: Mean cell volume. TB: Tuberculosis. WHO: World Health Organization.

For the table above, multivariate Poisson regression analysis with robust error variance was done first for all eligible patients then for all eligible patients on AZT-based regimens. After adjusting for potential confounders, the variables independently associated with poor adherence for all eligible patients were change in CD4 count at 6 months stratified by change in MCV at 6 months, CD4 count at baseline, WHO stage at baseline and MCV at baseline. The IRR of poor adherence increased with decreasing CD4 count and increasing WHO stage at baseline. Model fit was good as evidenced by a p-value of 0.914 on the link test. Change in CD4 count at 6 months and pregnancy during the first 6 months on ART were the only independent predictors of poor adherence in the final multivariate model for patients on AZT-based regimens. Model fit for this model was also good as evidenced by a p-value of 0.812 on the link test.

Table 3b: Univariate and multivariate Poisson regression models with robust error variance for all eligible patients and eligible patients on AZT-based regimens with a documented viral load result after multiple imputation

Variable	All eligible patients				All eligible patients on AZT-based regimens*			
	Univariate Analyses		Multivariate Analysis		Univariate Analyses		Multivariate Analysis	
	Crude IRR (95% CI)	P - value	Adjusted IRR (95% CI)	P- value	Crude IRR (95% CI)	P - value	Adjusted IRR (95% CI)	P- value
Gender								
Females	1	-	1	-	1	-	1	-
Males	1.16 (1.05 – 1.28)	0.002	1.07 (0.97 – 1.18)	0.166	1.54 (0.83 – 2.83)	0.168	1 .48 (0.81 – 2.73)	0.202
Age at initiation								
≤ 36.7 years	1	-	1	-	1	-	1	-
> 36.7 years	0.94 (0.85 – 1.04)	0.213	0.96 (0.87 – 1.06)	0.415	0.51 (0.28 – 0.93)	0.028	0.54 (0.30 – 0.98)	0.041
Education level								
≥ Secondary school	1	-	-	-	1	-	-	-
Primary school	0.89 (0.77 – 1.03)	0.107	-	-	1.09 (0.48 – 2.50)	0.835	-	-
< Primary school	0.82 (0.61 – 1.12)	0.217	-	-	0.68 (0.10 – 4.39)	0.684	-	-
Professional status								
Employed	1	-	-	-	1	-	-	-
Unemployed	1.09 (0.99 – 1.20)	0.072	-	-	0.88 (0.47 – 1.63)	0.677	-	-
BMI at initiation in kg/m^2								
< 18.5	1	-	-	-	1	-	-	-
18.5 – < 25	0.89 (0.78 – 1.02)	0.086	-	-	1.66 (0.46 – 6.00)	0.435	-	-
25 – < 30	0.86 (0.73 – 1.02)	0.082	-	-	1.39 (0.33 – 5.79)	0.652	-	-
≥ 30	0.79 (0.64 – 0.98)	0.030	-	-	0.45 (0.06 – 3.49)	0.447	-	-
Baseline haemoglobin								
≤ 8 g/dl	1	-	-	-	1	-	-	-
> 8g/dl	1.00 (0.80 – 1.26)	0.968	-	-	-‡	-	-	-
MCV at baseline								
< 80fL	1	-	-	-	1	-	-	-
80 – 100fL	1.20 (1.00 – 1.44)	0.056	-	-	0.65 (0.26 – 1.60)	0.347	-	-
> 100fl	1.24 (0.89 – 1.72)	0.196	-	-	0.45 (0.15 – 1.40)	0.169	-	-

Baseline CD4 count in cells/mm^3								
> 200	1	-	1	-	1	-	1	-
101 – 200	0.97 (0.83 – 1.13)	0.666	1.10 (0.94 – 1.29)	0.212	0.74 (0.30 – 1.82)	0.506	0.76 (0.31 – 1.86)	0.553
51 – 100	1.10 (0.92 – 1.31)	0.297	1.31 (1.09 – 1.56)	0.003	0.81 (0.31 – 2.15)	0.677	0.92 (0.35 – 2.44)	0.873
≤ 50	1.21 (1.02 – 1.41)	0.026	1.49 (1.27 – 1.74)	0.000	0.92 (0.37 – 2.29)	0.857	1.07 (0.45 – 2.54)	0.885
WHO stage at baseline								
I	1	-	-	-	1	-	-	-
II	1.21 (1.03 – 1.41)	0.018	-	-	0.83 (0.25 – 2.77)	0.765	-	-
III	1.24 (1.08 – 1.42)	0.002	-	-	1.17 (0.52 – 2.62)	0.701	-	-
IV	1.28 (1.09 – 1.51)	0.003	-	-	-‡	-	-	-
Self-reported adherence								
> 90%	1	-	-	-	1	-	-	-
> 60 – 90%	1.06 (0.83 – 1.35)	0.647	-	-	0.84 (0.16 – 4.45)	0.838	-	-
≤ 60%	1.64 (1.15 – 2.34)	0.006	-	-	2.64 (0.63 – 11.03)	0.184	-	-
Number of missed visits								
0	1	-	-	-	1	-	-	-
1	0.95 (0.85 – 1.07)	0.382	-	-	0.64 (0.25 – 1.61)	0.343	-	-
≥ 2	0.95 (0.80 – 1.14)	0.597	-	-	2.11 (0.99 – 4.49)	0.052	-	-
Number of missed ARV visits								
0	1	-	-	-	1	-	-	-
1	0.99 (0.88 – 1.12)	0.885	-	-	1.07 (0.50 – 2.31)	0.860	-	-
≥ 2	1.22 (0.98 – 1.51)	0.069	-	-	1.15 (0.32 – 4.09)	0.834	-	-
Number of missed medical visits								
0	1	-	-	-	1	-	-	-
1	1.04 (0.92 – 1.17)	0.547	-	-	1.37 (0.67 – 2.80)	0.393	-	-
≥ 2	1.24 (0.99 – 1.55)	0.063	-	-	3.54 (1.54 – 8.16)	0.003	-	-
BMI at 6 months in kg/m^2								
< 18.5	1	-	-	-	1	-	-	-
18.5 – < 25	0.82 (0.69 – 0.97)	0.024	-	-	1.43 (0.38 – 5.36)	0.592	-	-
25 – < 30	0.76 (0.63 – 0.92)	0.004	-	-	1.06 (0.26 – 4.25)	0.938	-	-
≥ 30	0.69 (0.56 – 0.86)	0.001	-	-	0.54 (0.09 – 3.26)	0.501	-	-

Haemoglobin at 6 months								
≤ 8 g/dl	1	-	-	-	1	-	-	-
> 8g/dl	0.72 (0.45 – 1.16)	0.178	-	-	0.19 (0.14 – 0.26)	0.000	-	-
MCV at 6 months								
< 80fL	1	-	-	-	1	-	-	-
80 – 100fL	1.32 (1.00 – 1.75)	0.054	-	-	0.92 (0.22 – 3.78)	0.908	-	-
> 100fL	0.60 (0.45 – 0.80)	0.001	-	-	0.22 (0.05 – 0.94)	0.041	-	-
Change in MCV at 6 months								
≥ 14.5 fL	1	-	-	-	1	-	-	-
< 14.5 fL	2.61 (2.22 – 3.04)	0.000	-	-	2.36 (1.09 – 5.08)	0.029	-	-
Change in CD4 count at 6 months in cells/mm^3								
≥ expected	1	-	-	-	1	-	1	-
< expected	1.76 (1.58 – 1.96)	0.000	-	-	2.50 (1.18 – 5.32)	0.017	2.34 (1.07 – 5.12)	0.033
Change in CD4 count at 6 months stratified by change in MCV at 6 months								
-Change in CD4 count ≥ expected and change in MCV ≥ 14.5fL:	1	-	1	-	1	-	-	-
-Change in CD4 count ≥ expected and change in MCV < 14.5fL:	2.46 (2.00 – 3.02)	0.000	2.52 (2.04 – 3.10)	0.000	2.16 (0.68 – 6.85)	0.192	-	-
-Change in CD4 count < expected and change in MCV ≥ 14.5fL:	1.53 (1.08 – 2.15)	0.017	1.61 (1.13 – 2..27)	0.008	2.34 (0.58 – 9.39)	0.229	-	-
-Change in CD4 count < expected and change in MCV < 14.5fL:	4.38 (3.59 – 5.34)	0.000	4.65 (3.79 – 5.70)	0.000	4.83 (1.71 – 13.66)	0.003	-	-
Regimen change								
No	1	-	-	-	1	-	-	-
Yes	0.87 (0.73 – 1.03)	0.107	-	-	2.76 (1.16 – 6.54)	0.021	-	-
TB: On TB treatment during first 6 months on ART:								
Never	1	-	-	-	1	-	-	-
At baseline	1.03 (0.91 – 1.15)	0.665	-	-	0.83 (0.32 – 2.16)	0.708	-	-
During follow-up	0.93 (0.73 – 1.20)	0.591	-	-	-‡	-	-	-

Pregnancy: Pregnancy during first 6 months on ART:								
Never	1	-	-	-	1	-	-	-
At baseline	1.03 (0.65 – 1.64)	0.889	-	-	-‡	-	-	-
During follow-up	1.23 (0.90 – 1.68)	0.195	-	-	4.6 (1.88 – 11.24)	0.001	-	-
Adverse events or side effects								
No	1	-	-	-	1	-	-	-
Yes	0.43 (0.27 – 0.68)	0.000	-	-	-‡	-	-	-

* On AZT-based regimens throughout the study period. . ‡ Cell in one or more imputations has no observations. ART: Antiretroviral Therapy. ARV: Antiretroviral. AZT: Zidovudine. CI: Confidence interval. IRR: Incidence rate ratio. MCV: Mean cell volume. TB: Tuberculosis. WHO: World Health Organization.

For the table above, multiple imputation (20 imputations) was done for all patients with a documented viral load result using multivariate normal regression. The imputation was done for explanatory variables with missing values. Multiple imputation was followed by multivariate Poisson regression analysis with robust error variance. After adjusting for potential confounders, the variables independently associated with poor adherence for all eligible patients were change in CD4 count at 6 months stratified by change in MCV at 6 months and CD4 count at baseline. The IRR of poor adherence increased with decreasing CD4 count. Model fit was good as evidenced by a p-value of < 0.05 on the FMI test. Change in CD4 count at 6 months and age at initiation were the only independent predictors of poor adherence in the final multivariate model for patients on AZT-based regimens. Model fit for this model was also good as evidenced by a p-value of < 0.05 on the FMI test.

3.3: MEASURES OF VALIDITY

3.3.1 Sensitivity, specificity, PPV and NPV:

In this section, sensitivity, specificity, positive and negative predictive values are presented in a series of tables and graphs.

Table 4a: Sensitivity, specificity, PPV and NPV of markers to identify poor adherence, using viral load as the gold standard, for all patients with a documented viral load result

VARIABLE	BEFORE MULTIPLE IMPUTATION				AFTER MULTIPLE IMPUTATION			
	Sensitivity	Specificity	PPV	NPV	Sensitivity	Specificity	PPV	NPV
Change in MCV at 6 months for patients on d4T or AZT-based regimens*	70.2%	61.4%	18.2%	94.4%	71.1%	58.5%	23.0%	92.1%
Change in MCV at 6 months for patients on TDF-based regimens**	97.4%	3.1%	24.4%	78.4%	89.2%	15.4%	27.2%	80.1%
Change in CD4 count at 6 months	33%	81.4%	26%	86%	33%	80.7%	28.5%	83.8%
Change in CD4 count at 6 months for patients on AZT-based regimens***	64.7%	75.2%	29.7%	92.9%	51.3%	75.5%	33.2%	86.7%
Change in CD4 count stratified by change in MCV	86.5%	37.3%	18.8%	94.3%	85.5%	35.6%	23.6%	91.3%
Change in CD4 count stratified by change in MCV for patients on d4T or AZT-based regimens*	76.8%	49.2%	15%	94.8%	79.3%	46.6%	20.6%	92.8%
Change in CD4 count stratified by change in MCV for patients on TDF-based regimens**	98.2%	2.4%	24.2%	80.8%	92.5%	13%	27.4%	83.1%
Pregnant during the first 6 months on ART for patients on AZT-based regimens***	20%	97.6%	60%	87%	20%	97.6%	60%	87%

*On d4T or AZT-based regimens throughout the study period. **On TDF-based regimens throughout the study period. ***On AZT-based regimens throughout the study period. ART: Antiretroviral Therapy. AZT: Zidovudine. d4T: Stavudine. MCV: Mean cell volume. NPV: Negative predictive value. PPV: Positive predictive value. TDF: Tenofovir.

Table 4b: Table showing sensitivity, specificity, PPV and NPV for each category of change in CD4 count stratified by change in MCV at 6 months, for all patients with a documented viral load result

VARIABLE CATEGORY	BEFORE MULTIPLE IMPUTATION				AFTER MULTIPLE IMPUTATION			
	Sensitivity	**Specificity**	**PPV**	**NPV**	**Sensitivity**	**Specificity**	**PPV**	**NPV**
Change in CD4 count ≥ expected and change in MCV< 14.5fL	79.3%	46.2%	16.6%	94.3%	78.3%	44.1%	21.3%	91.3%
Change in CD4 count < expected and change in MCV ≥ 14.5fL	20.6%	79.8%	5.9%	94.3%	28.7%	80%	13.3%	91.3%
Change in CD4 count < expected and change in MCV< 14.5fL	69.9%	79.2%	35.0%	94.3%	65.1%	77.5%	38.0%	91.3%

MCV: Mean cell volume. NPV: Negative predictive value. PPV: Positive predictive value.

3.3.2 Receiver operating characteristic (ROC) curves:

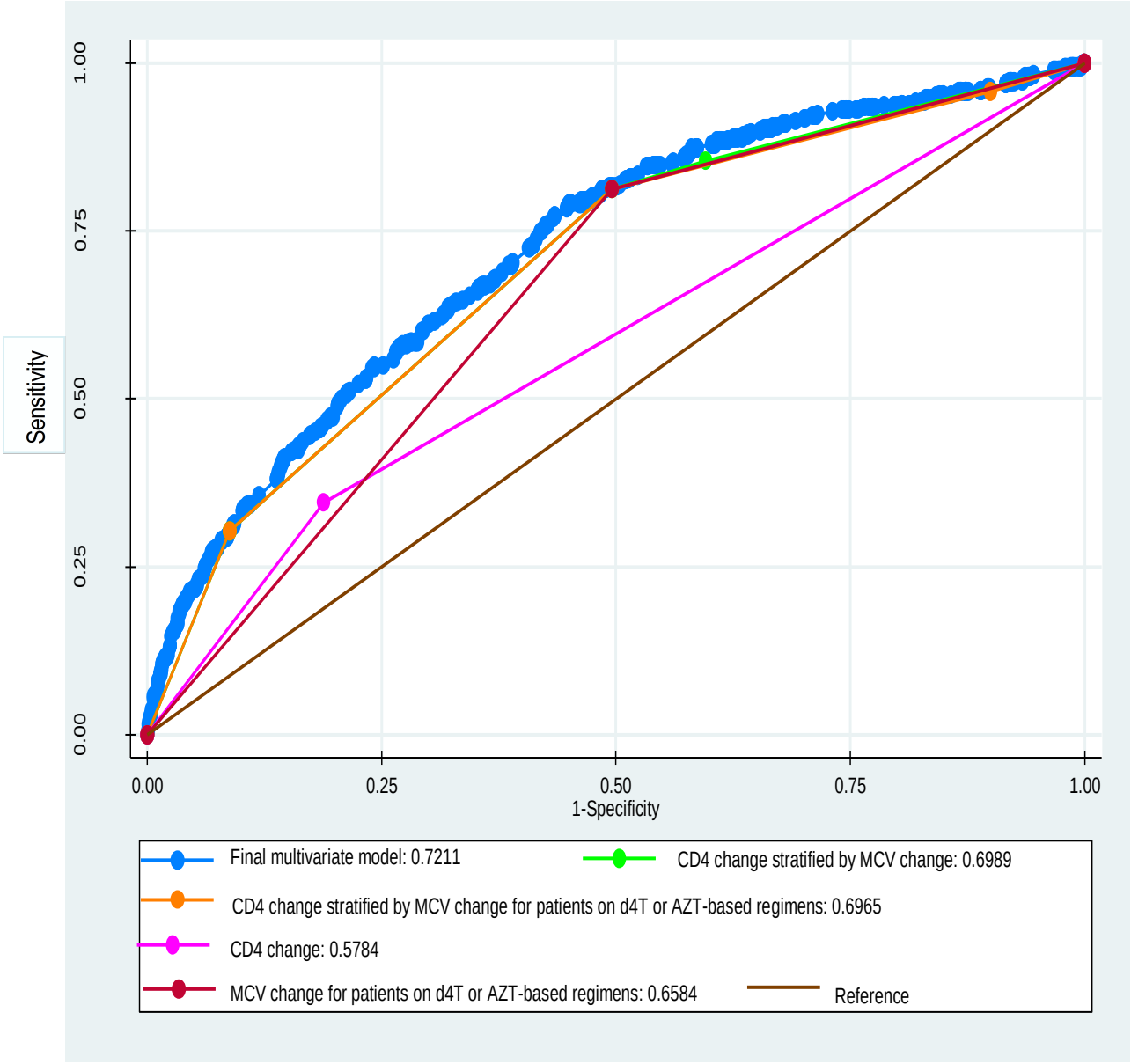


Figure 2a: ROC curves for different markers of adherence for eligible patients with a documented viral load result: The graph above shows ROC curves for individual markers of adherence as well as the ROC curve for the final multivariate model. The area under each curve is also shown.

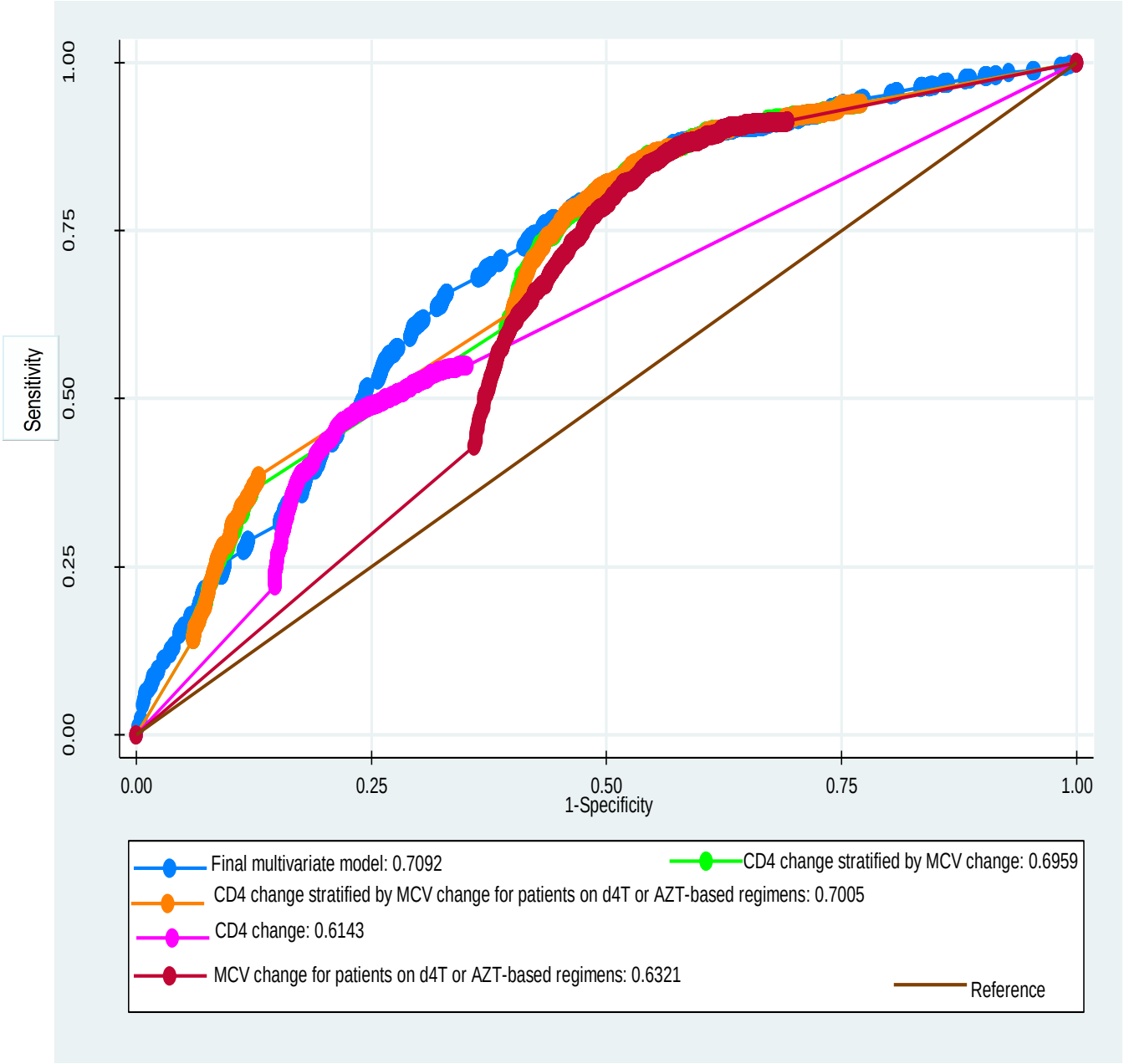


Figure 2b: ROC curves for different markers of adherence for eligible patients with a documented viral load result after multiple imputation: The graph above shows ROC curves for individual markers of adherence as well as the ROC curve for the final multivariate model. The area under each curve is also shown.

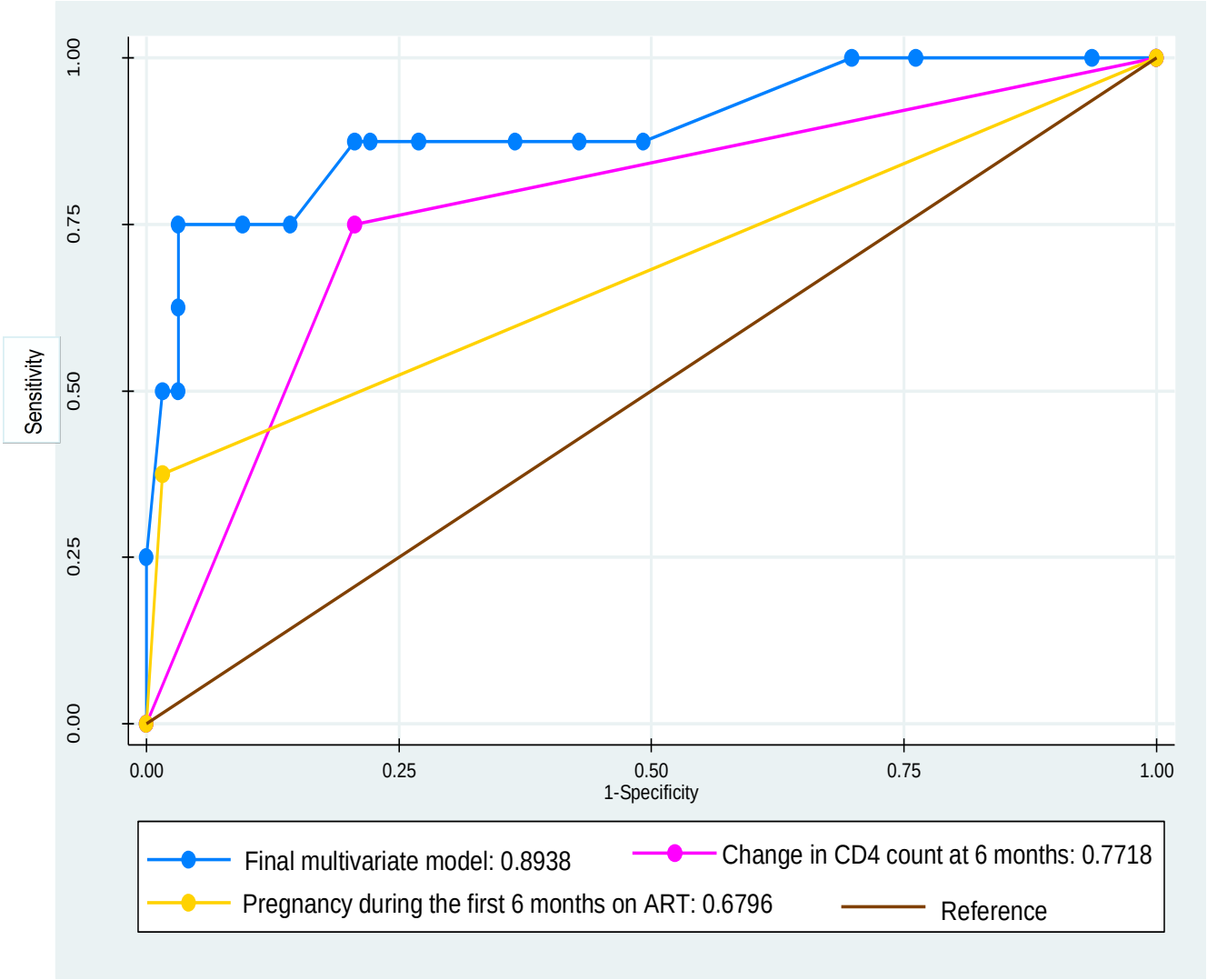


Figure 2c: ROC curves for different markers of adherence for all eligible patients on AZT-based regimens with a documented viral load result: The graph above shows different ROC curves for individual markers of adherence as well as the ROC curve for the final multivariate model for patients on AZT-based regimens. The area under each curve is also shown.

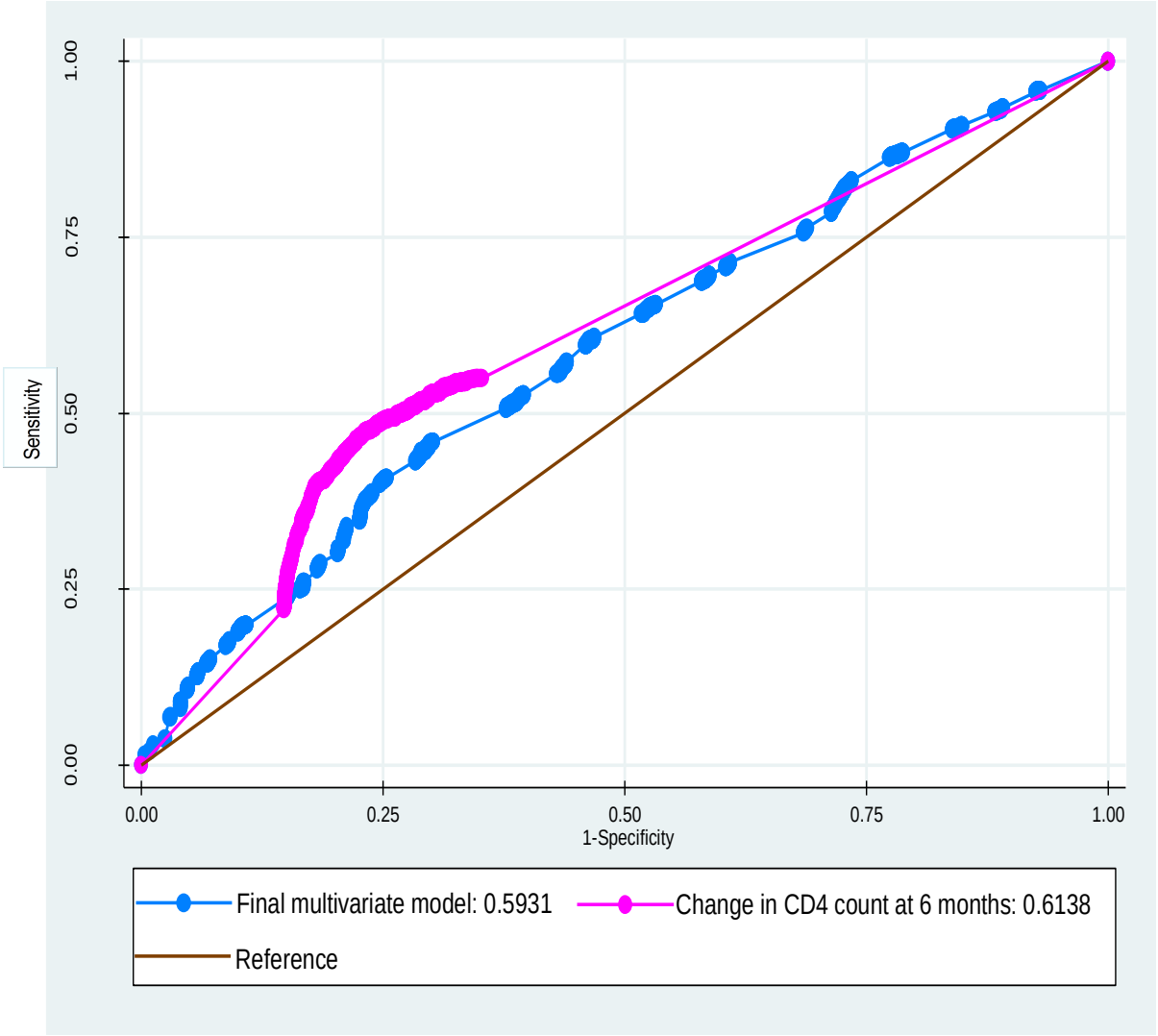


Figure 2d: ROC curves for different markers of adherence for all eligible patients on AZT-based regimens with a documented viral load result after multiple imputation: The graph above shows different ROC curves for individual markers of adherence as well as the ROC curve for the final multivariate model for patients on AZT-based regimens. The area under each curve is also shown.

3.4 SENSITIVITY ANALYSES:

3.4.1 Descriptive analysis:

In this section the baseline and 6 month characteristics of all patients including those without a documented viral load result at 6 months are presented.

Table 5a: Baseline demographic characteristics and their associations with poor adherence for all patients including those excluded due to missing viral load results assuming that those with missing viral load results had a viral load ≥ 400 copies/ml

Characteristics	All patients				All patients on AZT-based regimens*			
	Total N = 14384	Viral load < 400 copies/ml N = 5806	Viral load ≥ 400 copies/ml N = 8578	P-value	Total N = 370	Viral load < 400 copies/ml N = 143	Viral load ≥ 400 copies/ml N = 227	P-value
	n (%)	n (%)	n (%)		n (%)	n (%)	n(%)	
Gender								
Females	9051 (62.9%)	3716 (64%)	5335 (62.2%)	0.028	204 (55.1%)	82 (57.3%)	122 (53.7%)	0.498
Males	5333 (37.1%)	2090 (36%)	3243 (37.8%)		166 (44.9%)	61 (42.7%)	105 (46.3%)	
Age at initiation in years								
Median (IQR)	36.6 (31.4 – 43.2)	36.8 (31.6 – 43.5)	36.5 (31.3 – 43.1)	0.0240[b]	39.6 (33.6 – 45.7)	40.6 (35.6 – 48)	39.3 (32.7 – 45.5)	0.0198[b]
Age at initiation in years								
≤ 36.7	7263 (50.5%)	2901 (50.0%)	4362 (50.9%)	0.297	134 (36.2%)	43 (30.1%)	91 (40.1%)	0.051
> 36.7	7121 (49.5%)	2905 (50.0%)	4216 (49.1%)		236 (63.8%)	100 (69.9%)	136 (59.9%)	
Education level								
≥ Secondary school	8740 (60.8%)	3573 (61.5%)	5167 (60.2%)	0.638	197 (53.2%)	76 (53.1%)	121 (53.3%)	0.417
Primary school	1812 (12.6%)	744 (12.8%)	1068 (12.5%)		57 (15.4%)	18 (12.6%)	39 (17.2%)	
< Primary school	426 (3.0%)	184 (3.2%)	242 (2.8%)		12 (3.2%)	6 (4.2%)	6 (2.6%)	
Missing	3406 (23.6%)	1305 (22.5%)	2101 (24.5%)		104 (28.2%)	43 (30.1%)	61 (26.9%)	

Professional status								
Employed	7191 (50.0%)	3020 (52.0%)	4171 (48.6%)	0.000	180 (48.7%)	71 (49.7%)	109 (48.0%)	0.784
Unemployed	6973 (48.5%)	2700 (46.5%)	4273 (49.8%)		184 (49.7%)	70 (49.0%)	114 (50.2%)	
Missing	220 (1.5%)	86 (1.5%)	134 (1.6%)		6 (1.6%)	2 (1.3%)	4 (1.8%)	

*On AZT-based regimens throughout the study period. [p] Wilcoxon rank sum test. AZT: Zidovudine. IQR: Interquartile range.

The table above shows the baseline demographic characteristics of all patients and patients on AZT-based regimens. The characteristics were described using frequencies with percentages and medians with interquartile ranges. The table demonstrates that the baseline demographic characteristics of all patients including those excluded due to missing viral load results at the end-point were similar to those of the study population (Table 1a). This implies that the 7160 eligible patients in the main analysis, with a documented viral load result, are representative of the total study population. The majority of the patients in the sensitivity analysis (62.9%) were female. The median age at initiation was 36.6 years (IQR: 31.4 – 43.2) but the overall age range at initiation was 18 – 79.9 years. The table also shows p-values for the association between different variables and poor adherence at 6 months measured by viral load at 6 months. Except where otherwise stated, the chi-square test was used.

Table 5b: Baseline clinical characteristics and their associations with poor adherence for all patients including those excluded due to missing viral load results assuming that those with missing viral load results had a viral load ≥ 400 copies/ml

Characteristics	All patients				All patients on AZT-based regimens*			
	Total N = 14384	Viral load <400 copies/ml N = 5806	Viral load ≥400 copies/ml N = 8578	P-value	Total N = 370	Viral load <400 copies/ml N = 143	Viral load ≥400 copies/ml N = 227	P-value
	n (%)	n (%)	n (%)		n (%)	n (%)	n(%)	
BMI at initiation in kg/m^2								
Median (IQR)	22 (19 – 25)	22 (19.7 – 25)	22.0 (19 – 25)	0.000[M]	22.0 (19.3 – 28.0)	22.3 (20.0 – 29)	22 (19 – 26.5)	0.5301[Þ]
BMI at initiation in kg/m^2								
< 18.5	2179 (15.1%)	782 (13.5%)	1397 (16.3%)	0.000	45 (12.2%)	17 (11.9%)	28 (12.3%)	0.922
18.5 – < 25	6693 (46.5%)	2719 (46.8%)	3974 (46.3%)		129 (34.9%)	51 (35.7%)	78 (34.4%)	
25 – < 30	2290 (15.9%)	966 (16.6%)	1324 (15.4%)		40 (10.8%)	16 (11.2%)	24 (10.6%)	
≥ 30	1089 (7.6%)	462 (8.0%)	627 (7.3%)		52 (14.1%)	23 (16.1%)	29 (12.8%)	
Missing	2133 (14.9%)	877 (15.1%)	1256 (14.7%)		104 (28.0%)	36 (25.1%)	68 (29.9%)	
Baseline haemoglobin in g/dl								
Median (IQR)	11.7 (10.1 – 13.1)	11.7 (10.2 – 13.1)	11.6 (10.1 – 13)	0.1328[Þ]	12.6 (11.4 – 13.8)	12.5 (11.5 – 13.9)	12.6 (11.4 – 13.8)	0.8953[Þ]
Baseline haemoglobin								
≤ 8 g/dl	688 (4.8%)	273 (4.7%)	415 (4.8%)	0.929	1 (0.3%)	1 (0.7%)	0 (0%)	0.395[ð]
> 8g/dl	12298 (85.5%)	4901 (84.4%)	7397 (86.2%)		285 (77.0%)	112 (78.3%)	173 (76.2%)	
Missing	1398 (9.7%)	632 (10.9%)	766 (9.0%)		84 (22.7%)	30 (21.0%)	54 (23.8%)	
MCV at baseline in fL								
Median (IQR)	88.2 (83.7 – 92.2)	88.4 (83.9 – 92.3)	88.1 (83.6 – 92.2)	0.1727 [Þ]	90.8 (85.7 – 97.4)	91.9 (87.4 – 102.2)	90.3 (85 – 95.7)	0.0526 [Þ]
MCV at baseline								
< 80fL	1427 (9.9%)	557 (9.6%)	870 (10.1%)	0.300	26 (7.0%)	9 (6.3%)	17 (7.5%)	0.001
80 – 100fL	9757 (67.8%)	3749 (64.6%)	6008 (70.0%)		188 (50.8%)	60 (41.9%)	128 (56.4%)	
> 100fL	475 (3.3%)	199 (3.4%)	276 (3.2%)		55 (14.9%)	33 (23.1%)	22 (9.7%)	
Missing:	2725 (19.0%)	1301 (22.4%)	1424 (16.7%)		101 (27.3%)	41 (28.7%)	60 (26.4%)	

Baseline CD4 count in cells/mm^3 Median (IQR)	101 (39 – 172)	104 (40 – 176)	99 (39 – 169)	0.0025 [Þ]	128 (61 – 218)	154 (74 – 243)	122 (52 – 207)	0.0714 [Þ]
Baseline CD4 in cells/mm^3								
> 200	1858 (12.9%)	826 (14.2%)	1032 (12.0%)	0.001	87 (23.5%)	41 (28.7%)	46 (20.3%)	0.336
101 – 200	4443 (30.9%)	1749 (30.1%)	2694 (31.4%)		81 (21.9%)	30 (21.0%)	51 (22.5%)	
51 – 100	2489 (17.3%)	976 (16.8%)	1513 (17.6%)		51 (13.8%)	19 (13.3%)	32 (14.1%)	
≤ 50	3762 (26.2%)	1476 (25.4%)	2286 (26.6%)		60 (16.2%)	20 (14.0%)	40 (17.6%)	
Missing	1832 (12.7%)	779 (13.5%)	1053 (12.3%)		91 (24.6%)	33 (23.0%)	58 (25.5%)	
WHO stage at baseline								
I	4650 (32.3%)	1826 (31.5%)	2824 (32.9%)	0.031	133 (35.9%)	61 (42.7%)	72 (31.7%)	0.059
II	2376 (16.5%)	920 (15.8%)	1456 (17.0%)		27 (7.3%)	13 (9.1%)	14 (6.2%)	
III	3685 (25.6%)	1409 (24.3%)	2276 (26.5%)		70 (18.9%)	19 (13.3%)	51 (22.5%)	
IV	1337 (9.3%)	572 (9.9%)	765 (8.9%)		27 (7.3%)	11 (7.7%)	16 (7.0%)	
Missing	2336 (16.3%)	1079 (18.5%)	1257 (14.7%)		113 (30.6%)	39 (27.2%)	74 (32.6%)	

[*]On AZT-based regimens throughout the study period. [M] Median test. [Þ] Wilcoxon rank sum test [ð] Fisher's exact test. AZT: Zidovudine. BMI: Body Mass Index. IQR: Interquartile range. MCV: Mean cell volume. WHO: World Health Organization.

The table above shows the baseline clinical characteristics of all patients and patients on AZT-based regimens. The characteristics were described using frequencies with percentages and medians with interquartile ranges. The table demonstrates that the baseline clinical characteristics of all patients including those excluded due to missing viral load results at the end-point were similar to those of the study population (Table 1b). This implies that the 7160 eligible patients in the main analysis, with a documented viral load result, are representative of the total study population. The majority of patients in the sensitivity analysis were initiated in WHO stage I (32.3%) or III (25.6%). The MCV at baseline was normal in 67.8% of all patients. A median CD4 count of 101 cells/mm^3 (IQR: 39 – 172) was obtained for all eligible patients at baseline. A median BMI of 22 kg/m^2 (IQR: 19 – 25) was obtained for all eligible patients at baseline. The median haemoglobin level was 11.7 g/dl (IQR: 10.1 – 13.1) at the time of initiation. The table also shows p-values for the association between different variables and poor adherence at 6 months measured by viral load at 6 months. Except where otherwise stated, the chi-square test was used.

Table 5c: 6 month clinical characteristics and their associations with poor adherence for all patients including those excluded due to missing viral load results assuming that those with missing viral load results had a viral load ≥ 400 copies/ml

Characteristics	All patients				All patients on AZT-based regimens*			
	Total N = 14384	Viral load < 400 copies/ml N = 5806	Viral load ≥ 400 copies/ml N = 8578	P-Value	Total N = 370	Viral load < 400 copies/ml N = 143	Viral load ≥ 400 copies/ml N = 227	P-Value
	n (%)	n (%)	n (%)		n (%)	n (%)	n (%)	
Self-reported adherenceδ								
> 90%	6491 (45.1%)	2106 (36.3%)	4385 (51.1%)	0.394	147 (39.7%)	38 (26.6%)	109 (48.0%)	0.545ð
> 60 – 90%	738 (5.1%)	227 (3.9%)	511 (6.0%)		12 (3.2%)	5 (3.5%)	7 (3.1%)	
≤ 60%	171 (1.2%)	49 (0.8%)	122 (1.4%)		6 (1.6%)	1 (0.7%)	5 (2.2%)	
Missing:	6984 (48.6%)	3424 (59.0%)	3560 (41.5%)		205 (55.5%)	99 (69.2%)	106 (46.7%)	
Number of missed visits								
0	9697 (67.4%)	3731 (64.3%)	5966 (69.6%)	0.000	250 (67.6%)	88 (61.5%)	162 (71.4%)	0.064
1	3363 (23.4%)	1475 (25.4%)	1888 (22.0%)		78 (21.1%)	39 (27.3%)	39 (17.2%)	
≥ 2	989 (6.9%)	509 (8.8%)	480 (5.6%)		26 (7.0%)	10 (7.0%)	16 (7.0%)	
Missing	335 (2.3%)	91 (1.5%)	244 (2.8%)		16 (4.3%)	6 (4.2%)	10 (4.4%)	
Number of missed ARV visits								
0	10950 (76.1%)	4257 (73.3%)	6693 (78.0%)	0.000	273 (73.8%)	99 (69.2%)	174 (76.7%)	0.191
1	2622 (18.2%)	1219 (21.0%)	1403 (16.4%)		66 (17.8%)	30 (21.0%)	36 (15.9%)	
≥ 2	450 (3.1%)	233 (4.0%)	217 (2.5%)		15 (4.1%)	8 (5.6%)	7 (3.0%)	
Missing	362 (2.6%)	97 (1.7%)	265 (3.1%)		16 (4.3%)	6 (4.2%)	10 (4.4%)	
Number of missed medical visits								
0	11002 (76.5%)	4360 (75.1%)	6642 (77.4%)	0.000	281 (75.9%)	106 (74.1%)	175 (77.1%)	0.472ð
1	2580 (17.9%)	1139 (19.6%)	1441 (16.8%)		65 (17.6%)	29 (20.3%)	36 (15.9%)	
≥ 2	440 (3.0%)	210 (3.6%)	230 (2.7%)		8 (2.2%)	2 (1.4%)	6 (2.6%)	
Missing	362 (2.6%)	97 (1.7%)	265 (3.1%)		16 (4.3%)	6 (4.2%)	10 (4.4%)	
BMI at 6 months in kg/m^2								

Median (IQR)	23 (21 – 27)	23.7 (21 – 27)	20.8 (19 – 23)	0.0000[þ]	24 (21 – 28.0)	24 (21.7 – 28.2)	24 (20.7 – 28)	0.564[M]
BMI at 6 months in kg/m^2								
< 18.5	690 (4.8%)	351 (6.0%)	339 (4.0%)	0.001	21 (5.7%)	10 (7.0%)	11 (4.8%)	0.546
18.5 – < 25	4755 (33.1%)	2683 (46.2%)	2072 (24.2%)		102 (27.6%)	55 (38.5%)	47 (20.7%)	
25 – < 30	2341 (16.3%)	1375 (23.7%)	966 (11.3%)		54 (14.6%)	33 (23.1%)	21 (9.3%)	
≥ 30	1174 (8.2%)	695 (11.9%)	479 (5.6%)		42 (11.4%)	20 (14.0%)	22 (9.7%)	
Missing	5424 (37.6%)	702 (12.2%)	4722 (54.9%)		151 (40.7%)	25 (17.4%)	126 (55.5%)	
Haemoglobin at 6 months in g/dl								
Median (IQR) / Mean ± SD	13.2 (12 – 14.4)	13.3 (12.1 – 14.5)	13.1 (11.9 – 14.3)	0.0000[þ]	13.2 ± 1.8	13.4 ± 1.7	12.7 ± 2.0	0.0613[ł]
Haemoglobin at 6 months								
≤ 8 g/dl	66 (0.5%)	37 (0.6%)	29 (0.3%)	0.000	1 (0.3%)	0 (0%)	1 (0.4%)	0.243[ð]
> 8g/dl	5694 (39.6%)	4298 (74.0%)	1396 (16.3%)		135 (36.5%)	103 (72.0%)	32 (14.1%)	
Missing	8624 (59.9%)	1471 (25.4%)	7153 (83.4%)		234 (63.2%)	40 (28.0%)	194 (85.5%)	
MCV at 6 months in fL								
Median (IQR)	99.5 (92.9 – 106.5)	100.6 (94 – 107.5)	95.3 (90.5 – 102)	0.0000 [þ]	107.7 (102.5 – 113.5)	109.2 (103.3 – 114.2)	103.4 (96.4 – 112.6)	0.0043 [þ]
MCV at 6 months								
< 80fL	227 (1.6%)	177 (3.0%)	50 (0.6%)	0.000	2 (0.5%)	1 (0.7%)	1 (0.4%)	0.000[ð]
80 – 100fL	3030 (21.1%)	2203 (37.9%)	827 (9.6%)		24 (6.5%)	12 (8.4%)	12 (5.3%)	
> 100fL	3004 (20.9%)	2624 (45.2%)	380 (4.4%)		131 (35.4%)	113 (79.0%)	18 (7.9%)	
Missing	8123 (56.4%)	802 (13.9%)	7321 (85.4%)		213 (57.6%)	17 (11.9%)	196 (86.4%)	
Change in MCV at 6 months in fL								
Median (IQR)	12.3 (5.7 – 18.4)	13.6 (6.8 – 19)	7.1 (2.7 – 12.6)	0.0000[þ]	15.4 (4.6 – 20.6)	16.1 (5.9 – 22)	8.9 (3.7 – 17.2)	0.0600[þ]
Change in MCV at 6 months								
≥ 14.5fL	1952 (13.6%)	1785 (30.7%)	167 (1.9%)	0.000	56 (15.1%)	49 (34.3%)	7 (3.1%)	0.163
< 14.5fL	2737 (19.0%)	2073 (35.7%)	664 (7.7%)		53 (14.3%)	41 (28.7%)	12 (5.3%)	
Missing	9695 (67.4%)	1948 (33.6%)	7747 (90.4%)		261 (70.6%)	53 (37.0%)	208 (91.6%)	
Change in CD4 count at 6 months in cells/mm^3								
Median (IQR)	128 (69 – 204)	135 (76 – 208)	96 (35.5 – 177)	0.0000[þ]	95 (45 – 140)	102 (54 – 145)	52.5 (-9 – 89)	0.0050[þ]

Change in CD4 count at 6 months in cells/mm^3								
≥ expected	4463 (31.0%)	3738 (64.4%)	725 (8.5%)	0.000	86 (23.2%)	79 (55.2%)	7 (3.1%)	0.002
< expected	1191 (8.3%)	856 (14.7%)	335 (3.9%)		37 (10%)	26 (18.2%)	11 (4.8%)	
Missing	8730 (60.7%)	1212 (20.9%)	7518 (87.6%)		247 (66.8%)	38 (26.6%)	209 (92.1%)	
Change in CD4 count at 6 months stratified by change in MCV at 6 months								
-Change in CD4 count ≥ expected and change in MCV ≥ 14.5fL:	1428 (9.9%)	1331 (22.9%)	97 (1.1%)	0.003	37 (10%)	35 (24.5%)	2 (0.9%)	0.022$^{\delta}$
-Change in CD4 count ≥ expected and change in MCV < 14.5fL:	1906 (13.3%)	1553 (26.7%)	353 (4.1%)		29 (7.8%)	26 (18.2%)	3 (1.3%)	
-Change in CD4 count < expected and change in MCV ≥ 14.5fL:	362 (2.5%)	337 (5.8%)	25 (0.3%)		11 (3.0%)	9 (6.3%)	2 (0.9%)	
-Change in CD4 count < expected and change in MCV < 14.5fL:	555 (3.9%)	349 (6.0%)	206 (2.4%)		20 (5.4%)	13 (9.1%)	7 (3.1%)	
-Missing:	10133 (70.4%)	2236 (38.5%)	7897 (92.1%)		273 (73.8%)	60 (41.9%)	213 (93.8%)	
Regimen change								
No	13092 (91.0%)	5210 (89.7%)	7882 (91.9%)	0.000	358 (96.8%)	140 (97.9%)	218 (96.8%)	0.383$^{\delta}$
Yes	1292 (9.0%)	596 (10.3%)	696 (8.1%)		12 (3.2%)	3 (2.1%)	9 (3.2%)	
TB$^{\alpha}$: On TB treatment during first 6 months on ART:								
Never	10743 (74.7%)	4345 (74.8%)	6398 (74.6%)	0.597	287 (77.6%)	120 (83.9%)	167 (73.6%)	0.041$^{\delta}$
At baseline	3027 (21.0%)	1204 (20.7%)	1823 (21.3%)		76 (20.5%)	20 (14.0%)	56 (24.7%)	
During follow-up	614 (4.3%)	257 (4.5%)	357 (4.1%)		7 (1.9%)	3 (2.1%)	4 (1.7%)	
Pregnancy$^{\alpha}$: Pregnancy during first 6 months on ART:								
Never	8664 (60.2%)	3534 (60.9%)	5130 (59.8%)	0.007	195 (52.7%)	80 (55.9%)	115 (50.7%)	0.490$^{\delta}$
At baseline	163 (1.1%)	67 (1.2%)	96 (1.1%)		3 (0.8%)	0 (0%)	3 (1.3%)	
During follow-up	224 (1.6%)	115 (1.9%)	109 (1.3%)		6 (1.6%)	2 (1.4%)	4 (1.7%)	
Missing$^{\gamma}$	5333 (37.1%)	2090 (36.0%)	3243 (37.8%)		166 (44.9%)	61 (42.7%)	105 (46.3%)	

Adverse events or side effects								
No	13986 (97.2%)	5618 (96.8%)	8368 (97.6%)	0.005	368 (99.5%)	226 (99.6%)	142 (99.3%)	1.000[ð]
Yes	398 (2.8%)	188 (3.2%)	210 (2.4%)		2 (0.5%)	1 (0.4%)	1 (0.7%)	

[*] On AZT-based regimens throughout the study period. [Þ] Wilcoxon rank sum test. [ð] Fisher's exact test. [M] Median test. [ŧ] t-test. [δ] The last self-reported adherence levels on, or after, the 3 month visit (70 days post-initiation) and not more than 3 months (110 days) before the date of the viral load but at least 28 days before the viral load was done. [α] Variable has baseline and 6 month components. [γ] Male patients. ART: Antiretroviral Therapy. ARV: Antiretroviral. AZT: Zidovudine. BMI: Body Mass Index. IQR: Interquartile range. MCV: Mean cell volume. SD: Standard deviation. TB: Tuberculosis.

The table above shows the clinical characteristics of all patients and patients on AZT-based regimens at 6 months. The characteristics were described using frequencies with percentages and medians with interquartile ranges. Missing values for a missed medical visit, missed ARV drug pickup or missing either a medical or ARV visit were the result of the patient having only one documented scheduled visit during the study period (excluding the ART initiation visit) i.e there was no possibility of the patient falling into the group with ≥ 2 missed visits. The table demonstrates that some of the 6 month clinical characteristics of all patients including those excluded due to missing viral load results at the end-point are similar to those of the study population (Table 1c). Like the main analysis, the sensitivity analysis also showed a large proportion (almost 50%) of missing values for self-reported adherence. There were, however, differences in 6 month characteristics for the variables haemoglobin (59.9% of values missing), BMI, MCV, change in MCV and CD4 count due to the larger percentages of missing values at 6 months. This implies that most of the patients excluded from the main analysis due to missing viral load results also lack the scheduled haemoglobin, MCV and/or CD4 count at 6 months. Only 21.1% of all patients had a normal MCV at 6 months and only 20.9% had a high MCV at 6 months. A median change in CD4 count of 128 cells/mm^3 (IQR: 69 – 204) was obtained for all patients at 6 months. 4.3% of females were pregnant during the study period. About a quarter of the patients were on TB treatment during the study period but adverse events or side effects were rare. The table also shows p-values for the association between different variables and poor adherence at 6 months measured by viral load at 6 months. Except where otherwise stated, the chi-square test was used.

Table 5d: Baseline demographic characteristics and their associations with poor adherence for all patients including those excluded due to missing viral load results assuming that those with missing viral load results had a viral load < 400 copies/ml

Characteristics	All patients				All patients on AZT-based regimens*			
	Total N = 14384	Viral load < 400 copies/ml N = 13030	Viral load ≥ 400 copies/ml N = 1354	P-value	Total N = 370	Viral load < 400 copies/ml N = 336	Viral load ≥ 400 copies/ml N = 34	P-value
	n (%)	n (%)	n (%)		n (%)	n (%)	n (%)	
Gender								
Females	9051 (62.9%)	8244 (63.3%)	807 (59.6%)	0.008	204 (55.1%)	189 (56.3%)	15 (44.1%)	0.175
Males	5333 (37.1%)	4786 (36.7%)	547 (40.4%)		166 (44.9%)	147 (43.7%)	19 (55.9%)	
Age at initiation in years								
Median (IQR)	36.6 (31.4 – 43.2)	36.7 (31.4 – 43.3)	36.3 (30.7 – 42.9)	0.0675Þ	39.6 (33.6 – 45.7)	39.7 (33.9 – 45.8)	36.3 (30.7 – 44.9)	0.1222Þ
Age at initiation in years								
≤ 36.7	7263 (50.5%)	6561 (50.4%)	702 (51.8%)	0.296	134 (36.2%)	117 (34.8%)	17 (50%)	0.079
> 36.7	7121 (49.5%)	6469 (49.6%)	652 (48.2%)		236 (63.8%)	219 (65.2%)	17(50%)	
Education level								
≥ Secondary school	8740 (60.8%)	7831 (60.1%)	909 (67.1%)	0.143	197 (53.2%)	179 (53.3%)	18 (52.9%)	0.921
Primary school	1812 (12.6%)	1646 (12.6%)	166 (12.3%)		57 (15.4%)	51 (15.2%)	6 (17.6%)	
< Primary school	426 (3.0%)	390 (3.0%)	36 (2.7%)		12 (3.2%)	11 (3.3%)	1 (2.9%)	
Missing:	3406 (23.6%)	3163 (24.3%)	243 (17.9%)		104 (28.2%)	95 (28.2%)	9 (26.6%)	
Professional status								
Employed	7191 (50.0%)	6526 (50.1%)	665 (49.1%)	0.615	180 (48.7%)	162 (48.2%)	18 (52.9%)	0.539
Unemployed	6973 (48.5%)	6311 (48.4%)	662 (48.9%)		184 (49.7%)	169 (50.3%)	15 (44.1%)	
Missing	220 (1.5%)	193 (1.5%)	27 (2.0%)		6 (1.6%)	5 (1.5%)	1 (3.0%)	

* On AZT-based regimens throughout the study period. Þ Wilcoxon rank sum test. AZT: Zidovudine. IQR: Interquartile range.

The baseline demographic characteristics in the table above are the same as those shown in the table assuming that those with missing results had a viral load ≥ 400 copies/ml (Table 5a). They differ only in the distribution of patients in each variable by whether the viral load was < or ≥ 400 copies/ml. The table also shows p-values for the association between different variables and poor adherence at 6 months measured by viral load at 6 months. Except where otherwise stated, the chi-square test was used.

Table 5e: Baseline clinical characteristics and their associations with poor adherence for all patients including those excluded due to missing viral load results assuming that those with missing viral load results had a viral load < 400 copies/ml

Characteristics	All patients				All patients on AZT-based regimens*			
	Total N = 14384	Viral load < 400 copies/ml N = 13030	Viral load ≥ 400 copies/ml N = 1354	P-value	Total N = 370	Viral load < 400 copies/ml N = 336	Viral load ≥ 400 copies/ml N = 34	P-value
	n (%)	n (%)	n (%)		n (%)	n (%)	n (%)	
BMI at initiation in kg/m^2								
Median (IQR)	22 (19 – 25)	22 (19 – 25)	22 (19 – 25)	0.846[M]	22.0 (19.3 – 28.0)	22.0 (19.3 – 28.0)	23 (20 – 24.0)	0.6742[Þ]
BMI at initiation in kg/m^2								
< 18.5	2179 (15.1%)	1972 (15.1%)	207 (15.3%)	0.927	45 (12.2%)	43 (12.8%)	2 (5.9%)	0.062
18.5 – < 25	6693 (46.5%)	6082 (46.7%)	611 (45.1%)		129 (34.9%)	112 (33.3%)	17 (50%)	
25 – < 30	2290 (15.9%)	2080 (16.0%)	210 (15.5%)		40 (10.8%)	36 (10.7%)	4 (11.8%)	
≥ 30	1089 (7.6%)	993 (7.6%)	96 (7.1%)		52 (14.1%)	51 (15.2%)	1 (2.9%)	
Missing	2133 (14.9%)	1903 (14.6%)	230 (17.0%)		104 (28.0%)	94 (28.0%)	10 (29.4%)	
Baseline haemoglobin in g/dl								
Median (IQR) / mean ± SD	11.7 (10.1 – 13.1)	11.7 (10.1 – 13.1)	11.5 (10.1 – 13)	0.1364[Þ]	12.6 ± 1.8	12.5 ± 1.8	13.3 ± 1.4	0.0402[ł]
Baseline haemoglobin								
≤ 8 g/dl	688 (4.8%)	625 (4.8%)	63 (4.7%)	0.971	1 (0.3%)	1 (0.3%)	0 (0%)	1.000[ð]
> 8g/dl	12298 (85.5%)	11177 (85.8%)	1121 (82.8%)		285 (77.0%)	261 (77.7%)	24 (70.6%)	
Missing	1398 (9.7%)	1228 (9.4%)	170 (12.5%)		84 (22.7%)	74 (22.0%)	10 (29.4%)	
MCV at baseline in fL								
Median (IQR)	88.2 (83.7 – 92.2)	88.2 (83.6 – 92.2)	88.8 (84.3 – 92.6)	0.0064[Þ]	90.8 (85.7 – 97.4)	91.0 (85.7 – 97.5)	89.5 (82.9 – 97.4)	0.5637[Þ]
MCV at baseline								
< 80fL	1427 (9.9%)	1336 (10.3%)	91 (6.7%)	0.025	26 (7.0%)	22 (6.5%)	4 (11.8%)	0.358
80 – 100fL	9757 (67.8%)	8973 (68.9%)	784 (57.9%)		188 (50.8%)	174 (51.8%)	14 (41.2%)	
> 100fL	475 (3.3%)	428 (3.3%)	47 (3.5%)		55 (14.9%)	50 (14.9%)	5 (14.7%)	
Missing	2725 (19.0%)	2293 (17.5%)	432 (31.9%)		101 (27.3%)	90 (26.8%)	11 (32.3%)	
Baseline CD4 count in cells/mm^3								
Median (IQR)	101 (39 – 172)	102 (40 – 172)	90 (34 – 170)	0.0071[Þ]	128 (61 – 218)	126.5 (60 – 211.5)	169 (61 – 249)	0.4026[Þ]

Baseline CD4 in cells/mm^3								
> 200	1858 (12.9%)	1683 (12.9%)	175 (12.9%)	0.002	87 (23.5%)	75 (22.3%)	12 (35.3%)	0.539[ð]
101 – 200	4443 (30.9%)	4092 (31.4%)	351 (25.9%)		81 (21.9%)	75 (22.3%)	6 (17.6%)	
51 – 100	2489 (17.3%)	2258 (17.3%)	231 (17.1%)		51 (13.8%)	47 (14.0%)	4 (11.8%)	
≤ 50	3762 (26.2%)	3374 (25.9%)	388 (28.7%)		60 (16.2%)	55 (16.4%)	5 (14.7%)	
Missing	1832 (12.7%)	1623 (12.5%)	209 (15.4%)		91 (24.6%)	84 (25%)	7 (20.6%)	
WHO stage at baseline								
I	4650 (32.3%)	4315 (33.1%)	335 (24.7%)	0.000	133 (35.9%)	119 (35.4%)	14 (41.2%)	0.397
II	2376 (16.5%)	2161 (16.6%)	215 (15.9%)		27 (7.3%)	25 (7.4%)	2 (5.9%)	
III	3685 (25.6%)	3350 (25.7%)	335 (24.7%)		70 (18.9%)	64 (19.0%)	6 (17.6%)	
IV	1337 (9.3%)	1193 (9.2%)	144 (10.6%)		27 (7.3%)	27 (8.0%)	0 (0%)	
Missing	2336 (16.3%)	2011 (15.4%)	325 (24.1%)		113 (30.6%)	101 (30.2%)	12 (35.3%)	

[*] On AZT-based regimens throughout the study period. [M] Median test. [Þ] Wilcoxon rank sum test. [ŧ] t-test. [ð] Fisher's exact test. AZT: Zidovudine. BMI: Body Mass Index. IQR: Interquartile range. MCV: Mean cell volume. SD: Standard deviation. WHO: World Health Organization.

The baseline clinical characteristics in the table above are the same as those shown in the table assuming that those with missing results had a viral load ≥ 400 copies/ml (Table 5b). They differ only in the distribution of patients in each variable by whether the viral load was < or ≥ 400 copies/ml. The table also shows p-values for the association between different variables and poor adherence at 6 months measured by viral load at 6 months. Except where otherwise stated, the chi-square test was used.

Table 5f: 6 month clinical characteristics and their associations with poor adherence for all patients including those excluded due to missing viral load results assuming that those with missing viral load results had a viral load < 400 copies/ml

Characteristic	All patients				All patients on AZT-based regimens*			
	Total N = 14384	Viral load < 400 copies/ml N = 13030	Viral load ≥ 400 copies/ml N = 1354	P-value	Total N = 370	Viral load < 400 copies/ml N = 336	Viral load ≥ 400 copies/ml N = 34	P-value
	n (%)	n (%)	n (%)		n (%)	n (%)	n (%)	
Self-reported adherenceδ								
> 90%	6491 (45.1%)	5879 (45.1%)	612 (45.2%)	0.303	147 (39.7%)	140 (41.7%)	7 (20.6%)	0.255ð
> 60 – 90%	738 (5.1%)	665 (5.1%)	73 (5.4%)		12 (3.2%)	11 (3.3%)	1 (2.9%)	
≤ 60%	171 (1.2%)	149 (1.1%)	22 (1.6%)		6 (1.6%)	5 (1.5%)	1 (2.9%)	
Missing	6984 (48.6%)	6337 (48.7%)	647 (47.8%)		205 (55.5%)	180 (53.5%)	25 (73.6%)	
Number of missed visits								
0	9697 (67.4%)	8811 (67.6%)	886 (65.4%)	0.038	250 (67.6%)	231 (68.8%)	19 (55.9%)	0.020
1	3363 (23.4%)	3034 (23.3%)	329 (24.3%)		78 (21.1%)	73 (21.7%)	5 (14.7%)	
≥ 2	989 (6.9%)	875 (6.7%)	114 (8.5%)		26 (7.0%)	20 (5.9%)	6 (17.6%)	
Missing	335 (2.3%)	310 (2.4%)	25 (1.8%)		16 (4.3%)	12 (3.6%)	4 (11.8%)	
Number of missed ARV visits								
0	10950 (76.1%)	9969 (76.5%)	981 (72.5%)	0.000	273 (73.8%)	252 (75%)	21 (61.8%)	0.373ð
1	2622 (18.2%)	2344 (18.0%)	278 (20.5%)		66 (17.8%)	59 (17.6%)	7 (20.6%)	
≥ 2	450 (3.1%)	381 (2.9%)	69 (5.1%)		15 (4.1%)	13 (3.8%)	2 (5.8%)	
Missing	362 (2.6%)	336 (2.6%)	26 (1.9%)		16 (4.3%)	12 (3.6%)	4 (11.8%)	
Number of missed medical visits								
0	11002 (76.5%)	10009 (76.8%)	993 (73.3%)	0.000	281 (75.9%)	262 (78.0%)	19 (55.9%)	0.010ð
1	2580 (17.9%)	2308 (17.7%)	272 (20.1%)		65 (17.6%)	57 (17.0%)	8 (23.5%)	
≥ 2	440 (3.0%)	377 (2.9%)	63 (4.7%)		8 (2.2%)	5 (1.4%)	3 (8.8%)	
Missing	362 (2.6%)	336 (2.6%)	26 (1.9%)		16 (4.3%)	12 (3.6%)	4 (11.8%)	

BMI at 6 months in kg/m^2 Median (IQR)	23 (21 – 27)	23.2 (21 – 27)	23 (20.7 – 26)	0.0028[þ]	24 (21 – 28.0)	24 (21 – 29)	23 (20.7 – 25)	0.2020 [þ]
BMI at 6 months in kg/m^2 < 18.5 18.5 – < 25 25 – < 30 ≥ 30 Missing	690 (4.8%) 4755 (33.1%) 2341 (16.3%) 1174 (8.2%) 5424 (37.6%)	578 (4.4%) 4093 (31.4%) 2035 (15.6%) 1032 (7.9%) 5292 (40.7%)	112 (8.3%) 662 (48.9%) 306 (22.6%) 142 (10.5%) 132 (9.7%)	0.064	21 (5.7%) 102 (27.6%) 54 (14.6%) 42 (11.4%) 151 (40.7%)	19 (5.7%) 85 (25.3%) 48 (14.3%) 40 (11.9%) 144 (42.8%)	2 (5.9%) 17 (50%) 6 (17.6%) 2 (5.9%) 7 (20.6%)	0.267[ð]
Haemoglobin at 6 months in g/dl Median (IQR)	13.2 (12 – 14.4)	13.3 (12 – 14.5)	13.1 (12 – 14.3)	0.0539[þ]	13.1 (12.1 – 14.6)	12.9 (12 – 14.5)	13.7 (12.5 – 14.7)	0.4037[þ]
Haemoglobin at 6 months ≤ 8 g/dl > 8g/dl Missing	66 (0.5%) 5694 (39.6%) 8624 (59.9%)	53 (0.4%) 4802 (36.9%) 8175 (62.7%)	13 (1.0%) 892 (65.9%) 449 (33.2%)	0.371	1 (0.3%) 135 (36.5%) 234 (63.2%)	0 (0%) 118 (35.1%) 218 (64.9%)	1 (2.9%) 17 (50%) 16 (47.1%)	0.132[ð]
MCV at 6 months in fL Median (IQR)	99.5 (92.9 – 106.5)	100.5 (93.8 – 107.4)	95.2 (90.4 – 101.2)	0.0000 [þ]	107.7 (102.5 – 113.5)	109.3 (103.4 – 114.2)	100.2 (92.3 – 108.3)	0.0004 [þ]
MCV at 6 months < 80fL: 80 – 100fL: > 100fL: Missing:	227 (1.6%) 3030 (21.1%) 3004 (20.9%) 8123 (56.4%)	186 (1.4%) 2337 (17.9%) 2710 (20.8%) 7797 (59.9%)	41 (3.0%) 693 (51.2%) 294 (21.7%) 326 (24.1%)	0.000	2 (0.5%) 24 (6.5%) 131 (35.4%) 213 (57.6%)	1 (0.3%) 13 (3.9%) 119 (35.4%) 203 (60.4%)	1 (2.9%) 11 (32.4%) 12 (35.3%) 10 (29.4%)	0.000[ð]
Change in MCV at 6 months in fL Median (IQR)/ mean ± SD	12.3 (5.7 – 18.4)	13.5 (6.6 – 18.9)	6.8 (2.6 – 11.8)	0.0000 [þ]	13.5 ± 10.0	14.3 ± 10.0	8.0 ± 9.1	0.0216[ŧ]
Change in MCV at 6 months ≥ 14.5fL < 14.5fL Missing	1952 (13.6%) 2737 (19.0%) 9695 (67.4%)	1835 (14.1%) 2184 (16.8%) 9011 (69.1%)	117 (8.6%) 553 (40.8%) 684 (50.6%)	0.000	56 (15.1%) 53 (14.3%) 261 (70.6%)	52 (15.5%) 42 (12.5%) 242 (72.0%)	4 (11.8%) 11 (32.4%) 19 (55.8%)	0.052[ð]
Change in CD4 count at 6 months in cells/mm^3 Median (IQR)	128 (69 – 204)	133 (75 – 207)	97 (33 – 180)	0.0000[þ]	95 (45 – 140)	100.5 (54 – 145)	52 (-9 – 61)	0.0043[þ]

Change in CD4 count at 6 months in cells/mm^3								
≥ expected	4463 (31.0%)	3853 (29.6%)	610 (45.1%)	0.000	86 (23.2%)	80 (23.8%)	6 (17.6%)	0.001
< expected	1191 (8.3%)	891 (6.8%)	300 (22.2%)		37 (10%)	26 (7.7%)	11 (32.4%)	
Missing	8730 (60.7%)	8286 (63.6%)	444 (32.7%)		247 (66.8%)	230 (68.5%)	17 (50%)	
Change in CD4 count at 6 months stratified by change in MCV at 6 months								
-Change in CD4 count ≥ expected and change in MCV ≥ 14.5fL:	1428 (9.9%)	1347 (10.3%)	81 (6.0%)	0.000	37 (10%)	35 (10.4%)	2 (5.9%)	0.022^{ð}
-Change in CD4 count ≥ expected and change in MCV < 14.5fL:	1906 (13.3%)	1596 (12.2%)	310 (22.9%)		29 (7.8%)	26 (7.7%)	3 (8.8%)	
-Change in CD4 count < expected and change in MCV ≥ 14.5fL:	362 (2.5%)	341 (2.6%)	21 (1.6%)		11 (3.0%)	9 (2.7%)	2 (5.9%)	
-Change in CD4 count < expected and change in MCV < 14.5fL:	555 (3.9%)	367 (2.8%)	188 (13.9%)		20 (5.4%)	13 (3.9%)	7 (20.5%)	
-Missing:	10133 (70.4%)	9379 (72.1%)	754 (55.6%)		273 (73.8%)	253 (75.3%)	20 (58.9%)	
Regimen change								
No	13092 (91.0%)	11857 (91.0%)	1235 (91.2%)	0.794	358 (96.8%)	327 (91.3%)	31 (91.2%)	0.088^{ð}
Yes	1292 (9.0%)	1173 (9.0%)	119 (8.8%)		12 (3.2%)	9 (2.7%)	3 (8.8%)	
TB^{α}: On TB treatment during first 6 months on ART:								
Never	10743 (74.7%)	9733 (74.7%)	1010 (74.6%)	0.898	287 (77.6%)	257 (76.5%)	30 (88.2%)	0.321^{ð}
At baseline	3027 (21.0%)	2738 (21.0%)	289 (21.3%)		76 (20.5%)	72 (21.4%)	4 (11.8%)	
During follow-up	614 (4.3%)	559 (4.3%)	55 (4.1%)		7 (1.9%)	7 (2.1%)	0 (0%)	
Pregnancy^{α}: Pregnancy during first 6 months on ART:								
Never	8664 (60.2%)	7904 (60.7%)	760 (56.1%)	0.017	195 (52.7%)	183 (54.5%)	12 (35.3%)	0.011^{ð}
At baseline	163 (1.1%)	148 (1.1%)	15 (1.1%)		3 (0.8%)	3 (0.9%)	0 (0%)	
During follow-up	224 (1.6%)	192 (1.5%)	32 (2.4%)		6 (1.6%)	3 (0.9%)	3 (8.8%)	
Missing^{γ}	5333 (37.1%)	4786 (36.7%)	547 (40.4%)		166 (44.9%)	147 (43.7%)	19 (55.9%)	

Adverse events or side effects								
No	13986 (97.2%)	12649 (97.1%)	1337 (98.7%)	0.000	368 (99.5%)	334 (99.4%)	34 (100%)	1.000[ð]
Yes	398 (2.8%)	381 (2.9%)	17 (1.3%)		2 (0.5%)	2 (0.6%)	0 (0%)	

[*] On AZT-based regimens throughout the study period. [þ] Wilcoxon rank sum test. [ţ] t-test. [ð] Fisher's exact test. [δ] The last self-reported adherence levels on, or after, the 3 month visit (70 days post-initiation) and not more than 3 months (110 days) before the date of the viral load but at least 28 days before the viral load was done. [α] Variable has baseline and 6 month components. [γ] Male patients.
ART: Antiretroviral Therapy. ARV: Antiretroviral. AZT: Zidovudine. BMI: Body Mass Index. IQR: Interquartile range. MCV: Mean cell volume. SD: Standard deviation. TB: Tuberculosis.

The clinical characteristics of patients at 6 months, shown in the table above, are the same as those in the table assuming that those with missing results had a viral load ≥ 400 copies/ml (Table 5c). They differ only in the distribution of patients in each variable by whether the viral load was < or ≥ 400 copies/ml. The table also shows p-values for the association between different variables and poor adherence at 6 months measured by viral load at 6 months. Except where otherwise stated, the chi-square test was used.

Table 6: Table showing the point at which the ART adherence was reported and the points at which the MCV, CD4 count and viral load at 6 months were done for the sensitivity analyses

Variable	All patients including those excluded due to missing viral load results	All patients on AZT-based regimens including those excluded due to missing viral load results*
	Median (IQR)	**Median (IQR)**
Self-reported adherence	51 days (IQR: 42 – 75) before the viral load was done.	44 days (IQR: 40 – 71) before the viral load was done.
MCV at 6 months	198 days (IQR: 169 – 266) after ART initiation.	203 days (IQR: 173 – 260) after ART initiation.
CD4 count at 6 months	203 days (IQR: 170 – 270) after ART initiation.	207.5 days (IQR: 175 – 264) after ART initiation.
Viral load at 6 months	183 days (IQR: 183 – 203) after ART initiation.	183 days (IQR: 183 – 203) after ART initiation.

* On AZT-based regimens throughout the study period. AZT: Zidovudine. IQR: Interquartile range. MCV: Mean cell volume.

The table above shows that the MCV and CD4 count at 6 months for all patients and for patients on AZT were done at approximately the same time as in the main analysis (Table 2). The point at which the viral load was done was different from that for all eligible patients as the end-point was set at 183 days for all patients with a missing viral load. This is 20 days before the median end-point for eligible patients with a documented viral load result (Table 2). This may have compounded the problem of the sensitivity analyses having larger percentages of missing values for BMI, haemoglobin, MCV, change in MCV and CD4 count at 6 months than the main analysis.

3.4.2 Poisson regression with robust error variance for the sensitivity analysis

In this section, univariate and multivariate Poisson regression models (with robust error variance) for all patients including those without a documented viral load result at 6 months are presented.

Table 7a: Univariate and multivariate Poisson regression models with robust error variance assuming that those with missing viral load results had a viral load ≥ 400 copies/ml

Variable	All patients				All patients on AZT-based regimens*			
	Univariate Analyses		Multivariate Analysis		Univariate Analyses		Multivariate Analysis	
	Crude IRR (95% CI)	P - value	Adjusted IRR (95% CI)	P - value	Crude IRR (95% CI)	P - value	Adjusted IRR (95% CI)	P - value
Gender								
Females	1	-	1	-	1	-	-	-
Males	1.03 (1.00 – 1.06)	0.027	1.10 (0.96 – 1.26)	0.170	1.06 (0.90 – 1.24)	0.497	-	-
Age at initiation								
≤ 36.7 years	1	-	1	-	1	-	1	-
> 36.7 years	0.99 (0.96 – 1.01)	0.297	0.92 (0.80 – 1.05)	0.212	0.85 (0.72 – 1.00)	0.044	1.40 (0.49 – 3.98)	0.528
Education level								
≥ Secondary school	1	-	-	-	1	-	-	-
Primary school	1.00 (0.96 – 1.04)	0.888	-	-	1.11 (0.90 – 1.37)	0.311	-	-
< Primary school	0.96 (0.88 – 1.05)	0.356	-	-	0.81 (0.46 – 1.45)	0.485	-	-
Professional status								
Employed	1	-	-	-	1	-	-	-
Unemployed	1.06 (1.03 – 1.09)	0.000	-	-	1.02 (0.87 – 1.21)	0.784	-	-
BMI at initiation in kg/m^2								
< 18.5	1	-	-	-	1	-	-	-
18.5 – < 25	0.93 (0.89 – 0.96)	0.000	-	-	0.97 (0.74 – 1.27)	0.834	-	-
25 – < 30	0.90 (0.86 – 0.95)	0.000	-	-	0.96 (0.69 – 1.36)	0.834	-	-
≥ 30	0.90 (0.85 – 0.95)	0.000	-	-	0.90 (0.64 – 1.25)	0.519	-	-
Baseline haemoglobin								
≤ 8 g/dl	1	-	-	-	1	-	-	-

> 8g/dl	1.00 (0.94 – 1.06)	0.929	-	-	-‡	-	-	-
MCV at baseline								
< 80fL	1	-	1	-	1	-	-	-
80 – 100fL	1.01 (0.97 – 1.06)	0.661	1.28 (1.02 – 1.62)	0.037	1.04 (0.77 – 1.40)	0.789	-	-
> 100fL	0.95 (0.87 – 1.04)	0.278	0.89 (0.60 – 1.33)	0.578	0.61 (0.40 – 0.94)	0.025	-	-
Baseline CD4 count in cells/mm^3								
> 200	1	-	1	-	1	-	1	-
101 – 200	1.09 (1.04 – 1.14)	0.000	1.13 (0.91 – 1.40)	0.260	1.19 (0.92 – 1.54)	0.188	0.64 (0.07 – 5.88)	0.694
51 – 100	1.09 (1.04 – 1.15)	0.001	1.29 (1.02 – 1.63)	0.035	1.19 (0.89 – 1.59)	0.248	2.44 (0.50 – 11.88)	0.270
≤ 50	1.09 (1.04 – 1.15)	0.000	1.45 (1.17 – 1.79)	0.001	1.26 (0.96 – 1.65)	0.090	3.32 (0.44 – 24.77)	0.242
WHO stage at baseline								
I	1	-	-	-	1	-	-	-
II	1.01 (0.97 – 1.05)	0.655	-	-	0.96 (0.64 – 1.42)	0.831	-	-
III	1.02 (0.98 – 1.05)	0.336	-	-	1.35 (1.09 – 1.66)	0.006	-	-
IV	0.94 (0.89 – 0.99)	0.024	-	-	1.09 (0.77 – 1.55)	0.613	-	-
Self-reported adherence								
> 90%	1	-	-	-	1	-	-	-
> 60 – 90%	1.02 (0.97 – 1.08)	0.343	-	-	0.79 (0.48 – 1.28)	0.336	-	-
≤ 60%	1.06 (0.96 – 1.16)	0.267	-	-	1.12 (0.78 – 1.63)	0.538	-	-
Number of missed visits								
0	1	-	-	-	1	-	-	-
1	0.91 (0.88 – 0.94)	0.000	-	-	0.77 (0.61 – 0.98)	0.034	-	-
≥ 2	0.79 (0.74 – 0.84)	0.000	-	-	0.95 (0.69 – 1.30)	0.750	-	-
Number of missed ARV visits								
0	1	-	-	-	1	-	-	-
1	0.88 (0.84 – 0.91)	0.000	-	-	0.86 (0.67 – 1.09)	0.200	-	-
≥ 2	0.79 (0.72 – 0.87)	0.000	-	-	0.73 (0.42 – 1.27)	0.266	-	-
Number of missed medical visits								
0	1	-	-	-	1	-	-	-
1	0.93 (0.89 – 0.96)	0.000	-	-	0.89 (0.70 – 1.13)	0.331	-	-
≥ 2	0.87 (0.79 – 0.95)	0.002	-	-	1.20 (0.80 – 1.82)	0.375	-	-

BMI at 6 months in kg/m^2								
< 18.5	1	-	-	-	1	-	-	-
18.5 – < 25	0.89 (0.82 – 0.96)	0.004	-	-	0.88 (0.56 – 1.39)	0.585	-	-
25 – < 30	0.84 (0.77 – 0.92)	0.000	-	-	0.74 (0.44 – 1.26)	0.269	-	-
≥ 30	0.83 (0.75 – 0.92)	0.000	-	-	1 (0.61 – 1.65)	1.000	-	-
Haemoglobin at 6 months								
≤ 8 g/dl	1	-	-	-	1	-	-	-
> 8g/dl	0.56 (0.42 – 0.74)	0.000	-	-	0.24 (0.17 – 0.32)	0.000	-	-
MCV at 6 months								
< 80fL	1	-	-	-	1	-	-	-
80 – 100fL	1.24 (0.96 – 1.59)	0.095	-	-	1 (0.24 – 4.25)	1.000	-	-
> 100fL	0.57 (0.44 – 0.75)	0.000	-	-	0.27 (0.06 – 1.18)	0.082	-	-
Change in MCV at 6 months								
≥14.5fL	1	-	-	-	1	-	-	-
<14.5fL	2.84 (2.42 – 3.33)	0.000	-	-	1.81 (0.77 – 4.27)	0.174	-	-
Change in CD4 count at 6 months in cells/mm^3								
≥ expected	1	-	-	-	1	-	1	-
< expected	1.73 (0.55 – 1.94)	0.000	-	-	3.65 (1.53 – 8.71)	0.003	6.62 (1.35 – 32.35)	0.020
Change in CD4 count at 6 months stratified by change in MCV at 6 months								
-Change in CD4 count ≥ expected and change in MCV ≥ 14.5:	1	-	1	-	1	-	-	-
-Change in CD4 count ≥ expected and change in MCV < 14.5:	2.73 (2.20 – 3.34)	0.000	2.81 (2.27 – 3.48)	0.000	1.91 (0.34 – 10.80)	0.462	-	-
-Change in CD4 count < expected and change in MCV ≥ 14.5:	1.02 (0.67 – 1.55)	0.939	1.08 (0.70 – 1.65)	0.731	3.36 (0.53 – 21.40)	0.199	-	-
-Change in CD4 count < expected and change in MCV < 14.5:	5.46 (4.38 – 6.81)	0.000	5.92 (4.74 – 7.39)	0.000	6.48 (1.47 – 28.50)	0.013	-	-
Regimen change								
No	1	-	-	-	1	-	-	-
Yes	0.89 (0.85 – 0.94)	0.000	0.79 (0.62 – 1.01)	0.058	1.23 (0.88 – 1.73)	0.226	-	-
TB: On TB treatment during first 6 months on ART:								
Never	1	-	-	-	1	-	-	-

At baseline	1.01 (0.98 – 1.05)	0.505	-	-	1.27 (1.07 – 1.50)	0.005	-	-
During follow-up	0.98 (0.91 – 1.05)	0.495	-	-	0.98 (0.51 – 1.88)	0.956	-	-
Pregnancy: Pregnancy during first 6 months on ART:								
Never	1	-	-	-	1	-	-	-
At baseline	0.99 (0.87 – 1.13)	0.936	-	-	1.70 (1.51 – 1.91)	0.000	-‡	-
During follow-up	0.82 (0.72 – 0.94)	0.005	-	-	1.13 (0.63 – 2.02)	0.678	7.36 (2.12 – 25.60)	0.002
Adverse events or side effects:								
No:	1	-	-	-	1	-	-	-
Yes:	0.88 (0.81 – 0.97)	0.009	-	-	0.81 (0.20 – 3.27)	0.772	-	-

* On AZT-based regimens throughout the study period. ‡ Cell has no observations. ART: Antiretroviral Therapy. ARV: Antiretroviral. AZT: Zidovudine. BMI: Body Mass Index. CI: Confidence interval. IRR: Incidence rate ratio. MCV: Mean cell volume. TB: Tuberculosis. WHO: World Health Organization.

For the table above, multivariate Poisson regression analysis with robust error variance was done first for all eligible patients then for all eligible patients on AZT-based regimens. After adjusting for potential confounders, the variables independently associated with poor adherence for all patients were change in CD4 count stratified by change in MCV at 6 months, CD4 count at baseline and MCV at baseline. The IRR of poor adherence increased with decreasing CD4 count at baseline. The association between regimen change and poor adherence was marginally significant ($p < 0.1$). Model fit was good as evidenced by a p-value of 0.389 on the link test. Change in CD4 count at 6 months and pregnancy during the first 6 months on ART were the only independent predictors of poor adherence in the final multivariate model for patients on AZT-based regimens. Model fit for this model was also good as evidenced by a p-value of 0.649 on the link test.

Table 7b: Univariate and multivariate regression models with robust error variance assuming that those with missing viral load results had a viral load < 400 copies/ml

Variable	All patients				All patients on AZT-based regimens*			
	Univariate Analyses		Multivariate Analysis		Univariate Analyses		Multivariate Analysis	
	Crude IRR (95% CI)	P - value	Adjusted IRR (95% CI)	P - value	Crude IRR (95% CI)	P - value	Adjusted IRR (95% CI)	P - value
Gender								
Females	1	-	1	-	1	-	-	-
Males	1.15 (1.04 – 1.28)	0.008	1.09 (0.93 – 1.29)	0.283	1.56 (0.82 – 2.97)	0.179	-	-
Age at initiation								
≤ 36.7 years	1	-	1	-	1	-	1	-
> 36.7 years	0.95 (0.86 – 1.05)	0.296	0.95 (0.81 – 1.12)	0.567	0.57 (0.30 – 1.08)	0.082	1.17 (0.30 – 4.62)	0.823
Education level								
≥ Secondary school	1	-	-	-	1	-	-	-
Primary school	0.88 (0.75 – 1.03)	0.114	-	-	1.15 (0.48 – 2.77)	0.752	-	-
< Primary school	0.81 (0.59 – 1.12)	0.202	-	-	0.91 (0.13 – 6.29)	0.926	-	-
Professional status								
Employed	1	-	-	-	1	-	-	-
Unemployed	1.03 (0.93 – 1.14)	0.615	-	-	0.82 (0.42 – 1.57)	0.541	-	-
BMI at initiation in kg/m^2								
< 18.5	1	-	-	-	1	-	-	-
18.5 – < 25	0.96 (0.83 – 1.12)	0.603	-	-	2.97 (0.71 – 12.37)	0.136	-	-
25 – < 30	0.97 (0.80 – 1.16)	0.705	-	-	2.25 (0.43 – 11.67)	0.334	-	-
≥ 30	0.93 (0.74 – 1.17)	0.526	-	-	0.43 (0.04 – 4.64)	0.489	-	-
Baseline haemoglobin								
≤ 8 g/dl	1	-	-	-	1	-	-	-
> 8g/dl	1.00 (0.78 – 1.27)	0.971	-	-	-‡	-	-	-
MCV at baseline								
< 80fL	1	-	-	-	1	-	-	-
80 – 100fL	1.26 (1.02 – 1.55)	0.031	-	-	0.48 (0.17 – 1.36)	0.169	-	-
> 100fL	1.55 (1.11 – 2.17)	0.010	-	-	0.59 (0.17 – 2.02)	0.402	-	-

Baseline CD4 in cells/mm^3								
> 200	1	-	1	-	1	-	1	-
101 – 200	0.84 (0.71 – 1.00)	0.047	1.11 (0.84 – 1.46)	0.460	0.54 (0.21 – 1.37)	0.192	0.83 (0.07 – 9.43)	0.879
51 – 100	0.99 (0.82 – 1.19)	0.877	1.14 (0.85 – 1.54)	0.386	0.57 (0.19 – 1.67)	0.305	3.15 (0.47 – 21.20)	0.237
≤ 50	1.10 (0.92 – 1.30)	0.294	1.41 (1.07 – 1.85)	0.014	0.60 (0.22 – 1.63)	0.319	2.04 (0.11 – 38.88)	0.636
WHO stage at baseline								
I	1	-	1	-	1	-	-	-
II	1.26 (1.07 – 1.48)	0.006	1.15 (0.90 – 1.48)	0.257	0.70 (0.17 – 2.93)	0.629	-	-
III	1.26 (1.09 – 1.46)	0.002	1.28 (1.04 – 1.57)	0.018	0.81 (0.33 – 2.03)	0.660	-	-
IV	1.49 (1.24 – 1.80)	0.000	1.47 (1.15 – 1.89)	0.002	-‡	-	-	-
Self-reported adherence								
> 90%	1	-	-	-	1	-	-	-
> 60 – 90%	1.05 (0.83 – 1.32)	0.683	-	-	1.75 (0.23 – 13.15)	0.587	-	-
≤ 60%	1.36 (0.92 – 2.03)	0.125	-	-	3.5 (0.51 – 24.25)	0.205	-	-
Number of missed visits								
0	1	-	-	-	1	-	-	-
1	1.07 (0.95 – 1.21)	0.266	-	-	0.84 (0.33 – 2.19)	0.726	-	-
≥ 2	1.26 (1.05 – 1.52)	0.013	-	-	3.04 (1.33 – 6.93)	0.008	-	-
Number of missed ARV visits								
0	1	-	-	-	1	-	-	-
1	1.18 (1.04 – 1.34)	0.009	-	-	1.38 (0.61 – 3.11)	0.439	-	-
≥ 2	1.71 (1.37 – 2.14)	0.000	-	-	1.73 (0.45 – 6.73)	0.427	-	-
Number of missed medical visits								
0	1	-	-	-	1	-	-	-
1	1.17 (1.03 – 1.33)	0.017	-	-	1.82 (0.83 – 3.98)	0.133	-	-
≥ 2	1.59 (1.25 – 2.01)	0.000	-	-	5.55 (2.05 – 15.01)	0.001	-	-
BMI at 6 months in kg/m^2								
< 18.5	1	-	-	-	1	-	-	-
18.5 – < 25	0.86 (0.71 – 1.03)	0.101	-	-	1.75 (0.44 – 7.03)	0.430	-	-
25 – < 30	0.81 (0.66 – 0.98)	0.033	-	-	1.17 (0.25 – 5.35)	0.843	-	-
≥ 30	0.75 (0.59 – 0.94)	0.012	-	-	0.50 (0.07 – 3.32)	0.473	-	-

Haemoglobin at 6 months								
≤ 8 g/dl	1	-	-	-	1	-	-	-
> 8g/dl	0.80 (0.49 – 1.30)	0.361	-	-	0.13 (0.08 – 0.20)	0.000	-	-
MCV at 6 months								
< 80fL	1	-	-	-	1	-	-	-
80 – 100fL	1.27 (0.95 – 1.68)	0.104	-	-	0.92 (0.21 – 3.94)	0.907	-	-
> 100fL	0.54 (0.40 – 0.73)	0.000	-	-	0.18 (0.04 – 0.81)	0.026	-	-
Change in MCV at 6 months								
≥ 14.5fL	1	-	-	-	1	-	-	-
< 14.5fL	3.37 (2.79 – 4.08)	0.000	-	-	2.91 (0.98 – 8.61)	0.054	-	-
Change in CD4 count at 6 months in cells/mm^3								
≥ expected	1	-	-	-	1	-	1	-
< expected	1.84 (1.63 – 2.08)	0.000	-	-	4.26 (1.70 – 10.70)	0.002	7.73 (0.96 – 62.40)	0.055
Change in CD4 count at 6 months stratified by change in MCV at 6 months								
-Change in CD4count ≥ expected and change in MCV ≥ 14.5:	1	-	1	-	1	-	-	-
-Change in CD4 count ≥ expected and change in MCV < 14.5:	2.87 (2.27 – 3.63)	0.000	3.04 (2.35 – 3.93)	0.000	1.91 (0.34 – 10.80)	0.462	-	-
-Change in CD4 count < expected and change in MCV ≥ 14.5:	1.02 (0.64 – 1.63)	0.925	1.24 (0.77 – 2.01)	0.379	3.36 (0.53 – 21.40)	0.199	-	-
-Change in CD4 count < expected and change in MCV < 14.5:	5.97 (4.69 – 7.60)	0.000	6.73 (5.15 – 8.78)	0.000	6.48 (1.47 – 28.50)	0.013	-	-
Regimen change								
No	1	-	-	-	1	-	-	-
Yes	0.98 (0.82 – 1.17)	0.794	-	-	2.89 (1.02 – 8.15)	0.045	-	-
TB: On TB treatment during first 6 months on ART:								
Never	1	-	-	-	1	-	-	-
At baseline	1.02 (0.90 – 1.15)	0.804	-	-	0.50 (0.18 – 1.39)	0.184	-	-
During follow-up	0.95 (0.74 – 1.23)	0.714	-	-	-‡	-	-	-

Pregnancy: Pregnancy during first 6 months on ART:								
Never	1	-	-	-	1	-	1	-
At baseline	1.05 (0.64 – 1.71)	0.847	-	-	-‡	-	-‡	-
During follow-up	1.63 (1.17 – 2.26)	0.004	-	-	8.12 (3.07 – 21.48)	0.000	9.20 (2.16 – 39.18)	0.003
Adverse events or side effects								
No	1	-	-	-	1	-	-	-
Yes	0.45 (0.28 – 0.71)	0.001	-	-	-‡	-	-	-

* On AZT-based regimens throughout the study period. ‡Cell has no observations. ART: Antiretroviral Therapy. ARV: Antiretroviral. AZT: Zidovudine. BMI: Body Mass Index. CI: Confidence interval. IRR: Incidence rate ratio. MCV: Mean cell volume. TB: Tuberculosis. WHO: World Health Organization.

For the table above, multivariate Poisson regression analysis with robust error variance was done first for all eligible patients then for all eligible patients on AZT-based regimens. After adjusting for potential confounders, the variables independently associated with poor adherence for all patients were change in CD4 count stratified by change in MCV at 6 months, CD4 count at baseline and WHO stage at baseline. The IRR of poor adherence increased with decreasing CD4 count and increasing WHO stage at baseline. Model fit was good as evidenced by a p-value of 0.778 on the link test. The change in CD4 count at 6 months and pregnancy during the first 6 months on ART were the only independent predictors of poor adherence in the final multivariate model for patients on AZT-based regimens. Model fit for this model was also good as evidenced by a p-value of 0.805 on the link test.

3.5 POISSON REGRESSION WITH ROBUST ERROR VARIANCE, USING A VIRAL LOAD DONE 90 – 270 DAYS AFTER ART INITIATION

In this section, univariate and multivariate Poisson regression models (with robust error variance) for all patients with a documented viral load done 90 – 270 days after ART initiation are presented.

Table 8a: Univariate and multivariate Poisson regression models with robust error variance for all eligible patients and eligible patients on AZT-based regimens with a documented viral load done 90 to 270 after ART initiation

Variable	All eligible patients				All eligible patients on AZT-based regimens*			
	Univariate Analyses		Multivariate Analysis		Univariate Analyses		Multivariate Analysis	
	Crude IRR (95% CI)	P - value	Adjusted IRR (95% CI)	P- value	Crude IRR (95% CI)	P - value	Adjusted IRR (95% CI)	P- value
Gender								
Females	1	-	1	-	1	-	1	-
Males	1.17 (1.07 – 1.28)	0.000	1.07 (0.90 – 1.28)	0.434	1.78 (0.97 – 3.27)	0.064	1.80 (0.53 – 6.11)	0.343
Age at initiation								
≤ 36.7 years	1	-	1	-	1	-	1	-
> 36.7 years	0.93 (0.85 – 1.02)	0.117	0.86 (0.71 – 1.02)	0.088	0.99 (0.53 – 5.82)	0.962	1.10 (0.30 – 4.00)	0.890
Education level								
≥ Secondary school	1	-	-	-	1	-	-	-
Primary school	0.98 (0.85 – 1.11)	0.741	-	-	1.62 (0.75 – 3.50)	0.220	-	-
< Primary school	0.77 (0.57 – 1.04)	0.088	-	-	0.71 (0.12 – 5.43)	0.812	-	-
Professional status								
Employed	1	-	-	-	1	-	-	-
Unemployed	1.08 (0.99 – 1.18)	0.092	-	-	1.07 (0.58 – 1.95)	0.839	-	-
BMI at initiation in kg/m^2								
< 18.5	1	-	-	-	1	-	-	-
18.5 – < 25	0.95 (0.84 – 1.08)	0.449	-	-	3.05 (0.73 – 12.50)	0.121	-	-
25 – < 30	0.85 (0.73 – 1.00)	0.056	-	-	1.17 (0.17 – 7.82)	0.874	-	-

≥ 30	0.82 (0.67 – 1.01)	0.061	-	-	0.46 (0.04 – 4.89)	0.520	-	-
Baseline haemoglobin								
≤ 8 g/dl	1	-	-	-	1	-	-	-
> 8g/dl	1.01 (0.82 – 1.24)	0.958	-	-	-‡	-	-	-
MCV at baseline[L]								
< 80fL	1	-	1	-	1	-	-	-
80 – 100fL	1.25 (1.05 – 1.49)	0.013	1.46 (1.07 – 1.99)	0.016	1.01 (0.43 – 2.39)	0.977	-	-
> 100fL	1.46 (1.09 – 2.00)	0.011	1.03 (0.61 – 1.75)	0.902	0.87 (0.19 – 3.94)	0.857	-	-
Baseline CD4 count in cells/mm^3								
> 200	1	-	1	-	1	-	1	-
101 – 200	1.02 (0.88 – 1.20)	0.757	1.12 (0.83 – 1.51)	0.442	0.91 (0.40 – 2.10)	0.823	1.07 (0.35 – 3.27)	0.904
51 – 100	1.12 (0.95 – 1.33)	0.174	1.07 (0.77 – 1.49)	0.678	0.46 (0.14 – 1.59)	0.223	0.28 (0.04 – 1.79)	0.180
≤ 50	1.14 (0.98 – 1.33)	0.098	1.28 (0.95 – 1.72)	0.107	0.69 (0.25 – 1.89)	0.473	0.40 (0.04 – 3.86)	0.431
WHO stage at baseline								
I	1	-	1	-	1	-	-	-
II	1.18 (1.02 – 1.36)	0.023	1.18 (0.90 – 1.54)	0.232	1.11 (0.35 – 3.50)	0.858	-	-
III	1.26 (1.12 – 1.43)	0.000	1.41 (1.13 – 1.76)	0.002	1.88 (0.38 – 2.04)	0.762	-	-
IV	1.30 (1.10 – 1.53)	0.002	1.44 (1.10 – 1.90)	0.009	-‡	-	-	-
Self-reported adherence								
> 90%	1	-	-	-	1	-	-	-
> 60 – 90%	1.01 (0.83 – 1.23)	0.899	-	-	2.78 (0.72 – 10.63)	0.136	-	-
≤ 60%	1.51 (1.09 – 2.08)	0.012	-	-	4.44 (1.30 – 15.20)	0.018	-	-
Number of missed visits								
0	1	-	-	-	1	-	-	-
1	1.02 (0.92 – 1.13)	0.693	-	-	0.96 (0.42 – 2.17)	0.921	-	-
≥ 2	1.03 (0.87 – 1.22)	0.713	-	-	2.53 (1.19 – 5.38)	0.016	-	-
Number of missed ARV visits								
0	1	-	-	-	1	-	-	-
1	1.05 (0.94 – 1.18)	0.368	-	-	1.68 (0.84 – 3.33)	0.140	-	-
≥ 2	1.21 (0.96 – 1.51)	0.100	-	-	1.32 (0.34 – 5.06)	0.687	-	-
Number of missed medical visits								

0	1	-	-	-	1	-	-	-
1	1.12 (1.01 – 1.25)	0.039	-	-	1.96 (1.00 – 3.86)	0.051	-	-
≥ 2	1.30 (1.05 – 1.62)	0.018	-	-	4.44 (1.71 – 11.51)	0.002	-	-
BMI at 6 months in kg/m^2								
< 18.5	1	-	-	-	1	-	-	-
18.5 – < 25	0.84 (0.71 – 0.98)	0.031	-	-	1.74 (0.44 – 6.84)	0.429	-	-
25 – < 30	0.77 (0.64 – 0.92)	0.003	-	-	1.24 (0.28 – 5.42)	0.775	-	-
≥ 30	0.74 (0.60 – 0.90)	0.003	-	-	0.53 (0.08 – 3.46)	0.508	-	-
Haemoglobin at 6 months								
≤ 8 g/dl	1	-	-	-	1	-	-	-
> 8g/dl	0.79 (0.47 – 1.30)	0.349	-	-	0.12 (0.07 – 0.20)	0.000	-	-
MCV at 6 months[l]								
< 80fL	1	-	-	-	1	-	1	-
80 – 100fL	1.35 (1.00 – 1.84)	0.051	-	-	3.58 (1.56 – 8.59)	0.003	5.54 (1.70 – 18.02)	0.004
> 100fL	0.65 (0.48 – 0.90)	0.009	-	-	-[‡]	-	-[‡]	-
Change in MCV at 6 months								
≥ 14.5 fL	1	-	-	-	1	-	-	-
< 14.5 fL	2.82 (2.34 – 3.40)	0.000	-	-	1.94 (0.60 – 6.24)	0.266	-	-
Change in CD4 count at 6 months in cells/mm^3								
≥ expected	1	-	-	-	1	-	-	-
< expected	1.72 (1.51 – 1.96)	0.000	-	-	3.41 (1.16 – 10.01)	0.025	-	-
Change in CD4 count at 6 months stratified by change in MCV at 6 months								
-Change in CD4 count ≥ expected and change in MCV ≥ 14.5fL:	1	-	1	-	1	-	-	-
-Change in CD4 count ≥ expected and change in MCV < 14.5fL:	2.55 (2.01 – 3.23)	0.000	2.82 (2.16 – 3.67)	0.000	3.67 (0.40 – 33.70)	0.251	-	-
-Change in CD4 count < expected and change in MCV ≥ 14.5fL:	0.96 (0.60 – 1.54)	0.856	1.20 (0.73 – 1.97)	0.471	6.60 (0.66 – 66.31)	0.109	-	-
-Change in CD4 count < expected and change in MCV < 14.5fL:	4.83 (3.77 – 6.20)	0.000	5.49 (4.13 – 7.30)	0.000	6.2 (0.69 – 55.61)	0.104	-	-

Regimen change								
No	1	-	-	-	1	-	-	-
Yes	1.09 (0.94 – 1.26)	0.244	-	-	2.21 (0.66 – 7.42)	0.201	-	-
TB: On TB treatment during first 6 months on ART:								
Never	1	-	-	-	1	-	-	-
At baseline	1.06 (0.95 – 1.18)	0.298	-	-	0.95 (0.46 – 1.96)	0.887	-	-
During follow-up	0.98 (0.78 – 1.23)	0.865	-	-	-‡	-	-	-
Pregnancy: Pregnancy during first 6 months on ART:								
Never	1	-	-	-	1	-	-	-
At baseline	1.00 (0.65 – 1.55)	0.990	-	-	-‡	-	-	-
During follow-up	1.10 (0.79 – 1.54)	0.576	-	-	6.04 (2.29 – 15.93)	0.000	-	-
Adverse events or side effects								
No	1	-	-	-	1	-	-	-
Yes	0.39 (0.25 – 0.60)	0.000	-	-	-‡	-	-	-

* On AZT-based regimens throughout the study period. ⌊ The reference group for patients on AZT-based regimens is > 100fL, the following group is 80 – 100 fL and the last group is < 80fL. This is because < 80fL had no observations for MCV at 6 months in patients on AZT-based regimens. ‡ Cell has no observations. ART: Antiretroviral Therapy. ARV: Antiretroviral. AZT: Zidovudine. BMI: Body Mass Index. CI: Confidence interval. IRR: Incidence rate ratio. MCV: Mean cell volume. TB: Tuberculosis. WHO: World Health Organization.

For the table above, multivariate Poisson regression analysis with robust error variance was done first for all eligible patients then for all eligible patients on AZT-based regimens. After adjusting for potential confounders, the variables independently associated with poor adherence for all eligible patients were change in CD4 count at 6 months stratified by change in MCV at 6 months, WHO stage at baseline and MCV at baseline. The IRR of poor adherence increased with increasing WHO stage at baseline. Model fit was good as evidenced by a p-value of 0.794 on the link test. MCV at 6 months was the only independent predictor of poor adherence in the final multivariate model for patients on AZT-based regimens. Model fit for this model was also good as evidenced by a p-value of 0.897 on the link test.

Table 8b: Univariate and multivariate Poisson regression models with robust error variance for all eligible patients and eligible patients on AZT-based regimens with a documented viral load (done 90 to 270 days after ART initiation) after multiple imputation

Variable	All eligible patients				All eligible patients on AZT-based regimens*			
	Univariate Analyses		Multivariate Analysis		Univariate Analyses		Multivariate Analysis	
	Crude IRR (95% CI)	P - value	Adjusted IRR (95% CI)	P- value	Crude IRR (95% CI)	P - value	Adjusted IRR (95% CI)	P- value
Gender								
Females	1	-	1	-	1	-	1	-
Males	1.17 (1.07 – 1.28)	0.000	1.09 (0.99 – 1.19)	0.067	1.78 (0.97 – 3.27)	0.064	1.72 (0.93 – 3.19)	0.086
Age at initiation								
≤ 36.7 years	1	-	1	-	1	-	1	-
> 36.7 years	0.93 (0.85 – 1.02)	0.117	0.95 (0.87 – 1.03)	0.210	0.99 (0.53 – 5.82)	0.962	1.04 (0.56 – 1.92)	0.901
Education level								
≥ Secondary school	1	-	-	-	1	-	-	-
Primary school	0.97 (0.85 – 1.11)	0.660	-	-	1.40 (0.65 – 3.05)	0.391	-	-
< Primary school	0.79 (0.59 – 1.06)	0.114	-	-	0.71 (0.10 – 4.86)	0.725	-	-
Professional status								
Employed	1	-	-	-	1	-	-	-
Unemployed	1.08 (0.99 – 1.18)	0.096	-	-	1.07 (0.59 – 1.96)	0.814	-	-
BMI at initiation in kg/m^2								
< 18.5	1	-	-	-	1	-	-	-
18.5 – < 25	0.96 (0.85 – 1.10)	0.573	-	-	1.83 (0.56 – 6.05)	0.318	-	-
25 – < 30	0.88 (0.75 – 1.03)	0.119	-	-	0.83 (0.16 – 4.37)	0.822	-	-
≥ 30	0.82 (0.67 – 1.01)	0.061	-	-	0.52 (0.09 – 2.92)	0.460	-	-
Baseline haemoglobin								
≤ 8 g/dl	1	-	-	-	1	-	-	-
> 8g/dl	1.02 (0.83 – 1.25)	0.843	-	-	-‡	-	-	-
MCV at baseline[l]								
< 80fL	1	-	-	-	1	-	-	-
80 – 100fL	1.20 (1.03 – 1.40)	0.021	-	-	1.06 (0.46 – 2.42)	0.896	-	-
> 100fL	1.28 (0.97 – 1.71)	0.079	-	-	0.77 (0.17 – 3.43)	0.734	-	-

Baseline CD4 count in cells/mm^3								
> 200	1	-	1	-	1	-	1	-
101 – 200	1.02 (0.88 – 1.20)	0.733	1.12 (0.96 – 1.31)	0.136	0.93 (0.41 – 2.09)	0.859	0.83 (0.37 – 1.83)	0.640
51 – 100	1.11 (0.95 – 1.31)	0.195	1.24 (1.05 – 1.46)	0.012	0.66 (0.22 – 1.99)	0.464	0.57 (0.19 – 1.73)	0.323
≤ 50	1.14 (0.98 – 1.33)	0.090	1.26 (1.07 – 1.49)	0.005	0.95 (0.40 – 2.29)	0.911	0.72 (0.29 – 1.78)	0.473
WHO stage at baseline								
I	1	-	1	-	1	-	-	-
II	1.15 (1.00 – 1.33)	0.044	1.11 (0.97 – 1.28)	0.131	1.03 (0.35 – 3.05)	0.955	-	-
III	1.23 (1.09 – 1.38)	0.000	1.19 (1.06 – 1.34)	0.003	1.03 (0.48 – 2.23)	0.932	-	-
IV	1.24 (1.06 – 1.45)	0.006	1.19 (1.01 – 1.40)	0.033	-‡	-	-	-
Self-reported adherence								
> 90%	1	-	1	-	1	-	-	-
> 60 – 90%	1.08 (0.85 – 1.19)	0.925	1.01 (0.86 – 1.18)	0.941	1.60 (0.50 – 5.12)	0.428	-	-
≤ 60%	1.53 (1.11 – 2.10)	0.009	1.55 (1.13 – 2.13)	0.007	3.26 (1.05 – 10.16)	0.041	-	-
Number of missed visits								
0	1	-	-	-	1	-	-	-
1	1.02 (0.92 – 1.13)	0.693	-	-	0.96 (0.42 – 2.17)	0.921	-	-
≥ 2	1.03 (0.87 – 1.22)	0.713	-	-	2.53 (1.19 – 5.38)	0.016	-	-
Number of missed ARV visits								
0	1	-	-	-	1	-	-	-
1	1.05 (0.94 – 1.18)	0.358	-	-	1.66 (0.83 – 3.30)	0.150	-	-
≥ 2	1.20 (0.96 – 1.51)	0.104	-	-	1.23 (0.32 – 4.70)	0.762	-	-
Number of missed medical visits								
0	1	-	1	-	1	-	1	-
1	1.12 (1.01 – 1.25)	0.039	1.13 (1.01 – 1.26)	0.024	1.86 (0.97 – 3.59)	0.063	1.84 (0.94 – 3.61)	0.077
≥ 2	1.30 (1.04 – 1.61)	0.019	1.19 (0.96 – 1.48)	0.107	4.00 (1.54 – 10.33)	0.002	3.71 (1.50 – 9.15)	0.005
BMI at 6 months in kg/m^2								
< 18.5	1	-	-	-	1	-	-	-
18.5 – < 25	0.81 (0.68 – 0.96)	0.016	-	-	1.44 (0.36 – 5.74)	0.608	-	-
25 – < 30	0.74 (0.62 – 0.88)	0.001	-	-	1.05 (0.25 – 4.47)	0.946	-	-
≥ 30	0.70 (0.57 – 0.86)	0.001	-	-	0.39 (0.05 – 2.59)	0.331	-	-

Haemoglobin at 6 months								
≤ 8 g/dl	1	-	-	-	1	-	-	-
> 8g/dl	0.68 (0.41 – 1.13)	0.136	-	-	0.13 (0.10 – 0.18)	0.000	-	-
MCV at 6 months[L]								
< 80fL	1	-	-	-	1	-	1	-
80 – 100fL	1.24 (0.92 – 1.69)	0.161	-	-	2.09 (1.03 – 4.24)	0.041	2.00 (1.01 – 3.98)	0.047
> 100fL	0.64 (0.47 – 0.87)	0.005	-	-	-‡	-	-‡	-
Change in MCV at 6 months								
≥ 14.5 fL	1	-	-	-	1	-	-	-
< 14.5 fL	2.23 (1.90 – 2.61)	0.000	-	-	1.97 (0.84 – 4.61)	0.119	-	-
Change in CD4 count at 6 months in cells/mm^3								
≥ expected	1	-	-	-	1	-	-	-
< expected	1.63 (1.45 – 1.83)	0.000	-	-	2.24 (1.01 – 4.97)	0.048	-	-
Change in CD4 count at 6 months stratified by change in MCV at 6 months								
-Change in CD4 count ≥ expected and change in MCV ≥ 14.5fL:	1	-	1	-	1	-	-	-
-Change in CD4 count ≥ expected and change in MCV < 14.5fL:	2.14 (1.78 – 2.58)	0.000	2.18 (1.81 – 2.63)	0.000	2.39 (0.60 – 9.43)	0.211	-	-
-Change in CD4 count < expected and change in MCV ≥ 14.5fL:	1.47 (1.13 – 1.92)	0.004	1.51 (1.16 – 1.96)	0.002	3.10 (0.62 – 15.46)	0.166	-	-
-Change in CD4 count < expected and change in MCV < 14.5fL:	3.57 (2.98 – 4.27)	0.000	3.69 (3.07 – 4.42)	0.000	4.36 (1.15 – 16.52)	0.031	-	-
Regimen change								
No	1	-	-	-	1	-	-	-
Yes	1.09 (0.94 – 1.26)	0.244	-	-	2.21 (0.66 – 7.42)	0.201	-	-
TB: On TB treatment during first 6 months on ART:								
Never	1	-	-	-	1	-	-	-
At baseline	1.06 (0.95 – 1.18)	0.298	-	-	0.95 (0.46 – 1.96)	0.887	-	-
During follow-up	0.98 (0.78 – 1.23)	0.865	-	-	-‡	-	-	-

Pregnancy: Pregnancy during first 6 months on ART:								
Never	1	-	-	-	1	-	-	-
At baseline	1.00 (0.65 – 1.55)	0.990	-	-	-‡	-	-	-
During follow-up	1.10 (0.79 – 1.54)	0.576	-	-	6.04 (2.29 – 15.93)	0.000	-	-
Adverse events or side effects								
No	1	-	-	-	1	-	-	-
Yes	0.39 (0.25 – 0.60)	0.000	-	-	-‡	-	-	-

* On AZT-based regimens throughout the study period. ⌊ The reference group for patients on AZT-based regimens is > 100fL, the following group is 80 – 100 fL and the last group is < 80fL. This is because < 80fL had no observations for MCV at 6 months in patients on AZT-based regimens before multiple imputation. ‡ Cell has no observations. ART: Antiretroviral Therapy. ARV: Antiretroviral. AZT: Zidovudine. BMI: Body Mass Index. CI: Confidence interval. IRR: Incidence rate ratio. MCV: Mean cell volume. TB: Tuberculosis. WHO: World Health Organization.

For the table above, multivariate Poisson regression analysis with robust error variance was done first for all eligible patients then for all eligible patients on AZT-based regimens. After adjusting for potential confounders, the variables independently associated with poor adherence for all eligible patients were change in CD4 count at 6 months stratified by change in MCV at 6 months, CD4 count at baseline, WHO stage at baseline, self-reported adherence and number of missed medical visits. Model fit was good as evidenced by a p-value of < 0.05 on the FMI test. MCV at 6 months and the number of missed medical visits were the only independent predictors of poor adherence in the final multivariate model for patients on AZT-based regimens. Model fit for this model was also good as evidenced by a p-value of < 0.05 on the FMI test.

CHAPTER FOUR

DISCUSSION:

In this chapter, the results are examined in detail. This includes an examination of the strengths, limitations and generalizability of the study. The aim of this study is to identify markers that can be used routinely to assess adherence to ART at Themba Lethu Clinic in Johannesburg, South Africa. The first viral load after initiation (viral load at end-point), which is usually done at 6 months, is used to define adherence [9, 10].

The IRR is used to approximate the RR of poor adherence. Variables independently associated with poor adherence at 6 months were change in CD4 count stratified by change in MCV at 6 months, CD4 count at baseline, WHO stage at baseline and MCV at baseline. For patients on AZT-based regimens the variables independently associated with poor adherence at 6 months were change in CD4 count at 6 months and pregnancy during the first 6 months on ART.

The sensitivity analyses, subgroup analyses and analyses using a viral load done 90 to 270 days after ART initiation yielded the same predictors of poor adherence. After multiple imputation, however, other variables like self-reported adherence, number of missed medical visits and regimen change were also found to be predictors of poor adherence in some of the models.

The results are in keeping with the alternate hypothesis that adherence level (measured by viral load at 6 months) can be predicted from a combination of variables or markers and provides a basis to reject the null hypothesis.

4.1 Markers of adherence:

Previous studies have shown change in MCV to be a good marker of adherence [12, 17, 47-49]. In this study, there was an interaction between change in MCV and change in CD4 count at 6 months for all patients with a documented viral load result. In the presence of the interaction, change in MCV remained significantly associated with poor adherence ($p < 0.001$). Since most (62.9%) of the patients were on d4T or AZT-based regimens during the study period, it is not surprising that change in MCV at 6 months was a marker of adherence.

For patients on d4T or AZT-based regimens with a documented viral load result, the sensitivity, specificity, PPV and NPV of the change in MCV at 6 months were as follows: 70.2%, 61.4%, 18.2% and 94.4% respectively (Table 4a). For patients on TDF-based regimens with a documented viral load result, the sensitivity, specificity, PPV and NPV of the change in MCV at 6 months were as follows: 97.4%, 3.1%, 24.4% and 78.4% respectively (Table 4a). This implies that a change in MCV at 6 months of ≥ 14.5fL rules out poor adherence. However, change in MCV is a marker specific to patients on d4T or AZT so a change in MCV at 6 months of < 14.5fL does not necessarily imply poor adherence in patients on TDF hence the very low specificity for these patients.

To get a clearer picture of the interaction between change in MCV and change in CD4 count, change in CD4 count at 6 months was stratified by change in MCV at 6 months, showing how the two variables together can be used to predict poor adherence. The variable change in CD4 count stratified by change in MCV at 6 months was found to have a higher area under its ROC curve than change in CD4 count alone or change in MCV alone, indicating the strength of this stratified variable (Figure 2a, Figure 2b, Appendix 7A and Appendix 9A). The ROC curve for the stratified variable was also the closest to the ROC curve for the final multivariate model, indicating that it is the major contributor to the model (Figure 2a, Figure 2b, Appendix 7A and Appendix 9A). Patients who had a change in CD4 count ≥ or < expected at 6 months and a change in MCV that is < 14.5fL were at a significantly higher risk of having a detectable viral load at 6 months than their counterparts. The aRR (adjusted relative risk) of poor adherence was highest for patients with a change in CD4 count that was < expected and a change in MCV < 14.5fL at 6 months. Table 4b thus shows that of the 3 categories of the variable, this category had the lowest difference between sensitivity (69.9%) and specificity (79.2%).

MCV is known to increase, particularly in adherent patients on AZT[12, 16, 17, 47-49]. For patients on AZT, however, despite a $p < 0.2$ on univariate analysis, neither MCV at 6 months nor change in MCV at 6 months were significantly associated with adherence in any of the final multivariate models before multiple imputation, except where the viral load was done 90 to 270 days after ART initiation. This could be attributed to the large proportions of missing values for MCV at 6 months and MCV at baseline. After multiple imputation, the variable MCV at 6 months was found to be a predictor of poor adherence in patients on AZT-based

regimens only where the viral load was done 90 to 270 days after ART initiation. Change in MCV at 6 months was, however, still not a predictor of poor adherence in patients on AZT after multiple imputation. The small proportions of patients on AZT-based regimens (< 3%) could have also been a contributing factor. There was also no interaction between change in MCV and change in CD4 count at 6 months for patients on an AZT-based regimen.

For patients on AZT-based regimens, change in CD4 count at 6 months was one of the markers of adherence. The large aRR (7.66), marginal significance (p = 0.052) and wide CI (0.98 – 59.91) cast some doubt on the findings. This could be due to weak unmeasured confounders such as change in MCV at 6 months, which as mentioned above, had several missing values. The sensitivity and specificity of the change in CD4 count at 6 months for patients on AZT-based regimens in the main analysis were 64.7% and 75.2% respectively (Table 4a). The difference between the two was, however, not as high as that for all patients (33.0% and 81.4% respectively), which may explain why change in CD4 count at 6 months remained a significant predictor of poor adherence in patients on AZT-based regimens. Likewise, Figure 2c shows that change in CD4 count at 6 months also had a larger area under its ROC curve (0.7718) for patients on AZT-based regimens than it did for all patients as shown in Figure 2a (0.5784).

Self-reported adherence is expected to be associated with viral load at 6 months, which is a surrogate marker for adherence. The variable was, however, notably absent from the final multivariate models before multiple imputation. This is contrary to a meta-analysis by Nieuwkerk and colleagues (2005), where self-reported adherence was found to distinguish between poor and good adherence, as evidenced by detectable and undetectable viral loads respectively [(19)]. The absence of the self-reported adherence from the final multivariate models before multiple imputation may reflect the fact that the variable is not reliable as it is vulnerable to errors from patient recall and dishonesty[(11, 12)]. The results may also be testament to the fact that self-reported adherence is a measure of how ART is taken but does not encompass the full definition of adherence, which includes following other recommendations pertaining to food, etc[(13)]. Not following these other recommendations may lead to a detectable viral load at 6 months, despite taking the medication correctly. The absence of self-reported adherence from final multivariate models before multiple imputation may also be testament to the large proportion of patients (> 50%) without a self-reported adherence, making the results unreliable. After multiple imputation, self-reported adherence

appeared in only one of the final multivariate models (Table 8b). These results must be interpreted with caution as self-reported adherence had an area under the ROC curve of only 0.5247 (Appendix 7A).

Pregnancy during the first 6 months on ART was also found to increase the risk of having a detectable viral load at 6 months in patients on AZT-based regimens. Though this variable had a high specificity of 97.6% (Table 4a) in patients on AZT-based regimens, the sensitivity was very low (20%). This means that a pregnancy during follow up could mean poor adherence but no conclusions can be drawn if a patient is not pregnant. This implies that the variable is not a very reliable marker of adherence. The aRRs for this variable were also unusually high with very wide CIs. This may be due to weak unmeasured confounders, particularly as very few patients (< 3%) were on AZT-based regimens during the study period. Consequently a lot of the variables, such as change in MCV discussed earlier, either had large proportions of missing values or no observations for some categories in patients on AZT-based regimens. The variable pregnancy during the first 6 months on ART had no observations for the category pregnant at baseline. For all the above reasons, the appearance of the variable pregnancy during the first 6 months on ART as a marker of poor adherence should be interpreted with caution. However, when combined with an inadequate increase in CD4 count at 6 months, pregnancy during the first 6 months on ART may be a strong predictor of poor adherence. This is demonstrated by an area of 0.8938 under the ROC curve for the final multivariate model for patients on AZT-based regimens (Figure 2c). The strength of the variable, however, remains in question as the area under the ROC curve for this variable was only 0.5012 in the subgroup analysis (Appendix 9B). In addition, the variable was not found to be a predictor of poor adherence for patients with a viral load done 90 to 270 days after ART initiation or in any of the multiple imputation models.

Except for pregnancy during the first 6 months on ART (PPV 60% and NPV 87.0%), all other markers of adherence identified had low PPVs (< 40%) and high NPVs (> 75%). Conclusions are difficult to draw from these figures before multiple imputation as the prevalence of a detectable viral load at 6 months varies for the different variables and differs from the overall prevalence due to varying proportions of missing values. The figures do not change much after multiple imputation implying that the variables are better predictors of good adherence than poor adherence.

Lower CD4 count at baseline and higher WHO stage at baseline were also significantly associated with a detectable viral load at 6 months and hence poor adherence. These results are contrary to the findings of Maqutu and colleagues in South Africa (2010), where first month poor adherence (< 95%) measured by pill count was found to be associated with higher CD4 count at baseline[30]. The results are also contrary to a study by Heckman and colleagues in the USA (2004), which found good adherence by self-report to be associated with progression to AIDS[65]. The DART trial carried out by Muyingo and colleagues in Uganda and Zimbabwe (2008), however, had similar findings to this study. This was that complete (100%) adherence, measured by pill count, was associated with a higher baseline CD4 count[41].

4.2 Other variables:

In patients on AZT-based regimens after multiple imputation in the main analysis, those > 36.7 years at initiation had half (aRR 0.54, 95% CI 0.35 – 0.58; p = 0.041) the risk of virological failure that their counterparts aged 18 to 36.7 years had (Table 3b). Age was, however, not significantly associated with poor adherence in any of the other multivariate analyses. This is contrary to some previous studies that have associated younger age with poor adherence[33, 45, 58]. Gender was also not significantly associated with poor adherence in any of the final multivariate models. Godin and colleagues in a longitudinal study in the USA (2005), found male gender to be a predictor of good adherence[61]. In a retrospective review of patient records in South Africa (2011), Maqutu and colleagues, however, found that males have better adherence than females but discovered that after the age of 50, females have better adherence levels than males[32]. A prior baseline survey done by Maqutu and colleagues (2010) found first month good adherence to be associated with higher age in patients who owned cell phones or were a source of household income in the urban setting[30]. It would thus appear that there is interplay of several other factors with gender and age resulting in poor adherence and, hence, detectable viral load at 6 months.

Education level and professional status at ART initiation were also not significantly associated with poor adherence in any of the final multivariate models. A cohort study by El-Khatib and colleagues in South Africa (2011), had similar findings[21]. Average poor adherence measured by pill count, during the 6 month study period, was associated with less education and lack of support from a partner at initiation. Education level and lack of support from a partner at initiation were, however, not associated with viral load at 6 months. In the

same study, socioeconomic status at initiation was neither associated with average poor adherence measured by pill count nor viral load at 6 months. These findings imply that with education about treatment given as part of ART provision, adult patients on ART should be able to achieve a viral load < 400 copies/ml at 6 months despite their education level or professional status at ART initiation.

The numbers of missed ARV visits and missed medical visits were also not significantly associated with poor adherence in the final multivariate models before multiple imputation. Over 70% of the patients had no missed ARV or medical visits and about 20% of patients had 1 missed ARV or medical visit during the study period. The variables may be absent from the final multivariate model because on-time drug pick up does not guarantee adherence[46]. Furthermore, an adherent patient who misses an appointment due to having enough medication left over cannot be classified as non-adherent. For this reason, even though the number of missed medical visits was a marker of poor adherence after multiple imputation (Table 8b and Appendix 8B), the area under the ROC curve remained only around 0.5 for this variable (Appendices 7A & 7B: 0.5167 and Appendix 9B: 0.5059).

A quarter of the patients were on TB treatment during the study period. This is because TB co-infection is common among HIV positive patients[70]. This variable was expected to influence the outcome but it has no association with the viral load at 6 months in the final multivariate models for all patients. This is because TB, which may impair the ability to adhere due to severe ill health affecting memory or strength to take medication, is curable and thus transient. Likewise, TB medication increases the pill burden, with the possibility of reducing adherence only temporarily. Furthermore, treating active TB should put the virological response on a par with that of patients without TB. Note that the variable measures being on TB treatment as a proxy for having TB. It is also possible that some patients with TB and a detectable viral load at 6 months died or were lost to follow-up or transferred to other facilities before the end-point and were thus excluded from the study. This could be responsible for leaving the impression that the relative risk of having a detectable viral load at 6 months for patients on TB treatment is the same as that for patients not on TB treatment during the study period.

Haemoglobin level was also not significantly associated with poor adherence in any of the final multivariate models. This may be because only a small proportion of patients had a

haemoglobin ≤ 8g/dl (4.7% at baseline and 0.7 % at 6 months in the main analysis). It may also be due to the fact that there are measures that can be taken to improve haemoglobin levels hence the levels improved during the course of treatment. Russell and colleagues in South Africa (2010) found low haemoglobin levels to be a predictor of early mortality in adults starting ART[71]. This implies that patients with low haemoglobin levels in this study may have died before 6 months on ART, resulting in the variable not being a predictor of poor adherence. In keeping with the findings of Tedaldi and colleagues in the USA (2006), BMI was also not significantly associated with poor adherence[72].

Adverse events or side effects represent another variable that is not significantly associated with poor adherence in the final multivariate model despite being significantly associated with poor adherence on univariate analysis. This is contrary to the findings of Protopopescu and colleagues in a prospective cohort study done in France (2009), showing that side effects were significantly associated with poor adherence[58]. In that study, however, the follow up period was 1 to 10 years and adherence was measured by self-report. Adverse events or side effects are not in the final model in this study probably as they occur rarely (2.9% experienced side effects during the study period in the main analysis). The variable may also be inaccurate as 2 – 3% of patients in the different models had a documented side effect during the study period but 9 – 10% of patients had a change of regimen yet most substitutions during this time are due to side effects [68].

A subgroup analysis (Appendix 8A) of 6793 patients with a documented viral load result (excluding 95 patients with a history of documented prior ART use such as NVP prophylaxis and 272 patients changed to a non-standard or second-line ART regimen before 6 months on treatment) was done. This subgroup had 18.7% of patients exhibiting a detectable viral load at 6 months. The subgroup analysis yielded the same predictors of poor adherence as those in the main analysis (Table 3a). The only extra variable was regimen change (Appendix 8A). Adding this variable improved the model despite the marginal significance. It thus seems possible that drug substitutions within a 1st line regimen may lower the risk of having a detectable viral load at 6 months in this subgroup (No change in regimen; aRR 1 and change in regimen; aRR 0.68 95% CI 0.45 – 1.03, p = 0.066, ROC area: 0.5121). The variable was also in the final multivariate model for the subgroup analysis after multiple imputation (Appendix 8B: aRR 0.76 95% CI 0.59 – 0.97, p = 0.028, ROC area: 0.5114) and in the sensitivity analysis assuming that all patients with a missing viral load had a viral load ≥ 400

copies/ml (Table 7a: aRR 0.79 95% CI 0.62 – 1.01, p = 0.058). Most patients have substitutions within a 1st line regimen due to side effects[68]. Since drug substitution provides relief from these toxicities, adherence is expected to improve after a substitution. The ROC areas of about 0.5, however, cast doubt on the validity of regimen change as a predictor of poor adherence.

4.3 Strengths and limitations:

The study uses all available data in TherapyEdge-HIV™ and not a sample. This improves the power of the study. The study design also permits the ascertainment of the temporal relationship between the exposures and the outcome variable.

In addition, the results of the sensitivity analysis are almost exactly the same as those of the main analysis implying that the patients with a documented viral load at 6 months were representative of the entire study population. Furthermore, multiple imputation was done to cater for missing values.

Missing data for both exposure and outcome variables is a common limitation of the retrospective design. The proportion of missing data ranged from 1.5% for professional status among eligible patients to 70.0% for self-reported adherence in patients on AZT-based regimens with a documented viral load result in the main analysis. Half of the eligible patients did not have a viral load at 6 months. Other variables with large proportions of missing data included, MCV at 6 months, change in MCV at 6 months and change in CD4 count at 6 months. The missing data was thought to have contributed to all markers of adherence having ROC areas of < 0.8 so that no variable or marker could be considered a particularly accurate predictor of poor adherence. After multiple imputation, however, the baseline characteristics remained similar to those before multiple imputation. In addition, areas under the ROC curves remained < 0.8.

Furthermore, professional status and education level are recorded only at initiation so it is not possible to adjust for any changes in the two variables during the study period. In addition, viral load testing is not routinely done at baseline so it was not practical to adjust for this variable in the multivariate model.

Loss to follow up is another disadvantage of the retrospective design. Patients that were excluded due to loss to follow up during the study period may bias the results of the study if they are systematically different from the other patients (selection bias). Patients that died or were transferred out during the study period may also introduce selection bias into the results of the study if they are systematically different from those that did not die or get transferred out. This is because patients that were lost to follow up or died during the first 6 months on ART are likely to have had poor adherence. This early poor adherence may be a strong predictor of loss to follow up and ultimately progression to AIDS and mortality [73, 74].

Results may also be biased by the fact that the 6 months was defined as a wide range of days after ART initiation not a fixed point in time. Tables 2 and 6 show that the variation between the points at which MCV, CD4 count and viral load at 6 months were taken was minimal. The validity and reliability of the sensitivity analyses is, however, hampered by the dissimilarity between the proportions of patients with an MCV at 6 months, change in MCV at 6 months and change in CD4 count at 6 months in the main analysis and the sensitivity analysis. This dissimilarity also reduces the comparability value of the sensitivity analysis. In addition, self-reported adherence, defined as the last self-reported adherence value on, or after, 3 months (70 days) on ART and not more than 110 days before the viral load was done, but at least 28 days before the end-point, was taken at different points in time for the sensitivity analysis and the main analysis, limiting comparability (Tables 2 and 6).

An additional limitation is that the outcome variable viral load at end-point is itself a surrogate marker of adherence as poor adherence is the most common cause of a virological failure. Other factors, that are not routinely captured in TherapyEdge-HIV^{TM}, such as malabsorption, drug interactions, diabetes mellitus and the acquisition of a drug-resistant HIV strain may, however, contribute to a detectable viral load at 6 months, despite good adherence.

There were also a few statistical limitations. The first was that log-binomial regression using RRs had to be replaced with Poisson regression with robust error variance, where IRRs were used to approximate the RRs, due to the frequent occurrence of failure to converge. Another was that RRs could not be ascertained where cells had no observations, for patients on AZT-based regimens.

It is worth noting that there is no known surrogate clinical marker for the detection of adherence levels for TDF, 3TC, FTC, NVP and EFV. Change in MCV at 6 months is a marker only for patients on AZT-based regimens as well as the patients on d4T-based regimens, who are the majority (62.9%).

4.4 Generalizability

The results of the study are generalizable to adult patients on first-line ART at Themba Lethu Clinic, who were ART-naïve at the time of initiation, for their first 6 months on ART or up to the point of their first viral load. They should also, by extension, be generalizable to similar patients on first-line ART in other public urban CCMT sites in South Africa.

4.5 Conclusion and recommendations:

In conclusion, the marker of adherence to ART among HIV positive adults at Themba Lethu Clinic is change in CD4 count at 6 months stratified by change in MCV at 6 months. MCV, CD4 count and WHO stage at baseline are also predictors of viral load at 6 months and hence predictors of poor adherence. For adult patients on AZT-based regimens, the markers of adherence to ART are change in CD4 count at 6 months and pregnancy during the first 6 months on ART.

The fact that the aRR of a detectable viral load at 6 months increases with lower CD4 count and higher WHO stage at baseline emphasises the need for people to know their HIV status and start ART on time if HIV positive. The fact that almost one third of the patients were initiated in WHO stage I suggests that this message is getting to people. HIV testing and pre-ART service delivery need to be strengthened to increase this percentage as well as the baseline CD4 count at ART initiation as patients who start earlier are less likely to have a detectable viral load at 6 months.

In addition, attention should be paid to the above-mentioned markers of adherence. These markers could help health workers identify poor adherence in the absence of viral load testing and target patients for adherence interventions to prevent virological failure. Interventions such as intensified counselling, coupled with unannounced pill count or micro-electronic monitoring (MEM) could be used to improve the adherence of patients found to have poor adherence. These findings also suggest that provision of family planning services must be strengthened for all women of child-bearing age on ART. The latest (2013) National ART

Guidelines stipulate that an FBC, from which MCV is obtained be done only at baseline, 3 and 6 months for patients on AZT-based regimens to detect AZT toxicity[15]. However, the findings of this study make the case for MCV monitoring for patients on d4T as well, to gauge adherence. Beyond 6 months on ART, MCV monitoring may still be necessary for patients on AZT and d4T-based regimens to assess adherence. The interaction between change in MCV and change in CD4 count at 6 months also implies that CD4 count may need to be tested at 6 months and beyond the currently stipulated baseline and 1 year points on ART[15]. Further studies are, however, needed to verify whether the markers of adherence remain the same over longer periods on ART. It is worth noting that with TDF-based regimens currently being the first-line of treatment, the markers of adherence may change. In other words the association of variables such as MCV (markers of adherence for AZT and d4T) with viral load at 6 months may become less significant.

In future work, the above markers could be used in a composite score system to identify patients with poor adherence, using sensitivity, specificity, PPV and NPV of the composite score to determine how predictive it would be.

Further studies are also needed to assess, verify and explain the influence of self-reported adherence, number of missed medical visits and regimen change. This is because these variables appeared in one of the sensitivity analyses as well as in some of the final multivariate models after multiple imputation.

Another recommendation is that self-reported adherence be assessed and documented consistently at each visit as multivariate regression after multiple imputation suggests that it may also be a marker of poor adherence. Furthermore, additional emphasis needs to be placed on ensuring that all blood tests are done on schedule according to the National Guidelines for the benefit of the patient and for ease and accuracy of data analysis in case of future studies. More attention also needs to be paid to data accuracy on entry and data cleaning to rectify inaccuracies before they accumulate.

4.6 Dissemination:

Results will be published in relevant journals and presented orally or in poster format at scientific meetings.

REFERENCES:

1. Hovarth T, Azman H, Kennedy GE and Rutherford GW. Mobile phone text messaging for promoting adherence to antiretroviral therapy in patients with HIV infection. Cochrane Database of Syst Rev. 2012 Mar 14; **3(CD009756):** 1-38.
2. Katz IT, Essien T, Marinda ET, Gray GE, Bangsberg DR, Martinson NA et al. Antiretroviral therapy refusal among newly diagnosed HIV-infected adults. AIDS. 2011 Nov 13; **25(17):** 2177-81.
3. Reynolds C, de Koning CB, Pelly SC, van Otterlo WAL and Bode ML. In search of a treatment for HIV - current therapies and the role of non-nucleoside reverse transcriptase inhibitors (NNRTIs). Chem Soc Rev. 2012 Jul 7; **41(13):** 4657-70.
4. UNAIDS. UNAIDS Global report on global AIDS epidemic 2010 [Internet]. 2010 [updated 2010; cited 8 Dec 2012]. Available from: http://www.unaids.org/globalreport/global_report.htm.
5. Larson E, O'Bra H, Brown J, Mbengashe T and Klausner J. Supporting the massive scale-up of antiretroviral therapy: the evolution of PEPFAR-supported treatment facilities in South Africa, 2005-2009. BMC Public Health. 2012 Mar 9; **12(1):** 173.
6. World Health Organization. MDG 6: HIV/AIDS, malaria and other diseases, antiretroviral therapy coverage among people with advanced HIV [Internet]. 2011 [updated 2011; cited 8 Dec 2012]. Available from: http://apps.who.int/ghodata/?vid=360.
7. Fox MP, Maskew M, MacPhail AP, Long L, Brennan AT, Westreich D et al. Cohort Profile: The Themba Lethu Clinical Cohort, Johannesburg, South Africa. Int J Epidemiol. 2013 Apr; **42(2):** 430-9.
8. Westreich D, Cole SR, Nagar S, Maskew M, van der Horst C and Sanne I. Pregnancy and virologic response to antiretroviral therapy in South Africa. PLoS One. 2011 Aug 2; **6(8):** 1-8.
9. Allen D, Conradie F, Coovadia A, Cotton M, Duma MW, Fomundam H et al. National Antiretroviral Treatment Guidelines [Internet]. 2004 [updated 2004; cited 8 Dec 2012]. Available from: http://www.doh.gov.za/list.php?type=HIV%20and%20AIDS&year=2004.
10. Clinical Guidelines for the Management of HIV and AIDS in Adults and Adolescents. 2nd ed. [Internet]. 2010 [updated 2010; cited 8 Dec 2012]. Available from:http://www.sahivsoc.org/practise-guidelines/national-dept-of-health-guidelines.
11. Berg KM and Arnsten JH. Practical and conceptual challenges in measuring

antiretroviral adherence. JAIDS. 2006 Dec 1; **43:** S79-S87.

12. Segeral O, Madec Y, Ban B, Ouk V, Hak CR, Le Tiec C et al. Simplified assessment of antiretroviral adherence and prediction of virological efficacy in HIV-infected patients in Cambodia. AIDS Res Treat. 2010 Jan 1; **2010:** 1-6.
13. Tiyou A, Belachew T, Alemseged F and Biadgilign S. Predictors of adherence to antiretroviral therapy among people living with HIV/AIDS in resource-limited setting of southwest Ethiopia. AIDS Res Ther. 2010 Oct 30; **7(1):** 1-10.
14. Brennan AT, Maskew M, Sanne I and Fox MP. The importance of clinic attendance in the first six months on antiretroviral treatment: a retrospective analysis at a large public sector HIV clinic in South Africa. J Int AIDS Soc. 2010 Dec 7; **13(1):** 1-10.
15. The South African Antiretroviral Treatment Guidelines 2013. South Africa2013 [Internet]. 2013 [updated 2013; cited 27 Aug 2013]. Available from: http//www.sahivsoc.org/upload/documents/2013%20ART%20Treatment%20Guidelines%20Final%2025%20March%202013%20corrected.pdf.
16. Bartlett JG, Gallant JE and Pham PA. 2009-2010 Medical Management of HIV infection. 15th ed. Durham, North Carolina: Knowledge Source Solutions LLC; 2009.
17. Rivas P, Gorgolas M and Fernandez-Guerrero ML. Zidovudine and red-cell distribution width. NEJM. 2005 May 19; **352(20):** 2141-2.
18. Mugisha JO, Donegan K, Fidler S, Ramjee G, Hodson A, Dunn DT et al. Mean corpuscular volume as a marker for adherence to zidovudine-containing therapy in HIV-infected adults. Open AIDS J. 2012; **6:** 45-52.
19. Nieuwkerk PT and Oort FJ. Self-reported adherence to antiretroviral therapy for HIV-1 infection and virologic treatment response: a meta-analysis. JAIDS. 2005 Apr 1; **38(4):** 445-8.
20. Tuboi SH, Harrison LH, Sprinz E, Albernaz RKM and Schechter M. Predictors of virologic failure in HIV-1-infected patients starting highly active antiretroviral therapy in Porto Alegre, Brazil. JAIDS . 2005 Nov 1; **40(3):** 324-8.
21. El-Khatib Z, Ekstrom A, Coovadia A, Abrams E, Petzold M, Katzenstein D et al. Adherence and virologic suppression during the first 24 weeks on antiretroviral therapy among women in Johannesburg, South Africa - a prospective cohort study. BMC Public Health. 2011 Feb 8; **11(1):** 1-13.
22. Huynh TK, Luttichau HR, Roge BT and Gerstoft J. Natural history of hyperlactataemia in human immunodeficiency virus-1-infected patients during highly active antiretroviral therapy. Scand J Infect Dis. 2003 Apr 11; **35(1):** 62-6.

23. Levy A, McCandless L, Harrigan PR, Hogg R, Bondy G, Iloeje U et al. Changes in lipids over twelve months after initiating protease inhibitor therapy among persons treated for HIV/AIDS. Lipids Health Dis. 2005 Feb 10; **4(1):** 1-7.
24. Gandhi M, Ameli N, Bacchetti P, Anastos K, Gange SJ, Minkoff H et al. Atazanavir concentration in hair is the strongest predictor of outcomes on antiretroviral therapy. Clin Infect Dis. 2011 May 15; **52(10):** 1267-75.
25. van Zyl GU, van Mens TE, McIlleron H, Zeier M, Nachega JB, Decloedt E et al. Low lopinavir plasma or hair concentrations explain second-line protease inhibitor failures in a resource-limited setting. JAIDS. 2011 Apr 1; **56(4):** 333-9.
26. Podsadecki T, Vrijens B, Tousset E, Rode R and Hanna G. "White coat compliance" limits the reliability of therapeutic drug monitoring in HIV-1-infected patients. HIV Clin Trials. 2008 Aug 26; **9(4):** 238-46.
27. Pearson CR, Micek MA, Simoni JM, Hoff PD, Matediana E, Martin DP et al. Randomized control trial of peer-delivered, modified directly observed therapy for HAART in Mozambique. JAIDS. 2007 Oct 1; **46(2):** 238-44.
28. Sarna A, Luchters S, Geibel S, Chersich MF, Munyao P, Kaai S et al. Short- and long-term efficacy of modified directly observed antiretroviral treatment in Mombasa, Kenya: a randomized trial. JAIDS. 2008 Aug 15; **48(5):** 611-9.
29. Hill S and Kavookjian J. Motivational interviewing as a behavioral intervention to increase HAART adherence in patients who are HIV-positive: a systematic review of the literature. AIDS Care. 2012 May 1; **24(5):** 583-92.
30. Maqutu D, Zewotir T, North D, Naidoo K and Grobler A. Factors affecting first-month adherence to antiretroviral therapy among HIV-positive adults in South Africa. Afr J AIDS Res. 2010 Sep 22; **9(2):** 117-24.
31. Barclay TR, Hinkin CH, Castellon SA, Mason KI, Reinhard MJ, Marion SD et al. Age-associated predictors of medication adherence in HIV-positive adults: health beliefs, self-efficacy, and neurocognitive status. Health Psychol. 2007 Jan 11; **26(1):** 40-9.
32. Maqutu D, Zewotir T, North D, Naidoo K and Grobler A. Determinants of optimal adherence over time to antiretroviral therapy amongst HIV-positive adults in South Africa: a longitudinal study. AIDS Behav. 2011 Oct; **15(7):** 1465-74.
33. Hinkin CH HD, Mason KI, Castellon SA, Durvasula RS, Lam MN and Stefaniak M. Medication adherence in HIV-infected adults: effect of patient age, cognitive status, and substance abuse. AIDS. 2004 Jan 1; **18(1):** S19-S25.

34. Paterson DL, Swindells S, Mohr J, Brester M, Vergis EN, Squier C et al. Adherence to protease inhibitor therapy and outcomes in patients with HIV infection. Ann Intern Med. 2000 Jul 4; **133(1):** 21-30.
35. Hofer CB, Schechter M and Harrison LH. Effectiveness of antiretroviral therapy among patients who attend public HIV Clinics in Rio de Janeiro, Brazil. JAIDS. 2004 Aug 1; **36(4):** 967-71.
36. Bangsberg DR. Less than 95% adherence to non-nucleoside reverse transcriptase inhibitor therapy can lead to viral suppression. Clin Infect Dis. 2006 Oct 1; **43(7):** 939-41.
37. Nachega JB, Hislop M, Dowdy DW, Lo M, Omer SB, Regensberg L et al. Adherence to highly active antiretroviral therapy assessed by pharmacy claims predicts survival in HIV-infected South African adults. JAIDS. 2006 Sep; **43(1):** 78-84.
38. Nachega JB, Morroni C, Zuniga JM, Schechter M, Rockstroh J, Solomon S et al. HIV treatment adherence, patient health literacy, and health care provider-patient communication: results from the 2010 AIDS Treatment for Life International Survey. JIAPAC. 2012 Mar 30; **11(2):** 128-33.
39. Levine AJ, Hinkin CH, Castellon SA, Mason KI, Lam MN, Perkins A et al. Variations in patterns of highly active antiretroviral therapy (HAART) adherence. AIDS Behav. 2005 Sep 1; **9(3):** 355-62.
40. Johnson MO, Charlebois E, Morin SF, Catz SL, Goldstein RB, Remien RH et al. Perceived adverse effects of antiretroviral therapy. J Pain Symptom Manage. 2005 Feb; **29(2):** 193-205.
41. Muyingo SK, Walker AS, Reid A, Munderi P, Gibb DM, Ssali F et al. Patterns of individual and population-level adherence to antiretroviral therapy and risk factors for poor adherence in the first year of the DART trial in Uganda and Zimbabwe. JAIDS. 2008 Aug 1; **48(4):** 468-75.
42. Ramadhani HO, Thielman NM, Landman KZ, Ndosi EM, Gao F, Kirchherr JL et al. Predictors of incomplete adherence, virologic failure, and antiviral drug resistance among HIV-infected adults receiving antiretroviral therapy in Tanzania. Clin Infect Dis. 2007 Dec 1; **45(11):** 1492-8.
43. Mills EJ, Nachega JB, Buchan I, Orbinski J, Attaran A, Singh S et al. Adherence to antiretroviral therapy in sub-Saharan Africa and North America: a meta-analysis. JAMA. 2006 Aug 9; **296(6):** 679-90.
44. Weidle PJ, Wamai N, Solberg P, Liechty C, Sendagala S, Were W et al. Adherence to

antiretroviral therapy in a home-based AIDS care programme in rural Uganda. Lancet. 2006 Nov 4; **368(9547):** 1587-94.

45. Kunutsor S, Evans M, Thoulass J, Walley J, Katabira E, Newell JN et al. Ascertaining baseline levels of antiretroviral therapy adherence in Uganda: a multimethod approach. JAIDS. 2010 Oct 1; **55(2):** 221-4.
46. Harrigan PR, Hogg RS, Dong WWY, Yip B, Wynhoven B, Woodward J et al. Predictors of HIV drug-resistance mutations in a large antiretroviral-naive cohort initiating triple antiretroviral therapy. J Infect Dis. 2005 Feb 1; **191(3):** 339-47.
47. Steele RH, Keogh GL, Quin J, Fernando SL and Stojkova V. Mean cell volume (MCV) changes in HIV-positive patients taking nucleoside reverse transcriptase inhibitors (NRTIs): a surrogate marker for adherence. Int J STD AIDS. 2002 Nov 1; **13(11):** 748-54.
48. Romanelli F, Empey K and Pomeroy C. Macrocytosis as an indicator of medication (zidovudine) adherence in patients with HIV infection. AIDS Patient Care STDs. 2002 Sep; **16(9):** 405-11.
49. Kamya MR, Mayanja-Kizza H, Kambugu A, Bakeera-Kitaka S, Semitala F, Mwebaze-Songa P et al. Predictors of long-term viral failure among Ugandan children and adults treated with antiretroviral therapy. JAIDS. 2007 Oct 1; **46(2):** 187-93.
50. Kantor R, Diero L, DeLong A, Kamle L, Muyonga S, Mambo F et al. Misclassification of first-line antiretroviral treatment failure based on immunological monitoring of HIV infection in resource-limited settings. Clin Infect Dis. 2009 Aug 1; **49(3):** 454-62.
51. van Oosterhout JJG, Brown L, Weigel R, Kumwenda JJ, Mzinganjira D, Saukila N et al. Diagnosis of antiretroviral therapy failure in Malawi: poor performance of clinical and immunological WHO criteria. Trop Med Int Health. 2009 Aug; **14(8):** 856-61.
52. Reynolds SJ, Nakigozi G, Newell K, Ndyanabo A, Galiwongo R, Boaz I et al. Failure of immunologic criteria to appropriately identify antiretroviral treatment failure in Uganda. AIDS. 2009 Mar 27; **23(6):** 697-700.
53. Mee P, Fielding KL, Charalambous S, Churchyard GJ and Grant AD. Evaluation of the WHO criteria for antiretroviral treatment failure among adults in South Africa. AIDS. 2008 Oct 1; **22(15):** 1971-7.
54. Oyugi JH, Byakika-Tusiime J, Charlebois ED, Kityo C, Mugerwa R, Mugyenyi P et al. Multiple validated measures of adherence indicate high levels of adherence to generic HIV antiretroviral therapy in a resource-limited setting. JAIDS. 2004 Aug 15;

36(5): 1100-2.

55. Mocroft A, Phillips AN, Gatell J, Ledergerber B, Fisher M, Clumeck N et al. Normalisation of CD4 counts in patients with HIV-1 infection and maximum virological suppression who are taking combination antiretroviral therapy: an observational cohort study. Lancet. 2007 Aug 4; **370(9585):** 407-13.
56. Moore RD and Keruly JC. CD4+ cell count 6 years after commencement of highly active antiretroviral therapy in persons with sustained virologic suppression. Clin Infect Dis. 2007 Feb 1; **44(3):** 441-6.
57. García F, de Lazzari E, Plana M, Castro P, Mestre G, Nomdedeu M et al. Long-term CD4+ T-cell response to highly active antiretroviral therapy according to baseline CD4+ T-cell count. JAIDS. 2004 Jun 1; **36(2):** 702-13.
58. Protopopescu C, Raffi F, Roux P, Reynes J, Dellamonica P, Spire B et al. Factors associated with non-adherence to long-term highly active antiretroviral therapy: a 10 year follow-up analysis with correction for the bias induced by missing data. J Antimicrob Chemother. 2009 Sep 1; **64(3):** 599-606.
59. Nachega JB, Hislop M, Nguyen H, Dowdy DW, Chaisson RE, Regensberg L et al. Antiretroviral therapy adherence, virologic and immunologic outcomes in adolescents compared with adults in Southern Africa. JAIDS. 2009 May 1; **51(1):** 65-71.
60. Sherr L, Harding R, Lampe F, Johnson M, Anderson J, Zetler S et al. Clinical and behavioural aspects of aging with HIV infection. Psychol Health Med. 2009 May 1; **14(3):** 273-9.
61. Godin G, Côté J, Naccache H, Lambert LD and Trottier S. Prediction of adherence to antiretroviral therapy: a one-year longitudinal study. AIDS Care. 2005 May 1; **17(4):** 493-504.
62. Nachega JB, Stein DM, Lehman DA, Hlatshwayo D, Mothopeng R, Chaisson RE et al. Adherence to antiretroviral therapy in HIV-infected adults in Soweto, South Africa. AIDS Res Hum Retroviruses. 2004 Oct; **20(10):** 1053-6.
63. Mills EJ, Nachega JB, Bangsberg DR, Singh S, Rachlis B, Wu P et al. Adherence to HAART: A systematic review of developed and developing nation patient-reported barriers and facilitators. PLoS Med. 2006 Nov 21; **3(11):** 2039-2063.
64. Buscher A, Hartman C, Kallen MA and Giordano TP. Impact of antiretroviral dosing frequency and pill burden on adherence among newly diagnosed, antiretroviral-naïve HIV patients. Int J STD AIDS. 2012 May 1; **23(5):** 351-5.

65. Heckman BD, Catz SL, Heckman TG, Miller JG and Kalichman SC. Adherence to antiretroviral therapy in rural persons living with HIV disease in the United States. AIDS Care. 2004 Feb 1; **16(2):** 219-30.
66. Schneider J, Kaplan S, Greenfield S, Li W and Wilson I. Better physician-patient relationships are associated with higher reported adherence to antiretroviral therapy in patients with HIV infection. J Gen Intern Med. 2004 Nov; **19(11):** 1096-1103.
67. Rosen S, Fox MP and Gill CJ. Patient retention in antiretroviral therapy programs in Sub-Saharan Africa: a systematic review. PLoS Med. 2007 Oct 16; **4(10):** 1691-1700.
68. Sanne I, Westreich D, Macphail A, Rubel D, Majuba P and Van Rie A. Long term outcomes of antiretroviral therapy in a large HIV/AIDS care clinic in urban South Africa: a prospective cohort study. J Int AIDS Soc. 2009 Dec 17; **12(1):** 1-11.
69. Geis S, Maboko L, Saathoff E, Hoffmann O, Geldmacher C, Mmbando D et al. Risk factors for HIV-1 infection in a longitudinal, prospective cohort of adults from the Mbeya region, Tanzania. JAIDS. 2011 Apr 15; **56(5):** 453-9.
70. Naidoo S,Taylor M and Jinabhai CC. Critical risk factors driving the tuberculosis epidemic in Kwazulu-Natal South Africa. SAJEI. 2007; **22(2,3):** 45-9.
71. Russell EC, Charalambous S, Pemba L, Churchyard GJ, Grant AD and Fielding K. Low haemoglobin predicts early mortality among adults starting antiretroviral therapy in an HIV care programme in South Africa: a cohort study. BMC Public Health. 2010 Jul 23; **10(1):** 1-8
72. Tedaldi EM, Brooks JT, Weidle PJ, Richardson JT, Baker RK, Buchacz K et al. Increased Body Mass Index does not alter response to initial Highly Active Antiretroviral Therapy in HIV-1 – infected patients. J Acquir Immune Defic Syndr. 2006 Sep; **43(1):** 35-41.
73. Hambisa MT, Ali A and Dessie Y. Determinants of mortality among HIV-positives after initiating antiretroviral therapy in Western Ethiopia: A hospital-based retrospective cohort Study. ISRN AIDS. 2013; **2013(2013):** 1-7.
74. Tran DA, Ngo AD, Shakeshaft A, Wilson DP, Doran C and Zhang L. Trends in and determinants of loss to follow up and early mortality in a rapid expansion of the antiretroviral treatment program in Vietnam: Findings from 13 outpatient clinics. PLoS One. 2013; **8(9):** 1-8.

APPENDIX 1: Summary of different types of adherence measurement methods

Method	Surrogate or non-surrogate	Description	Advantages	Disadvantages
Appointment keeping	Surrogate	Adherence measured by whether a patient attends a visit on schedule or not.	1. Cheap and easy method of assessing adherence.	1. On-time drug pick up does not guarantee adherence[46]. 2. An adherent patient who misses an appointment due to having enough medication left over cannot be classified as non-adherent.
CD4 count	Surrogate	A measure of cluster of differentiation number 4 helper T lymphocytes/mm^3 of blood.	1. Comparison with a previous value tells us if the patient is responding well to treatment (immunologically) or not. 2. Cheaper than a viral load test.	1. Detects long-term adherence but is poor at detecting recent adherence levels. 2. There is no consensus on the expected rate of CD4 count rise for patients on ART[16, 55 – 57]. 3. Poor predictor of virological treatment failure[50 – 53]. 4. Other factors such as malabsorption, drug interactions, diabetes mellitus and the acquisition of a drug-resistant HIV strain may contribute to a persistently low CD4 despite good adherence. 5. Invasive.

Directly observed therapy	Non-surrogate	A second person observes the patient taking medication.	1. Direct / non-surrogate measure of adherence.	1. May not be practical as it requires the guaranteed presence of a constant treatment supporter who is willing and able to directly observe medication being taken correctly.
Drug levels in hair	Non-surrogate	A small thatch of hair is cut as close as possible to the scalp from the occiput[24]. The hair sample is then tested for drug levels.	1. Direct / non-surrogate measure of adherence. 2. May allow for detection or prevention of drug toxicity, which can lead to poor adherence[11].	1. Expensive. 2. Detect long-term adherence but are poor at detecting recent adherence levels. 3. Levels may be low for reasons other than poor adherence e.g diet, drug interactions[11].
Mean cell volume (MCV)	Surrogate	Mean volume of red blood cells, a parameter obtained from a full blood count (FBC).	1. Cheaper than a viral load test.	1. Useful only for patients on d4T or AZT-based regimens. 2. Detects long-term adherence but is poor at detecting recent adherence levels. 3. Other factors such as vitamin B12 and/or folic acid deficiency, hepatic disease and alcoholism may also cause an elevated MCV. 4. Invasive.

Micro-electronic monitoring (MEM)	Surrogate	Adherence calculated by electronic caps attached to the medication bottle to digitally record adherence based on how often and at what times of the day, the medication bottle was opened.	1. Adherence measures correlate closely with the viral load[11].	1. Expensive. 2. Vulnerable to technological malfunction[11]. 3. Tells you the time the pill bottle was opened but not whether medication was actually taken or whether food restrictions were adhered to. 4. May underestimate adherence rates among patients who remove an additional dose from their pill bottle with the intention of taking it later, making the use of pill boxes problematic[11, 33].
Pill count	Surrogate	Adherence calculated as the percentage of pills expected to be taken that were actually taken.	1. Cheap method of assessing adherence. 2. Quantifiable measure of adherence.	1. Time-consuming. 2. Prone to error as the patient may remove medication and throw it away so this measure may be more accurate if it is unannounced[36, 44]. 3. Does not tell you whether the medication was taken at the correct time with the appropriate dietary requirements.

Self-reported adherence	Surrogate	Patient's assessment of pills taken over a certain period e.g in the last week computed to a percentage.	1. Cheap and easy method of assessing adherence. 2. Computation of patient's assessment to a percentage adherence does not take as long as pill count followed by calculation of the adherence percentage.	1. Subject to errors from patient recall (recall bias) and dishonesty[11, 12]. 2. Does not encompass the full definition of adherence, which includes following other recommendations pertaining to food, etc. 3. No standardized questions[11].
Serum drug levels	Non-surrogate	Serum drug levels measured from a sample of the patient's blood.	1. Direct / non-surrogate measure of adherence. 2. Plasma concentration directly determines virologic response[11]. 3. May allow for detection or prevention of drug toxicity, which can lead to poor adherence[11].	1. Expensive. 2. Detect only recent adherence levels hence vulnerable to white coat adherence[11, 24, 26]. 3. Cannot routinely measure nucleoside reverse transcriptase inhibitor (NRTI) levels because active moieties are intracellular[11, 16]. 4. There is substantial intra-patient variation in drug concentrations[11, 16]. 5. Levels may be low for reasons other than poor adherence e.g diet, drug interactions[11]. 6. Invasive[11].

Serum lactate levels	Surrogate	Measurement of lactic acid levels from a sample of the patient's blood (plasma).	1. Assists in the prevention and detection of drug toxicity.	1. Useful only for patients on d4T or ddI. 2. Serum lactate elevation does not occur in all patients. 3. Detect long-term adherence but are poor at detecting recent adherence levels. 4. Stringent lab requirements needed for accurate results i.e the patient must be well hydrated and must not have exercised within the past 24 hours and blood must be collected without a tourniquet or clenched feast into a pre-chilled fluoride-oxalate tube and transported on ice for processing within 4 hours[16]. 5. Other factors such as uncontrolled diabetes mellitus, renal and hepatic disease may contribute to elevated lactate levels. 6. Invasive.
Serum lipid levels	Surrogate	Measurement of cholesterol and triglyceride levels from a sample of the patient's blood (plasma). Lipid	1. Assists in the prevention and detection of drug toxicity.	1. Useful only for PIs other than unboosted ATV[16, 23]. 2. Detect long-term adherence but are poor at detecting recent adherence levels. 3. Other factors such as metabolic

		levels rise over time in patients on PIs[16, 23].		syndrome may contribute to elevated lipid levels. 4. Invasive.
Viral load	Surrogate	A measure of HIV viral ribonucleic acid copies/ml of blood (plasma).	1. Most reliable surrogate marker of adherence[14, 20, 21, 54]. 2. Tells us if the patient is responding well to treatment (virologically) or not. 3. Benchmark against which the efficacy of other routinely collected and more affordable markers can be assessed as measures of adherence.	1. Expensive. 2. Other factors such as malabsorption, drug interactions, diabetes mellitus and the acquisition of a drug-resistant HIV strain may contribute to a persistently detectable viral load despite good adherence. 3. Invasive.
Visual analogue scale (VAS)	Surrogate	Patients estimate their adherence by placing a mark on the scale.	1. Cheap and easy method of assessing adherence. 2. Adherence assessment does not take as long as pill count followed by calculation of the adherence percentage.	1. Subject to errors from patient recall (recall bias) and dishonesty. 2. Does not encompass the full definition of adherence, which includes following other recommendations pertaining to food etc.

The table above shows the different types of adherence measurement methods, their advantages and disadvantages. All reference numbers (superscripts) in the table above correspond to references in the reference list above the table.

APPENDIX 2: Adherence by self report at Themba Lethu Clinic

Adherence Category	Response to "How many of your pills have you taken in the past week?"
> 90%	All
> 60-90%	Most
> 30-60%	About half
10-30%	A few
< 10%	None

APPENDIX 3A

Tables showing formulae for calculating sensitivity, specificity, PPV and NPV of change in MCV at 6 months and change in CD4 count at 6 months

Change in MCV at 6 months	**Viral load < 400 copies/ml**	**Viral load ≥ 400 copies/ml**
≥ 14.5fL	a	b
< 14.5fL	c	d

Change in CD4 count at 6 months	**Viral load < 400 copies/ml**	**Viral load ≥ 400 copies/ml**
≥ expected	a	b
< expected	c	d

Sensitivity = d/(b+d)
Specificity = a/(a+c)
PPV = d/(c+d)
NPV = a/(a+b)

APPENDIX 3B

Table showing formulae for calculating sensitivity, specificity, PPV and NPV of each category of change in CD4 count stratified by change in MCV at 6 months for all patients with a documented viral load result

	Viral load < 400 copies/ml	**Viral load ≥ 400 copies/ml**
Change in CD4 count ≥ expected and change in MCV ≥ 14.5fL	a	b
Change in CD4 count ≥ expected and change in MCV< 14.5fL	c	d
Change in CD4 count < expected and change in MCV ≥ 14.5fL	e	f
Change in CD4 count < expected and change in MCV< 14.5fL	g	h

-Sensitivity for change in CD4 count ≥ expected and change in MCV < 14.5fL: d/(b+d)
-Sensitivity for change in CD4 count < expected and change in MCV ≥ 14.5fL: f/(b+f)
-Sensitivity for change in CD4 count < expected and change in MCV < 14.5fL: h/(b+h)

-Specificity for change in CD4 count ≥ expected and change in MCV < 14.5fL: a/(a+c)
-Specificity for change in CD4 count < expected and change in MCV ≥ 14.5fL: a(a+e)
-Specificity for change in CD4 count < expected and change in MCV < 14.5fL: a/(a+g)

-PPV for change in CD4 count ≥ expected and change in MCV < 14.5fL: d/(c+d)
-PPV for change in CD4 count < expected and change in MCV ≥ 14.5fL: f/(e+f)
-PPV for change in CD4 count < expected and change in MCV < 14.5fL: h/(g+h)

-NPV for change in CD4 stratified by change in MCV: a/(a+b)

APPENDIX 3C

Table showing formulae for calculating overall sensitivity, specificity, PPV and NPV of change in CD4 count stratified by change in MCV at 6 months for all patients with a documented viral load result

	Viral load < 400 copies/ml	**Viral load ≥ 400 copies/ml**
Change in CD4 count ≥ expected and change in MCV ≥ 14.5fL	a	b
Change in CD4 count ≥ expected and change in MCV< 14.5fL or Change in CD4 count < expected and change in MCV ≥ 14.5fL or Change in CD4 count < expected and change in MCV< 14.5fL	c	d

Sensitivity: d/(b+d)
Specificity: a/(a+c)
PPV: d/(c+d)
NPV: a/(a+b)

APPENDIX 3D

Table showing formulae for calculating sensitivity, specificity, PPV and NPV for pregnancy during the first 6 months on ART for patients on AZT with a documented viral load result

Pregnant during the first 6 months on ART	**Viral load < 400 copies/ml**	**Viral load ≥ 400 copies/ml**
Never	a	b
At baseline	No observations	No observations
During follow-up	c	d

Sensitivity: d/(b+d)
Specificity: a/(a+c)
PPV: d/(c+d)
NPV: a/(a+b)

APPENDIX 4A

Baseline demographic characteristics after multiple imputation for all eligible patients and eligible patients on AZT-based regimens with a documented viral load result

Characteristics	All eligible patients			All eligible patients on AZT-based regimens*		
	Total N = 143200	**Viral load < 400 copies/ml N = 116120**	**Viral load ≥ 400 copies/ml N = 27080**	**Total N = 3540**	**Viral load < 400 copies/ml N = 2860**	**Viral load ≥ 400 copies/ml N = 680**
	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)
Gender						
Females	90460 (63.2%)	74320 (64%)	16140 (59.6%)	1940 (54.8%)	1640 (57.3%)	300 (44.1%)
Males	52740 (36.8%)	41800 (36%)	10940 (40.4%)	1600 (45.2%)	1220 (42.7%)	380 (55.9%)
Age at initiation in years						
≤ 36.7	72060 (50.3%)	58020 (50.0%)	14040 (51.8%)	1200 (33.9%)	860 (30.1%)	340 (50%)
> 36.7	71140 (49.7%)	58100 (50.0%)	13040 (48.2%)	2340 (66.1%)	2000 (69.9%)	340 (50%)
Education level						
≥ Secondary school	113849 (79.5%)	91769 (79.0%)	22080 (81.5%)	2656 (75.0%)	2149 (75.1%)	507 (74.6%)
Primary school	24480 (17.1%)	20259 (17.5%)	4221 (15.6%)	723 (20.4%)	571 (20.0%)	152 (22.3%)
< Primary school	4871 (3.4%)	4092 (3.5%)	779 (2.9%)	161 (4.6%)	140 (4.9%)	21 (3.1%)
Professional status						
Employed	74835 (52.3%)	61283 (52.8%)	13552 (50.0%)	1810 (51.1%)	1440 (50.3%)	370 (54.4%)
Unemployed	68365 (47.7%)	54837 (47.2%)	13528 (50.0%)	1730 (48.9%)	1420 (49.7%)	310 (45.6%)

* On AZT-based regimens throughout the study period. AZT: Zidovudine.

APPENDIX 4B

Baseline clinical characteristics after multiple imputation for all eligible patients and eligible patients on AZT-based regimens with a documented viral load result

Characteristics	All eligible patients			All eligible patients on AZT-based regimens*		
	Total N = 143200	**Viral load < 400 copies/ml N = 116120**	**Viral load ≥ 400 copies/ml N = 27080**	**Total N = 3540**	**Viral load < 400 copies/ml N = 2860**	**Viral load ≥ 400 copies/ml N = 680**
	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)
BMI at initiation in kg/m^2						
< 18.5	23714 (16.6%)	18691 (16.1%)	5023 (18.5%)	509 (14.4%)	433 (15.1%)	76 (11.2%)
18.5 – < 25	78567 (54.9%)	63783 (54.9%)	14784 (54.6%)	1880 (53.1%)	1433 (50.1%)	447 (65.7%)
25 – < 30	27884 (19.5%)	22799 (19.6%)	5085 (18.9%)	588 (16.6%)	470 (16.4%)	118 (17.4%)
≥ 30	13035 (9.0%)	10847 (9.4%)	2188 (8.0%)	563 (15.9%)	524 (18.4%)	39 (5.7%)
Baseline haemoglobin						
≤ 8 g/dl	7122 (5.0%)	5781 (5.0%)	1341 (5.0%)	35 (1.0%)	29 (1.0%)	6 (0.9%)
> 8g/dl	136078 (95.0%)	110339 (95.0%)	25739 (95.0%)	3505 (99.0%)	2831 (99.0%)	674 (99.1%)
MCV at baseline						
< 80fL	17741 (12.4%)	14881 (12.8%)	2860 (10.6%)	341 (9.6%)	238 (8.3%)	103 (15.2%)
80 – 100fL	119706 (83.6%)	96636 (83.2%)	23070 (85.2%)	2413 (68.2%)	1943 (67.9%)	470 (69.1%)
> 100fL	5753 (4.0%)	4603 (4.0%)	1150 (4.2%)	786 (22.2%)	679 (23.7%)	107 (15.7%)
Baseline CD4 count in cells/mm^3						
> 200	23157 (16.2%)	19064 (16.4%)	4093 (15.1%)	1178 (33.3%)	922 (32.2%)	256 (37.6%)
101 – 200	48899 (34.1%)	40550 (34.9%)	8349 (30.8%)	986 (27.8%)	826 (28.9%)	160 (23.5%)
51 – 100	27984 (19.6%)	22545 (19.4%)	5439 (20.1%)	618 (17.5 %)	508 (17.8%)	110 (16.2%)
≤ 50	43160 (30.1%)	33961 (29.3%)	9199 (34.0%)	758 (21.4%)	604 (21.1%)	154 (22.7%)
WHO stage at baseline						
I	54522 (38.1%)	45530 (39.2%)	8992 (33.2%)	1943 (54.9%)	1557 (54.4%)	386 (56.8%)
II	28769 (20.1%)	23047 (19.9%)	5722 (21.1%)	513 (14.5%)	424 (14.8%)	89 (13.1%)
III	42657 (29.8%)	33940 (29.2%)	8717 (32.2%)	760 (21.5%)	583 (20.4%)	177 (26.0%)
IV	17252 (12.0%)	13603 (11.7%)	3649 (13.5%)	324 (9.1%)	296 (10.4%)	28 (4.1%)

* On AZT-based regimens throughout the study period. AZT: Zidovudine. BMI: Body Mass Index. MCV: Mean cell volume. WHO: World Health Organization.

APPENDIX 4C

6 month clinical characteristics after multiple imputation for all eligible patients and eligible patients on AZT-based regimens with a documented viral load result

Characteristics	All eligible patients			All eligible patients on AZT-based regimens*		
	Total N = 143200	**Viral load < 400 copies/ml N = 116120**	**Viral load ≥ 400 copies/ml N = 27080**	**Total N = 3540**	**Viral load < 400 copies/ml N = 2860**	**Viral load ≥ 400 copies/ml N = 680**
	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)
Self-reported adherence[δ]						
> 90%	127078 (88.7%)	103344 (89%)	23734 (87.6%)	3126 (88.3%)	2533 (88.6%)	593 (87.2%)
> 60 – 90%	14670 (10.2%)	11769 (10.1%)	2901 (10.7%)	374 (10.5%)	307 (10.7%)	67 (9.9%)
≤ 60%	1452 (1.1%)	1007 (0.9%)	445 (1.7%)	40 (1.2%)	20 (0.7%)	20 (2.9%)
Number of missed visits						
0	92340 (64.5%)	74620 (64.3%)	17720 (65.4%)	2140 (60.5%)	1760 (61.5%)	380 (55.9%)
1	36080 (25.2%)	29500 (25.4%)	6580 (24.3%)	880 (24.9%)	780 (27.3%)	100 (14.7%)
≥ 2	12460 (8.7%)	10180 (8.8%)	2280 (8.5%)	320 (9.0%)	200 (7.0%)	120 (17.7%)
Missing	2320 (1.6%)	1820 (1.5%)	500 (1.8%)	200 (5.6%)	120 (4.2%)	80 (11.7%)
Number of missed ARV visits						
0	106534 (74.4%)	86538 (74.5%)	19996 (73.8%)	2543 (71.8%)	2067 (72.3%)	476 (70%)
1	30576 (21.4%)	24888 (21.4%)	5688 (21%)	790 (22.3%)	631 (22.1%)	159 (23.4%)
≥ 2	6090 (4.2%)	4694 (4.1%)	1396 (5.2%)	207 (5.9%)	162 (5.6%)	45 (6.6%)
Number of missed medical visits						
0:	108874 (76.0%)	88623 (76.3%)	20251 (74.8%)	2640 (74.6%)	2202 (77.0%)	438 (64.4%)
1:	28828 (20.1%)	23264 (20.0%)	5564 (20.5%)	796 (22.5%)	615 (21.5%)	181 (26.6%)
≥ 2:	5498 (3.9%)	4233 (3.7%)	1265 (4.7%)	104 (2.9%)	43 (1.5%)	61 (9.0%)

BMI at 6 months in kg/m^2						
< 18.5	10261 (7.2%)	7831 (6.7%)	2430 (9.0%)	275 (7.8%)	229 (8.0%)	46 (6.8%)
18.5 – < 25	75959 (53.0%)	61253 (52.8%)	14706 (54.3%)	1769 (50.0%)	1351 (47.2%)	418 (61.5%)
25 – < 30	38405 (26.8%)	31516 (27.1%)	6889 (25.4%)	958 (27.1%)	791 (27.7%)	167 (24.6%)
≥ 30	18575 (13.0%)	15520 (13.4%)	3055 (11.9%)	538 (15.1%)	489 (17.2%)	49 (7.1%)
Haemoglobin at 6 months						
≤ 8 g/dl	1002 (0.7%)	741 (0.6%)	261 (1.0%)	20 (0.6%)	0 (0%)	20 (2.9%)
> 8g/dl	142198 (99.3%)	115379 (99.4%)	26819 (99.0%)	3520 (99.4%)	2860 (100%)	660 (97.1%)
MCV at 6 months						
< 80fL	4640 (3.2%)	3725 (3.2%)	915 (3.4%)	42 (1.2%)	20 (0.7%)	22 (3.2%)
80 – 100fL	69146 (48.3%)	51182 (44.1%)	17964 (66.3%)	718 (20.3%)	377 (13.2%)	341 (50.2%)
> 100fL	69414 (48.5%)	61213 (52.7%)	8201 (30.3%)	2780 (78.5%)	2463 (86.1%)	317 (46.6%)
Change in MCV at 6 months						
≥ 14.5fL	57281 (40%)	51753 (44.6%)	5528 (20.4%)	1765 (49.9%)	1562 (54.6%)	203 (29.9%)
< 14.5fL	85919 (60%)	64637 (55.4%)	21552 (79.6%)	1775 (50.1%)	1298 (45.4%)	477 (70.1%)
Change in CD4 count at 6 months in cells/mm^3						
≥ expected	111903 (78.1%)	93757 (80.7%)	18146 (67.0%)	2489 (70.3%)	2158 (75.5%)	331 (48.7%)
< expected	31297 (21.9%)	22363 (19.3%)	8934 (33.0%)	1051 (29.7%)	702 (24.5%)	349 (51.3%)
Change in CD4 count at 6 months stratified by change in MCV at 6 months						
-Change in CD4 count ≥ expected and change in MCV ≥ 14.5fL:	45323 (31.7%)	41384 (35.6%)	3939 (14.6%)	1346 (38.0%)	1229 (43.0%)	117 (17.2%)
-Change in CD4 count ≥ expected and change in MCV < 14.5fL:	66580 (46.5%)	52373 (45.1%)	14207 (52.5%)	1143 (32.3%)	929 (32.5%)	214 (31.5%)
-Change in CD4 count < expected and change in MCV ≥ 14.5fL:	11958 (8.4%)	10369 (8.9%)	1589 (5.9%)	419 (11.8%)	333 (11.6%)	86 (12.7%)
-Change in CD4 count < expected and change in MCV < 14.5fL:	19339 (13.4)	11994 (10.4%)	7345 (27.0%)	632 (17.9%)	369 (12.9%)	263 (38.6%)
Regimen change						
No	128900 (90.0%)	104200 (89.7%)	24700 (91.2%)	3420 (96.6%)	2800 (97.9%)	620 (91.2%)
Yes	14300 (10.0%)	11920 (10.3%)	2380 (8.8%)	120 (3.4%)	60 (2.1%)	60 (8.8%)

TB[α]: On TB treatment during first 6 months on ART:						
Never	107100 (74.8%)	86900 (74.8%)	20200 (74.6%)	3000 (84.7%)	2400 (83.9%)	600 (88.2%)
At baseline	29860 (20.9%)	24080 (20.7%)	5780 (21.3%)	480 (13.6%)	400 (14.0%)	80 (11.8%)
During follow-up	6240 (4.3%)	5140 (4.5%)	1100 (4.1%)	60 (1.7%)	60 (2.1%)	0 (0%)
Pregnancy[α]: Pregnancy during first 6 months on ART:						
Never	85880 (60.0%)	70680 (60.9%)	15200 (56.1%)	1840 (52.0%)	1600 (55.9%)	240 (35.3%)
At baseline	1640 (1.1%)	1340 (1.2%)	300 (1.1%)	0 (0%)	0 (0%)	0 (0%)
During follow-up	2940 (2.1%)	2300 (1.9%)	640 (2.4%)	100 (2.8%)	40 (1.4%)	60 (8.8%)
Missing[γ]	52740 (36.8%)	41800 (36.0%)	10940 (40.4%)	1600 (45.2%)	1220 (42.7%)	380 (55.9%)
Adverse events or side effects						
No	139100 (97.1%)	112360 (96.8%)	26740 (98.7%)	3520 (99.4%)	2840 (99.3%)	680 (100%)
Yes	4100 (2.9%)	3760 (3.2%)	340 (1.3%)	20 (0.6%)	20 (0.7%)	0 (0%)

[*] On AZT-based regimens throughout the study period. [δ] The last self-reported adherence levels on, or after, the 3 month visit (70 days post-initiation) and not more than 3 months (110 days) before the date of the viral load but at least 28 days before the viral load was done. [α] Variable has baseline and 6 month components. [γ] Male patients. ART: Antiretroviral Therapy. ARV: Antiretroviral. AZT: Zidovudine. BMI: Body Mass Index. MCV: Mean cell volume. TB: Tuberculosis.

APPENDIX 5A

Baseline demographic characteristics and their associations with poor adherence for all eligible patients and eligible patients on AZT-based regimens with a documented viral load done 90 – 270 days after ART initiation

Characteristics	All eligible patients				All eligible patients on AZT-based regimens*			
	Total N = 11724	**Viral load < 400 copies/ml N = 10011**	**Viral load ≥ 400 copies/ml N = 1713**	**P-value**	**Total N = 285**	**Viral load < 400 copies/ml N = 247**	**Viral load ≥ 400 copies/ml N = 38**	**P-value**
	n (%)	n (%)	n (%)		n (%)	n (%)	n (%)	
Gender								
Females	7409 (63.2%)	6391 (63.8%)	1018 (59.4%)	0.000	153 (53.7%)	138 (55.9%)	15 (39.5%)	0.059
Males	4315 (36.8%)	3620 (36.2%)	695 (40.6%)		132 (46.3%)	109 (44.1%)	23 (60.5%)	
Age at initiation in years								
Median (IQR)	36.7 (31.5 – 43.3)	36.8 (31.5 – 43.3)	36.3 (31.1 – 43)	0.0644[þ]	39.3 (33.6 – 45.7)	39.3 (33.8 – 45.8)	39.9 (31.1 – 45.6)	0.5245[þ]
Age at initiation in years								
≤ median (36.7)[α]	5893 (50.3%)	5002 (50.0%)	891 (52.0%)	0.117	104 (36.5%)	90 (36.4%)	14 (36.8%)	0.962
> median (36.7)	5831 (49.7%)	5009 (50.0%)	822 (48.0%)		181 (63.5%)	157 (63.6%)	24 (63.2%)	
Education level								
≥ Secondary school	7062 (60.2%)	5996 (59.9%)	1066 (62.2%)	0.215	148 (51.9%)	131 (53.0%)	17 (44.7%)	0.441[ð]
Primary school	1484 (12.7%)	1265 (12.6%)	219 (12.8%)		43 (15.1%)	35 (14.2%)	8 (21.1%)	
< Primary school	343 (2.9%)	303 (3.0%)	40 (2.3%)		11 (3.9%)	10 (4.1%)	1 (2.6%)	
Missing	2835 (24.2%)	2447 (24.5%)	388 (22.7%)		83 (29.1%)	71 (28.7%)	12 (31.6%)	
Professional status								
Employed	5913 (50.4%)	5086 (50.8%)	827 (48.3%)	0.092	133 (46.7%)	116 (47.0%)	17 (44.7%)	0.839
Unemployed	5639 (48.1%)	4788 (47.8%)	851 (49.7%)		147 (51.6%)	127 (51.4%)	20 (52.6%)	
Missing	172 (1.5%)	137 (1.4%)	35 (2.0%)		5 (1.7%)	4 (1.6%)	1 (2.7%)	

* On AZT-based regimens throughout the study period. [α] The median age at ART initiation for all eligible patients with a documented viral load result. [þ] Wilcoxon rank sum test. [ð] Fisher's exact test. AZT: Zidovudine. IQR: Interquartile range.

Note: Except where otherwise stated, the chi-square test was used.

APPENDIX 5B

Baseline clinical characteristics and their associations with poor adherence for all eligible patients and eligible patients on AZT-based regimens with a documented viral load done 90 – 270 days after ART initiation

Characteristics	All eligible patients				All eligible patients on AZT-based regimens*			
	Total N = 11724	**Viral load < 400 copies/ml N = 10011**	**Viral load ≥ 400 copies/ml N = 1713**	**P-value**	**Total N = 285**	**Viral load < 400 copies/ml N = 247**	**Viral load ≥ 400 copies/ml N = 38**	**P-value**
	n (%)	n (%)	n(%)		n (%)	n (%)	n (%)	
BMI at initiation in kg/m^2								
Median (IQR)	22 (19 – 25)	22 (19 – 25)	21.4 (19 – 25)	0.0041[þ]	22.0 (19.5 – 26.3)	22.0 (19.5 – 28)	21.5 (19.8 – 24)	0.4221 [þ]
BMI at initiation in kg/m^2								
< 18.5	1736 (14.8%)	1473 (14.7%)	263 (15.4%)	0.103	35 (12.3%)	33 (13.4%)	2 (5.3%)	0.042[ð]
18.5 – < 25	5549 (47.3%)	4749 (47.4%)	800 (46.7%)		109 (38.2%)	90 (36.4%)	19 (50%)	
25 – < 30	1923 (16.4%)	1674 (16.7%)	249 (14.5%)		30 (10.5%)	28 (11.3%)	2 (5.3%)	
≥ 30	900 (7.7%)	788 (7.9%)	112 (6.5%)		38 (13.3%)	37 (15.0%)	1 (2.6%)	
Missing	1616 (13.8%)	1327 (13.3%)	289 (16.9%)		73 (25.7%)	59 (23.9%)	14 (36.8%)	
Baseline haemoglobin in g/dl								
Median (IQR)	11.7 (10.1 – 13.1)	11.7 (10.2 – 13.1)	11.6 (10.1 – 13)	0.2566[þ]	12.6 (11.6 – 14)	12.6 (11.2 – 14)	12.9 (12 – 14.2)	0.3854[þ]
Baseline haemoglobin								
≤ 8 g/dl	579 (4.9%)	498 (5.0%)	81 (4.7%)	0.958	1 (0.4%)	1 (0.4%)	0 (0%)	1.000[ð]
> 8g/dl	10108 (86.2%)	8686 (86.8%)	1422 (83.0%)		227 (79.6%)	198 (80.2%)	29 (76.3%)	
Missing	1037 (8.9%)	827 (8.2%)	210 (12.3%)		57 (20%)	48 (19.4%)	9 (23.7%)	
MCV at baseline in fL								
Median (IQR)	88.2 (83.6 – 92.2)	88.4 (83.9 – 92.3)	88.1 (83.5 – 92.1)	0.0001[þ]	90.9 (86.6 – 99.0)	90.5 (86.6 – 99)	92.5 (86.9 – 97.5)	0.5783[þ]
MCV at baseline								
< 80fL	1196 (10.2%)	1072 (10.7%)	124 (7.2%)	0.015	18 (6.3%)	16 (6.5%)	2 (5.3%)	1.000[ð]
80 – 100fL	8012 (68.3%)	6974 (69.7%)	1038 (60.6%)		147 (51.6%)	128 (51.8%)	19 (50%)	
> 100fL	376 (3.2%)	319 (3.2%)	57 (3.3%)		47 (16.5%)	41 (16.6%)	6 (15.8%)	
Missing	2140 (18.3%)	1646 (16.4%)	494 (28.9%)		73 (25.6%)	62 (25.1%)	11 (28.9%)	

Baseline CD4 count in cells/mm^3 Median (IQR)	99 (39 – 171)	100 (39 – 172)	94 (36 – 168)	0.0376[þ]	124 (55 – 209)	120 (55 – 207)	168 (61 – 244)	0.2628[þ]
Baseline CD4 count in cells/mm^3	1473 (12.6%)	1281 (12.8%)	192 (11.2%)	0.167	65 (22.8%)	55 (22.3%)	10 (26.3%)	0.619
> 200	3638 (31.0%)	3152 (31.5%)	486 (28.4%)		64 (22.5%)	55 (22.3%)	9 (23.7%)	
101 – 200	2069 (17.6%)	1766 (17.6%)	303 (17.7%)		42 (14.7 %)	39 (15.8%)	3 (7.9%)	
51 – 100	3154 (26.9%)	2685 (26.8%)	469 (27.4%)		47 (16.5%)	42 (17%)	5 (13.2%)	
≤ 50	1390 (11.9%)	1127 (11.3%)	263 (15.3%)		67 (23.5%)	56 (22.6%)	11 (28.9%)	
Missing								
WHO stage at baseline								
I	3850 (32.8%)	3397 (33.9%)	453 (26.4%)	0.000	111 (38.9%)	96 (38.9%)	15 (39.5%)	0.366[ð]
II	1926 (16.4%)	1659 (16.6%)	267 (15.6%)		20 (7.0%)	17 (6.9%)	3 (7.9%)	
III	3066 (26.2%)	2610 (26.1%)	456 (26.6%)		59 (20.7%)	52 (21.1%)	7 (18.4%)	
IV	1084 (9.2%)	918 (9.2%)	166 (9.7%)		20 (7.0%)	20 (8.1%)	0 (0%)	
Missing	1798 (15.4%)	1427 (14.2%)	371 (21.7%)		75 (26.4%)	62 (25.1%)	13 (34.2%)	

[*] On AZT-based regimens throughout the study period. [þ] Wilcoxon rank sum test. [ð] Fisher's exact test. AZT: Zidovudine. BMI: Body Mass Index. IQR: Interquartile range. MCV: Mean cell volume. WHO: World Health Organization.

Note: Except where otherwise stated, the chi-square test was used.

APPENDIX 5C

6 month clinical characteristics and their associations with poor adherence for all eligible patients and eligible patients on AZT-based regimens with a documented viral load done 90 – 270 days after ART initiation

Characteristics	All eligible patients				All eligible patients on AZT-based regimens*			
	Total N = 11724	Viral load < 400 copies/ml N = 10011	Viral load ≥ 400 copies/ml N = 1713	P-value	Total N = 285	Viral load < 400 copies/ml N = 247	Viral load ≥ 400 copies/ml N = 38	P- value
	n (%)	n (%)	n (%)		n (%)	n (%)	n (%)	
Self-reported adherence								
> 90%	5381 (45.9%)	4589 (45.8%)	792 (46.2%)	0.054	111 (38.9%)	101 (40.9%)	10 (26.3%)	0.045ð
> 60 – 90%	644 (5.5%)	548 (5.5%)	96 (5.6%)		8 (2.8%)	6 (2.4%)	2 (5.3%)	
≤ 60%	135 (1.1%)	105 (1.0%)	30 (1.8%)		5 (1.9%)	3 (1.2%)	2 (5.3%)	
Missing	5564 (47.5%)	4769 (47.7%)	795 (46.4%)		161 (56.4%)	137 (55.5%)	24 (63.1%)	
Number of missed visits								
0	7944 (67.8%)	6798 (67.9%)	1146 (66.9%)	0.882	181 (63.5%)	161 (65.2%)	20 (52.6%)	0.048
1	2790 (23.8%)	2379 (23.8%)	411 (24.0%)		66 (23.2%)	59 (23.9%)	7 (18.4%)	
≥ 2	846 (7.2%)	720 (7.2%)	126 (7.4%)		25 (8.8%)	18 (7.3%)	7 (18.4%)	
Missing	144 (1.2%)	114 (1.1%)	30 (1.7%)		13 (4.5%)	9 (3.6%)	4 (10.6%)	
Number of missed ARV visits								
0	9042 (77.1%)	7749 (77.4%)	1293 (75.5%)	0.203	203 (71.2%)	181 (73.3%)	22 (57.9%)	0.277ð
1	2145 (18.3%)	1822 (18.2%)	323 (18.9%)		55 (19.3%)	45 (18.2%)	10 (26.3%)	
≥ 2	382 (3.3%)	316 (3.2%)	66 (3.8%)		14 (5.0%)	12 (4.9%)	2 (5.2%)	
Missing	155 (1.3%)	124 (1.2%)	31 (1.8%)		13 (4.5%)	9 (3.6%)	4 (10.6%)	
Number of missed medical visits								
0:	9016 (76.9%)	7747 (77.4%)	1269 (74.1%)	0.013	207 (72.6%)	187 (75.7%)	20 (52.6%)	0.013ð
1:	2171 (18.5%)	1828 (18.3%)	343 (20.0%)		58 (20.4%)	47 (19.0%)	11 (28.9%)	
≥ 2:	382 (3.3%)	312 (3.1%)	70 (4.1%)		7 (2.5%)	4 (1.7%)	3 (7.9%)	
Missing:	155 (1.3%)	124 (1.2%)	31 (1.8%)		13 (4.5%)	9 (3.6%)	4 (10.6%)	

BMI at 6 months in kg/m^2 Median (IQR)	23.0 (21 – 27)	23.4 (21 – 27)	23 (20.9 – 26)	0.0002[þ]	24 (21.2 – 28)	24 (21.3 – 28.4)	23 (21 – 25)	0.2831[þ]
BMI at 6 months in kg/m^2 < 18.5 18.5 – < 25 25 – < 30 ≥ 30 Missing	606 (5.2%) 4277 (36.5%) 2102 (17.9%) 1050 (9.0%) 3689 (31.4%)	471 (4.7%) 3479 (34.8%) 1742 (17.4%) 878 (8.8%) 3441 (34.3%)	135 (7.9%) 798 (46.6%) 360 (21.0%) 172 (10.0%) 248 (14.5%)	0.010	17 (6.0%) 88 (30.9%) 48 (16.8%) 32 (11.2%) 100 (35.1%)	15 (6.1%) 70 (28.3%) 41 (16.6%) 30 (12.1%) 91 (36.9%)	2 (5.3%) 18 (47.4%) 7 (18.4%) 2 (5.3%) 9 (23.6%)	0.292[ð]
Haemoglobin at 6 months in g/dl Median (IQR)	13.2 (12 – 14.4)	13.3 (12.1 – 14.5)	13.1 (12 – 14.3)	0.0000[þ]	13.1 (12.0 – 14.6)	13.1 (12 – 14.6)	13.4 (11 – 14.6)	0.7142[þ]
Haemoglobin at 6 months ≤ 8 g/dl > 8g/dl Missing	56 (0.5%) 4990 (42.6%) 6678 (56.9%)	44 (0.4%) 4150 (41.5%) 5817 (58.1%)	12 (0.7%) 840 (49.0%) 861 (50.3%)	0.361	1 (0.4%) 115 (40.4%) 169 (59.2%)	0 (0%) 101 (40.9%) 146 (59.1%)	1 (2.6%) 14 (36.8%) 23 (60.6%)	0.129[ð]
MCV at 6 months in fL Median (IQR)	99.4 (92.9 – 106.4)	100.3 (93.7 – 107.2)	95.5 (90.5 – 102.1)	0.0000 [þ]	107.5 (102.4 – 113.5)	108.6 (103.2 – 113.7)	102 (96.4 – 112.5)	0.0158[þ]
MCV at 6 months < 80fL 80 – 100fL > 100fL Missing	213 (1.8%) 2786 (23.8%) 2721 (23.2%) 6004 (51.2%)	177 (1.8%) 2148 (21.5%) 2420 (24.2%) 5266 (52.5%)	36 (2.1%) 638 (37.2%) 301 (17.6%) 738 (43.1%)	0.000	2 (0.7%) 21 (7.4%) 118 (41.4%) 144 (50.5%)	2 (0.8%) 14 (5.7%) 107 (43.3%) 124 (50.2%)	0 (0%) 7 (18.4%) 11 (28.9%) 20 (52.7%)	0.026[ð]
Change in MCV at 6 months in fL Median (IQR)	12.2 (5.7 – 18.3)	13.2 (6.4 – 18.8)	7.1 (2.7 – 12.6)	0.0000[þ]	15.7 (4.6 – 20.3)	16.1 (4.2 – 21.6)	8.9 (6.2 – 17.6)	0.2722[þ]
Change in MCV at 6 months ≥ 14.5fL < 14.5fL Missing	1756 (15.0%) 2512 (21.4%) 7456 (63.6%)	1633 (16.3%) 2016 (20.1%) 6362 (63.6%)	123 (8.6%) 496 (40.8%) 1094 (50.6%)	0.000	51 (17.9%) 46 (16.1%) 188 (66.0%)	47 (19.0%) 39 (15.8%) 161 (65.2%)	4 (10.5%) 7 (18.4%) 27 (71.1%)	0.205[ð]

Change in CD4 count at 6 months in cells/mm^3								
Median (IQR)	128 (69 – 204)	133 (74 – 207)	99 (34 – 188)	0.0000 [þ]	96 (48 – 142)	103 (49 – 146)	52.5 (5.5 – 84.5)	0.0333[þ]
Change in CD4 count at 6 months in cells/mm^3								
≥ expected	4108 (35.0%)	3522 (35.2%)	586 (34.2%)	0.000	78 (27.4%)	73 (29.6%)	5 (13.2%)	0.018
< expected	1079 (9.2%)	814 (8.1%)	265 (15.5%)		32 (11.2%)	25 (10.1%)	7 (18.4%)	
Missing	6537 (55.8%)	5675 (56.7%)	862 (50.3%)		175 (61.4%)	149 (60.3%)	26 (68.4%)	
Change in CD4 count at 6 months stratified by change in MCV at 6 months								
-Change in CD4 count ≥ expected and change in MCV ≥ 14.5fL:	1283 (10.9%)	1202 (12.0%)	81 (4.7%)	0.000	33 (11.6%)	32 (13.0%)	1 (2.6%)	0.146[ð]
-Change in CD4 count ≥ expected and change in MCV < 14.5fL:	1776 (15.1%)	1490 (14.9%)	286 (16.7%)		27 (9.5%)	24 (9.7%)	3 (7.9%)	
-Change in CD4 count < expected and change in MCV ≥ 14.5fL:	331 (2.8%)	311 (3.1%)	20 (1.2%)		10 (3.5%)	8 (3.2%)	2 (5.3%)	
-Change in CD4 count < expected and change in MCV < 14.5fL:	495 (4.2%)	344 (3.4%)	151 (8.8%)		16 (5.6%)	13 (5.3%)	3 (7.9%)	
Missing:	7839 (67.0%)	6664 (66.6%)	1175 (68.6%)		199 (69.8%)	170 (68.8%)	29 (76.3%)	
Regimen change								
No	10648 (90.8%)	9105 (90.9%)	1543 (90.1%)	0.247	278 (97.5%)	242 (98.0%)	36 (94.7%)	0.236[ð]
Yes	1076 (9.2%)	906 (9.1%)	170 (9.9%)		7 (2.5%)	5 (2.0%)	2 (5.3%)	
TB[α]: On TB treatment during first 6 months on ART:								
Never	8777 (74.9%)	7509 (75.0%)	1268 (74.0%)	0.560	217 (76.1%)	187 (75.7%)	30 (78.9%)	0.808[ð]
At baseline	2460 (21.0%)	2084 (20.8%)	376 (21.9%)		61 (21.4%)	53 (21.5%)	8 (21.1%)	
During follow-up	487 (4.1%)	418 (4.2%)	69 (4.1%)		7 (2.5%)	7 (2.8%)	0 (0%)	
Pregnancy[α]: Pregnancy during first 6 months on ART:								
Never	7079 (60.4%)	6109 (61.0%)	970 (56.6%)	0.857	145 (50.9%)	133 (53.8%)	12 (31.6%)	0.022[ð]
At baseline	131 (1.1%)	113 (1.1%)	18 (1.1%)		2 (0.7%)	2 (0.8%)	0 (0%)	

During follow-up	199 (1.7%)	169 (1.7%)	30 (1.7%)		6 (2.1%)	3 (1.2%)	3 (7.9%)	
Missing[γ]	4315 (36.8%)	3620 (36.2%)	695 (40.6%)		132 (46.3%)	109 (44.2%)	23 (60.5%)	
Adverse events or side effects								
No	11379 (97.1%)	9686 (96.8%)	1693 (98.8%)	0.000	284 (99.6%)	246 (99.6%)	38 (100%)	1.000[ð]
Yes	345 (2.9%)	325 (3.2%)	20 (1.2%)		1 (0.4%)	1 (0.4%)	0 (0%)	

[*] On AZT-based regimens throughout the study period. [þ] Wilcoxon rank sum test. [ð] Fisher's exact test. [α] Variable has baseline and 6 month components. [γ] Male patients. ART: Antiretroviral Therapy. ARV: Antiretroviral. AZT: Zidovudine. BMI: Body Mass Index. IQR: Interquartile range. MCV: Mean cell volume. TB: Tuberculosis.

Note: Except where otherwise stated, the chi-square test was used.

APPENDIX 6A

Baseline demographic characteristics after multiple imputation for all eligible patients and eligible patients on AZT-based regimens with a documented viral load done 90 – 270 days after ART initiation

Characteristics	All eligible patients			All eligible patients on AZT-based regimens*		
	Total N = 234480	**Viral load < 400 copies/ml N = 200220**	**Viral load ≥ 400 copies/ml N = 34260**	**Total N = 5700**	**Viral load < 400 copies/ml N = 4940**	**Viral load ≥ 400 copies/ml N = 760**
	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)
Gender						
Females	148180 (63.2%)	127820 (63.8%)	20360 (59.4%)	3060 (53.7%)	2760 (55.9%)	300 (39.5%)
Males	86300 (36.8%)	72400 (36.2%)	13900 (40.6%)	2640 (46.3%)	2180 (44.1%)	460 (60.5%)
Age at initiation in years						
≤ 36.7	117860 (50.3%)	100040 (50.0%)	17820 (52.0%)	2080 (36.5%)	1800 (36.4%)	280 (36.8%)
> 36.7	116620 (49.7%)	100180 (50.0%)	16440 (48.0%)	3620 (63.5%)	3140 (63.6%)	480 (63.2%)
Education level						
≥ Secondary school	185465 (79.1%)	158037 (78.9%)	27428 (80.1%)	4236 (74.3%)	3711 (75.1%)	525 (69.1%)
Primary school	41301 (17.6%)	35372 (17.7%)	5929 (17.3%)	1209 (21.2%)	997 (20.2%)	212 (27.9%)
< Primary school	7714 (3.3%)	6811 (3.4%)	903 (2.6%)	255 (4.5%)	232 (4.7%)	23 (3.0%)
Professional status						
Employed	120011 (51.2%)	103122 (51.5%)	16889 (49.3%)	2720 (47.7%)	2371 (48%)	349 (45.9%)
Unemployed	114469 (48.8%)	97098 (48.5%)	17371 (50.7%)	2980 (52.3%)	2569 (52%)	411 (54.1%)

* On AZT-based regimens throughout the study period. AZT: Zidovudine.

APPENDIX 6B

Baseline clinical characteristics after multiple imputation for all eligible patients and eligible patients on AZT-based regimens with a documented viral load done 90 – 270 days after ART initiation

Characteristics	All eligible patients			All eligible patients on AZT-based regimens*		
	Total N = 234480	**Viral load < 400 copies/ml N = 200220**	**Viral load ≥ 400 copies/ml N = 34260**	**Total N = 5700**	**Viral load < 400 copies/ml N = 4940**	**Viral load ≥ 400 copies/ml N = 760**
	n (%)	n (%)	n(%)	n (%)	n (%)	n (%)
BMI at initiation in kg/m^2						
< 18.5	40809 (17.4%)	34475 (17.2%)	6334 (18.5%)	951 (16.7%)	853 (17.3%)	98 (12.9%)
18.5 – < 25	128489 (54.8%)	109275 (54.6%)	19214 (56.1%)	2985 (52.4%)	2445 (49.5%)	540 (71.1%)
25 – < 30	44724 (19.1%)	38623 (19.3%)	6101 (17.8%)	863 (15.1%)	790 (16.0%)	73 (9.6%)
≥ 30	20458 (8.7%)	17847 (8.9%)	2611 (7.6%)	901 (15.8%)	852 (17.2%)	49 (6.4%)
Baseline haemoglobin						
≤ 8 g/dl	12121 (5.2%)	10384 (5.2%)	1737 (5.1%)	41 (0.7%)	33(0.7%)	8 (1.1%)
> 8g/dl	222359 (94.8%)	189836 (94.8%)	32523 (94.9%)	5659 (99.3%)	4907 (99.3%)	752(98.9%)
MCV at baseline						
< 80fL	29925 (12.8%)	26216 (13.1%)	3709 (10.8%)	521 (9.1%)	467(9.5%)	54 (7.1%)
80 – 100fL	196384 (83.8%)	167140 (83.5%)	29244 (85.4%)	4206 (73.8%)	3627 (73.4%)	579 (76.2%)
> 100fL	8171 (3.4%)	6864 (3.4%)	1307 (3.8%)	973 (17.1%)	846 (17.1%)	127 (16.7%)
Baseline CD4 count in cells/mm^3						
> 200	33157 (14.1%)	28654 (14.3%)	4503 (13.1%)	1474 (25.9%)	1257 (25.4%)	217 (28.6%)
101 – 200	82533 (35.2%)	71025 (35.5%)	11508 (33.6%)	1747 (30.6%)	1507 (30.5%)	240 (31.6%)
51 – 100	46999 (20.0%)	39888 (19.9%)	7111 (20.8%)	1139 (20.0 %)	1025 (20.7%)	114 (15%)
≤ 50	71791 (30.7%)	60653 (30.3%)	11138 (32.5%)	1340 (23.5%)	1151 (23.4%)	189 (24.8%)
WHO stage at baseline						
I	91921 (39.2%)	79986 (40.0%)	11935 (34.8%)	2874 (50.4%)	2474 (50.1%)	400 (52.6%)
II	46000 (19.6%)	39106 (19.5%)	6894 (20.1%)	681 (12.0%)	581 (11.8%)	100 (13.2%)
III	71844 (30.6%)	60402 (30.2%)	11442 (33.4%)	1620 (28.4%)	1386 (28.1%)	234 (30.8%)
IV	24715 (10.6%)	20726 (10.3%)	3989 (11.7%)	525 (9.2%)	499 (10.0%)	26 (3.4%)

* On AZT-based regimens throughout the study period. AZT: Zidovudine. BMI: Body Mass Index. MCV: Mean cell volume. WHO: World Health Organization.

APPENDIX 6C

6 month clinical characteristics after multiple imputation for all eligible patients and eligible patients on AZT-based regimens with a documented viral load done 90 – 270 days after ART initiation

Characteristics	All eligible patients			All eligible patients on AZT-based regimens*		
	Total N = 234480	**Viral load < 400 copies/ml N = 200220**	**Viral load ≥ 400 copies/ml N = 34260**	**Total N = 5700**	**Viral load < 400 copies/ml N = 4940**	**Viral load ≥ 400 copies/ml N = 760**
	n (%)	n (%)	n (%)	n (%)	n (%)	n (%)
Self-reported adherence						
> 90%	206040 (87.9%)	176147 (88.0%)	29893 (87.3%)	5108 (89.6%)	4487 (90.8%)	621 (81.7%)
> 60 – 90%	25727 (11.0%)	21961 (11.0%)	3766 (11.0%)	491 (8.6%)	392 (7.9%)	99 (13.0%)
≤ 60%	2713 (1.1%)	2112 (1.0%)	601 (1.7%)	101 (1.8%)	61 (1.2%)	40 (5.3%)
Number of missed visits						
0	158880 (67.8%)	135960 (67.9%)	22920 (66.9%)	3620 (63.5%)	3220 (65.2%)	400 (52.6%)
1	55800 (23.8%)	47580 (23.8%)	8220 (24.0%)	1320 (23.2%)	1180 (23.9%)	140 (18.4%)
≥ 2	16920 (7.2%)	14400 (7.2%)	2520 (7.4%)	500 (8.8%)	360 (7.3%)	140 (18.4%)
Missing	2880 (1.2%)	2280 (1.1%)	600 (1.7%)	260 (4.5%)	180 (3.6%)	80 (10.6%)
Number of missed ARV visits						
0	183237 (78.1%)	156907 (78.4%)	26330 (76.9%)	4257 (74.7%)	3762 (76.2%)	495 (65.1%)
1	43573 (18.6%)	36971 (18.5%)	6602 (19.3%)	1163 (20.4%)	938 (19.0%)	225 (29.6%)
≥ 2	7670 (3.3%)	6342 (3.1%)	1328 (3.8%)	280 (4.9%)	240 (4.8%)	40 (5.3%)
Number of missed medical visits						
0:	182705 (77.9%)	156856 (78.3%)	25849 (75.5%)	4331 (76.0%)	3873 (78.4%)	458 (60.3%)
1:	44110 (18.8%)	37107 (18.5%)	7003 (20.4%)	1227 (21.5%)	985 (19.9%)	242 (28.9%)
≥ 2:	7665 (3.3%)	6257 (3.2%)	1408 (4.1%)	142 (2.5%)	82 (1.7%)	60 (7.9%)

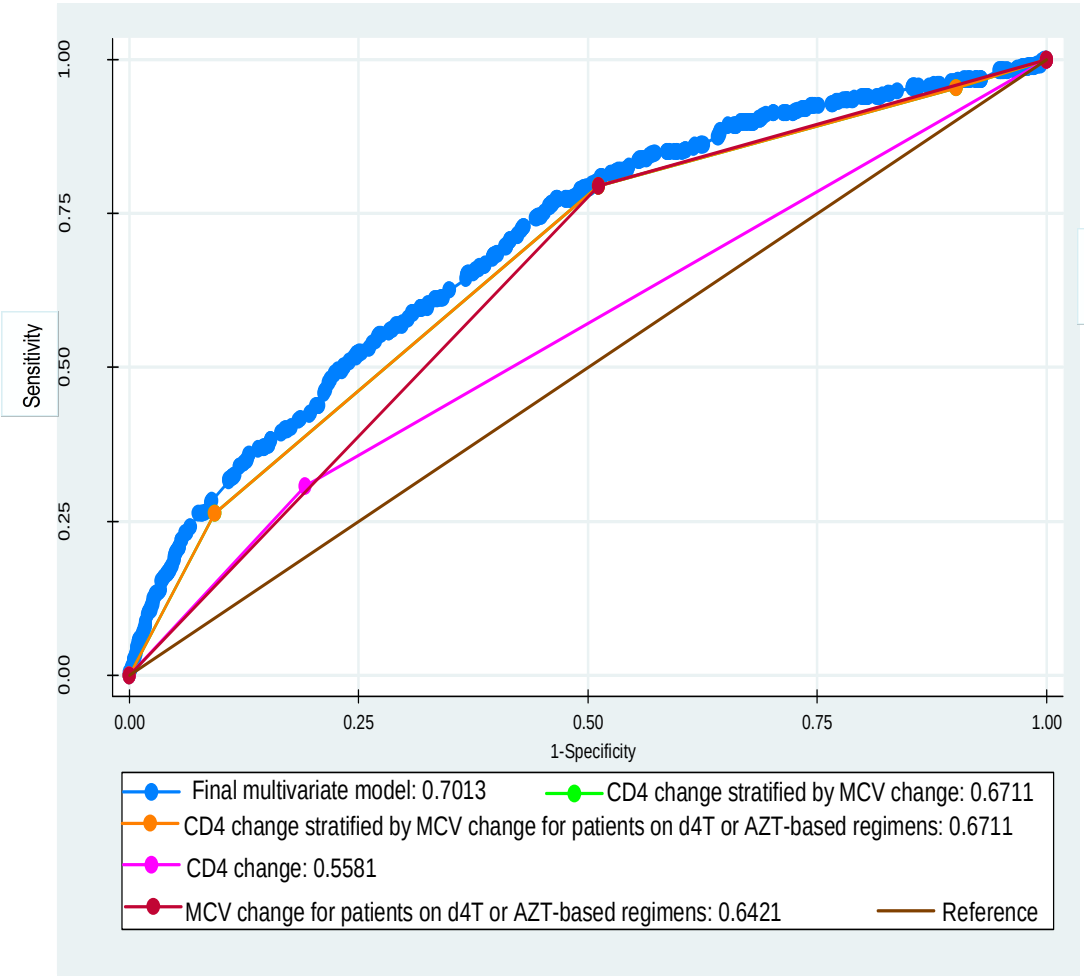
BMI at 6 months in kg/m^2						
< 18.5	16502 (7.0%)	13449 (6.7%)	3053 (8.9%)	448 (7.9%)	393 (8.0%)	55 (7.2%)
18.5 – < 25	125949 (53.7%)	107068 (53.5%)	18881 (55.1%)	2804 (49.2%)	2329 (47.1%)	475 (62.5%)
25 – < 30	62609 (26.7%)	54077 (27.0%)	8532 (24.9%)	1480 (26.0%)	1296 (26.2%)	184 (24.2%)
≥ 30	29420 (12.6%)	25626 (12.8%)	3794 (11.1%)	968 (16.9%)	922 (18.7%)	46 (6.1%)
Haemoglobin at 6 months						
≤ 8 g/dl	1122 (0.5%)	882 (0.4%)	240 (0.7%)	20 (0.4%)	0 (0%)	20 (2.6%)
> 8g/dl	233358 (99.5%)	199338 (99.6%)	34020 (99.3%)	5680 (99.6%)	4940 (100%)	740 (97.4%)
MCV at 6 months						
< 80fL	5816 (2.5%)	4946 (2.5%)	870 (2.5%)	65 (1.1%)	64 (1.3%)	1 (0.1%)
80 – 100fL	113518 (48.4%)	91526 (45.7%)	21992 (64.2%)	1726 (30.3%)	1366 (27.7%)	360 (47.4%)
> 100fL	115146 (49.1%)	103748 (51.8%)	11398 (33.3%)	3909 (68.6%)	3510 (71.0%)	399 (52.5%)
Change in MCV at 6 months						
≥ 14.5fL	96253 (41.0%)	88088 (44%)	8165 (23.8%)	2669 (46.8%)	2432 (49.2%)	237 (31.2%)
< 14.5fL	138227 (59.0%)	112132 (56%)	26095 (76.2%)	3031 (53.2%)	2508 (50.8%)	523 (68.8%)
Change in CD4 count at 6 months in cells/mm^3						
≥ expected	182951 (78.0%)	159463 (79.6%)	23488 (68.6%)	4206 (73.8%)	3783 (76.6%)	423 (55.7%)
< expected	51529 (22.0%)	40757 (20.4%)	10772 (31.4%)	1494 (26.2%)	1157 (23.4%)	337 (44.3%)
Change in CD4 count at 6 months stratified by change in MCV at 6 months						
-Change in CD4 count ≥ expected and change in MCV ≥ 14.5fL:	75412 (32.2%)	69613 (34.8%)	5799 (16.9%)	2068 (36.3%)	1942 (39.3%)	126 (16.6%)
-Change in CD4 count ≥ expected and change in MCV < 14.5fL:	107539 (45.9%)	89850 (44.9%)	17689 (51.6%)	2138 (37.5%)	1841 (37.3%)	297 (39.1%)
-Change in CD4 count < expected and change in MCV ≥ 14.5fL:	20841 (8.9%)	18475 (9.2%)	2366 (6.9%)	601 (10.5%)	490 (9.9%)	111 (14.6%)
-Change in CD4 count < expected and change in MCV < 14.5fL:	30688 (13.0%)	22282 (11.1%)	8406 (24.6%)	893 (15.7%)	667 (13.5%)	226 (29.7%)
Regimen change						
No	212960 (90.8%)	182100 (90.9%)	30860 (90.1%)	5560 (97.5%)	4840 (98.0%)	720 (94.7%)
Yes	21520 (9.2%)	18120 (9.1%)	3400 (9.9%)	140 (2.5%)	100 (2.0%)	40 (5.3%)

TB[α]: On TB treatment during first 6 months on ART:						
Never	175540 (74.9%)	150180 (75.0%)	25360 (74.0%)	4340 (76.1%)	3740 (75.7%)	600 (78.9%)
At baseline	49200 (21.0%)	41680 (20.8%)	7520 (21.9%)	1220 (21.4%)	1060 (21.5%)	160 (21.1%)
During follow-up	9740 (4.1%)	8360 (4.2%)	1380 (4.1%)	140 (2.5%)	140 (2.8%)	0 (0%)
Pregnancy[α]: Pregnancy during first 6 months on ART:						
Never	141580 (60.4%)	122180 (61.0%)	19400 (56.6%)	2900 (50.9%)	2660 (53.8%)	240 (31.6%)
At baseline	2620 (1.1%)	2260 (1.1%)	360 (1.1%)	40 (0.7%)	40 (0.8%)	0 (0%)
During follow-up	3980 (1.7%)	3380 (1.7%)	600 (1.7%)	120 (2.1%)	60 (1.2%)	60 (7.9%)
Missing[γ]	86300 (36.8%)	72400 (36.2%)	13900 (40.6%)	2640 (46.3%)	2180 (44.2%)	460 (60.5%)
Adverse events or side effects						
No	227580 (97.1%)	193720 (96.8%)	33860 (98.8%)	5680 (99.6%)	4920 (99.6%)	760 (100%)
Yes	6900 (2.9%)	6500 (3.2%)	400 (1.2%)	20 (0.4%)	20 (0.4%)	0 (0%)

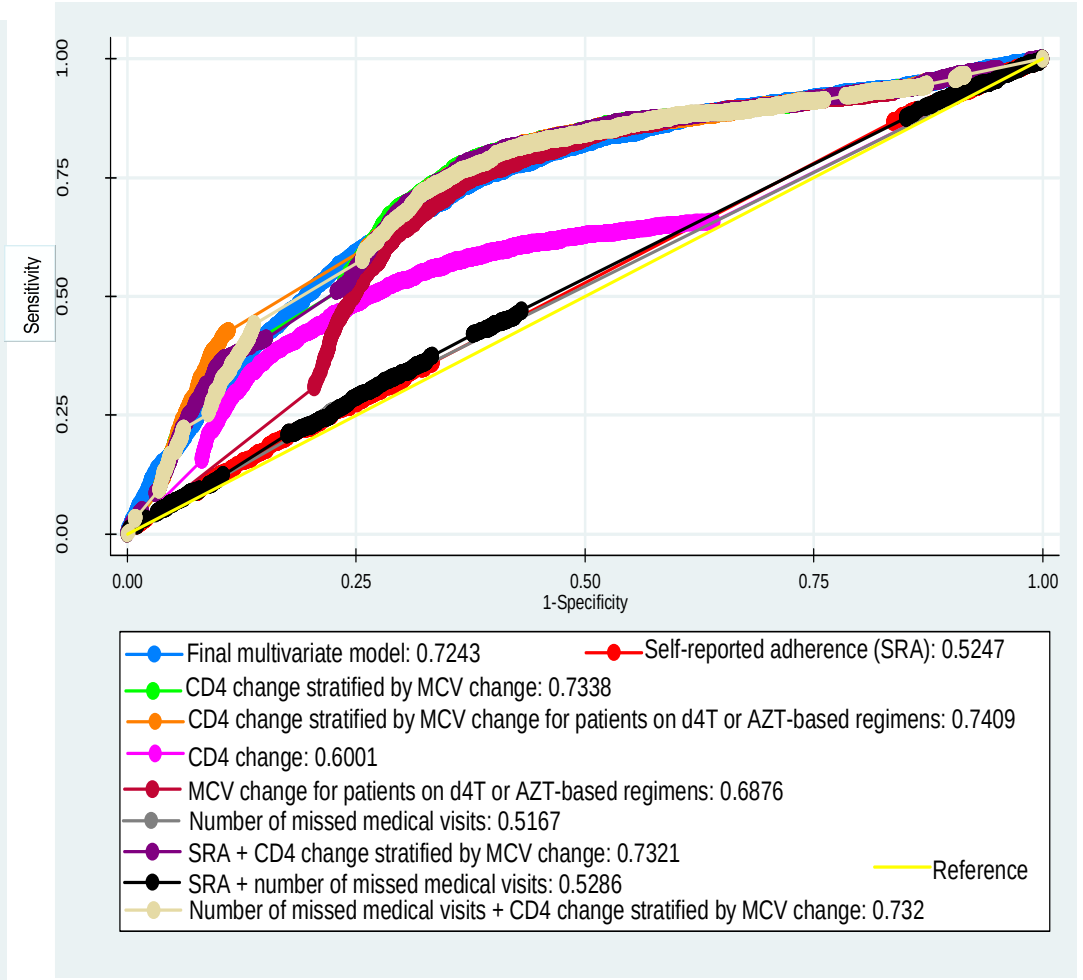
[*] On AZT-based regimens throughout the study period. [α] Variable has baseline and 6 month components. [γ] Male patients. ART: Antiretroviral Therapy. ARV: Antiretroviral. AZT: Zidovudine. BMI: Body Mass Index. MCV: Mean cell volume. TB: Tuberculosis.

APPENDIX 7A: ROC curves for different markers of adherence for all eligible patients with a documented viral load, done 90 to 270 days after ART initiation

Before multiple imputation

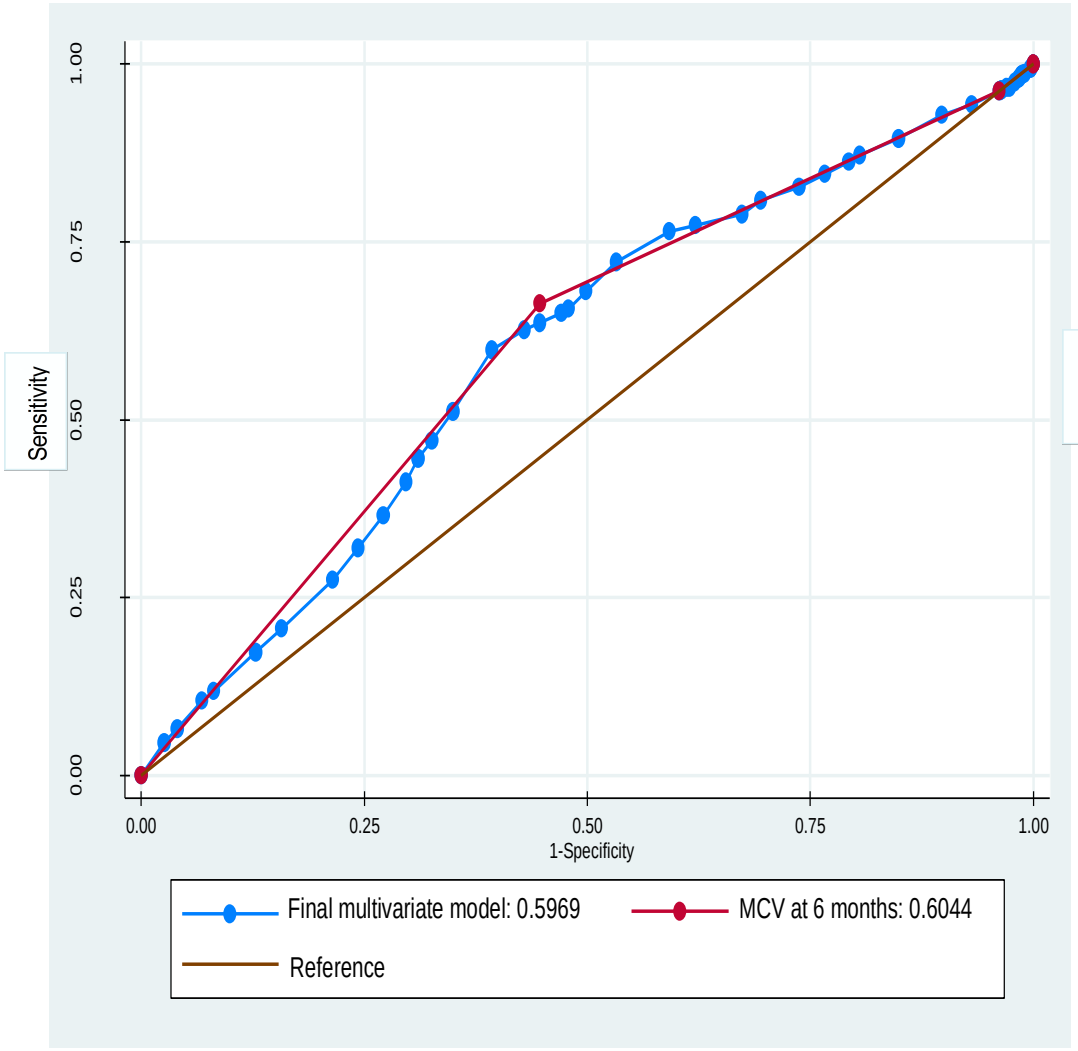


After multiple imputation

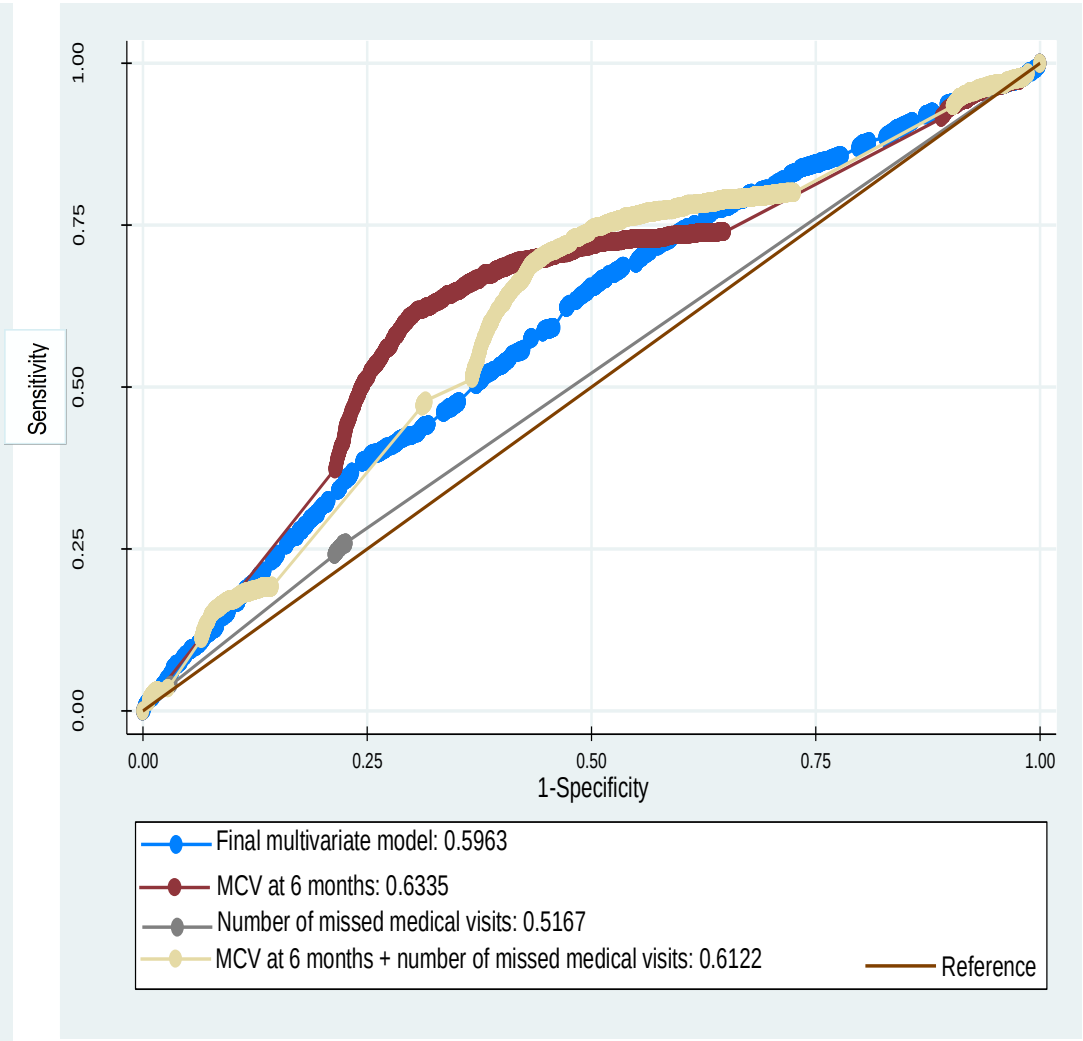


APPENDIX 7B: ROC curves for different markers of adherence for eligible patients on AZT-based regimens with a documented viral load (done 90 to 270 days after ART initiation)

Before multiple imputation



After multiple imputation



APPENDIX 8A

Univariate and multivariate Poisson regression models with robust error variance for all eligible patients and eligible patients on AZT-based regimens, with a documented viral load result, excluding patients with a history of prior ART use such as NVP prophylaxis and excluding patients switched to second-line or a non-standard ART regimen before 6 months on treatment

Variable	All eligible patients				All eligible patients on AZT-based regimens*			
	Univariate Analyses		Multivariate Analysis		Univariate Analyses		Multivariate Analysis	
	Crude IRR (95% CI)	P - value	Adjusted IRR (95% CI)	P - value	Crude IRR (95% CI)	P - value	Adjusted IRR (95% CI)	P - value
Gender								
Females	1	-	1	-	1	-	-	-
Males	1.17 (1.05 – 1.29)	0.003	1.09 (0.92 – 1.30)	0.293	1.44 (0.75 – 2.76)	0.271	-	-
Age at initiation								
≤ 36.7 years	1	-	1	-	1	-	1	-
> 36.7 years	0.93 (0.84 – 1.03)	0.157	0.91 (0.77 – 1.07)	0.260	0.53 (0.28 – 1.00)	0.051	2.72 (0.78 – 9.40)	0.115
Education level								
≥ Secondary school	1	-	-	-	1	-	-	-
Primary school	0.92 (0.79 – 1.07)	0.297	-	-	1.41 (0.61 – 3.26)	0.414	-	-
< Primary school	0.81 (0.59 – 1.11)	0.197	-	-	0.81 (0.12 – 5.30)	0.826	-	-
Professional status								
Employed	1	-	-	-	1	-	-	-
Unemployed	1.09 (0.99 – 1.21)	0.079	-	-	0.93 (0.48 – 1.81)	0.838	-	-
MCV at baseline								
< 80fL	1	-	1	-	1	-	-	-
80 – 100fL	1.29 (1.04 – 1.60)	0.020	1.40 (1.04 – 1.87)	0.024	0.69 (0.23 – 2.08)	0.506	-	-
> 100fl	1.50 (1.07 – 2.08)	0.017	1.07 (0.66 – 1.74)	0.775	0.46 (0.12 – 1.77)	0.256	-	-
Baseline CD4 count in cells/mm^3								
> 200	1	-	1	-	1	-	1	-

101 – 200	0.97 (0.82 – 1.15)	0.724	1.08 (0.81 – 1.44)	0.585	1 (0.38 – 2.64)	1.000	1.03 (0.07 – 14.84)	0.982
51 – 100	1.10 (0.91 – 1.32)	0.317	1.07 (0.78 – 1.47)	0.687	1.04 (0.35 – 3.12)	0.939	3.97 (0.48 – 32.89)	0.202
≤ 50	1.22 (1.03 – 1.45)	0.019	1.35 (1.01 – 1.79)	0.042	1.43 (0.53 – 3.87)	0.483	6.93 (0.52 – 92.81)	0.143
WHO stage at baseline								
I	1	-	1	-	1	-	-	-
II	1.23 (1.05 – 1.45)	0.010	1.20 (0.93 – 1.54)	0.164	0.74 (0.18 – 3.00)	0.678	-	-
III	1.25 (1.08 – 1.44)	0.002	1.26 (1.02 – 1.56)	0.031	1.34 (0.56 – 3.20)	0.509	-	-
IV	1.34 (1.18 – 1.61)	0.002	1.53 (1.19 – 1.98)	0.001	-‡	-	-	-
Self-reported adherence								
> 90%	1	-	-	-	1	-	-	-
> 60 – 90%	1.11 (0.89 – 1.37)	0.351	-	-	1.4 (0.20 – 9.58)	0.732	-	-
≤ 60%	1.29 (0.88 – 1.90)	0.192	-	-	3.5 (0.72 – 17.12)	0.122	-	-
Number of missed visits								
0	1	-	-	-	1	-	-	-
1	0.93 (0.83 – 1.05)	0.239	-	-	0.80 (0.31 – 2.04)	0.635	-	-
≥ 2	0.92 (0.77 – 1.11)	0.404	-	-	2.45 (1.08 – 5.59)	0.033	-	-
Number of missed ARV visits								
0	1	-	-	-	1	-	-	-
1	0.97 (0.86 – 1.10)	0.646	-	-	1.20 (0.52 – 2.77)	0.673	-	-
≥ 2	1.22 (0.98 – 1.53)	0.075	-	-	1.60 (0.45 – 5.73)	0.472	-	-
Number of missed medical visits								
0	1	-	-	-	1	-	-	-
1	1.00 (0.88 – 1.13)	0.970	-	-	1.49 (0.67 – 3.34)	0.329	-	-
≥ 2	1.24 (0.99 – 1.56)	0.067	-	-	4.35 (1.86 – 10.19)	0.001	-	-
MCV at 6 months								
< 80fL	1	-	-	-	1	-	-	-
80 – 100fL	1.35 (1.00 – 1.82)	0.050	-	-	0.78 (0.17 – 3.51)	0.744	-	-
> 100fL	0.56 (0.41 – 0.77)	0.000	-	-	0.20 (0.04 – 0.88)	0.034	-	-

Change in MCV at 6 months								
≥ 14.5 fL	1	-	-	-	1	-	-	-
< 14.5 fL	3.48 (2.86 – 4.24)	0.000	-	-	1.97 (0.61 – 6.36)	0.255	-	-
Change in CD4 count at 6 months in cells/mm^3								
≥ expected	1	-	-	-	1	-	1	-
< expected	1.85 (1.63 – 2.10)	0.000	-	-	3.31 (1.24 – 8.85)	0.017	6.09 (1.10 – 33.75)	0.039
Change in CD4 count at 6 months stratified by change in MCV at 6 months								
-Change in CD4 count ≥ expected and change in MCV ≥ 14.5fL:	1	-	1	-	1	-	-	-
-Change in CD4 count ≥ expected and change in MCV < 14.5fL:	3.06 (2.39 – 3.90)	0.000	3.16 (2.42 – 4.13)	0.000	1.98 (0.35 – 11.18)	0.438	-	-
-Change in CD4 count < expected and change in MCV ≥ 14.5fL:	1.10 (0.69 – 1.76)	0.684	1.31 (0.81 – 2.14)	0.271	3.36 (0.53 – 21.41)	0.199	-	-
-Change in CD4 count < expected and change in MCV < 14.5fL:	6.26 (4.86 – 8.06)	0.000	6.90 (5.23 – 9.10)	0.000	4.35 (0.87 – 21.69)	0.073	-	-
Regimen change								
No	1	-	1	-	1	-	-	-
Yes	0.70 (0.55 – 0.89)	0.004	0.68 (0.45 – 1.03)	0.066	-‡‡	-	-	-
TB: On TB treatment during first 6 months on ART:								
Never	1	-	-	-	1	-	-	-
At baseline	1.00 (0.89 – 1.13)	0.972	-	-	0.89 (0.34 – 2.33)	0.814	-	-
During follow-up	0.89 (0.68 – 1.16)	0.371	-	-	-‡	-	-	-
Pregnancy: Pregnancy during first 6 months on ART:								
Never	1	-	-	-	1	-	1	-
At baseline	1.10 (0.68 – 1.76)	0.705	-	-	-‡	-	-‡	-

During follow-up	1.19 (0.81 – 1.75)	0.381	-	-	7.5 (4.42 – 12.74)	0.000	35.66 (4.49 – 283.24)	0.001
Adverse events or side effects								
No	1	-	-	-	1	-	-	-
Yes	0.34 (0.19 – 0.58)	0.000	-	-	-‡		-	-

* On AZT-based regimens throughout the study period. ‡ Cell has no observations. ‡‡ Cells have no observations. ART: Antiretroviral Therapy. ARV: Antiretroviral. AZT: Zidovudine. CI: Confidence interval. IRR: Incidence rate ratio. MCV: Mean cell volume. TB: Tuberculosis. WHO: World Health Organization.

For the table above, multivariate Poisson regression analysis with robust error variance was done first for all eligible patients then for all eligible patients on AZT-based regimens. After adjusting for potential confounders, the variables independently associated with poor adherence for all patients were change in CD4 count at 6 months stratified by change in MCV at 6 months, CD4 count at baseline, WHO stage at baseline and MCV at baseline. The association between regimen change and poor adherence was marginally significant ($p < 0.1$). Change in CD4 count at 6 months and pregnancy during the first 6 months on ART were the only independent predictors of poor adherence in the final multivariate model for patients on AZT-based regimens. Model fit for the multivariate models was good as evidenced by a p-value of 0.827 on the link test for the model for all patients and 0.622 on the link test for the model for patients on AZT-based regimens.

APPENDIX 8B

Univariate and multivariate Poisson regression models with robust error variance after multiple imputation, for all eligible patients and eligible patients on AZT-based regimens, with a documented viral load result, excluding patients with a history of prior ART use such as NVP prophylaxis and excluding patients switched to second-line or a non-standard ART regimen before 6 months on treatment

Variable	All eligible patients				All eligible patients on AZT-based regimens*			
	Univariate Analyses		Multivariate Analysis		Univariate Analyses		Multivariate Analysis	
	Crude IRR (95% CI)	P - value	Adjusted IRR (95% CI)	P - value	Crude IRR (95% CI)	P - value	Adjusted IRR (95% CI)	P - value
Gender								
Females	1	-	1	-	1	-	1	-
Males	1.17 (1.05 – 1.29)	0.003	1.07 (0.97 – 1.18)	0.195	1.44 (0.75 – 2.76)	0.271	1.36 (0.70 – 2.64)	0.360
Age at initiation								
≤ 36.7 years	1	-	1	-	1	-	1	-
> 36.7 years	0.93 (0.84 – 1.03)	0.157	0.96 (0.87 – 1.05)	0.359	0.53 (0.28 – 1.00)	0.051	0.56 (0.29 – 1.07)	0.078
Education level								
≥ Secondary school	1	-	-	-	1	-	-	-
Primary school	0.92 (0.79 – 1.08)	0.315	-	-	1.19 (0.51 – 2.73)	0.688	-	-
< Primary school	0.84 (0.61 – 1.15)	0.282	-	-	0.87 (0.13 – 5.71)	0.888	-	-
Professional status								
Employed	1	-	-	-	1	-	-	-
Unemployed	1.09 (0.99 – 1.21)	0.072	-	-	0.93 (0.48 – 1.79)	0.821	-	-
MCV at baseline								
< 80fL	1	-	-	-	1	-	-	-
80 – 100fL	1.26 (1.02 – 1.55)	0.031	-	-	0.78 (0.28 – 2.22)	0.649	-	-
> 100fl	1.35 (0.99 – 1.85)	0.057	-	-	0.50 (0.14 – 1.87)	0.306	-	-

Baseline CD4 count in cells/mm^3								
> 200	1	-	1	-	1	-	1	-
101 – 200	0.98 (0.83 – 1.16)	0.802	1.12 (0.95 – 1.32)	0.181	1.04 (0.42 – 2.63)	0.920	1.03 (0.42 – 2.50)	0.945
51 – 100	1.09 (0.90 – 1.32)	0.359	1.30 (1.08 – 1.57)	0.006	1.11 (0.38 – 3.23)	0.841	1.11 (0.36 – 3.42)	0.857
≤ 50	1.22 (1.03 – 1.45)	0.020	1.52 (1.28 – 1.80)	0.000	1.46 (0.51 – 4.13)	0.479	1.31 (0.47 – 3.71)	0.603
WHO stage at baseline								
I	1	-	-	-	1	-	-	-
II	1.21 (1.04 – 1.42)	0.016	-	-	0.83 (0.24 – 2.90)	0.768	-	-
III	1.24 (1.07 – 1.43)	0.004	-	-	1.22 (0.55 – 2.69)	0.624	-	-
IV	1.29 (1.08 – 1.54)	0.005	-	-	-‡	-	-	-
Self-reported adherence								
> 90%	1	-	-	-	1	-	-	-
> 60 – 90%	1.09 (0.87 – 1.37)	0.459	-	-	1.21 (0.33 – 4.50)	0.769	-	-
≤ 60%	1.56 (1.06 – 2.29)	0.024	-	-	2.90 (0.69 – 12.21)	0.147	-	-
Number of missed visits								
0	1	-	-	-	1	-	-	-
1	0.93 (0.83 – 1.05)	0.239	-	-	0.80 (0.31 – 2.04)	0.635	-	-
≥ 2	0.92 (0.77 – 1.11)	0.404	-	-	2.45 (1.08 – 5.59)	0.033	-	-
Number of missed ARV visits								
0	1	-	-	-	1	-	-	-
1	0.97 (0.86 – 1.10)	0.635	-	-	1.20 (0.53 – 2.70)	0.662	-	-
≥ 2	1.22 (0.98 – 1.53)	0.076	-	-	1.53 (0.43 – 5.49)	0.510	-	-
Number of missed medical visits								
0	1	-	-	-	1	-	1	-
1	1.00 (0.88 – 1.13)	0.973	-	-	1.43 (0.65 – 3.14)	0.376	1.34 (0.61 – 2.97)	0.467
≥ 2	1.24 (0.98 – 1.56)	0.070	-	-	3.76 (1.58 – 8.94)	0.003	2.85 (1.35 - 6.02)	0.006

MCV at 6 months								
< 80fL	1	-	-	-	1	-	-	-
80 – 100fL	1.42 (1.05 – 1.92)	0.024	-	-	0.89 (0.19 – 4.10)	0.880	-	-
> 100fL	0.64 (0.47 – 0.87)	0.005	-	-	0.26 (0.05 – 1.23)	0.089	-	-
Change in MCV at 6 months								
≥ 14.5 fL	1	-	-	-	1	-	-	-
< 14.5 fL	2.62 (2.18 – 3.15)	0.000	-	-	2.09 (0.84 – 5.18)	0.111	-	-
Change in CD4 count at 6 months in cells/mm^3								
≥ expected	1	-	-	-	1	-	-	-
< expected	1.76 (1.56 – 1.98)	0.000	-	-	2.03 (0.96 – 4.32)	0.065	-	-
Change in CD4 count at 6 months stratified by change in MCV at 6 months								
-Change in CD4 count ≥ expected and change in MCV ≥ 14.5fL:	1	-	1	-	1	-	-	-
-Change in CD4 count ≥ expected and change in MCV < 14.5fL:	2.46 (1.96 – 3.08)	0.000	2.51 (2.00 – 3.15)	0.000	2.10 (0.62 – 7.16)	0.232	-	-
-Change in CD4 count < expected and change in MCV ≥ 14.5fL:	1.50 (1.07 – 2.09)	0.019	1.58 (1.12 – 2.22)	0.009	2.11 (0.47 – 9.39)	0.326	-	-
-Change in CD4 count < expected and change in MCV < 14.5fL:	4.39 (3.53 – 5.47)	0.000	4.64 (3.70 – 5.81)	0.000	3.72 (1.18 – 11.77)	0.026	-	-
Regimen change								
No	1	-	1	-	1	-	-	-
Yes	0.70 (0.55 – 0.89)	0.004	0.76 (0.59 – 0.97)	0.028	-‡‡	-	-	-
TB: On TB treatment during first 6 months on ART:								
Never	1	-	-	-	1	-	-	-
At baseline	1.00 (0.89 – 1.13)	0.972	-	-	0.89 (0.34 – 2.33)	0.814	-	-
During follow-up	0.89 (0.68 – 1.16)	0.371	-	-	-‡	-	-	-

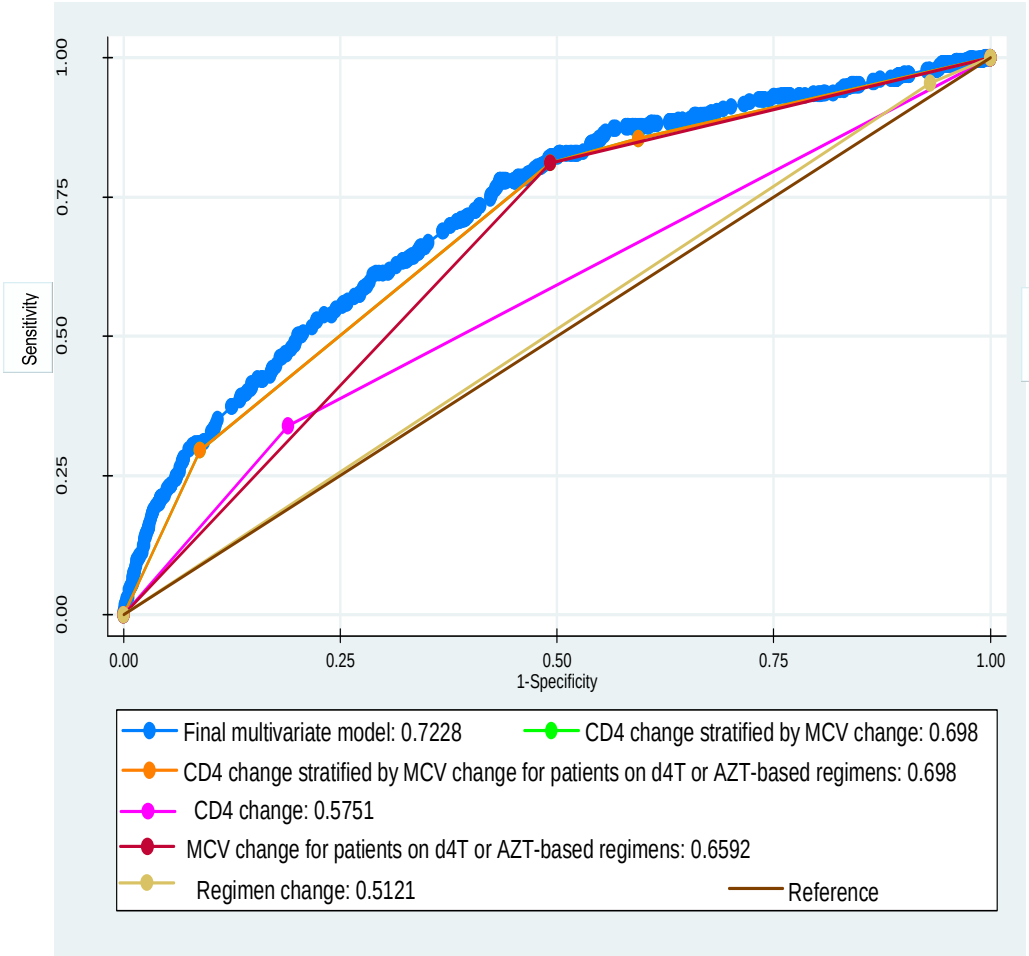
Pregnancy: Pregnancy during first 6 months on ART:								
Never	1	-	-	-	1	-	-	-
At baseline	1.10 (0.68 – 1.76)	0.705	-	-	-‡	-	-	-
During follow-up	1.19 (0.81 – 1.75)	0.381	-	-	7.5 (4.42 – 12.74)	0.000	-	-
Adverse events or side effects								
No	1	-	-	-	1	-	-	-
Yes	0.34 (0.19 – 0.58)	0.000	-	-	-‡		-	-

* On AZT-based regimens throughout the study period. ‡ Cell has no observations. ‡‡ Cells have no observations. ART: Antiretroviral Therapy. ARV: Antiretroviral. AZT: Zidovudine. CI: Confidence interval. IRR: Incidence rate ratio. MCV: Mean cell volume. TB: Tuberculosis. WHO: World Health Organization.

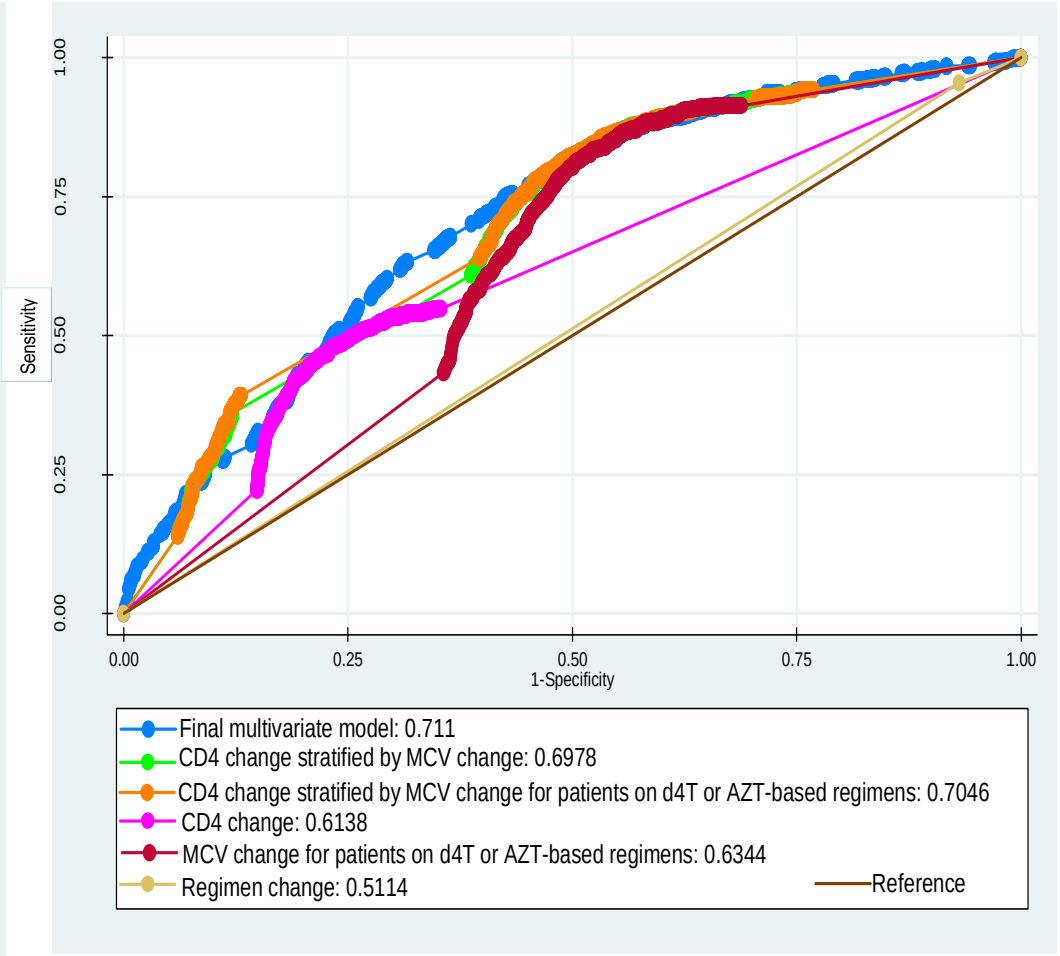
For the table above, multiple imputation (20 imputations) was done for all patients with a documented viral load result using multivariate normal regression. The imputation was done for explanatory variables in the final multivariate models with missing values. Multiple imputation was followed by multivariate Poisson regression analysis with robust error variance. After adjusting for potential confounders, the variables independently associated with poor adherence for all eligible patients were change in CD4 count at 6 months stratified by change in MCV at 6 months, regimen change and CD4 count at baseline. Model fit was good as evidenced by a p-value of < 0.05 on the FMI test. Number of missed medical visits was the only independent predictor of poor adherence in the final multivariate model for patients on AZT-based regimens. Model fit for this model was also good as evidenced by a p-value of < 0.05 on the FMI test.

APPENDIX 9A: ROC curves for different markers of adherence for eligible patients with a documented viral load, excluding patients with a history of prior ART use such as NVP prophylaxis and excluding patients switched to second-line or a non-standard ART regimen before 6 months on treatment

Before multiple imputation

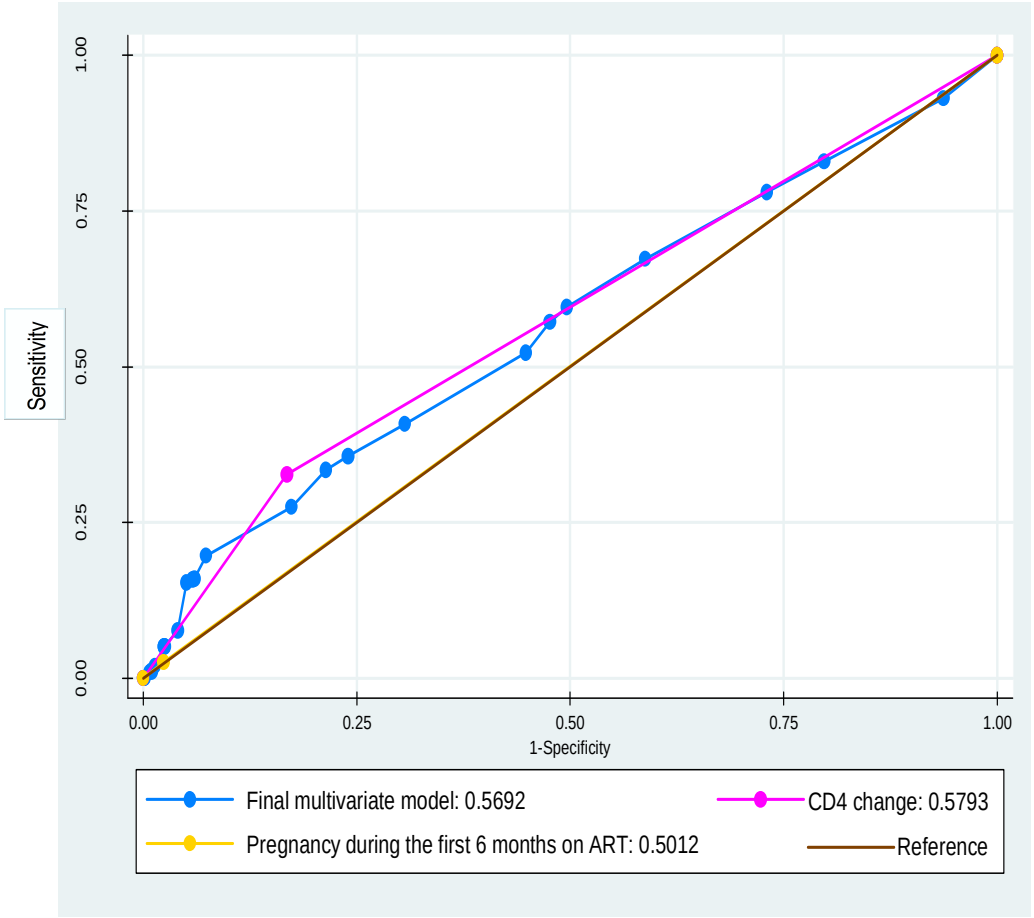


After multiple imputation

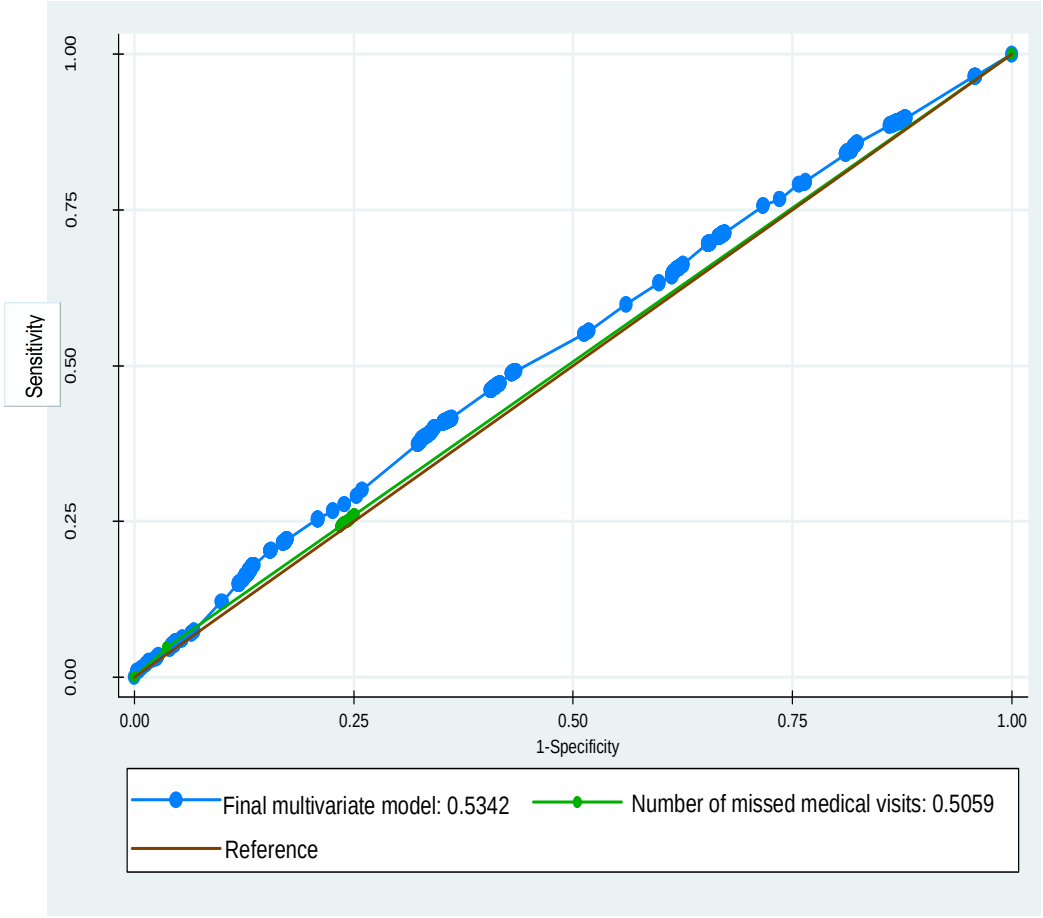


APPENDIX 9B: ROC curves for different markers of adherence for eligible patients on AZT-based regimens with a documented viral load, excluding patients with a history of prior ART use such as NVP prophylaxis and excluding patients switched to second-line or a non-standard ART regimen before 6 months on treatment

Before multiple imputation



After multiple imputation



APPENDIX 10: Permission letter to use TherapyEdge-HIV™

University of the Witwatersrand, Johannesburg

Clinical HIV Research Unit

SECRETARIAT: Suite 176, Private Bag x2600, Houghton 2041, South Africa • Tel: +27-11-276-8800 • Fax: +27-11-482-2130

Date: 30th July, 2012

This is to certify that Maria Nnambalirwa has been granted permission to use data from the Themba Lethu Clinic cohort stored on TherapyEdge™ for a research project titled "Markers of Adherence among HIV-positive Adults on Antiretroviral Therapy at Themba Lethu Clinic".

The standard operating procedures for the Clinical HIV Unit have been complied with.

……

Professor AP MacPhail.

APPENDIX 11: Ethics clearance certificate

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG

Division of the Deputy Registrar (Research)

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)
R14/49 Dr Maria Nnambalirwa

CLEARANCE CERTIFICATE M120858

PROJECT Markers of Adherence among HIV-Positive Adults on Antiretroviral Therapy at Themba Lethu Clinic

INVESTIGATORS Dr Maria Nnambalirwa

DEPARTMENT School of Public Health

DATE CONSIDERED 31/08/2012

DECISION OF THE COMMITTEE* Approvedunconditionally

Unless otherwise specified this ethical clearance is valid for 5 years and may be renewed upon application.

DATE **CHAIRPERSON** .....
(Professor P E Cleaton Jones)

*Guidelines for written 'informed consent' attached where applicable

cc: Supervisor : Dr Peter Nyasulu

DECLARATION OF INVESTIGATOR(S)

To be completed in duplicate and **ONE COPY** returned to the Secretary at Room 10004, 10th Floor, Senate House, University.

~~I/We~~ fully understand the conditions under which I am~~/we are~~ authorized to carry out the abovementioned research and ~~I/we~~ guarantee to ensure compliance with these conditions. Should any departure ~~to be~~ contemplated from the research procedure as approved ~~I/we~~ undertake to resubmit the protocol to the Human Research Ethics Committee (Medical).

Dr Maria Nnambalirwa.

APPENDIX 12



Confidentiality Agreement

Patient's health records and personal information are confidential. I understand that I may become aware of patient information in the course of my association with any of the Right to Care Clinics and I am prohibited from divulging or communicating this information both during and after such association. I agree to respect the patient's right to confidentiality and privacy. I agree to access patient's personal health information only as permitted in the performance of my duties/course of my research or as otherwise directed. I agree to preserve the confidentiality of all clinical or patient information and not to divulge this in any form. Any breach of this agreement during or after my association with the Clinic will be taken seriously. Any violation can or may result in legal or disciplinary action.

I……., acknowledge that I have read this confidentiality agreement and understand my responsibilities as they pertain to confidentiality of personal information and agree to the principles as stated above.

……. ……

Signature of Employee/Investigator Date