Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Peng Wu Date

Adherence Measure Methods to Antiretroviral Therapy in HIV-Infected Patients in South Africa

By

Peng Wu

MSPH

Emory University

Rollins School of Public Health

Department of Biostatistics

Brent A. Johnson

John J. Hanfelt

Adherence Measure Methods to Antiretroviral Therapy in HIV-Infected Patients in South Africa

By

Peng Wu

B.S., Sichuan University, 2011

MSPH, Emory University

Rollins School of Public Health

2013

Thesis Committee Chair: Brent Johnson, PhD

An abstract of

A thesis submitted to the Faculty of the

Rollins School of Public Health of Emory University

in partial fulfillment of the requirements for the degree of

Master of Science in Public Health

in Biostatistics

2013

Adherence Measure Methods to Antiretroviral Therapy in HIV-Infected Patients in South Africa

By: Peng Wu

Abstract:

Background
Previous studies have shown that highly active antiretroviral therapy (HAART) has a positive impact in reducing the HIV-related death. Adherence to therapy is a strong predictor of virologic failure (VF) among patients. A few methods such as pill count have been utilized to monitor the adherence. However, in resource-limited settings such as in South Africa, more trials and analyses should be done due to inadequate research.
Objective
In the retrospective case-control study we conducted, we sought several adherence assessment methods including pill count, medication possession ratio (MPR), adherence score and aimed at identifying risk factors that may predict VF at 6 months.
Methods
Several smooth splines corresponding to each of these adherence assessment methods were fit in a logistic model to evaluate the association between the estimated probability and those predictors. We also used different combinations of those adherence assessment methods as new factors to predict VF. Additionally, we split the pharmacy refill period into several intervals according to each patient's pharmacy refill dates to monitor the habits of the patients' follow-to-prescription. We adopted a Generalized Estimating Equation (GEE) approach to investigate the relationship between the virologic failure and the pick-up times as well as pills left at each visit.
Results
458 patients participated in antiretroviral therapy from October 2010 to June 2012. Of these, 158 (34.50%) had virologic failure (cases) and 300 (65.50%) did not (controls). The mean adherence score is 98.15% in case group and 94.80% in control group. Adherence score may be not the best in retrospective study. ROC analysis shows the combination of pharmacy refill ratio plus self-reported question can be a strong tool to predict VF. The number of visits have strong association with refill pattern ($p<0.0001$) using Generalized Estimating Equation (GEE) approach.

Adherence Measure Methods to Antiretroviral Therapy in HIV-Infected Patients in South Africa

By

Peng Wu

B.S., Sichuan University, 2011

MSPH, Emory University

Rollins School of Public Health

2013

Thesis Committee Chair: Brent Johnson, PhD

A thesis submitted to the Faculty of the

Rollins School of Public Health of Emory University

in partial fulfillment of the requirements for the degree of

Master of Science in Public Health

in Biostatistics

2013

ACKNOWLEDGEMENT

First and foremost, I wish to thank my thesis advisor Dr. Brent A. Johnson of the Biostatistics and Bioinformatics Department at Emory University for his persistent help. Without his guidance, ideas and support, the thesis would not have been possible. I learned a lot in working with Dr. Johnson and it is a great help to my future study.

I am grateful to Dr. Vincent Marconi, School of Medicine, who provided us with valuable data of this study and insight into the adherence problem. I would like to express my immense gratitude to Dr. Marconi for his support and inspiration throughout my thesis. Collaboration with him is very impressive and his hard work is greatly appreciated.

In addition, a thank you to all the team members of RFVF study group. I wish to thank Baohua Wu for his assistance in some coding problems. I also feel grateful to Claudia Ordonez for the study background materials. And I would like to thank Rachel Kearns and Anna Quinn for their advices on the project and information about the literature review.

I spent a great time in the past two years in the Department of Biostatistics and Bioinformatics at Emory University. It gives me great pleasure in acknowledging the support from all the professors and instructors such as Dr. John J. Hanfelt and Dr. Robert H. Lyles. I could not complete the two years' study successfully without them.

Most of all, I want to thank my parents for their support and encouragement that motivated me to pursue an advanced degree in the graduate school.

Table of Contents

CHAPTER 1: INTRODUCTION & BACKGROUND

HIV/AIDS and Antiretroviral Therapy (ART)

Human Immunodeficiency virus (HIV) is a retrovirus that compromises human immune system, which causes Acquired Immunodeficiency Syndrome (AIDS). It was first clinically observed in 1981 in the United States [1]. HIV infection will cause the reduction of CD4+ T cells in the human immune system and thus leads to diseases that endanger life.

According to UNAIDS report in 2010, 33.3 million people were living with HIV in the world and AIDS has the highest prevalence in South Africa. The estimated number of infected patients is 5.6 million in South Africa, which remains the largest in the world [2]. Moreover, AIDS is the largest cause of maternal mortality and 35% of under-5-year children deaths are due to AIDS [3]. From 2001 the prevalence of AIDS is increasing, however, the rate of new HIV infections shows the sign of declining (<0.10% globally). In South Africa, the incidence rate in adults decreased from 2.35% in 2001 to 1.49% in 2009 [2].

The reason why the overall prevalence is not declining is that HIV treatments like antiretroviral therapy (ART) allows people infected with HIV to have a longer life. ART is a treatment for HIV-infected patients using anti-retrovirus drugs. Usually, the combination of over three drugs that block and suppress the virus replication is included in ART procedures. ART has been found to be the most effective in reducing the mortality and morbidity of HIV-infected patients as well as improving their life quality

[4]. As retroviral virus has damaged the immune system of HIV-infected patients, they have to take antiretroviral drugs regularly for the rest of their life for ART to work [5].

The ART programs have been initiated since a decade ago in sub-Saharan Africa with the support from international organizations. Since the scale-up of ART programs, they have had a positive impact in reducing AIDS-related deaths especially in sub-Saharan Africa. Compared to 2004, there were 32,000 fewer deaths of AIDS-related patients in 2009 [2]. The reduction in incidence rate can be offset by the effect of ART.

Marconi Study (Risk Factors of Virological Failure Study)

The Risk Factors of Virological Failure (RFVF) Study [6] is an international research study currently conducted by Dr. Marconi and investigators from Durban, South Africa. The main focus of this study is to identify the prevalence of HIV drug resistance after antiretroviral therapy as well as the risk factors that are related to the virologic failure. Clinical outcomes after subsequent therapy, adherence and pharmacy refill patterns are the major concern of the study. This research also aims at comparing the disparities between urban and rural settings in South Africa and examining the influence of minority resistance variants on treatment response for patients.

Investigators conducted a 6-month retrospective case-control study investigating risk factors for virologic failure (RFVF) among 458 patients at McCord Hospital (MCH) in Durban, South Africa. In the study at the Sinikithemba Clinic at McCord Hospital conducted by Richard Murphy et al., the patients underwent genotypic resistance testing and among which who have HIV-1 RNA viral load $\geq$ 1000 copies/ml were considered virologic failure [7]. Marconi used the same criteria to define case and control in RFVF

study. The viral load (VL) was monitored at 5 months after initialing first line ART in RFVF study. Patients with virologic failure, i.e. with viral load (VL) ≥ 1000 copies/mL were defined as cases. Correspondingly, controls were defined as patients with VL < 1000 copies/mL and received following treatment and monitoring thereafter.

Adherence Measure Methods

In the meeting of 2011, World Health Organization defined "adherence" as "the extent to which patients follow the medical instruction". Adherence to antiretroviral therapy is considered to be a good predictor of virus suppression and illness control [8]. Adherence is utilized to monitor whether the patients maintain in the medication regimen and follow the prescription. Patients who do not take medications at all, take reduced amount of doses and do not take at prescribed frequencies are considered non-adherent [9]. Non-adherence leads to the development of drug-resistant virus and the increase of viral load copies. Therefore, improving the adherence is the key to preventing virologic failure and achieving positive therapeutic effect [9].

In order to assess the patient medication adherence, a set of discrete measures including self-report and pill count may be used. Adherence may also be calculated as a continuous measure, such as medication possession ratio (MPR) and proportion of days covered (PDC). MPR represents the ratio of days pharmacy supplied to days in a time interval (usually the general study period) [10].

Different adherence measure methods have different definitions. The different adherence measure methods may have their own pros and cons, so there is no "gold standard" for measuring adherence to medication in antiretroviral therapy [16]. For

example, (1) Nachega et al. conducted a cross-sectional study of adherence among HIV-infected individuals at Chris Hani Baragwanath Hospital's Adult HIV Clinic in Soweto, South Africa [11]. They assess adherence by utilizing the survey result of a 1-month, self-report questionnaire. The ratio of pills taken to pills prescribed was defined as adherence. (2) A. Palepu et al. calculated adherence as the ratio of number of days the HIV-infected patients who received HAART refills to total number of days of medication follow-up period which lasted for 12 months in the Vancouver Injection Drug User Study (VIDUS) [8]. (3) Robert Grossberg et al. defined self-reported adherence as (1-(missed doses/prescribed doses))*100% in the observational cohort study of HIV-infected individuals in HAART at a Veteran's Affairs Medical Center in Philadelphia, Pennsylvania [12].

Access to ART in Resource-Limited Settings

Since the beginning of the 21st century, the international support to ensure access to HAART has been expanding. Despite the effort contributed by Global Fund, the US President's Emergency Plan for AIDS Relief (PEPFAR), the World Bank, the Bill and Melinda Gates and the Clinton Foundations, the coverage of HAART varies significantly in resource-limited countries [13]. According to the statistics of WHO 2005, the HAART coverage of December 2004 in Latin America and Caribbean is 65%, much higher than the coverage in sub-Saharan Africa, which is 8%. The conclusion is drawn in Nattrass's study that the access to HAART has main correlation with per capita income along with other political and regional variables [13].

In developing countries having poor resources, there are two primary hindrances in implementing HAART programs [14]. One is the high price of the medications. The monthly retail cost for the three drugs in ART can be as high as $768 from US manufacturers [15]. The other main objection is due to the lack of necessary local health infrastructures to effectively deliver ART [14]. The capacity of the local health infrastructures such as health clinics is relatively weak.

CHAPTER 2: METHODOLOGY

Study Question

Despite several adherence measure methods were used in different study, for specific data, choosing one or two best methods is a major concern. Our main goal of the RVFV study is to determine whether and how the adherence measure methods we chose are associated with virologic failure and find most predictive combination of adherence measures using ROC analysis. Besides, we aim at assessing the effect of last refill and number of refills in predicting virologic outcome.

Study Site

The RFVF study was conducted at the McCord Hospital in Durban, South Africa, the largest city in the KwaZulu-Natal province. The McCord Hospital is a referral center for ART and receives funding from President's Emergency Plan for AIDS and South Africa government [6].

Data Collection and Cleaning

All HIV-infected patients enrolled into the retrospective study underwent a single, semi-structured interview that consisted of a questionnaire, a neurocognitive assessment, and a pill count at enrollment. The research coordinator who was blinded to the study assignment conducted the interview. The questionnaire contained information regarding

ART adherence and pharmacy refills. The demographics, pharmacy and laboratory data were stored in the electronic records. The data used in this paper were abstracted from Redcap (electronic data capture) at Emory University.

Missing values are common in pharmacy refill claims databases. In our study the missing data proportion is relatively small and no missing data handling methods should be used. Moreover, a few extreme values corresponding to error have been identified and the database has been updated.

Adherence Measurement

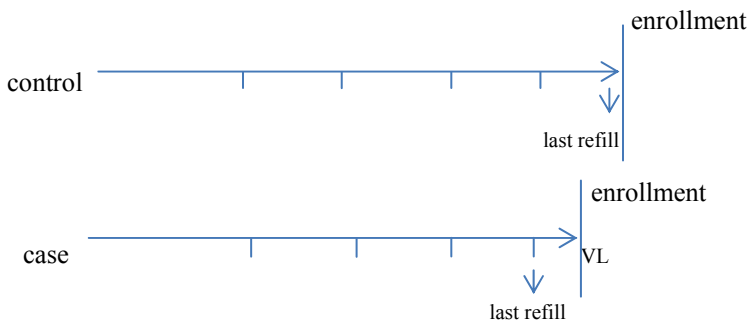
Pill Count:

At the enrollment of the study, we counted the pills left in the bottle, and defined it as *pill count at enrollment*. The pharmacy refill claims were depicted in a simple graph below. Generally, the patients in case group were enrolled when detecting VL≥1000 copies/ml. Then, their pills left in the bottle were recorded. For patients in control group, their pills left were counted usually at the end of the refill interval. Thus, the patients in case group tend to have larger pill count than those in control group.

After the 6-month follow-up, we calculated *expected pills* left by subtracting total pills dispensed from number of days between refills. We called the number of days between refills the *pillday*. Finally, we used *pill count at enrollment* minus *expected pills* as surrogate for pill count adherence.

In brief, expected pills = (pills dispensed – pillday) * pill / dose * dose / day;

pill count adherence = pill count at enrollment – expected pills.



Pharmacy Refill Adherence Ratio**:**

The pharmacy refill ratio is similar to the measure of clinical attendance (access). We defined the pharmacy refill adherence as the ratio of dispensed pills (pills prescribed) to the total number of days over the study period (from the enrollment date to the first refill date). Most of the values lie within 0.5 to 1.5. Values close to 1 are viewed favorably.

Briefly, pharmacy refill adherence ratio = pills dispensed / pillday.

Adherence Score**:**

Adherence score was calculated as one minus the ratio of pill count adherence divided by total number of pills dispensed during the study period.

That is, based to the previous definition and our algorithm, if expected pills left are greater than 0, then adherence score

= (1 – (pill count at enrollment – expected pills) / pills dispensed/(pill/dose*dose/day)). Otherwise, the adherence score

= (1 – (pill count at enrollment – expected pills) / (pills dispensed*pill/dose*dose/day – expected pills).

Self-report questions:

During the follow-up, patients were asked questions like this:

How many pills have you missed last week? How many pills did you take more than 1 hour late in the last week? How many pills have you missed last month? How many pills did you take more than 1 hour late in the last month? How do you remember to take pills (choice=cell phone/choice=media)? How do you remember to come for your drug collection appt (choice=cell phone)? How often you were away from home? How often you were busy with other things? How often you fell asleep through the dose time? How often you ran out of pills? How often you forgot to take pills? How often you had wanted to avoid side effects? The pharmacist recorded the answer in each visit. All the questions listed above can be referred to Table 3.

Model Specification

A multivariate logistic regression model is constructed since the outcome (case/control) is binary. Covariates contain the adherence measurement methods that are considered to predict virologic failure well in previous study such as adherence score and self-reported questions.

$$\log\frac{P(Y=1|X)}{P(Y=0|X)} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \sum_{i=3}^{n} \beta_i X_i$$

where Y=1=case group, Y=0=control group, X_1=adherence score, X_2=pharmacy refill ratio, X_i=self-reported questions (i=3, …, n).

The unknown parameters are $\beta_0, \beta_1, \ldots, \beta_n$. The intercept β_0 in this model corresponds to the log odds of being in VF group when $X_1, X_2, \ldots, X_n$ are at the hypothetical value of zero. The coefficient β_1 for X_1 is interpreted as difference in the log

odds. In other words, holding other covariates at fixed values, for a one-unit increase in the adherence score, the expected change in log odds is β_1. Likewise, if other predictors remain unchanged, for a one-unit increase in the pharmacy refill ratio, the expected change in log odds is β_2. The coefficients $\beta_3, \dots, \beta_n$ for self-reported predictors represent the log of odds ratio between the particular group and reference group. By taking the exponential of the different coefficients, we can calculate the corresponding odds ratios.

The log likelihood function is written as:

$$l(\beta) = \sum_{k=1}^{N} \log P(X_k;\beta) = \sum_{k=1}^{N} y_k \log P\left(X_k;\beta\right) + (1 - y_k)\log\left(1 - P(X_k;\beta)\right).$$

Based on Newton-Raphson method, n+1nonlinear equations $\frac{\partial l(\beta)}{\partial \beta_{1i}}$ $(i = 0,1, \dots, n)$ are solved. The parameter estimates are found by maximizing the conditional likelihood using RFVF data.

Receiver Operating Characteristic (ROC)

ROC curve is a graphical plot that illustrates the true positive rate against the false positive rate at various cutpoints. The true positive rate is called sensitivity while the false positive rate is called specificity. Ahead of generating ROC curve, we need to make assumptions that the data we used is normally distributed random variables [17].

The ROC curve is an effective statistical tool to diagnose the accuracy and fitness of the model. The area under the ROC curve (AUC) indicates the power of the model to discriminate the outcome. Under normal assumptions, AUC is equal to the probability that a randomly chosen positive instance ranks higher than a randomly chosen negative one [18]. Usually, the model with larger AUC will be more predictive.

In order to assess the fitness of the adherence measurement methods in predicting VF, different models were constructed using different predictors. The diagnostic accuracy of these models was evaluated by comparing the AUCs of corresponding models. The model that achieves high sensitivity in predicting VF should have large AUC and the ROC curve should be away from 45-degree line and close to the upper-left part of the plot.

Generalized Estimating Equation (GEE)

A generalized linear model (GLM) formulated by John Nelder and Robert Wedderburn [19] can be applied to the data with a single observation for each subject. However, for the subject with repeated observations, correlation between the values within the subjects should be addressed. Liang and Zeger introduced GEEs in 1986 to estimate the parameters of a generalized linear model with unknown correlation within subjects [20]. The occurrence of GEE model brings the improvement in data analysis dealing with repeated measures, nested and cluster structures, particularly in longitudinal study [21].

The assumed marginal regression model is: $g(E[Y_{ij}|x_{ij}]) = \beta^T x_{ij}$, Where g(.) is the link function, x_{ij} is a n*1 vector of covariates, β consists of the n regression parameters, and Y_{ij} denotes the j^{th} outcome for the i^{th} subject. Generally, the link function contains "logit" link for binary data and "log" link for count data. The covariance matrix is modeled as: $V_i = \Phi A_i^{\frac{1}{2}} R(\alpha) A_i^{\frac{1}{2}}$, Where Φ is a dispersion parameter of generalized linear

model, A is a diagonal matrix of variance functions, and $R(\alpha)$is the working correlation matrix.

GEE has very good property because it provides consistent parameter estimates even the working correlation matrix is misspecified. The robust variance estimator for the regression coefficient estimates is often called “sandwich” estimator. The correlation matrix we specified represents within-subject correlation, and it can have structures like “independent”, “auto-regression” and “exchangeable”.

CHAPTER 3: RESULTS

Descriptive Analysis

Between October 2010 and June 2012, 458 HIV-infected individuals experienced first line ART and received genotypic testing. The cohort demographics as well as adherence score, pill count at enrollment and MPR are displayed in Table 1. Also the mean of days supplied, pills dispensed and pharmacy refill ratio are computed.

About 35% of the patients are male and the black is almost the only race (over 99%). The average age is 39.58 years, and the age of case group is 37.07 years, slightly younger than control group, which is 40.90 years. The average adherence score is 95.90%, ranging from 33.33% to 134.44%. Adherence score is higher in case group (98.48%) than in control group (94.97%). The medication possession ratio ranges from 3.89% to 100%. MPR is slightly higher in control (96.58%) than case (95.25%). The average pharmacy refill rate is closer to 1.0 in control group (1.02) than in case group (1.15). The overall days supplied during study period are 160.92 for cases and 165.39 for controls. The average pills dispensed for cases are 173.92 and for cases are 168.51.

"Drug Hierarchy"

The medication regimen of ART requires every patient took three or more drugs or their combinations. We compute the adherence score and pharmacy refill ratio based on just one drug according to the "drug hierarchy" below. For a patient, we choose the one drug that ranks highest in the following "drug hierarchy".

***EFV* > TDF > *FTC*/TDF > *FTC* > *3TC*> *3TC/ZDV* >*3TC/ABC* > *D4T* > *DDI* > *LPV/r* > *ABC*.**

EFV:efavirenz; TDF: tenofovir disoproxil fumarate;

FTC/TDF: emtricitabine and tenofovir; FTC: Emtricitabine;

3TC: lamivudine; 3TC/ZDV: lamivudine and zidovudine;

ZDV: zidovudine (azidothymidine) 3TC/ABC: lamivudine and abacavir;

D4T: stavudine; DDI: didanosine;

LPV/r: lopinavir and rionavir; ABC: abacavir.

For instance, if a patient takes FTC, D4T and LPV/r, we will choose FTC to calculate our adherence score and pharmacy refill rate since FTC ranks higher.

Table 2A-2C displays the adherence score, pharmacy refill ratio and raw pill count based on the priority drug according to the "drug hierarchy". As we can see, EFV is the most prevalent drug in our study, that is, the majority of the patients used EFV. Over 110 cases and 245 controls used EFV and EFV ranks highest in the "drug hierarchy".

For 3TC users, the average adherence score (0.988 vs. 0.956) and raw pill count (24.5 vs. 18) is higher in cases than controls. The mean pharmacy refill ratio (1.071 vs. 1.065) for controls is closer to 1. Similarly, for EFV users, the average adherence score (0.992 vs. 0.961) and raw pill count (13 vs. 9) is higher in cases than controls. The mean pharmacy refill ratio (1.078 vs. 1.040) for controls is closer to 1. The statistics for TDF drug users shows something different. The average adherence scores (1.017 vs. 1.009) for cases and controls of the TDF users are slightly greater than 1. The mean raw pill count at

enrollment (14/8) for cases is larger. The average pharmacy refill ratio (1.144 vs. 1.071) for controls is closer to 1.

Parametric Splines

The smooth splines of the probability of being virologic failure versus respectively three adherence measurement methods are shown in Figure 2A-2C.

Except the sparse distributed outliers, when adherence scores lies within 0.7 to 0.8, the probability that the patients were in virologic failure group is the lowest. That means, based on our algorithm, if adherence score is too small (<0.6) or too large (>1.1), the patients have a higher probability of being VF. The smooth spline of MPR indicates the similar characteristics in Figure 2B. Based on our calculation, if the medication possession ratio is in (0.65, 0.85), the probability that the patients were VF is relatively lower. Particularly, if the patients have a MPR of less than 0.6, he or she is more likely to be VF.

From the smooth splines of pharmacy refill ratio, we can roughly see that the lowest probability (close to 0.2) appears when the pharmacy refill ratio is between 0.9 and 1.0. For the patients who have pharmacy refill ratio less than 0.7 or greater than 1.1, the probability of VF increases dramatically. The conclusion that when the ratio tends to be close to 1, the probability of the patients of being VF is low can be drawn. In other words, if the patients follow the schedule to take the medication, his/her viral load will be more likely to be less than 1000 copies/ml. Additionally, the pharmacy refill ratio >1.5 can be considered as extreme values.

Bar plots in Figure 1A-1C display the distribution of adherence score, medication possession ratio and pharmacy refill ratio. The frequencies of cut offs of the three adherence measure methods are shown in Table 3A-3C respectively. Most of adherence scores are within 0.9 to 1.1 (305/458). Apparently, adherence scores in case group tend to have more extreme values i.e. >1.10 or <0.50. Likewise, most of MPR lie into the interval of 0.9 to 1.0 (337/458) and it is difficult to distinguish the case and group from the characteristics of these frequencies. For pharmacy refill ratio, most values are within 0.8 to 1.2 (416/458). The proportion of extreme values (>1.20) is larger in case group (9.49% vs. 3.33%).

ROC Analysis

Besides the measurement methods such as adherence score and pharmacy refill ratio, all the patients were asked questions concerning sexual behavior or pill use. We chose 13 of all the questions interviewed to be included in the final model since they are significant in the univariate logistic model. The 13 self-reported questions are listed in Table 4.

The logistic regression consists of 7 models:

1. A model including self-reported questions (QS) only.
2. A model including adherence score only.
3. A model including pharmacy refill ratio only.
4. A model including the combination of adherence score and QS.
5. A model including the combination of pharmacy refill ratio and QS.

6. A model including the combination of adherence score and pharmacy refill ratio.
7. The full model including the combination of adherence score, pharmacy refill ratio and QS.

Due to ROC curves in Figure 3A and 3B, the full model (the blue solid line) including pharmacy refill ratio, self-reported questions and adherence score has AUC of 0.7684 with AIC of 471.30 and -2LogLik of 439.30. The combination of pharmacy refill ratio and self-reported questions (the dashed purple line in Figure 3A) is the most predictive (AUC: 0.7614, AIC: 482.69, -2LogLik:452.69). Adherence alone (AUC: 0.5994, AIC: 539.51, -2LogLik: 535.51) and self-reported questions only (AUC: 0.6188, AIC: 517.70, -2LogLik: 489.70) do not predict the VF well since the AUC is smaller and AIC is larger. As well, the pharmacy refill ratio alone (AUC: 0.6816, AIC: 516.94, -2LogLik: 512.94) is not as predictive as its combination with self-reported questions (in Figure 3B).

In sum, the combination of pharmacy refill ratio and self-reported questionnaires is the most appropriate method to predict virologic outcome in RFVF study. Adherence alone cannot provide enough information in predicting virologic failure based on ROC analysis. Among the 13 self-reported questions, the strongest predictors are: how many doses missed in last week ($p=0.044$), whether used media to remember to take medications ($p=0.0065$), how often were busy with other things ($p=0.0094$).

Refill Gap Analysis

Repeated Measurements

In our study, refill gap was defined as the absolute value of pharmacy refill minus the days between two visits. For each patient, we consider refill gap as repeated outcomes since there are more than one visit and there might be unknown correlation among different visits. In the unadjusted GEE model, the status of patients (case/control) and number of visits are covariates of interest. As shown in Table 6A, the number of visits is statistically in the model ($p<0.0001$) and the assumption is satisfied that the number of pharmacy visits is negatively associated with the refill gaps. Whether the patients are in case or control group does not provide us with any information about the refill gaps ($p=0.3822$).

In addition, the average refill gap in case is 4.38, while the average refill gap in control is 4.67 (Table 6B). There is no difference of the refill gap in the two groups ($p=0.3461$). The patients in case and control group show similarity in refill pattern.

Generalized Linear Models

Let X_i denote the vector of covariates of interest for the *i*th patients. Here, adherence score is utilized as the covariate to illustrate the modeling approach. We compute the mean and variance of the refill gap for each patient. Since we are interested in studying how the adherence scores are associated with refill pattern, a linear and a quadratic model, respectively for mean and log-scaled variance are used.

The models are then $\mu_i = \beta_0 + \beta_1 X_i$ and $\log(\sigma_i^2) = \alpha_0 + \alpha_1 X_i + \alpha_2 X_i^2$.

Based on the observations from Table 7A and Table 7B, the mean of refill gap is found to be negatively associated with adherence score ($p<0.0001$). When the patients are more adherent to the medications, their average refill gaps tend to be smaller. Similar result holds for the quadratic model of log-scaled variance ($p<0.0001$). When the patients have higher adherence, variance of their refill gaps decrease. Therefore, an adherent patient are more likely to have a better refill pattern -- with gap mean close to 0 and smaller gap variance.

CHAPTER 4: CONLUSION & DISCUSSION

Descriptive analysis and smooth splines in our study shows adherence score and medication possession ratio do not perform well in predicting VF. Using t-test, we can find the adherence score ($p=0.4446$) is not significantly different in case and control group. MPR ($p=0.0067$) is statistically higher in case than in control. Pharmacy refill ratio ($p<0.0001$) of case is significantly larger than control. 60% -70% can be a cutoff point in adherence score, but the evidence is not strong enough. The poor adherence (<50%) and over-adherence (>110%) will indicate the higher probability of being VF. In addition, pharmacy refill ratio is another measurement method of adherence. We found the closer it is to 1, the higher the chances of the patients being VF, demonstrating it to be a valid adherence measure.

ROC analysis results in providing us a different view of the accuracy of the adherence measurement methods. The combination of pharmacy refill ratio and self-reported question come up to be a better tool in logistic model with its relatively higher AUC (0.7614), lower AIC (482.69) and smaller -2logLik (452.69). As the self-reported method is inexpensive to carry out and the pharmacy refill ratio is easier to calculate than adherence score and MPR, this finding is important in further study of pharmacy refill data.

The number of visits is strongly associated with the refill gap. It is statistically significant ($p<0.0001$) in GEE model. Recording the patients' visits can help us better

understanding their follow-up to the ART. However, whether or not the patients are in case or control group does not influence the refill pattern significantly (p=0.3822).

An interesting aspect of this study is that the adherence score and medication possession ratio are higher in case group than in control group based on our algorithm. The question is raised that whether the over-adherence implies virologic failure? However, this is not determined. An explanation of this could be the over-adherent patients are more likely to pour the pills in their bottles in order to look like "adherent". Another reason may be the choice of endpoint of our study. According to the enrollment criteria, the patients in case are more likely to have higher pill count. Therefore, this may lead to inaccurate information and calculation.

Our study's strengths include comparing the several adherence measure methods. Although it is hard to establish a gold standard for adherence measure, our findings suggest pharmacy refill ratio has some advantages in current ART study. Moreover, we computed an actual adherence score. Unlike many previous studies in which the measures were dichotomized as good versus poor adherence, our adherence score is able to include the information like pill count in retrospective study and quantify the adherence.

Although our study shows some informative results, there are a few limitations. Whether the missing data is missing at random should be justified, otherwise we should be cautious when using GEE model in order to satisfy the assumptions of covariance matrix. Second, the relatively short duration of ART (6 months) of the 458 patients may indicate that the patients who have difficulties consistently following to the ART may not

be selected [12]. Additionally, the self-reported questionnaires may bring reporting bias into the results since some people may not be honest with their answers.

To sum up, pharmacy refill ratio and self-reported questionnaires are helpful and achievable in predicting virologic outcome of ART in resource-limited countries. As well, the number of visits to pharmacy for each patient provides us with additional information. In further research, we can pay more attention on how to optimize the algorithm of adherence score. Furthermore, improving the study design may make us better understanding the refill pattern.

CHAPTER 5: TABLES & FIGURES

Table 1. Summarized Characteristics of Variables of Interest

Variables (Mean)	Case (n=158)	Control (n=300)	Overall	Missing
Gender (Male)	47.47 %	29.00%	35.37%	n=0
Age	37.07	40.90	39.58	n=0
Race (Black)	99.36%	98.66%	98.90%	n=1
MPR	95.25%	96.58%	96.11%	n=0
Adherence Score	98.48%	94.97%	96.13%	n=29
Pill Count at Enrollment	16.07	11.93	13.30	n=27
Days Supplied	160.92	165.39	163.87	n=4
Pills Dispensed	173.92	168.51	170.35	n=4
Pill Count Adherence	3.33	8.30	6.66	n=29
Pharmacy Refill Ratio	1.15	1.02	1.06	n=4
Last Refill Days	29.35	32.76	31.61	n=6

Table 2A. Adherence Score

drug	group	# N	Adherence Score mean	Adherence Score SD	Adherence Score median
3TC	case/control	18/24	0.9151/0.9527	0.2420/0.1934	0.9881/0.9556
EFV	case/control	112/245	0.9988/0.9453	0.0889/0.1082	0.9916/0.9611
TDF	case/control	11/16	1.0162/0.9863	0.1210/0.0962	1.0167/1.0088

Table 2B. Pill Count at Enrollment

drug	group	# N	Pill Count mean	Pill Count SD	Pill Count median
3TC	case/control	18/24	26.00/23.58	20.12/16.59	24.5/18
EFV	case/control	113/245	14.33/10.87	9.95/7.95	13/9
TDF	case/control	12/16	18.25/10	12.44/6.69	14/8

Table 2C. Pharmacy Refill Ratio

drug	group	# N	Pharmcy Refill Ratio mean	Pharmcy Refill Ratio SD	Pharmcy Refill Ratio median
3TC	case/control	19/26	1.0369/1.0484	0.1935/0.1091	1.0714/1.0651
EFV	case/control	123/254	1.1676/1.0134	0.8154/0.1068	1.0778/1.0404
TDF	case/control	12/17	1.1152/1.0571	0.1339/0.08407	1.1439/1.0714

Table 3A. Adherence Score Frequencies

level	Overall (n=454)	Control (n=299)	Case (n=155)
<0.50	29(6.39%)	12(4.01%)	17(10.97%)
0.50-0.70	10(2.20%)	8(2.68%)	2(1.29%)
0.70-0.90	81(17.84%)	63(21.07%)	18(11.61%)
0.90-1.10	305(67.18%)	203(67.89%)	102(65.81%)
>1.10	29(6.39%)	13(4.35%)	16(10.32%)

Table 3B. Medication Possession Ratio Frequencies

level	Overall (n=458)	Control (n=300)	Case (n=158)
<0.50	6(1.31%)	3(1.00%)	3(1.90%)
0.50-0.75	24(5.24%)	18(6.00%)	6(3.80%)
0.75-0.90	91(19.87%)	72(24.00%)	19(12.03%)
0.90-1.0	337(73.58%)	207(69.00%)	130(82.28%)

Table 3B. Pharmacy Refill Ratio Frequencies

level	Overall (n=458)	Control (n=300)	Case (n=158)
0.50-0.80	17(3.71%)	11(3.67%)	6(3.80%)
0.80-1.20	416(90.83%)	279(93.00%)	137(86.71%)
>1.20	25(5.46%)	10(3.33%)	15(9.49%)

Table 4. Self-Reported Questions

Questions	Response options (score recorded)			
1. How many doses have you missed in the last week? 2. How many doses have you missed in the last month? 3. How many doses did you take more than 1 hour late in the last week? 4. How many doses did you take more than 1 hour late in the last month?				
5.How do you remember to take your meds? (choice=Cell phone (3)) 6.How do you remember to take your meds? (choice=Media (TV/Radio)(7)) 7.How do you remember to come for your drug collection appt? (choice=Cellphone (3))	No (0) No (0) No (0)		Yes (1) Yes (1) Yes (1)	
8.How often you were away from home?	Never (0)	Rarely (1)	Sometimes (2)	Frequently (3)
9.How often you were busy with other things?	Never (0)	Rarely (1)	Sometimes (2)	Frequently (3)
10.How often you fell asleep through the dose time?	Never (0)	Rarely (1)	Sometimes (2)	Frequently (3)
11.How often you ran out of pills?	Never (0)	Rarely (1)	Sometimes (2)	Frequently (3)
12.How often you forgot to take pills?	Never (0)	Rarely (1)	Sometimes (2)	Frequently (3)
13.How often you had wanted to avoid side effects?	Never (0)	Rarely (1)	Sometimes (2)	Frequently (3)

Table 5A. ROC Statistics

ROC Models	Estimate	Standard Error	95% Wald Confidence Limits		Chi-Square	Pr > ChiSq
Self-reported questions	-0.1496	0.0260	-0.2006	-0.0986	33.0465	<.0001*
Pharmacy Refill ratio	-0.0868	0.0248	-0.1354	-0.0381	12.2284	0.0005*
Adherence score	-0.1690	0.0325	-0.2328	-0.1053	27.0052	<.0001*
Pharmcy refill ratio+QS	-0.00696	0.0110	-0.0286	0.0147	0.3976	0.5283
Adherence score+QS	-0.0704	0.0236	-0.1167	-0.0242	8.9013	0.0028*
PRR+adherence	-0.0862	0.0228	-0.1309	-0.0415	14.2627	0.0002*

Table 5B. ROC Fitness

ROC Models	AUC	AIC	-2LogL
Self-reported questions (QS) only	0.6188	517.70	489.70
Pharmacy refill ratio only	0.6816	516.94	512.94
Adherence score only	0.5994	539.51	535.51
Pharmcy refill ratio+QS	0.7614	482.69	452.69
adherence score+QS	0.6980	509.50	479.50
pharmcy refill ratio+adherence	0.6822	506.77	500.77

Table 6A. Analysis of GEE Parameters

Parameter	Estimate	Standard Error	95% CI		Z	Pr > |Z|
Intercept	10.8869	0.9919	8.9428	12.8309	10.98	<.0001*
Case/Control	0.2920	0.3342	-0.3629	0.9469	0.87	0.3822
# of visits	-1.3014	0.1875	-1.6689	-0.9338	-6.94	<.0001*

Table 6B. Refill Gap Statistics

Group	N	Mean	95% CI		Min	Max	Pvalue
Control	300	4.67	4.28	5.07	1.5	26.0	0.3461
Case	152	4.38	3.79	4.98	1.5	26.3	-

Table 7A. Refill Gap Model 1

Parameter	Estimate	Standard Error	t Value	Pr > |t|
Intercept	20.7264	1.3198	15.70	<.0001*
adherence	-16.8501	1.3645	-12.35	<.0001*

Table 7B. Refill Gap Model 2

Parameter	Estimate	Standard Error	t Value	Pr > |t|
Intercept	15.885	1.8147	8.75	<.0001*
adherence	-28.9161	3.9606	-7.30	<.0001*
adherence²	13.7594	2.1555	6.38	<.0001*

Figure 1A. Frequency of Adherence Score

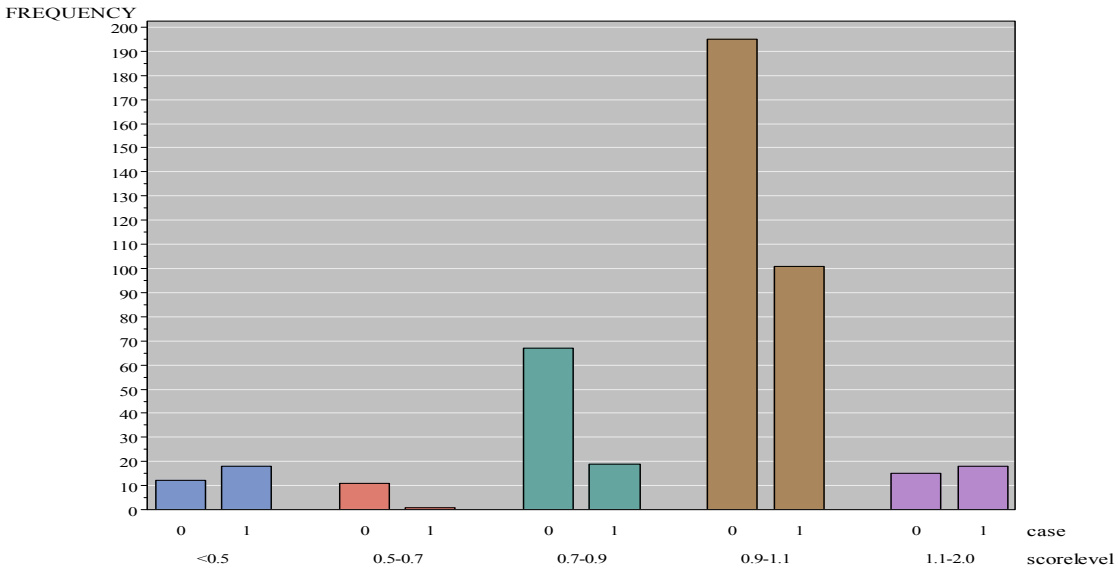


Figure 1B. Frequency of Medication Possession Ratio

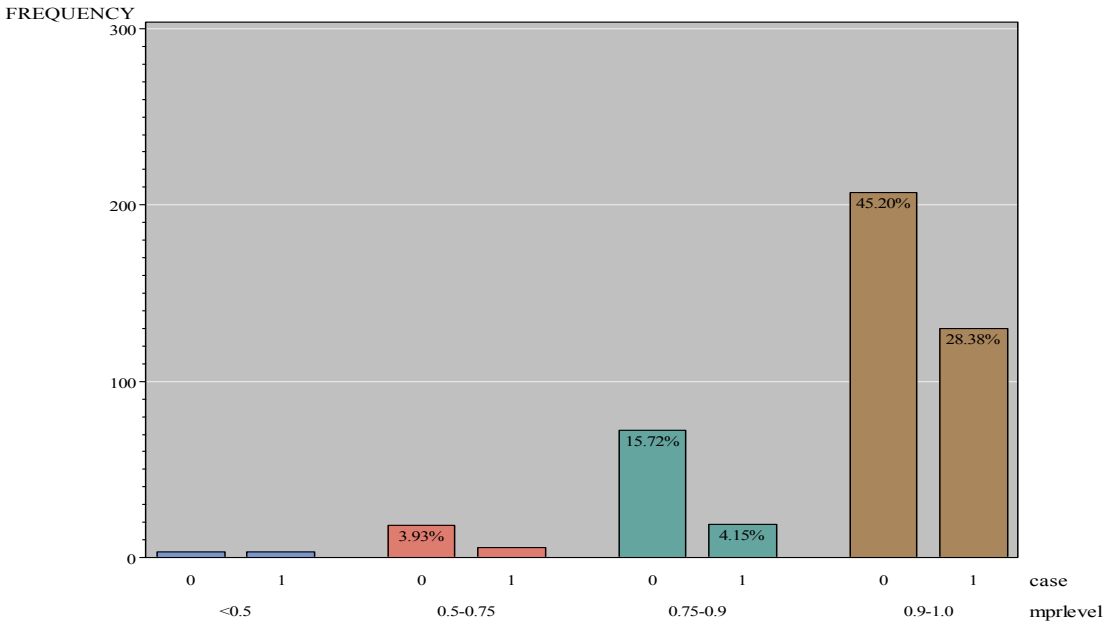


Figure 1C. Frequency of Pharmacy Refill Ratio

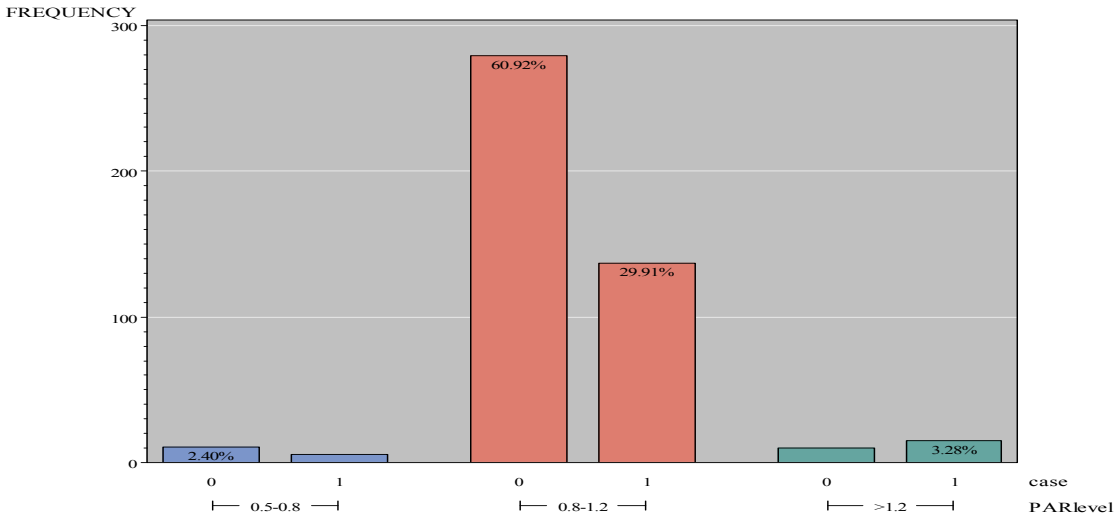


Figure 2A. Smooth Splines of Adherence Score

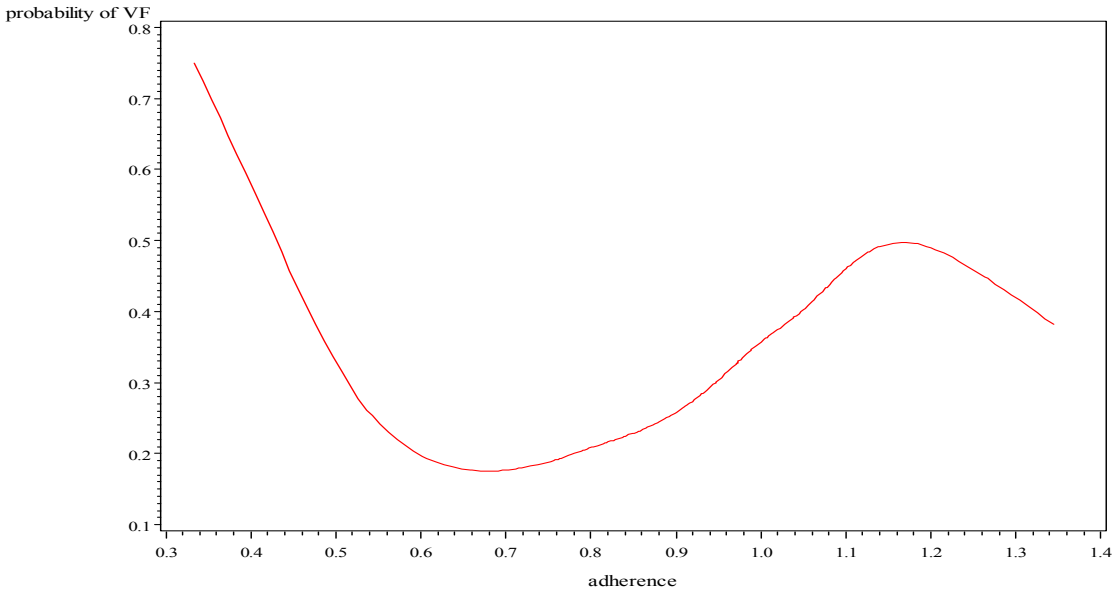


Figure 2B. Smooth Splines of Medication Possession Ratio

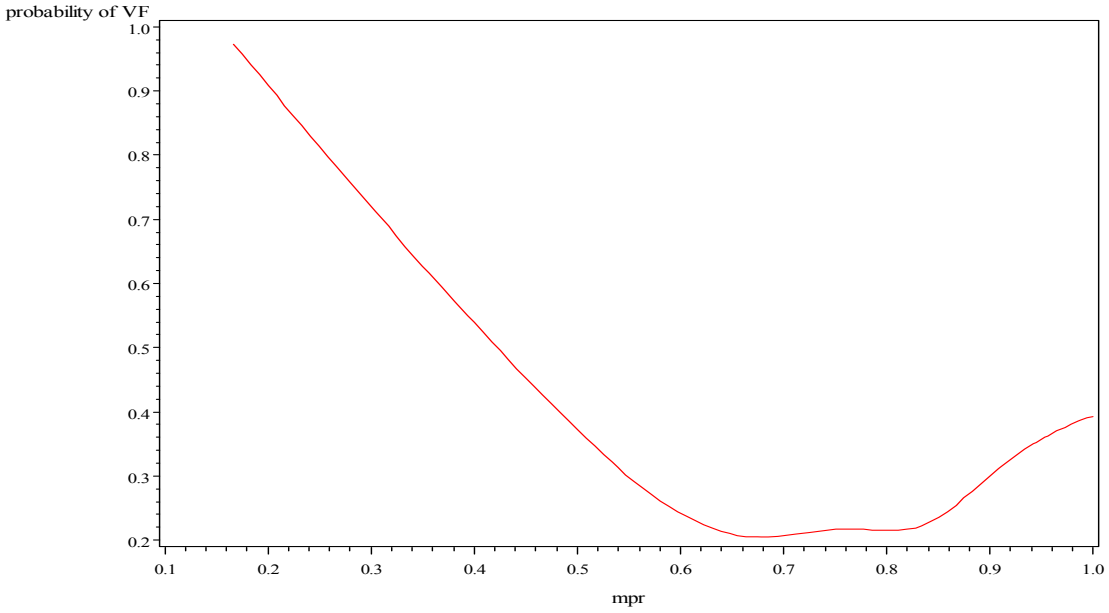


Figure 2C. Smooth Splines of Pharmacy Refill Ratio

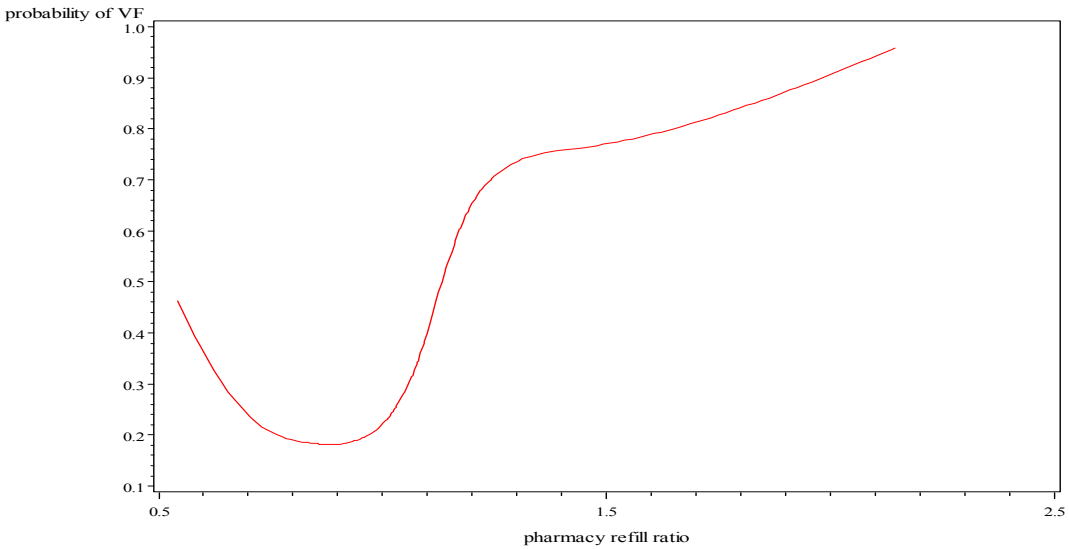


Figure 3A. ROC AUCs

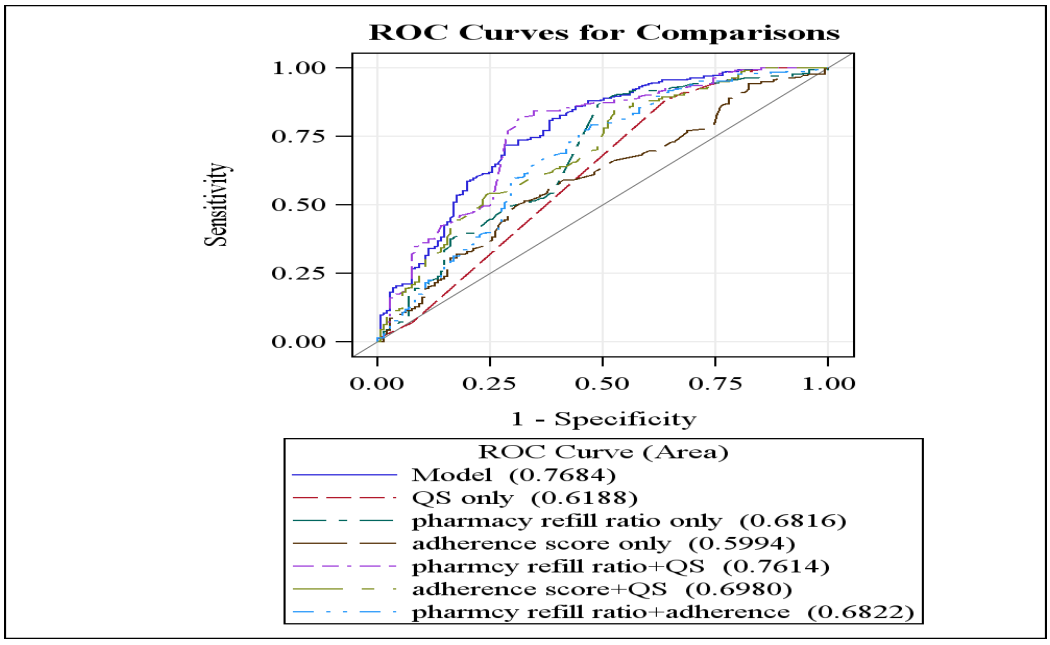
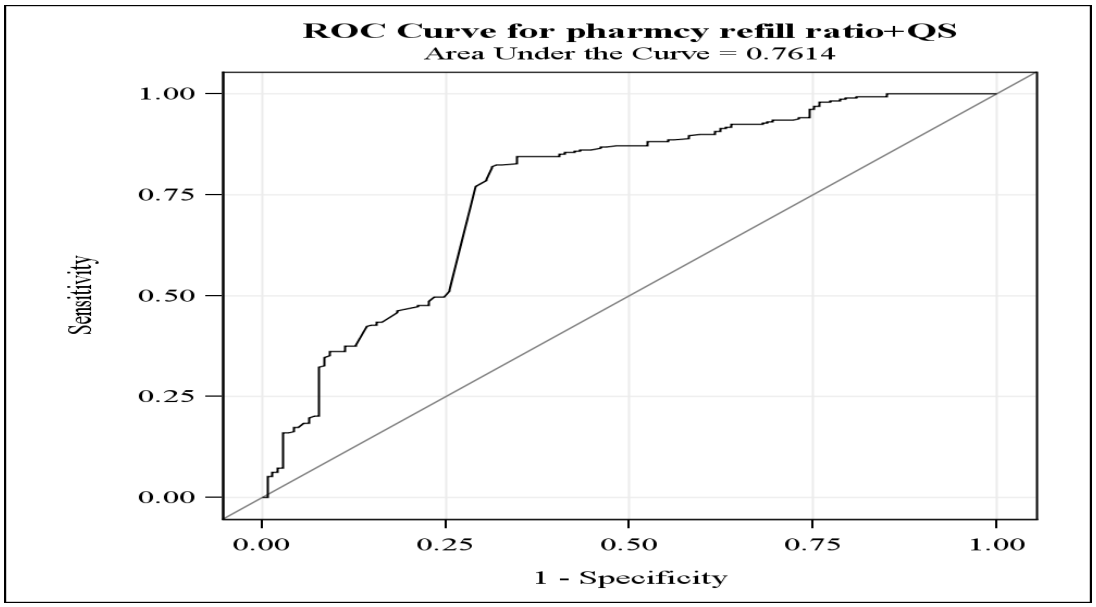


Figure 3B. ROC AUCs



References

[1] Dolin, Gerald L. Mandell, John E. Bennett, Raphael. Mandell, Douglas, and Bennett's principles and practice of infectious diseases, 2010.

[2] Joint United Nations Programme on HIV/AIDS (UNAIDS), Global report: UNAIDS report on the global AIDS epidemic 2010. 2010: Geneva, Switzerland.

[3] Shisana O et al. South African National HIV Prevalence, Incidence, Behavior and Communication Survey, 2008: the health of our children. Cape Town, HSRC Press, 2010.

[4] WHO 2013

[5] Sydney Rosen et al. Patient Retention in Antiretroviral Therapy Programs in Sub-Saharan Africa: A Systematic Review. PLOS Medicine, 2007.4(10): p. 1691-1701.

[6] Marconi, V.C., et al., Prevalence of HIV-1 drug resistance after failure of a first highly active antiretroviral therapy regimen in KwaZulu Natal, South Africa. Clin Infect Dis, 2008. 46(10): p. 1589-97.

[7] Murphy, R.A., et al., Outcomes after virologic failure of first-line ART in South Africa. AIDS, 2010. 24(7): p. 1007-1012.

[8] Anita Palepu et al. Antiretroviral adherence and HIV treatment outcomes among HIV/HCV co-infected injection drug users: The role of methadone maintenance therapy. Drug and Alcohol Dependence. 84(2006) 188-194

[9] Ann Intern Med. 1998 Sep 15; 129(6):503-5. The era of adherence to HIV therapy. Altice FL, Friedland GH. PMID: 9735090

[10] Calculating medication compliance, adherence, and persistence in administrative pharmacy claims databases. R. Sco Leslie, MedImpact Helathcare Systems, Inc., San Diego, CA

[11] J.B. Nachega et al. Adherence to Antiretroviral Therapy in HIV-Infected Adults in Soweto, South Africa. AIDS Research and Human Retroviruses. Vol 20, No.10, 2004. 1053-1056

[12] Robert Grossberg, Yawei Zhang, Robert Gross. A time-to-prescription-refill measure of antiretroviral adherence predicted changes in viral load in HIV. Journal of Clinical Epidemiology 57 (2004) 1107-1110.

[13] Nicoli Nattrass. What Determines Cross-Country Access to Antiretroviral Treatment? Development Policy Review, 2006, 24(3): 321-337.

[14] Paul Farmer et al. Community-Based Approaches to HIV Treatment in Resource-Poor Settings. The Lancet Vol 358, Aug 4, 2001.

[15] McNeil DG Jr. Selling Cheap 'Generic' Drugs, India's Copycats Irk Industry. New York Times. Dec 1, 2000: 1(A).

[16] Chesney MA. The Elusive Gold Standard: Future Perspectives for HIV Adherence Assessment and Intervention. JAIDS. Dec 2006; Vol 43 Issue pp S149-S155.

[17] Donna Katzman McCLish. Analyzing a portion of the ROC Curve. Med Decis Making.

[18] Fawcett Tom. An introduction to ROC analysis, Pattern Recognition Letters, 2006, 27, 861–874

[19] Nelder, John; Wedderburn, Robert (1972). "Generalized Linear Models". Journal of the Royal Statistical Society. Series A (General) (Blackwell Publishing) 135 (3): 370–384. doi:10.2307/2344614. JSTOR 2344614.

[20] Kung-Yee Liang and Scott Zeger (1986). "Longitudinal data analysis using generalized linear models". Biometrika 73 (1): 13–22.

[21] Hardin, James; Hilbe, Joseph (2003). Generalized Estimating Equations. London: Chapman and Hall/CRC. ISBN 978-1-58488-307-4.