ADDIS ABABA UNIVERSITY

OFFICE OF GRADUATE PROGRAMS

FACULTY OF SCIENCE

DEPARTMENT OF STATISTICS

EVALUATION OF FACTORS AFFECTING THE CHANCE OF SURVIVAL/DEATH STATUS AMONG HIV POSITIVE PEOPLE UNDER THE ANTI RETROVIRAL TREATMENT PROGRAM: THE CASE OF ADAMA HOSPITAL

By

Nuredin Ibrahim

A Thesis submitted to the Office of Graduate Programs of Addis Ababa University in Partial fulfillment of the requirement for the Degree of Master of Science in Statistics

AUGUST 2007

ADDIS ABABA UNIVERSITY

OFFICE OF GRADUATE PROGRAMS

FACULTY OF SCIENCE

DEPARTMENT OF STATISTICS

EVALUATION OF FACTORS AFFECTING THE CHANCE OF SURVIVAL/DEATH STATUS AMONG HIV POSITIVE PEOPLE UNDER THE ANTI RETROVIRAL TREATMENT PROGRAM: THE CASE OF ADAMA HOSPITAL

By

Nuredin Ibrahim

Approve by the Board of Examiners:

_____.	----------
Department Head	**Signature**
_____.	----------
Internal Examiner	**Signature**
_____.	----------
External Examiner	**Signature**

Tel. 251-1-239465 **P.O.Box 1176 Addis Ababa, Ethiopia**

ACKNOWLEDGEMENT

It is a pleasure to express my appreciation to those who have influenced this work.
My special gratitude to Professor Eshetu Wencheko continues for his encouragement, guidance and his kind correspondence up to the completion of this work.

I must thank the staff members of Adama Hospital ART Clinic, W/t Selam, Ato Semere and Sister Tibka for their cooperation during the time of data collection. Ato Seid and his friends also made my stay at Adama easier. I also thank the Addis Ababa University for allowing me to join the postgraduate program and I am very grateful to the staff members Department of Statistics for all their cooperation. The materials I got from Haj Adane Mamuye and his colleagues also helped me very much in the process of literature review. I also thank Ato Hussien for allowing me to use his Laptop computer throughout the course of the research.

Finally, I convey my thankful enjoyment to my family for their continuous encouragement in my studies.

ABSTRACT

Antiretroviral Treatment (ART), although not a cure, can help people from becoming ill for many years and this has improved the survival of HIV patients. In Ethiopia about 55,000 people and globally more than 1.3 million people have been taking ART until the end of 2006. Scaling up of ART treatment is planned by WHO and the Ministry of Health of the Ethiopian government. Although ART treatment has decreased HIV associated mortality and morbidity, a number of patients still die after the start of ART. A sample of 259 patients has been taken from Adama Hospital ART clinic for the study. The analysis of the data using the multivariate logistic regression method shows that some health, economic and risk behavior factors influence survival of patients. The study shows that survival of an HIV patient under ART follow up is affected by partners' HIV status, Base Line Weight, Base Line CD4 Counts, Condom use, Drinking Alcohol, taking Soft/Hard Drugs, and number of rooms the patients live in. The last factor is used as an indicator of economic status. Thus, the performance of ART programs can be improved if: we can bring behavioral change among HIV patients under ART follow up, take appropriate clinical and non-clinical measures like providing medicine and support (can be home-based) to patients and if health workers and other stakeholders find ways of supporting those patients of low economic status with respect to improving their nutrition in particular, and other assistance in general.

TABLE OF CONTENTS

CHAPTER ONE

1. INTRODUCTION

1.1 Background

The rate of spread of the HIV/AIDS and the damages accompanying it have reached a level which shock economists, health workers, politicians etc that it has now become a worldwide issue in general and developing countries in particular. The disease being one without any cure is still accountable for economic, social and health crises especially in developing countries. Its high prevalence and/or distribution among the youth made things even more complicated.

In just 25 years, HIV has spread relentlessly from a few widely scattered “hot spots” to virtually every country in the world, infecting 65 million people and killing 25 million. An estimated 38.6 million [33.4 million–46.0 million] people worldwide were living with HIV in 2005. An estimated 4.1 million [3.4 million–6.2 million] became newly infected with HIV and an estimated 2.8 million [2.4 million–3.3 million] lost their lives due to AIDS. In sub-Saharan Africa, the region with the largest burden of the AIDS epidemic, data also indicate that the HIV incidence rate has peaked in most countries. However, the epidemics in this region are highly diverse and especially severe in southern Africa, where the epidemic is still expanding. New survey data underscore the disproportionate impact of the AIDS epidemic on women, especially in sub-Saharan Africa where, on average, three women are HIV-infected for every two men. Among young people (15–24 years), that ratio widens considerably, to three young women for every young man [1].

In Ethiopia HIV was first detected in stored sera collected in 1984 and the first two AIDS cases were reported in 1986. In 2005 it was estimated that a total of 1,320,000 were living with HIV/AIDS. Of these, 634,000 were living in rural areas and 686,000 in urban areas. It was estimated that in 2005, a total of 137,500 new AIDS cases, 128,900 new HIV infections (353 a day) including 30,300 HIV positive births and 134,500 (368 a day) AIDS deaths (including 20,900 in children under 15 years) occurred. The estimated total number of persons requiring Antiretroviral Treatment (ART) in 2005 was 277,800

(including 43,100 children). AIDS accounted for an estimated 34% of all young adult deaths 15-49 in Ethiopia and 66.3% of all young adult deaths between 15 and 49 years of age in urban Ethiopia [2].

The 2006 Report on the Global AIDS Epidemic also indicates wide differences between countries in implementing the response pledged in the Declaration of Commitment on HIV/AIDS. While some have reached key targets and milestones, many have fallen short of the pledges made in 2001. According to this report if we do not urgently strengthen the AIDS response, neither the 2010 targets nor the Millennium Development Goal of halting the spread of AIDS and rolling back HIV infections by 2015 will be met. Failure to meet this Goal will also seriously endanger progress towards the Millennium Development Goals to reduce poverty, hunger and childhood mortality, as each of these is inextricably tied to our response-or lack of response-to AIDS. National economies and international security are at risk.

1.2 Motivation /Statement of the problem

Although the current HIV/AIDS Surveillance estimates indicate some encouraging signs in that the epidemic is stabilizing, the observed changes are not sufficient enough compared to the desired goals of the response against the epidemic. Given the size of the population and the magnitude of the damage inflicted, it will take us a number of years to see significant declines in HIV prevalence and incidence with concerted and sustained efforts. Although there are advances in the availability, accessibility and utilization of HIV/AIDS prevention, care, support and treatment services and improvements in the management of the epidemic and the increasing resource availability, we still face a situation unlikely to give us respite in the near future. Despite all the challenges, both governmental and non-governmental organizations are working hard to contain the epidemic and the achievements so far are encouraging. One of such endeavors comes from researchers and students from different universities. The results found from these researches have contributed their own share in the prevention, care, support and treatment services around the epidemic.

However, most of the researches in our country focused on the prevention, on the factors that increase the chance of contracting the disease etc, all dealing on how to prevent it before a person is HIV positive. For example, [3] worked on the latter case. It can be said that less attention was given for researches dealing with improving the situation of HIV positives taking ART. Unfortunately all our prevention programs will not realize their targets unless otherwise we give due consideration for people living with HIV/AIDS (PLWHA). It is this fact that little has been done on the factors that influence the survival/death status of a person given s/he is already HIV positive and is under the follow up of ART that motivated this study. The rationale behind such a research is to improve the achievements of ART programs run by different health institutions of the country in order to minimize HIV related mortality. The question we want to address here is "Which social, demographic, economic, health etc factors/variables affect the chance of survival/death among HIV-positive people taking ART?"

This is important as health policy workers on the disease must be able to determine the relative importance of various causes of death among HIV positives in order to do their jobs effectively.

It is hoped that this research will be useful because it is assumed to serve as an input for policy makers and concerned health specialists who work on providing care, support and treatment aspect of the HIV/AIDS programs of the country.

One of such programs available in the country is the so called ART which started in July 2003 on fee basis in 35 hospitals across the country. The objective of this program is mainly giving health related support/treatment for HIV positive persons. This includes different medical examinations, providing medicine and follow up of the patient/carrier. Here every improvement and/or bad conditions in the person's health status together with some social, demographic, economic and health variables will be recorded. Currently the program is inaugurated at 96 public and 13 private hospitals, and 77 health centers around the country under the support of the Federal Ministry of Health (FMOH) and

several non-governmental organizations. Among the largest ART service providers which have been integrated under this program in the Oromiya regional state is the Adama

Hospital ART clinic. And the research will cover this hospital which is located in one of the largest cities of the country where the problem of HIV/AIDS is more pronounced.

1.3 GOAL OF THE STUDY

To give ideas/suggestions that may improve the success of the ART program in our country.

1.4 OBJECTIVES OF THE STUDY

General objective:

To study some socio-economic, demographic and health factors that influence the survival/death status of HIV-Positives under ART follow up, i.e. to evaluate the association of the factors with survival/death status.

Specific objectives:

- **i)** To determine or get some clue on the relative importance of the factors,
- **ii)** to develop a statistical model that predicts the chance of survival/dying among HIV-positives taking ART,
- **iii)** to provide an estimate of the probability of dying under certain predictor variables,
- **iv)** to provide information for policy makers on the factors affecting survival/death status of HIV positives taking ART.

1.5 APPLICATION OF RESULTS

The outcome of the research will help health care workers to anticipate and inform patients about the possible causes of death they might encounter. Moreover, clinicians can decrease mortality among HIV positives by early diagnosis and appropriate intervention. On top of this, the result of the study will enable clinicians and policy makers to enhance the awareness of the society about factors which increase the probability of death in HIV

patients. The result of this study can also be used as a source of information to other researchers in the future.

Finally, this study will also contribute its part to the achievements of WHO and Ethiopian FMOH HIV/AIDS treatment goal. We can also list the applications of the research as follows:

♠ The major findings from this research are expected to benefit different stakeholders/activists working in the areas of giving care, support and treatment for HIV positives specifically and HIV/AIDS prevention in general.

♠ The mission of the above group would be achieved through getting the factors to put special focus on (those highly accountable for the death among HIV positives).

♠ The results obtained will serve as input in policy design regarding HIV/AIDS

♠ It is also believed that results of the research will be used as a basis for further study on HIV positives.

1.6 Limitation of the study

- The study does not cover HIV-positive individuals who take ART outside Adama ART Clinic.
- Lost and transferred cases from the sample which were later replaced by new sample units.
- Lack of literature on our country related to the subject under study.
- Poor data recording on the different patient charts.

CHAPTER TWO

2. CONCEPTS AND DEFINITIONS

2.1 Introduction to HIV/AIDS Treatment

Potent combination antiretroviral therapy (ART), consisting of 3 or more antiretroviral drugs (ARVs), has greatly improved the health and survival rates of HIV-infected patients in areas of the world with access to ARVs.

More than 20 individual ARVs are available in the resource-sufficient world, in addition to several fixed-dose combination preparations. These can be combined to construct a number of effective regimens for initial and subsequent therapy. ART is not without limitations, however. ART does not cure HIV infection and it requires that multiple medications be taken for very long periods of time (usually for the duration of life). It is expensive, may cause a variety of adverse effects (some severe), requires close adherence to be effective and to prevent the emergence of resistance, and often fails (because of the patient's imperfect adherence or other factors). The failure of an ARV regimens when accompanied by drug resistance usually means that subsequent regimens are less likely to succeed.

Greatly overshadowing the limitations of ART, however, is the overwhelming evidence that ART saves lives and improves or restores immune system function. Mortality and morbidity benefits are particularly obvious in patients with relatively advanced immune suppression or with symptoms related to HIV infection. For asymptomatic patients with relatively high CD4 cell counts (>350 cells/μL), it is less clear whether or when to start ART. In deciding when to start ART for any patient, practitioners must weigh the expected benefits of ART for that individual (in terms of morbidity and mortality) against the possible risks like toxicity, drug resistance, adverse drug interactions [4].

In the next sections some concepts about antiretroviral therapy and related issues are presented based on the information available on [4] and [5].

2.2 HIV antiretroviral treatment

This is the main type of treatment for HIV or AIDS. It is not a cure, but it can stop people from becoming ill for many years. The treatment consists of drugs that have to be taken every day for the rest of someone's life. ART for HIV infection consists of drugs which work against HIV infection itself by slowing down the replication of HIV in the body. The drugs are often referred to as antiretrovirals or anti-HIV drugs or HIV antiviral drugs.

2.3 Combination Therapy and HAART

For antiretroviral treatment to be effective for a long time, it has been found that the patient needs to take more than one antiretroviral drug at a time. This is what is known as Combination Therapy. The term Highly Active Antiretroviral Therapy (HAART) is used to describe a combination of three or more anti-HIV drugs.

Taking two or more antiretrovirals at the same time vastly reduces the rate at which resistance develops. HAART consists of a combination of three or more drugs. The most common combination given to those beginning treatment consists of two NRTIs combined with either an NNRTI or a "boosted" protease inhibitor. Ritonavir (in small doses) is the drug most commonly used to boost a protease inhibitor. An example of a common combination is the two NRTIs zidovudine and lamivudine combined with the NNRTI efavirenz.

2.3 Absence of HAART

When a person's immune system is damaged by HIV, then certain infections or cancers will develop that the body would normally "fight off" quite easily. These are known as opportunistic infections. Treatment for opportunistic infections is usually provided when antiretrovirals are not available, or when the antiretrovirals drugs are no longer effective as the HIV strain has become resistant to them.

The antiretroviral HIV drugs that are currently available can improve the quality of life of someone infected with HIV, helping them to stay well much longer than they otherwise

would. The drugs slow down the replication of HIV within the body, but it must be remembered that they are a treatment and not a cure.

When to start treatment

Deciding when to start treatment can be difficult as there is no proven ‘right’ time. There are different views of the benefits of starting HIV treatment earlier or later, though most guidelines recommend not starting treatment until the advanced stages of HIV infection. This is an important decision with long term consequences.

There are certain tests available that will help determine when to start treatment, in particular the CD4 test and the viral load test.

The CD4 Test

HIV attacks a type of immune system cell called the T-helper cell. This cell carries on its surface a protein called CD4, which HIV uses to attach itself to the cell before gaining entry.

The T-helper cell plays an important part in the immune system by helping to co-ordinate all the other cells to fight illnesses. A major reduction in the number of T-helper cells can have a serious effect on the immune system.

HIV causes many T-helper cells to be damaged or destroyed. As a result, there are fewer cells available to help the immune system to fight illnesses.

The CD4 test measures the number of T-helper cells in your blood. The more cells you have per cubic millimeter of blood, the stronger is your immune system. The stronger your immune system, the better your body can fight illnesses. A low CD4 count does not mean that you will certainly become ill, but it makes it more likely.

Generally, treatment is started when the CD4 test shows between 200 to 350 T-helper cells per cubic millimeter of blood, although advice varies slightly between countries.

The Viral Load Test

Viral load refers to the amount of HIV in your blood. Like the CD4 test, the viral load test can provide important information about the likely course of HIV infection. There are different viral load tests available, which use a variety of techniques to measure the amount of virus. The results of these tests tell you whether your viral load is low, medium or high.

Opportunistic Infections

As the immune system becomes increasingly damaged by HIV it becomes susceptible to opportunistic infections. These infections would usually be fought off easily by a healthy immune system, but a low T-helper cell count means opportunistic infections such as PCP (a type of pneumonia) can be life-threatening. Another common example is TB. If one of these illnesses has become a serious problem then HIV antiretroviral treatment may be advised straight away.

Choosing the best combination

For most people, there are a number of drug combinations available to choose from. There are more than 20 approved drugs belonging to four different groups. It is not always easy to tell which will be the best option, since a combination that suits one person might not suit another.

The first time you use antiretrovirals is when they are most effective. This is why you should try to get the combination right first time and strictly follow the guidelines on taking the therapy.

It is important that the drugs can be taken properly and on time. It is therefore necessary to think ahead about the restrictions and limitations that the drugs may impose on your lifestyle.

CHAPTER THREE

3. LITERATURE REVIEW

3.1 Background

HIV/AIDS pandemic has brought significant morbidities and mortalities all over the world. A total of 65 million people have been infected over the last two decades out of which 25 million people have died [1].

Fortunately, the introduction of combination ARV therapy in 1996 made HIV/AIDS to be a chronic manageable disease. For example, within the prison system in the United States, mortality due to AIDS has dropped dramatically since the advent of effective combination antiretroviral therapy, with the number of AIDS-related deaths decreasing by 72% in state prisons between 1995 and 2002 [7].

The goal of ARV therapy is to improve survival; to reduce HIV associated morbidity and mortality, to increase quality of life, to restore immune function and to achieve maximal and sustained suppression of viral replication [5].

On September 01, 2003 WHO/UNAIDS “3 by 5” initiative had set a target of putting 3 million people on ART by the end of 2005 in low income countries, but only 1.3 million people have started ART which is 55% of the target. By 2010 WHO has planed to put 9.8 million people on ART with the goal of providing universal access to HIV care and ART [1].

FMOH of Ethiopia has estimated 277,000 people to be eligible for treatment and it had planned to put 74, 000 people on ARV treatment by the end of 2006. But only about 55,000 people are getting ART at the end of 2006 [2]. The ARV drugs approved by FMOH for use in Ethiopia as first line drugs are NRTI and NNRTIs (d4T, 3TC, AZT, EFV and NVP). These drugs are combined to form the following four first line regimens taken by more than 95 % of adult HIV/AIDS patients in the Adama Hospital.

1.d4T/3TC/NVP 2.d4T/3TC/EFV 3.AZT/3TC/NVP 4 .AZT/3TC/EFV

In addition to the first line therapies few patients are taking TDF, NFVand LPV/r due to toxicity to the first line drugs or on whom the first line drugs have failed.

In the Ethiopian setting, even though broad study about the impact of ART on patient survival is inadequate, a cohort study from Arba Minch Hospital has shown that ARV administration has decreased HIV/AIDS mortality by 65 % [8]. It also implied some health factors like HG, TLC, BMI etc as potential predictors of early death in patients treated with HAART. Although these patients may survive many years after the diagnosis of AIDS if treated with HAART, some still die during treatment [25]. It is believed that, in resource-poor countries like Ethiopia, economic, demographic and health factors can also have considerable impact on the mortality rate of patients treated with ART. In other words, even if ARV treatment has shown significant clinical importance by meeting the goal of therapy, we are still facing a number of deaths that can otherwise be avoided by appropriate interventions on certain socio-economic, demographic and health factors.

3.2 General Literature

3.2.1 Socio-Economic factors and HIV death

The main focus of this literature review is to collect information from previous studies about the determinants of AIDS mortality, distribution of HIV/AIDS, socioeconomic and demographic status of different groups of individuals dying of AIDS. The review covered the tools that have been used to conduct related studies. There is evidence that the prevalence rate of HIV-infection is highest among most advanced social classes [9]. In [10] it is suggested that the gap between rural and urban HIV rates which was previously substantial is now narrowing rapidly in many countries. AIDS prevalence in rural Kenya is now higher than in the nation's urban areas.

Initially, most AIDS patients were found in Kenya's urban centers, but the situation reversed such that 62 percent of the country's HIV/AIDS cases are now in rural areas [11].

According to the report, the shift is because most organizations fighting AIDS concentrate their activities in towns. Rural communities bear a higher burden of the cost of HIV/AIDS as many urban dwellers and migrant laborers return to their villages when they become sick. This is also true for Ethiopia [1].

People living in developing countries especially in sub-Saharan Africa are the most affected by the disease. Sub-Saharan Africa continues to have the highest HIV prevalence and HIV death in the world [6].

Table 3.1: Adult (aged 15–49 years) HIV prevalence (%) in countries in sub-Saharan Africa which have conducted population-based HIV surveys in recent years

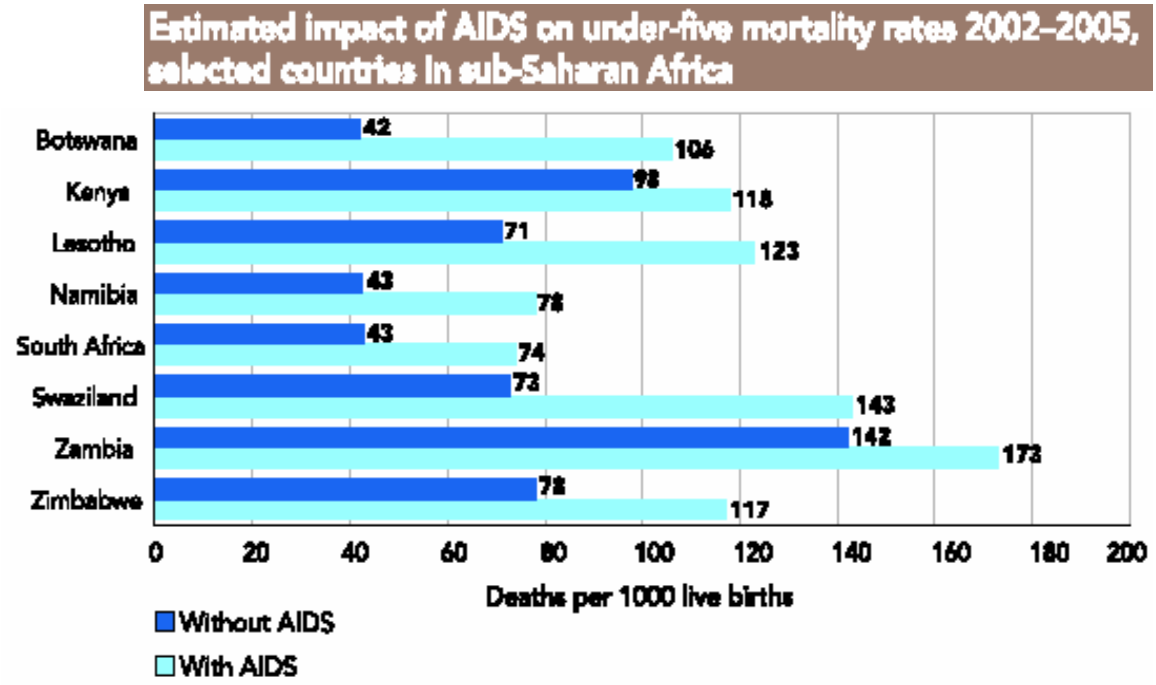
Country	Median HIV prevalence (%) among women attending antenatal clinics 2003–2004*	Population-based survey prevalence (%) (year)	2003 HIV prevalence (%) reported in 2004 Report on the global AIDS epidemic	Adjusted 2003 HIV prevalence (%) in current report	2005 HIV prevalence (%) in current report	Trend in prevalence
Botswana	38.5	25.2 (2004)	38	24	24.1	Stable
Burkina Faso	2.5	1.8 (2003)	4.2	2.1	2	Decline in urban areas
Burundi	4.8	3.6 (2002)	6	3.3	3.3	Decline in capital city
Cameroon	7.3†	5.5 (2004)	7	5.5	5.4	Stable
Ethiopia	8.5	1.6 (2005)	4.4	1.0–3.5	0.9–3.5	Decline in urban areas
Ghana	3.1	2.2 (2003)	3.1	2.3	2.3	Stable
Guinea	4.2	1.5 (2005)	2.8	1.6	1.5	Stable
Lesotho	28.4	23.5 (2004)	29.3	23.7	23.2	Stable
Rwanda	4.6	3.0 (2005)	5.1	3.8	3.1	Decline in urban areas
Senegal	1.9	0.7 (2005)	0.8	0.9	0.9	Stable
Sierra Leone	3	1.5 (2005)	–	1.6	1.6	Stable
South Africa	29.5	16.2 (2005)	20.9	18.6	18.8	Increasing
United Republic of Tanzania	7	7.0 (2004)	9	6.6	6.5	Stable
Uganda	6.2‡	7.1 (2004–5)	4.1	6.8	6.7	Stable

*WHO Africa (2005). HIV/AIDS epidemiological surveillance report for the WHO African Region, 2005 Update, Harare.

†Estimate based on country report for 2002 (2003), Ministry of Public Health Cameroon. National HIV sentinel surveillance report 2002.

‡Estimate based on country report 2002 (2003), Ministry of Health Uganda. STD/HIV/AIDS Surveillance report. STD/AIDS control programme, Kampala.

The above table (Table 3.1) shows that there is a spatial variation in HIV prevalence rates among countries in sub-Saharan Africa. In general, the HIV prevalence rates appeared to be rising in urban areas with time with levels above national averages. Rural areas showed lower incidences of HIV. Empirical studies support the view that HIV/AIDS is emerging as a leading cause of adult death and a substantial number of child deaths. HIV-infected children suffer from other severe infections like measles more than non-HIV-infected children [12]. HIV/AIDS also increases under-five mortality because of the diminishing care the children receive during the period of the parents' illness, as well as because of orphanhood. The extent to which orphans are under-enrolled relative to other children is country-specific, at least in part because the correlation between orphan status and poverty is consistent across countries [13].



Source: UNICEF (2005). United Nations Population Division, World Population Prospects: The 2004 Revision, database.

In developing countries affected by the HIV virus, the epidemic is deeply embedded in the cycle of poverty [14]. Poverty combines with gender inequality to fuel the spread of the virus – and the epidemic in turn impoverishes families. Gender differences in access to education, training, and paid employment are factors that disenfranchise women and increase their vulnerability to poverty, thus contributing to women's economic

dependence on men. Poverty has made many women and girls to engage in unsafe sexual practices and commercial sex as a way of earning a living. On the other hand, migration, urbanization, and social dislocation have given the virus an easy transmission route, and have increased men's vulnerability to the HIV virus [15]. Women in poorer households have less access to knowledge and are more likely not to have changed their behavior, meaning that they are relatively more vulnerable to infection than women from more affluent households [16]. The paper suggests that continued efforts aimed at uplifting women and addressing gender discrimination will be crucial in addressing HIV/AIDS.

The association between socioeconomic characteristics of a society and the provision of health care has become a point of interest among several authors. Inequalities have been reported, with particular contrasts between urban and rural areas [17]. Levels of urban health are at times viewed as being worse than in rural areas, but contrary evidence do also exist [18]. Health is believed to be influenced by ecological, social environment, genetics and individual characteristics. Urban and rural differences in mortality are determined by many factors, such as differences in life style, ecological situation, and access to social and health services, unequal distribution of incomes and resources [19]. In [19] it was also indicated that lower levels of rural poverty are associated with improved health for all indicators used. Rural people are less likely to know how to protect themselves against HIV. If they fall ill, they are less likely to receive adequate care. Poverty, widespread in rural areas, leads to poor nutrition and poor health, which make a person more vulnerable to HIV death. Poor health can also shorten the incubation period of the virus, causing symptoms to appear sooner. In addition, there is reduced access to medical care in rural areas. There have been dramatic improvements in health and mortality for most of the less developed countries during the past several decades, especially for infants and children. Cross-sectional studies have highlighted important degree of differences in infant and child mortality according to social, economic, and demographic factors with rural-urban differentials among the most commonly examined variables [20].

Some previous studies identified problems confronting HIV-infected rural residents as long distances to medical facilities and increased discrimination [21]. These studies indicated that compared with HIV-infected urban residents, people with HIV in rural areas would report lower life quality, more barriers to care, elevated perceptions of loneliness, and more frequent incidents of AIDS related discrimination, elevated fear that their HIV sero-status might become public knowledge, and less social support.

Regional differences in social cohesion and "social capital" have been put forward as a potential explanation of the variations in mortality experienced across areas of Hungary [22]. Social capital has been defined as the assets and resources available to individuals through their connections to their communities and society at large. It is hypothesized that more socially cohesive communities (i.e. communities richer in stocks of social capital) are better able to buffer the stresses and uncertainties associated with economic transformation. There is a close association between premature mortality and material deprivation [23]. One geographical manifestation of this relationship is the large contrast between the health status of residents of urban and rural areas, the former harboring pockets of poor health and the latter relatively low average rates of mortality (by most causes) and morbidity. In [22] it is stated that several studies have confirmed that this pattern of variation is replicated for all major disease classes, and in particular, heart disease, strokes, most cancers and respiratory diseases, with well known exceptions occurring in the case of traffic accidents and suicides, for which standardized rates tend to be higher in rural areas. According to [22] inequalities in all-cause premature mortality are widest in the cities, narrowest in the deeper rural areas. This is largely a reflection of the different distributions of material deprivation in these areas. Two thirds of Africans live in rural area which means that though the infection rates are still lower than in urban areas, the absolute numbers of HIV infected persons in rural communities will surpass that of urban populations [24]. At the same time, there has been an increase in urbanization with general migration from rural areas to cities. The changes in societal make-up, increased overcrowding, breakdown in cultural barriers due to disruption of social network has increased the vulnerability of urban population. This will contribute to the expected increase in rates of HIV infection. The study shows that there is a

correlation between urbanization and HIV/AIDS death, particularly in the peri-urban areas.

Despite its effective lowering effect on HIV/AIDS mortality, the coverage of Community Support/HIV Support groups for Patients is currently very low in Africa and likely to reduce drastically in the near future [25]. The low coverage is attributable to the limited government involvement in providing home-care services or just strengthening support to institutions providing home-based care. Further research and development is needed to develop affordable, feasible and sustainable home-care programs that can be implemented by staff working for the government, NGO and religion based health facilities. In addition, innovative strategies are required to establish effective partnership between the NGO, religion based and government health facilities.

3.2.2 Demographic and Health factors in relation to HIV/AIDS death

The notion that both demographic and health factors will have significant relationship is supported by many researchers. The impact of HIV on an individual's probability of dying is complex and depends on many factors including sex, mode of infection, number of infections, age at infection(s), immune competence, overall health, and treatment(s), among many factors [26]. This source [26] concluded that natural history of an HIV infection in an adult is reasonably well characterized but the population-level sex-age-specific mortality rate associated with HIV-related deaths is a very complex cumulation of many individual disease progressions. The number and sex-age composition of the infected subpopulation is in turn strongly influenced by population-specific behavioral factors such as sexual networking preferences.

Ethiopian girls and young women are particularly vulnerable to infection; a high age differential between regular male and female partners is often cited as a contributing factor to rising HIV incidence among Ethiopian women [2]. The prevalence of HIV among women was 4% while the prevalence among their men counterparts was 3% in 2005. According to the 2005 sentinel surveillance data, 106,000 HIV positive pregnancies occurred throughout country. There were also 134,450 (368 per day) AIDS-

related deaths including 20,929 children 0-14 years (83.6% under age five). Females accounted for 54.5% of AIDS-related deaths. The number of AIDS orphans aged 0-17 years reached 744,100 [2].

According to [20] indicated that as the epidemic ages, the sex-age composition of the infected and uninfected subpopulations change significantly and may eventually stabilize at some “equilibrium” level in the absence of interventions or other external impacts on the epidemic. As a result of trend of HIV epidemics and the number and sex-age composition of the infected and uninfected sub-populations vary significantly according to the specific characteristics of the individual populations. Consequently, a single universal HIV-related age-pattern of mortality is unlikely, and even a single infected population is likely to reveal many different HIV-related age-patterns of mortality as the epidemic ages.

There is generally little reliable national data describing the distribution of deaths by cause for sub- Saharan Africa. The data for one of sub-Saharan nations, South Africa, indicates that in 1996 roughly four percent of all deaths were attributed to HIV-related causes. These data have been superseded by work done by the South Africa Medical Research Council estimating that roughly 40 percent of deaths between ages 15 and 49 and 20 percent of all adult deaths in 2000 were attributable to HIV-related causes [27]. Given the lack of reliable national-level data on cause of death, the South African study instead relied on indirect methods to estimate the levels of HIV-related mortality, and it is among very few credible national-level studies of HIV-related mortality in sub-Saharan Africa [28].When we come to the availability of similar data on people treated with ART the situation is more or less the same.

In sub-Saharan Africa, mortality rates in the 15-54 year age cohort have risen dramatically since the onset of HIV/AIDS. The estimated life expectancy in sub-Saharan Africa is now 47 years, down by five years since 1993, and an estimated 15 years shorter than it would have been in the absence of AIDS [29]. In Ethiopia, the life expectancy lost due to AIDS alone had reached 5 years in 2005 [2]. The increase in adult mortality

associated with HIV/AIDS reduces the average number of years a newborn can expect to live.

While only a certain proportion of prime-age deaths can be attributed to AIDS, a review of recent epidemiological studies in Eastern and Southern Africa indicates that HIV is one of the leading causes of disease-related death among adults between 15-49 years in all cases [30].

Moreover, a growing emphasis among development planners on understanding the dynamics of poverty requires a better understanding of the effects of prime-age adult mortality. The effects of adult mortality can be readily assessed through standard nationally representative socio-demographic and economic surveys [20].

A recent study made in Belgrade [31] indicated that Long-term survival was associated with PI-based HAART regimens and lower viraemia, but not with the immunological status either at baseline or at the end of follow up. Structured treatment interruption (STI) when CD4 counts reach 350 cells/μL, along with undetectable viraemia, was a strong predictor of long-term survival.The study underscored that strong medical and social infrastructure is necessary to support and sustain both ongoing basic healthcare and the introduction of antiretroviral drug combinations that can transform HIV from a certain death sentence into a manageable disease.

In the case of Ethiopia, there is little research and literature on what affects the survival of patients taking ART. However, a cohort study made by [8] at Arbaminch Hospital exhibits WHO clinical stage IV, total lymphocyte count (TLC), Body mass index (BMI) and Weight loss as predictors of early death in HIV-positives using Kaplan Meier and Cox regression survival analyses methods. Another national symposium organized by Walta Information Center on ” The success and challenges of ART in Ethiopia” indicated behavioral factors such as drug or alcohol abuse, sexual behavior, etc determine the success of ART and drug adherence [38]. In addition to this, traditional and religious factors, socio-economic factors including poor living conditions, low level of literacy and

gender inequity are among the challenges of ART success in Ethiopia. For instance, the summary statistics from the papers presented on the workshop illustrate that 65% of those taking ART at ALERT hospital do not have any kind of income. They do not have access to shelter, food and transport. About 20% of the patients enrolled in the treatment cannot read and write, have no watches or cannot listen to radio to know the time for taking their drugs, which makes ART counterproductive.

CHAPTER FOUR

4. DATA AND METHODOLOGY

4.1 The Data

This is a retrospective study, which reviews the patient intake forms and follow up charts of adult HIV/AIDS patients taking combined antiretroviral therapy in Adama Hospital ART clinic which is located at Adama city in Oromiya Region. The clinic started its service in 2003 and it has currently two physicians, three nurses and two data clerks attending HIV/AIDS patients regularly and filling the follow up charts more or less appropriately. A total 4121 patients have visited the ART clinic until February 2007 of which 205 have died while some 670 are either transferred to other similar clinics or lost to follow up.

The patient charts include the patient intake forms and follow up cards, which are prepared by FMOH to be uniformly used by clinicians to early identify and document clinical and laboratory variables. Thus, in this research we will use secondary data which will be collected from patient follow up records (ART clinic patient record) based on the questionnaire designed to extract only those variables to be considered in this study.

4.2 Sampling Design

4.2.1 Sampling frame and Procedure

The target population for this study will be patients under the follow up of ART at Adama hospital until February 2007. Here patients who are transferred to other health institutions or those lost to follow up cases are excluded from this study. This is because we cannot have information about their survival/death status. We have totally 3651 patients in our sampling frame, which is the list of all patients who visited the hospital since the initiation of the ART program. Here each patient has a chart/record with distinctive identification number which is known as ART unique identification number.

In this study systematic random sampling method is adopted as an appropriate sampling design for selecting a representative sample of the patients based on their ART unique identification number.

4.2.2 Sample Size Determination

In conducting researches that require taking a sample, we always have the stage of deciding the sample size. The decision is important because taking too large sample implies waste of resources while too small sample reduces the usefulness of the results. In order to have an optimum sample size, there are a number of issues/points one has to take into account. Some of the issues are:

- Objective of the research
- Design of the research
- Cost constraint
- Plan for statistical analysis
- Degree of precision required for generalization
- Degree of confidence with which to conclude etc.

Based on the above information, there are several formulas developed for sample size calculation that conform to different research situations. Accordingly, the sample size determination formula adopted for this study is:

$$n_0 = \frac{Z_{\frac{\alpha}{2}}^{2}\, p(1-p)}{d^2}$$

where n_0 is the sample size, p the population proportion of death and d is the absolute precision defined as:

d= $Z_{\frac{\alpha}{2}}$ SE, where SE is the standard error.

And if sampling is from a finite population of size N, then

$$n = \frac{n_0}{\left(1+\frac{n_0}{N}\right)} \quad .$$

Further discussions on sampling methods are available in detail in [32] and [33].

In this study, the sample size is determined by using the proportion of death (p), 5% of statistical level of significance (alpha=0.05) and an absolute precision of 0.03, i.e., d= 0.03. The population proportion of death (p) is estimated to be about 7% by the clinicians in the ART clinic. Hence, the sample size with N=3651, together with above specifications will be,

n_0 = 278 implying that n = 259.

Thus a total of some 259 patient charts have been reviewed for the data collection. The researcher and two data collectors (both data clerks of the ART clinic) collected the data for the study.

4.2.3 Variables considered in the research

I-The dependent variable is the dichotomous random variable "Survival/Death Status" (dead=1, alive=0) of an HIV positive person under ART follow up. It is denoted by DEATH-STAT.

II- Independent variables

The variables that are assumed to influence the prediction of survival/death status are presented in the table below.

Table 4.1: Independent variables

No.	Description and Name	Categories
1	Age in complete years(AGE)	(0) < 26 (1) Between 26 and 45 (2) > 45
2	Sex(SEX)	(0) Female (1) Male
3	Marital Status(MARIT_STAT)	(0) Never married (1) Married (2) Divorced/Separated/Widowed

4	Educational level(EDUC_LEV)	(0) No education (1) Primary (2) Secondary or above
5	Previous attendance of HIV counseling Session(s)(COUNS)	(0) No (1)Yes
6	Residence(RESID)	(0)Urban (1) Rural
7	Employment Status(EMP_STAT)	(0) Working full/ part time (1) Not working due to ill health (2) Unemployed
8	Number of rooms(NO_ROOMS)	(0) <2 (1)>=2
9	Running water(RUN_WAT)	(0) No (1)Yes
10	Running electricity(RUN_ELECT)	(0) No (1)Yes
11	TB(TB_STAT)	(0)Positive (1) Negative/UK
12	Partners HIV status(PART_HIV)	(0) Negative (1)Positive (2) Unknown/NA
13	Base Line Weight(BASE_WEIG)	(0)<45 (1)>=45
14	Base Line CD4 Counts(BASE_CD4)	(0)<150 (1)>=150
15	WHO Clinical Stage(STAGE)	(0) Stage I (1) Stage II (2) Stage III (3) Stage IV
16	Antiretroviral regimen(REGIMEN)	(0) d4T (30 or 40)/3TC/NVP-(1a 30 or 40) (1) d4T(30 or 40)/3TC/EFV-(1b 30 or 40)

		(2) AZT/3TC/NVP- (1C) OR AZT/3TC/EFV-(1d)
17	Condom use(CONDOM)	(0) Mostly (1) Never/ Rarely/No response
18	Tobacco(TOBAC)	(0) No (1) Yes (2) No response
19	Alcohol(ALCOH)	(0) No (1) Yes (2) No response
20	Soft/Hard Drugs(DRUG)	(0) Yes (1) No

III-Interactions

Two interaction terms are also proposed for this study. The first is interaction between Base Line Weight (BASE_WEIG) and WHO Clinical Stage (STAGE) denoted by BASE_WEIG*STAGE while the other is interaction between Educational level (EDUC_LEV) and Employment Status (EMP_STAT) represented by EDUC_LEV * EMP_STAT.

4.3 Methodology

4.3.1 Introduction

There are many situations in which the response of interest is dichotomous rather than continuous. Examples of variables that assume only two possible values are disease status (the disease is either present or absent), survival following surgery (a patient is either alive or dead), survival after knowledge of being HIV positive (alive or dead) etc. In general the value 1 is used to represent a "success" or the outcome we are most interested in, and 0 represents a "failure". Just as we estimate the mean of a response variable Y when it is continuous, we would like to be able to estimate the probability of a dichotomous response (which of course is also its mean) for various values of explanatory variables. To do this, we use a technique known as Logistic Regression.
Since all the above explanation matches exactly with the case in this research, it implies the logistic regression as the methodology to be adopted [34].

Logistic regression can be binary or multinomial. The binary or Binomial logistic regression is the type of regression which is used when the dependent variable is a dichotomous and the independent variables are of any type while Multinomial logistic regression is used when the dependent variable has more than two categories. When multiple classes of the dependent variable can be ranked, then ordinal logistic regression is preferred to multinomial logistic regression. Continuous variables are not used as dependent variables in logistic regression.

Logistic regression can be used to predict a dependent variable on the basis of continuous and/or categorical independent variables and to determine the percent of variance in the dependent variable explained by the independent variables; to rank the relative importance of independent variables; to assess interaction effects; and to understand the impact of covariate control variables. The logistic regression applies maximum likelihood estimation after transforming the dependent into a logit variable (the natural log of the odds of the dependent variable occurring or not). In this way, logistic regression estimates the probability of a certain event occurring. Note that logistic regression calculates changes in the log odds of the dependent variable, not changes in the

dependent variable itself as OLS regression does. However, logistic regression has many analogies to OLS regression: logit coefficients correspond to b coefficients in the logistic regression equation, the standardized logit coefficients correspond to beta weights, and a pseudo R2 statistic is available to summarize the strength of the relationship. Unlike OLS regression, however, logistic regression does not assume linearity of relationship between the independent variables and the dependent variable, does not require normally distributed variables, does not assume homoscedasticity, and in general has less stringent requirements [35].

The logistic regression is also preferred from multiple regression and discriminant analysis as it results in a biologically meaningful interpretation, it is mathematically flexible and easily used distribution and it requires fewer assumptions [36].

4.3.2 The Logistic Function

Consider the model

$$p = \alpha + \beta x$$

where p is the probability of "success", α is the intercept of the line and β is its slope. This model is not feasible even on inspection. Since p is a probability, it is restricted to taking values between 0 and 1. The term $\alpha + \beta x$, in contrast, could easily yield a value that lies outside this range.

We might try to solve this problem by fitting the model as

$$p = e^{\alpha+\beta x}$$

This equation/model guarantees merely the estimate of p is positive. Otherwise the term $e^{\alpha+\beta x}$ although cannot be negative, it can result in a value that is greater than 1.

To accommodate this final constraint, we fit a model of the form

$$p = \frac{e^{\alpha+\beta x}}{1+e^{\alpha+\beta x}}.$$

The expression on the right, called a logistic function, cannot yield a value that is negative or greater than 1;consequently restricting the estimated value of p to the required range (between 0 and 1).

If an event occurs with probability p, then the odds in favor of the event are $\frac{p}{1-p}$ to 1.

Thus, if a success occurs with probability

$$p = \frac{e^{\alpha+\beta x}}{1+e^{\alpha+\beta x}},$$

the odds in favor of success are

$$\frac{p}{1-p} = \frac{e^{\alpha+\beta x} \big/ (1+e^{\alpha+\beta x})}{1 \big/ (1+e^{\alpha+\beta x})}$$

$$= e^{\alpha+\beta x}$$

Taking the natural logarithm of each side of this equation,

$$\ln\left[\frac{p}{1-p}\right] = \ln\left[e^{\alpha+\beta x}\right]$$

$$= \alpha + \beta x$$

Thus, modeling the probability p with logistic function is equivalent to fitting a linear regression model in which the continuous response y has been replaced by the logarithm of the odds of success for a dichotomous random variable. Instead of assuming that the relationship between p and x is linear, we assume that the relationship between ln $\left[p/(1-p)\right]$ and x is linear. The technique of fitting a model of this form is known as logistic regression.

4.3.3 The Multiple Logistic Regression Model

Consider a collection of k independent variables which will be denoted by the vector $X' = (x_1, x_2, \ldots, x_k)$. Let the conditional probability that the outcome is present be denoted by $P(Y=1|\ X) = p(X)$. Then the logit of the multiple logistic regression is given by the equation

$$g(X) = \alpha + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k$$

and the odds in favor of success for the multivariate logistic regression will be

$$\ln\left[\frac{p}{1-p}\right] = \ln\left[e^{g(X)}\right]$$

$$= \alpha + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k$$

in which case

$$p(X) = \frac{e^{g(X)}}{1+e^{g(X)}}$$

$$= \frac{e^{\alpha+\beta_1 x_1+\beta_2 x_2+\ldots+\beta_k x_k}}{1+e^{\alpha+\beta_1 x_1+\beta_2 x_2+\ldots+\beta_k x_k}}$$

4.3.4 Fitting the Logistic Regression Model

Suppose we have a sample of n independent observations of the pair (x_i, y_i), i=1,2, …,n, where y_i denotes the value of the dichotomous outcome variable and x_i is the value of the independent variables for the i^{th} subject. Fitting the model requires that we obtain estimates of the values of α and β_1 (represented by a vector β').

In linear regression the method used most often for estimating unknown parameters is least squares. In that method we choose those values of β' which minimize the sum of squared deviations of the observed values of Y from the predicted values based upon the model. Under the usual assumptions for linear regressions the method of least squares yields estimators with a number of desirable statistical properties. Unfortunately when the method of least squares is applied to a model with a dichotomous outcome the estimators no longer have these same properties [36].

The general method of estimation that leads to the least squares function under the linear regression model (when the error terms are normally distributed) is called maximum likelihood. It is this method that provides the foundation for our approach to estimation with the logistic regression model. In a very general sense the method of maximum likelihood yields values for the unknown parameters which maximize the probability of obtaining the observed set of data. In order to apply this method we must first construct a function called likelihood function. The maximum likelihood estimators of these parameters are chosen to be those values which maximize this function. Thus, the resulting estimators are those which agree most closely with the observed data.

If Y is coded as zero or one then the expression for p(x) above provides (for an arbitrary value of β', the vector of parameters) the conditional probability that Y is equal to 1 given x (i.e., P(Y=1|x)). It follows that 1- p(x) gives the conditional probability that Y is equal to 0 given x, P(Y= 0|x). Thus, for those pairs (x_i, y_i), where y_i=1 the contribution to the likelihood function is p(x_i), and for those pairs where y_i=0 the contribution to the likelihood function is 1-p(x_i), where the quantity p(x_i) denotes the value of p(x) computed at x_i.A convenient way to express the contribution to the likelihood function for the pair (x_i, y_i) is through the Bernoulli distribution,

$\theta(x_i) = \mathrm{p}(x_i)^{y_i}\left[1-\mathrm{p}(x_i)\right]^{1-y_i}$, where $p\left(x_i\right) = \dfrac{e^{\alpha+\beta x_i}}{1+e^{\alpha+\beta x_i}}$, i=1, 2, ..., n

Since the observations are assumed to be independent, the likelihood function is obtained as the product of the terms given in the above expression as follows:

$l(\beta) = \prod_{i=1}^{n}\theta(x_i)$ …….(*)

The principle of maximum likelihood states that we use as our estimate of β the value which maximizes the expression in equation (*). However, it is easier mathematically to work with the log of equation (*). This expression, the log likelihood is defined as

$L(\beta) = \ln\left[l(\beta)\right] = \sum_{i=1}^{n}\left\{y_i \ln[p(x_i)] + (1-y_i)\ln[1-p(x_i)]\right\}$

To find the value of β that maximizes $L(\beta)$ we differentiate $L(\beta)$ with respect to α and β_1 and set the resulting equation to zero. These equations are as follows:

$\sum_{i=1}^{n}[y_i - \mathrm{p}(\mathrm{x}_i)] = 0$ ……. (a) and

$\sum_{i=1}^{n} x_i[y_i - \mathrm{p}(\mathrm{x}_i)] = 0$ ……. (b)

And these equations are called the likelihood equations. For logistic regression the expressions in (a) and (b) are non linear in α and β_1, and thus require special methods for their solution. These methods are iterative(like Newton Raphson) in nature and have been programmed into available logistic regression packages like SPSS, STATA, SAS etc.

We can proceed to the multivariate case with the same steps followed in the univariate case above.

4.3.5 Model Building Strategies/Variable Selection

In modeling with many independent variables, one is usually concerned with the goal of selecting those variables that result in the "best" model within the scientific context of the problem. Having a basic plan to follow in selecting the variables for the model and assessing the adequacy of the model both in terms of the individual variables and from the point of view of the overall fit of the model is required for achieving this "best" model. It is also highlighted in [36] that successful modeling of a complex data set is part science, part statistical methods, and part experience and common sense.

The traditional approach to statistical model building involves seeking the most parsimonious model that still explains the data. Recently researchers are shifting to including all scientifically relevant variables in the model, irrespective of their contribution to the model. This is based on the fact that it is possible for individual variables not to exhibit strong association while they do show considerable association when taken collectively. Both approaches have their merits and demerits and details of this can be found in [36].

In general the following steps are recommended by [36] to aid in the selection of variables for a logistic regression. Firstly, the selection process should begin with univariate analysis of each variable. Secondly, selection of variables for the multivariate analysis will follow based on the results in the univariate analysis along with all variables of known biologic importance. Finally, the importance of each variable included in the multivariate model should be verified by different model assessment techniques.

If we have a large number of possible independent variables then it may be worthwhile to employ stepwise selection procedures like forward stepwise or backward stepwise in fitting the models as they are available in different statistical software like SPSS and SAS.

4.3.6 Assessing the Fit of the Model

Once a model has been developed through the various steps indicated in the above section, we now would like to know how effective the model is in describing the outcome variable. This is what is referred to as goodness-of-fit. We say our model fits when summary measures of the distance between the observed response variable $Y' = (y_1, y_2, y_3, ..., y_n)$ and its fitted value $\hat{Y}' = (\hat{y}_1, \hat{y}_2, \hat{y}_3, ..., \hat{y}_n)$ are small and the contribution of each pair $(y_i, \hat{y}_i)$, i=1,2,3, …, n to these summary measures is unsystematic and is small.

In testing the hypothesis that the model fits the data, the two common approaches are Pearson's X^2 statistic and the likelihood-ratio statistic(G^2) which are based on the comparison of the fitted and the observed counts. The large values of X^2 and G^2 indicate lack of fit of the model. When the fit is poor, residuals and other diagnostic measures describe the influence of individual observations on the model fit and highlight reasons for the inadequacy [37].

The likelihood-ratio statistic (G^2) is given by

$$G^2 = 2\sum\left[(Observed)\log\left(\frac{Observed}{Fitted}\right)\right] = -2LL_R - (-2LL_F) = -2\ln\left(\frac{likelihood_R}{likelihood_F}\right).$$

The likelihood ratio is -2ln (likelihood R) for a restricted (smaller) model minus (-2 likelihood F) for a full (larger) model that is the same as the log of the ratio of the two likelihoods, which is distributed as chi-square. The full or larger model has all the parameters of interest in it. The restricted is said to be nested in the larger model. The restricted model has one or more of parameters in the full model restricted to some value (usually zero). The parameters in the nested model must be a proper subset of the parameters in the full model. The chi-square is used to statistically test whether including a variable reduces badness-of-fit measure. This is analogous to producing an increment in R-square in hierarchical regression. If chi-square is significant, the variable is considered to be a significant predictor in the equation.

The Pearson's X^2 statistic is given by

$$X^2 = \sum\left[\frac{\left(Observed - Fitted\right)^2}{Fitted}\right] = \sum\left[\frac{\left(y_i - m_i\hat{p}_i\right)^2}{m_i\hat{p}_i(1-\hat{p}_i)}\right]$$

where m_i is number of subjects with $X=X_i$, $\hat{p}_i$=Fitted probability and $\sum m_i = n$.

The Hosmer-Lemeshow Test is another alternative in checking model fitness. This is based on the work of Hosmer and Lemeshow who proposed grouping based on the values of the estimated probabilities [36]. The grouping can be either based on the percentiles of the estimated probabilities or fixed values of the estimated probabilities. In either case, the Hosmer-Lemeshow goodness-of-fit statistic, $\hat{C}$, is obtained by calculating the Pearson chi-square statistic from the 2xg table of observed and expected frequencies and its formula is given as:

$$\hat{C} = \sum_{k=1}^{g}\frac{\left(O_k - n_k'\bar{p}_k\right)^2}{n_k'\bar{p}_k(1-\bar{p}_k)}$$

where g is the number of groups, n_k' is the number of covariate patterns in the k^{th} group, $O_k = \sum_{j=1}^{n_k'} y_j$ the number of responses among the n_k' covariate patterns, and $\bar{p}_k = \sum_{j=1}^{n_k'} m_j\bar{p}_j / n_k'$ the average estimated probability.

If the Hosmer-Lemeshow goodness-of-fit test statistic is greater than 0.05, we will not reject the null hypothesis that there is no difference between observed and model-predicted values, implying that the model estimates are adequate to fit the data at an acceptable level.

The leverage statistic, h, is available to identify cases which influence the logistic regression model more than others. The leverage statistic varies from 0 (no influence on the model) to 1 (completely determines the model). The leverage of any given case may be compared to the average leverage, which equals p/n, where p = (k+1)/n, where k= the number of independent variables and n=the sample size. Note that influential cases may nonetheless have small leverage values when predicted probabilities are <0.1 or >0.9.

The leverage of the i^{th} case, h_i, is the i^{th} diagonal element of the matrix

$\hat{V}^{\frac{1}{2}} X (X' C \hat{V} X) X' \hat{V}^{\frac{1}{2}}$

where $\hat{V} = Diag\{\hat{p}_1(1-\hat{p}_1),..., \hat{p}_n(1-\hat{p}_n)\}$, $C =$ the asymptotic covariance matrix for $\hat{\beta}$ [39].

Cook's distance is another measure of influence of a case and it depends on the standardized residual for a case as well as its leverage. It is a measure of how much deleting a given case affects residuals for all cases.

Cook' s $D_i = z_i h_i / (1-h_i)^2$ where z_i is the standardized residual and h_i is the leverage.

Finally, we conclude our discussion of model assessment by giving a diagnostic check for the significance of individual model estimates-the Wald test.

The Wald statistic is an alternative test which is commonly used to test the significance of individual logistic regression coefficients for each independent variable (that is, to test the null hypothesis in logistic regression that a particular logit (effect) coefficient is zero,i.e, $\beta_i = 0$ against $\beta_i \neq 0$). For a dichotomous independent variable, the Wald statistic (W) is the squared ratio of the unstandardized logit coefficient to its standard error, that is,

$$W = Z^2 = \frac{\hat{\beta}_i^2}{\text{var}(\hat{\beta}_i)}$$

W has χ^2 distribution with one degree of freedom.

4.3.7 Assumptions regarding the logistic regression model

As indicated in the previous sections, one advantage of the logistic regression is it gives some relaxation with respect to the usual OLS assumptions. There are, however, other assumptions one should consider for the efficient use of logistic regression as detailed in [35]:

1. Meaningful coding: Logistic coefficients will be difficult to interpret if not coded meaningfully. The convention for binomial logistic regression is to code the

dependent class of greatest interest as 1 and the other class as 0, and to code its expected correlates also as +1 to assure positive correlation. For multinomial logistic regression, the class of greatest interest should be the last class. Logistic regression is predicting the log odds of being in the class of greatest interest.

2. Inclusion of all relevant variables in the regression model: If relevant variables are omitted, the common variance they share with included variables may be wrongly attributed to those variables, or the error term may be inflated.
3. Exclusion of all irrelevant variables: If causally irrelevant variables are included in the model, the common variance they share with included variables may be wrongly attributed to the irrelevant variables. The stronger the correlation of the irrelevant variable(s) with other independent variables, the greater the standard errors of the regression coefficients for these independent variable.
4. Error terms are assumed to be independent (independent sampling). Violations of this assumption can have serious effects. Violations will occur, for instance, in correlated samples and repeated measures designs, such as before-after or matched-pairs studies, cluster sampling, or time-series data. That is, subjects cannot provide multiple observations at different time points.

CHAPTER FIVE

5. STATISTICAL DATA ANALYSIS AND DISCUSSION

5.1 Introduction

The response variable, death status, is binary assuming only two values 0 and 1. That is, if the response is “yes”, the individual patient has died under ART follow up, and “no” indicates survival of the patient taking ART. In this study logistic regression is used to see the relationship between the proposed independent variables and the response variable. In logistic regression there is no assumption of normality instead we have the assumption of binomial variability. The overall proportion of death in the Adama Hospital ART Clinic, (Y=1) is around 0.07(small). This may affect the fitted model to produce unstable estimates. The estimated coefficients may be large with large standard error. The corresponding odds ratio will be too small or too large. And it is with this understanding that the univariate and multivariate analyses are made in this study. We start our data analysis by giving the summary statistics for the variables considered in the study; we then proceed to the univariate analysis and complete the final model in the multivariate analysis.

5.2 Summary Statistics

I-Socio-Economic factors and HIV death

In this study a total of 259 patient cards have been reviewed of which 13.1% (totally 34) are death cases. Of the patients who didn’t attend Previous HIV counseling Session(s) 27.6% died while the death percentage is lower (11.3%) for those who attended. The urban residents show higher proportion of death (13.9%) in the sample as compared to rural residents (6.9%). The Unemployed and Not working due to ill health groups have higher proportion than the Working group. It is apparent from the sample that people with more than two rooms have lower death proportion (5.8% against 19.4%). The sample data also reveal small differences among individuals who have electricity access at their home (11.7%) and those who don’t have the access to electricity (13.7%). The proportion of death is found to be almost the same for the groups who have tap water access at their home (13.1%) and those who don’t have such access (13.2%).

Table 5.1 **Socio-Economic factors by HIV death**

Socio-Economic factors		**No. of deaths**	**Percentage of deaths**
Previous attendance of HIV counseling Session(s)	(0) No	8	27.6
	(1)Yes	26	11.3
Residence	(0)Urban	32	13.9
	(1) Rural	2	6.9
Employment Status	(0) Working full/ part time	4	4.8
	(1) Not working due to ill health	10	14.9
	(2) Unemployed	20	18.5
Number of rooms	(0) <2	27	19.4
	(1)>=2	7	5.8
Running water	(0) No	17	13.2
	(1)Yes	17	13.1
Running electricity	(0) No	9	11.7
	(1)Yes	25	13.7

II-Demographic and Health factors in relation to HIV/AIDS death

Death proportions appear to be very similar for all age groups among patients taking ART at Adama Hospital as shown by their percentage 12.5%, 13.4% and 12.5% for the three groupings used in this study. Female gender seems to have lower proportion (10.5%) than male (16.8%). The married group showed the smallest percentage (10.9%) with respect to death proportions than the other two groups and more educated groups revealed the highest proportion of death (20%) with in the educational level. HIV-death also shows increasing effect on TB positive cases, on individuals with HIV-positive partner, and those with lower baseline weight and CD4 counts. The WHO clinical stage IV cases and patients who took the second group of antiretroviral regimen seem have

larger proportions with respect to HIV-death. All the results have been summarized in Table 5.2 below.

Table 5.2 **Demographic and Health factors by HIV/AIDS death**

Demographic and Health factors		**No. of deaths**	**Percentage of Deaths**
Age in complete years	(0) < 26	5	12.5
	(1) Between 26 and 45	24	13.4
	(2) > 45	5	12.5
Sex	(0) Female	16	10.5
	(1) Male	18	16.8
Marital Status	(0) Never married	9	21.4
	(1) Married	12	10.9
	(2)Divorced/Separated/ Widowed	13	12.1
Educational level	(0) No education	8	11.6
	(1) Primary	7	7.4
	(2) Secondary or above	19	20.0
TB	(0)Positive	24	14.1
	(1) Negative/UK	10	11.2
Partners HIV status	(0) Negative	9	3.6
	(1)Positive	6	38.7
	(2) Unknown/NA	19	11.6
Base Line Weight	(0)<45	15	23.8
	(1)>=45	19	8.0
Base Line CD4 Counts	(0)<150	26	18.8
	(1)>=150	8	6.6
WHO Clinical Stage	(0) Stage I	1	6.7
	(1) Stage II	2	4.4
	(2) Stage III	19	12.7
	(3) Stage IV	12	24.5

Antiretroviral regimen	(0) d4T (30 or 40)/3TC/NVP-(1a 30 or 40)	21	11.1
	(1) d4T(30 or 40)/3TC/EFV-(1b 30 or 40)	12	25.5
	(2) AZT/3TC/NVP- (1C) OR AZT/3TC/EFV-(1d)	1	4.5

III-Risk Behavior factors and HIV/AIDS death

The results of summary statistics for the Risk Behavior factors reveal interesting results that coincide with the usual instructions of physicians to their patients. The data indicate that there is a very high proportion in HIV-deaths among individuals who experience risk behaviors like not using condom during intercourse, smoking Tobacco and drinking Alcohol and drug abuse. Table 5.3 below reveals this in detail.

Table 5.3 **Risk Behavior factors by HIV/AIDS death**

Risk Behavior factors		**No. of deaths**	**Percentage of Deaths**
Condom use	(0) Mostly	7	4.2
	(1) Never/ Rarely/ No response	27	29.3
Tobacco	(0) No	19	9.5
	(1) Yes	10	37.0
	(2) No response	5	16.1
Alcohol	(0) No	11	6.8
	(1) Yes	18	26.1
	(2) No response	5	17.9
Soft/Hard Drugs	(0) Yes	18	38.3
	(1) No	16	7.5

5.3 Univarite Analysis

The selection process will begin with a careful analysis of each variable. It is known that the likelihood ratio chi-square test with k-1 degrees of freedom is exactly equal to the value of the likelihood ratio test for the significance of the coefficients for the k-1 design variables in a univariate logistic regression model that contains that single independent variable. Since the Pearson chi-square test is asymptotically equivalent to the likelihood ratio chi-square test, it can also be used to test the significance of univarite relationships.

In univariate analysis, using Pearson chi-square test, the variables that are found to be significant are Sex(SEX),Marital Status(MARIT_STAT), Educational level(EDUC_LEV), Previous attendance of HIV counseling Session(s)(COUNS), Employment Status(EMP_STAT), Number of rooms(NO_ROOMS), Partners HIV status(PART_HIV),Base Line Weight(BASE_WEIG), Base Line CD4 Counts(BASE_CD4),WHO Clinical Stage(STAGE), Antiretroviral regimen(REGIMEN), Condom use(CONDOM), Tobacco(TOBAC), Alcohol(ALCOH), Soft/Hard Drugs(DRUG). Here the p-value used as a criterion for significance is 0.25 [see 41].

Relatively stronger associations (at α=0.05 level of significance) of HIV-death under ART follow up are depicted for Educational level, Previous attendance of HIV counseling Session(s), Employment Status, Number of rooms, Base Line CD4 Counts, WHO Clinical Stage, Antiretroviral regimen, Condom use, Tobacco, Alcohol, Soft/Hard Drugs. Table 5.4 summarizes the findings of the univarite analysis.

Table 5.4 **Variables in the univarite analysis**

Variable	**Pearson chi-square**	**DF**	**P-value (Assymptotic)**
Age in complete years(AGE)	.040	2	.980
Sex(SEX)	2.183(b)	1	.140
Marital Status(MARIT_STAT)	3.102	2	.212
Educational level(EDUC_LEV)	6.840	2	.033
Previous attendance of HIV	5.986	1	.014

counseling Session(s)(COUNS)			
Residence(RESID)	1.112	1	.292
Employment Status(EMP_STAT)	8.097	2	.017
Number of rooms(NO_ROOMS)	10.431	1	.001
Running water(RUN_WAT)	.001	1	.981
Running electricity(RUN_ELECT)	.199	1	.656
TB(TB_STAT)	.425	1	.514
Partners HIV status(PART_HIV)	3.721	2	.156
Base Line Weight(BASE_WEIG)	3.423	1	.064
Base Line CD4 Counts(BASE_CD4)	8.454	1	.004
WHO Clinical Stage(STAGE)	9.099	3	.028
Antiretroviral regimen(REGIMEN)	8.597	3	.035
Condom use(CONDOM)	32.918	1	.000
Tobacco(TOBAC)	16.160	2	.000
Alcohol(ALCOH)	16.416	2	.000
Soft/Hard Drugs(DRUG)	31.899	1	.000

5.4 Multivariate Analysis

One problem with any univariate approach is that it ignores the possibility that a collection of variables, each of which is weakly associated with the outcome, can become an important predictor of the outcome when taken together. If this is thought to be a possibility, then we should choose a significance level large enough to allow the suspected variables to become candidates for inclusion in the multivariate model. It is for this reason that we used p-value of 0.25 for selection of variables that are candidates for the multivariate analysis from univariate findings.

Multivariate analysis is done using the significant variables in the univariate analysis. Forward stepwise likelihood ratio method is used to select variables. The stepwise selection proceeded by entering the Variable CONDOM on step 1; ALCOH on step 2; BASE_CD4 on step 3; BASE_WEIG on step 4; DRUG on step 5; NO_ROOMS on step 6; PART_HIV on step 7. To facilitate computation and interpretation, the coding scheme used in SPSS is given below in Table 5.5.

Table 5.5 **Categorical Variables Codings**

			Parameter coding		
		Frequency	(1)	(2)	(3)
WHO Clinical Stage	Stage I	15	1.000	.000	.000
	Stage II	45	.000	1.000	.000
	Stage III	150	.000	.000	1.000
	Stage IV	49	.000	.000	.000
ALCOH	No	162	1.000	.000	
	Yes	69	.000	1.000	
	No Response	28	.000	.000	
TOBAC	No	201	1.000	.000	
	Yes	27	.000	1.000	
	N0 Response	31	.000	.000	
Marital Status	Never Married	42	1.000	.000	
	Married	110	.000	1.000	
	Divorced/Separated/Widowed	107	.000	.000	
Antiretroviral regimen	d4T (30 or 40)/3TC/NVP-(1a 30 or 40)	190	1.000	.000	
	d4T(30 or 40)/3TC/EFV-(1b 30 or 40)	47	.000	1.000	
	AZT/3TC/NVP- (1C) OR AZT/3TC/EFV-(1d)	22	.000	.000	

Educational level	No Education	69	1.000	.000	
	Primary	95	.000	1.000	
	Secondary or above	95	.000	.000	
Partners HIV status	Negative	55	1.000	.000	
	Positive	31	.000	1.000	
	Unknown/NA	173	.000	.000	
Employment Status	Working Full/PartTime	84	1.000	.000	
	Not Working due to ill Health	67	.000	1.000	
	Unemployed	108	.000	.000	
Soft/Hard Drugs	Yes	47	1.000		
	No/UK	212	.000		
Previous attendance HIV related counseling session(s)	No	29	1.000		
	Yes	230	.000		
Number of rooms	<2	139	1.000		
	>=2	120	.000		
Condom use	Mostly	167	1.000		
	Never/Rarely/No response	92	.000		
Base Line CD4 Counts(/ml)	<150	138	1.000		
	>=150	121	.000		
Base Line Weight(in kg)	<45	84	1.000		
	>=45	175	.000		
SEX	Female	152	1.000		
	Male	107	.000		

The variables that are found to be significant in the multivariate analysis are CONDOM, ALCOH, BASE_CD4, BASE_WEIG, DRUG, NO_ROOMS, and PART_HIV. And this is in effect in line with the results obtained from the univariate analysis. The values of the Wald statistic for individual β coefficients support that the estimated values ($\hat{\beta}_i's$) are significantly different from zero at an α=0.05 level of significance for all the above seven covariates. The remaining variables which were used in the univariate analysis, including the interaction terms and the constant term are found to be non significant.

The estimated coefficients ($\hat{\beta}_i's$) for the covariates in the final model, their standard error and the odds ratio corresponding each estimated coefficient ($\hat{\beta}_i$) is given in the following table.

Table 5.6 **Variables in the Final Model**

Covariates	$\hat{\beta}$	**S.E.**	**Wald**	**DF**	**Sig.**	Exp($\hat{\beta}$)
NO_ROOMS(1)	1.458	.649	5.053	1	.025	4.297
PART_HIV			14.689	2	.001	
PART_HIV(1)	-.581	.967	.360	1	.548	.560
PART_HIV(2)	2.475	.690	12.863	1	.000	11.881
BASE_WEIG(1)	1.268	.555	5.219	1	.022	3.555
BASE_CD4(1)	1.752	.636	7.586	1	.006	5.769
CONDOM(1)	-3.603	.787	20.932	1	.000	.027
ALCOH			14.668	2	.001	
ALCOH(1)	-3.447	.998	11.938	1	.001	.032
ALCOH(2)	-1.531	.947	2.615	1	.106	.216
DRUG(1)	1.404	.619	5.142	1	.023	4.073

5.5 Model Checking and Diagnostics

In our logistic regression analysis of the data by SPSS package, results of several goodness-of-fit tests accompany the SPSS output if we check the appropriate options during the model fitting process. The goodness-of-fit tests discussed in the previous chapter are all available in the package output. We will use some of them here below in order to check the adequacy of our model.

The Likelihood ratio test

Recall the **Likelihood ratio** statistic $-2LL_R - (-2LL_F) = -2\ln\left(\frac{likelihood_R}{likelihood_F}\right)$ given in the previous chapter. The $-2LL_R$ for the restricted (smaller) model is 201.399 and that of $-2LL_F$ for full (larger) model is 94.157 which implies a model chi square of 107.242 with p-value 0.00 which signifies significant decrease in deviance thereby implying a good fit of the model.

Hosmer - Lemeshow Test

The values of Hosmer –Lemeshow statistic and its contingency table for the final step is computed in SPSS as

Chi-square	Df	Sig.
4.361	8	.823

Table 5.7 **Contingency Table for Hosmer - Lemeshow Test**

	Death Status = Alive		Death Status = Dead		Total
	Observed	Expected	Observed	Expected	
1	27	26.991	0	.009	27
2	25	24.977	0	.023	25
3	23	22.948	0	.052	23
4	28	27.879	0	.121	28
5	25	25.717	1	.283	26
6	27	26.420	0	.580	27
7	24	23.815	1	1.185	25
8	22	22.933	4	3.067	26
9	20	17.814	6	8.186	26
10	4	5.505	22	20.495	26

And the final stage result of the Hosmer –Lemeshow statistic leads to non rejection of the null hypothesis saying the model fits the data which is in agreement with the previous likelihood ratio test.

The classification table, Cox & Snell R Square and Nagelkerke R Square Provided the same conclusion as the above two tests. Their values are given below.

Model Summary

-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
108.109	.302	.560

Classification Table

Observed		Predicted		
		Death Status		Percentage Correct
		Alive	Dead	
Death Status	Alive	220	5	97.8
	Dead	16	18	52.9
Overall Percentage				91.9

The cut value is .500

Checking For Influential cases

The three measures influence, the leverage statistic, Cook's distance and dfbeta indicated that there are no suspicious influential observations (see Appendix). Thus, all in all we can say that our model fits the data very well.

5.6 Discussion and Interpretation of results

Although most of the results obtained from the summary statistics matches with the findings of both univariate and multivariate analysis, it has some deviations with respect to certain variables. For example, it showed that patients of the age group between 26 and 45, who are urban dwellers and diagnosed with TB positive are at higher risk of HIV-death. But the three variables are not significant even for 0.25 significance levels in the univariate analysis (and hence are not candidates in the multivariate case). This means that HIV-death is equally distributed in all age groups both in rural and urban areas. One unexpected result from this study is that TB doesn't affect HIV-death in both the univariate and multivariate analysis. This can be attributed to the fact that large proportion HIV patients are TB positive irrespective of their survival status and now at clinical level, the disease is controlled. The damage TB caused to HIV patients is given due attention by government and non government organizations and appropriate interventions (like supply of medicine) have been made.

As we proceed from the univariate analysis to multivariate analysis, we will find out that several variables are non significant; these were significant in the univariate case (see Section 5.3). Only the variables CONDOM, ALCOH, BASE_CD4, BASE_WEIG, DRUG, NO_ROOMS, and PART_HIV remain in the multivariate model as the factors affecting the survival/death status of HIV patients at Adama Hospital. And there is enough theory to support the claim that these variables explain the survival/death status of HIV patients very well.

If we investigate these variables closely, they can be classified into three groups;
CONDOM, ALCOH and DRUG may be termed as risk behavior group indicating the behavior of HIV patients towards factors associated with non-adherence, BASE_CD4, BASE_WEIG, and PART_HIV will fall under the group of health indicators while NO_ROOMS may be used as indicator of Economic status.

Thus, we can now interpret the effect of each covariate using the estimated odds ratio given above. The negative sign for the odds ratio of the variable condom use implies that death risk is lower for those patients who use condom during sexual intercourse. Its magnitude indicates that, the odds of being at risk of death for those patients who use condom during sexual intercourse is 2.7% less than for those who do not use it or for those in the non response group. This is due to the scientific reason that, the condom prevents the transfer of stronger HIV virus to the patient. Alcohol and drug abusers are usually accused of ARV non-adherence by clinicians in addition to the complications they bring in to one's health. The results of this study support this claim: Non alcoholics are 3.2% less likely for the risk of death than those who didn't explain their behavior towards alcohol. On the other hand patients who responded they take more alcohol are still at a lower risk of death than those who didn't explain their alcoholic behavior. This lesser risk might be attributed to the fact that those who explain their status will get counseling services on how to avoid their bad addictions from the clinics.
The odds of death risk for Drug (Soft/Hard) abusers is much higher (odds ratio=4.073) than those who don't use drugs. The increased risk here arises again due to the problem of non-adherence of patients with this behavior to ART.

Partners' HIV status affects survival since the wellbeing of one's partner has economic, social and psychological advantages that may influence health positively. The odds of death risk for patients with HIV negative Partner is less by 56.0% than those who don't know their partner's HIV status. Compared to those who don't know their partner's HIV status, patients with HIV-positive partner are less likely to survive (odds ratio=11.881). The other two health factors, base line CD4 and base line weight determine one's resistance to different opportunistic diseases. Thus, the larger their number/value, the lower the danger of being at risk of HIV death. The outcomes of this study support this fact as the odds of death risk are high for those with lower CD4 count and weight.

The last variable, number of rooms, is used here as an indicator of economic status. Patients with more number of rooms have better economic status than those who live perhaps in a room with their families. Patients living in less than 2 rooms are more likely not to survive than those living in more than 2 rooms with their family (odds ratio=4.297). Patients of higher economic class have better chance of survival under ART follow up. From clinical point of view, better economic class implies better nutrition which increases the resistance of patients to opportunistic diseases thus lowering the patient's risk of death.

CHAPTER SIX

6. CONCLUSION AND RECOMENDATION

6.1 Conclusion

Since the aim ART is to improve the health of HIV-positive individuals, it would be essential to study the factors that can improve the performance of ART.

The multivariate logistic regression analysis based on this empirical study from the ART Clinic of Adama Hospital showed that the factors that affect the survival/death status of HIV patients who take ART include Condom use (CONDOM), Alcohol (ALCOH), Soft/Hard Drugs (DRUG), Number of rooms (NO_ROOMS), Base Line Weight (BASE_WEIG), Base Line CD4 Counts (BASE_CD4) and Partners HIV status (PART_HIV).

The results of this study also indicated that survival/death status doesn't show differences for different age groups and educational levels. It also doesn't depend on gender and employment status. Also, TB is not a risk factor for patients taking ART at Adama Hospital perhaps due to the consideration given to control the disease through the supply of drugs and treatment facilities. At clinical level, 65.6% of the patients who were taking ART are TB-positive and about 86% of them have survived.

The factors that influence survival/death status can be grouped as risk behavior factors (CONDOM, ALCOH, and DRUG), health factors (BASE_WEIG, BASE_CD4 and PART_HIV) and an economic factor (NO_ROOMS). Under such grouping we can state that patients involved in risky behaviors will have higher risk of death perhaps due to their tendency of non-adherence. Similarly, patients with poor health indicators like small baseline CD4 and weight who have HIV-positive partner are less likely to survive. The variable used as economic status indicator, number of rooms, indicates patients living in a single room with their family have also less chance of survival. Poor nutrition as a consequence of low economic status makes patients vulnerable to opportunistic diseases thereby decreasing the chance of survival.

6.2 Recommendations

The high prevalence of HIV/AIDS in Ethiopia calls for an extensive demand for ART services. This demand is given due attention from government and non-government organizations. As a consequence, ART clinics are now widespread in all corners of the country. And it would be useful to initiate programs that emphasize coping strategies for improving the performance over and above testing and educational campaigns. Patients are being treated in the ART clinics to extend their lives. However, a number of patients are still dying under ART follow up. This paper tries to identify the factors affecting survival status of the patients. And the following recommendations are made for health policy workers and clinicians:

- The results of the study underlined behavioral (risk) factors as important predictors of survival/death status. Thus, clinicians are expected to work hard to bring about behavioral change. Currently the emphasis is on bringing behavioral change for prevention purposes. But we should consider it with respect to HIV-positive persons taking ART. ART programs will not be successful unless we can change the behavior of patients under ART follow up.

- Health workers should be cautious when a patient has lower CD4 and weight at baseline and an HIV-positive partner. When this is the case appropriate clinical and non-clinical measures like medicine and support (can be home-based) should be provided.

- For those patients with low economic status, health workers/stakeholders need to find ways of supporting the patients with respect to improving their nutrition in particular and other assistance in general.

REFERENCES

[1]. UNAIDS (2006). Report on the Global HIV/AIDS Epidemic. Geneva.

[2]. Ministry of Health (2006).AIDS in Ethiopia, Sixth Report, Addis- Ababa, Ethiopia.

[3]. Alemtsehai Abate and Eshetu Wencheko (2006). Association of HIV Infection with some Selected Factors and Modeling the Chance of Contracting HIV: The Case of Hawassa and Its Surrounding. Journal of Ethiopian Statistical Association; 15, 27-48

[4]. AVERT, AVERTing HIV and AIDS(2006). Introduction to antiretroviral therapy. www.avert.org.

[5]. Office of AIDS Research Advisory Council (OARAC) (2006). Guidelines for the Use of Antiretroviral Agents in HIV-1-Infected Adults and Adolescents

[6]. UNAIDS (2002). Report on the Global HIV/AIDS Epidemic. Geneva.

[7]. Maruschak L. (2000). HIV in Prisons. Bureau of Justice Statistics Bulletin, NCJ 196023. Washington, DC.

[8]. Degu J., Aschalew E.,Yewubnesh H., and Bernt L.(2006). Predictors of early death in a cohort of Ethiopian patients treated with HAART (Abstract). BMC Infect Dis. 6: 136.

[9]. Cohen, D. (2002). Poverty and HIV/AIDS in sub-Saharan Africa. UNDP, Issue paper number 27.

[10]. U.S. Centers for Disease Control and Prevention, HIV/AIDS Newsroom (2003).

[11]. National AIDS Control Council of Kenya (2003). Kenya National AIDS Control Report.

[12]. Walraven, G., A. Nicoll, N. Njau and I. Timeaus (1996). The Impact of HIV-1 on child health in sub-Saharan Africa: The burden on the health services. Tropical Medicine and international Health; 1, 1:3-14.

[13]. Ainsworth, M. and Filmer D. (2002). Poverty, AIDS and Children's Schooling: A Targeting Dilemma. World Bank Policy Research Working Paper 2885.

[14]. World Bank (1996). Why all the Talk about Women and AIDS?

[15]. UNIFEM (2000).Gender, HIV and Human Rights. UNIFEM, New York

[16]. Booysen, F. (2001). HIV/AIDS, Poverty, and Risky Sexual Behavior: Evidence from South Africa.

[17]. Williams C.C, and Windebank J. (2000). Helping people to help themselves: Policy lessons from a study of deprived urban neighborhoods in Southampton. Journal of Social Policy 29: 355-373.

[18]. Buve, A., Carael, M. and Hayes, RJ (2001). Multi-center Study on Factors Determining Differences in Rate of Spread of HIV in sub-Saharan Africa: Methods and Prevalence of HIV Infection. AIDS 15, (4): 5-14.

[19]. Verheij, RA. (1996). Explaining urban-rural variations in health: A review of interactions between individual and environment. Social Science and Medicine 42: 923-935.

[20]. Johnson S. A. (2004). HIV/AIDS Mortality Differential Across Provinces in Kenya and Through Time. Texas Tech University.

[21]. Berry, D.E., McKinney, M.M., and McClain, M. (1996). Rural HIV service networks: Patterns of care and policy issues. AIDS & Public Policy Journal, 11, 36-46.

[22]. Coleman, JS. (1990). Foundations of Social Theory. Harvard University Press: Cambridge.

[23]. Williams H, Martyn S and Garry H. (1991). Urban-rural mortality differentials: A case study of Wales.

[24]. Congressional Research Service (CRS) Reports and Issue Briefs (2001). AIDS in Africa, HIV/AIDS IN AFRICA UPDATE.

[25]. Nsutebu, E.F, Walley,J.D. Mataka, E. and Simon, C.F. (2001). Scaling-up HIV/AIDS and TB home-based care: Lessons from Zambia. Health Policy and Planning, 16(3): 240-247.

[26]. Bajaria, S.H. (2002). Dynamics of Naïve and Memory CD4+ T Lymphocytes in HIV-1 Disease Progression. Journal of Acquired Immune Deficiency Syndrome, 30:41-58.

[27]. Dorrington, R. (2001). The Impact of HIV/AIDS on Adult Mortality in South Africa. Report, South African Medical Research Council.

[28]. Timeaus, I.M. (1998). Impact of the HIV epidemic on Mortality in sub-Saharan Africa: Evidence from National surveys and censuses. AIDS, 12(1):15-27.

[29]. UNAIDS and World Health Organization. 2001. AIDS Epidemic Update. Geneva.

[30]. Kahn K, Tollman SM, Garenne M, Gear JS. (1999). Who dies from what: determining cause of death in South Africa's rural northeast. Tropical Medicine and International Health.

[31]. Jevtovic, D.O., Salemovic, D., Ranin, J., Pesic, I., Zerjav S. and Djurkovic-Djakovic O. (2006). Long-term survival of HIV-infected patients treated with highly active antiretroviral therapy in Serbia and Montenegro. Medscape Today.

[32]. Colton T. (1974). Statistics in Medicine. Lippincot Williams & Wilkins, Philadelphia.

[33]. Cochran W.G. (1977).Sampling Techniques. New York. John Wiley & Sons, Inc.

[34]. Pagano, M. and Gauvreau, K. (2000), Principles of Biostatistics. Thomson Learning, Boston.

[35]. StatNotes (2006). Logistic Regression. www2.chass.ncsu.edu/garson/PA765/logistic.

[36]. Hosmer D. and Lemeshow S. (1989). Applied Logistic Regression. John Wiley and Sons. Inc., New York.

[37]. Agresti, A. (1996). An Introduction to Categorical data Analysis. John Wiley and Sons, Inc., New York.

[38]. Walta Information Center (2006).Proceeding of the national symposium on" The success and challenges of ART in Ethiopia". Addis Ababa.

[39]. SPSS (2006). Statistical Support Algorithms. http://support.spss.com/Student/Documentation/algorithms/14.0/logistic_regression.pdf

APPENDIX

Logistic Regression Output

Case Processing Summary

Unweighted Cases(a)		N	Percent
Selected Cases	Included in Analysis	259	100.0
	Missing Cases	0	.0
	Total	259	100.0
Unselected Cases		0	.0
Total		259	100.0

a If weight is in effect, see classification table for the total number of cases.

Dependent Variable Encoding

Original Value	Internal Value
Alive	0
Dead	1

Iteration History(a,b,c)

Iteration		-2 Log likelihood	Coefficients
			Constant
Step 0	1	207.022	-1.475
	2	201.481	-1.837
	3	201.399	-1.889
	4	201.399	-1.890

a Constant is included in the model.
b Initial -2 Log Likelihood: 201.399
c Estimation terminated at iteration number 4 because parameter estimates changed by less than .001.

Classification Table(a,b)

			Observed		Predicted
			Death Status		Percentage Correct
			Alive	Dead	
Step 0	Death Status	Alive	225	0	100.0
		Dead	34	0	.0
	Overall Percentage				86.9

a Constant is included in the model.
b The cut value is .500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 0	Constant	-1.890	.184	105.479	1	.000	.151

Variables not in the Equation

			Score	df	Sig.
Step 0	Variables	SEX(1)	2.183	1	.140
		MARIT_STAT	3.102	2	.212
		MARIT_STAT(1)	3.029	1	.082
		MARIT_STAT(2)	.825	1	.364
		EDUC_LEV	6.840	2	.033
		EDUC_LEV(1)	.194	1	.660
		EDUC_LEV(2)	4.363	1	.037
		COUNS(1)	5.986	1	.014
		EMP_STAT	8.097	2	.017
		EMP_STAT(1)	7.629	1	.006
		EMP_STAT(2)	.256	1	.613
		NO_ROOMS(1)	10.431	1	.001
		PART_HIV	22.507	2	.000
		PART_HIV(1)	5.516	1	.019
		PART_HIV(2)	20.209	1	.000
		BASE_WEIG(1)	12.439	1	.000
		BASE_CD4(1)	8.454	1	.004
		STAGE	9.099	3	.028
		STAGE(1)	.583	1	.445
		STAGE(2)	3.601	1	.058
		STAGE(3)	.066	1	.797
		REGIMEN	8.480	2	.014
		REGIMEN(1)	2.692	1	.101

	REGIMEN(2)	7.747	1	.005
	CONDOM(1)	32.918	1	.000
	TOBAC	16.160	2	.000
	TOBAC(1)	10.628	1	.001
	TOBAC(2)	15.110	1	.000
	ALCOH	16.416	2	.000
	ALCOH(1)	15.233	1	.000
	ALCOH(2)	13.852	1	.000
	DRUG(1)	31.899	1	.000
	EDUC_LEV * EMP_STAT	6.441	4	.169
	EDUC_LEV(1) by EMP_STAT(1)	1.112	1	.292
	EDUC_LEV(1) by EMP_STAT(2)	.749	1	.387
	EDUC_LEV(2) by EMP_STAT(1)	2.347	1	.126
	EDUC_LEV(2) by EMP_STAT(2)	.435	1	.510
	BASE_WEIG * STAGE	4.209	3	.240
	BASE_WEIG(1) by STAGE(1)	.502	1	.479
	BASE_WEIG(1) by STAGE(2)	.355	1	.552
	BASE_WEIG(1) by STAGE(3)	3.399	1	.065
Overall Statistics		119.885	31	.000

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	31.924	1	.000
	Block	31.924	1	.000
	Model	31.924	1	.000
Step 2	Step	22.254	2	.000
	Block	54.178	3	.000
	Model	54.178	3	.000
Step 3	Step	21.642	2	.000
	Block	75.820	5	.000
	Model	75.820	5	.000
Step 4	Step	13.565	1	.000
	Block	89.385	6	.000
	Model	89.385	6	.000
Step 5	Step	8.467	1	.004
	Block	97.851	7	.000

	Model	97.851	7	.000
Step 6	Step	4.201	1	.040
	Block	102.053	8	.000
	Model	102.053	8	.000
Step 7	Step	5.190	1	.023
	Block	107.243	9	.000
	Model	107.243	9	.000

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	169.475(a)	.116	.215
2	147.221(a)	.189	.349
3	125.579(b)	.254	.470
4	112.014(b)	.292	.540
5	103.548(b)	.315	.582
6	99.346(b)	.326	.603
7	94.157(b)	.339	.627

a Estimation terminated at iteration number 6 because parameter estimates changed by less than .001.
b Estimation terminated at iteration number 7 because parameter estimates changed by less than .001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	.000	0	.
2	.249	3	.969
3	4.006	6	.676
4	4.523	8	.807
5	4.744	8	.785
6	4.423	8	.817
7	4.361	8	.823

Contingency Table for Hosmer and Lemeshow Test

		Death Status = Alive		Death Status = Dead		
		Observed	Expected	Observed	Expected	Total
Step 1	1	160	160.000	7	7.000	167
	2	65	65.000	27	27.000	92
Step 2	1	100	99.929	1	1.071	101
	2	37	37.591	3	2.409	40
	3	23	22.479	3	3.521	26
	4	51	51.071	10	9.929	61
	5	14	13.929	17	17.071	31
Step 3	1	24	23.974	0	.026	24
	2	68	68.647	1	.353	69
	3	25	24.440	0	.560	25
	4	34	32.681	0	1.319	34
	5	15	16.145	3	1.855	18
	6	34	33.888	5	5.112	39
	7	15	15.490	10	9.510	25
	8	10	9.735	15	15.265	25
Step 4	1	14	13.996	0	.004	14
	2	27	26.973	0	.027	27
	3	17	16.952	0	.048	17
	4	41	41.721	1	.279	42
	5	25	24.714	0	.286	25
	6	29	28.856	1	1.144	30
	7	26	24.331	0	1.669	26
	8	27	28.467	8	6.533	35
	9	16	15.623	11	11.377	27
	10	3	3.366	13	12.634	16
Step 5	1	26	25.991	0	.009	26
	2	25	24.958	0	.042	25
	3	25	24.908	0	.092	25
	4	24	23.857	0	.143	24
	5	27	27.583	1	.417	28
	6	21	20.542	0	.458	21
	7	26	25.651	1	1.349	27
	8	22	24.216	5	2.784	27
	9	20	17.678	5	7.322	25
	10	9	9.616	22	21.384	31
Step 6	1	29	28.988	0	.012	29
	2	26	25.967	0	.033	26
	3	24	23.932	0	.068	24
	4	29	28.823	0	.177	29
	5	26	26.600	1	.400	27
	6	26	25.187	0	.813	26
	7	24	23.397	1	1.603	25
	8	20	22.326	6	3.674	26
	9	16	15.928	10	10.072	26

	10	5	3.852	16	17.148	21
Step 7	1	27	26.991	0	.009	27
	2	25	24.977	0	.023	25
	3	23	22.948	0	.052	23
	4	28	27.879	0	.121	28
	5	25	25.717	1	.283	26
	6	27	26.420	0	.580	27
	7	24	23.815	1	1.185	25
	8	22	22.933	4	3.067	26
	9	20	17.814	6	8.186	26
	10	4	5.505	22	20.495	26

Classification Table(a)

Observed			Predicted		
			Death Status		Percentage Correct
			Alive	Dead	
Step 1	Death Status	Alive	225	0	100.0
		Dead	34	0	.0
	Overall Percentage				86.9
Step 2	Death Status	Alive	211	14	93.8
		Dead	17	17	50.0
	Overall Percentage				88.0
Step 3	Death Status	Alive	211	14	93.8
		Dead	14	20	58.8
	Overall Percentage				89.2
Step 4	Death Status	Alive	218	7	96.9
		Dead	17	17	50.0
	Overall Percentage				90.7
Step 5	Death Status	Alive	217	8	96.4
		Dead	14	20	58.8
	Overall Percentage				91.5
Step 6	Death Status	Alive	219	6	97.3
		Dead	15	19	55.9
	Overall Percentage				91.9
Step 7	Death Status	Alive	219	6	97.3
		Dead	11	23	67.6
	Overall Percentage				93.4

a The cut value is .500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1(a)	CONDOM(1)	-2.251	.449	25.136	1	.000	.105
	Constant	-.879	.229	14.724	1	.000	.415
Step 2(b)	CONDOM(1)	-2.898	.577	25.238	1	.000	.055
	ALCOH			19.351	2	.000	
	ALCOH(1)	-2.682	.773	12.029	1	.001	.068
	ALCOH(2)	-.894	.733	1.486	1	.223	.409
	Constant	1.044	.744	1.969	1	.161	2.842
Step 3(c)	PART_HIV			18.472	2	.000	
	PART_HIV(1)	-1.564	.848	3.404	1	.065	.209
	PART_HIV(2)	2.029	.559	13.168	1	.000	7.608
	CONDOM(1)	-3.379	.664	25.901	1	.000	.034
	ALCOH			19.675	2	.000	
	ALCOH(1)	-3.107	.851	13.317	1	.000	.045
	ALCOH(2)	-1.037	.767	1.825	1	.177	.355
	Constant	1.215	.805	2.278	1	.131	3.372
Step 4(d)	PART_HIV			17.201	2	.000	
	PART_HIV(1)	-1.297	.895	2.101	1	.147	.273
	PART_HIV(2)	2.170	.598	13.161	1	.000	8.759
	BASE_CD4(1)	1.895	.565	11.229	1	.001	6.650
	CONDOM(1)	-3.834	.730	27.587	1	.000	.022
	ALCOH			19.180	2	.000	
	ALCOH(1)	-3.179	.913	12.117	1	.000	.042
	ALCOH(2)	-.916	.825	1.233	1	.267	.400
	Constant	.111	.930	.014	1	.905	1.118
Step 5(e)	PART_HIV			14.425	2	.001	
	PART_HIV(1)	-1.017	.995	1.045	1	.307	.362
	PART_HIV(2)	2.133	.619	11.884	1	.001	8.437
	BASE_WEIG(1)	1.500	.531	7.967	1	.005	4.480
	BASE_CD4(1)	2.040	.608	11.268	1	.001	7.689
	CONDOM(1)	-3.692	.751	24.152	1	.000	.025
	ALCOH			19.601	2	.000	
	ALCOH(1)	-3.145	.969	10.529	1	.001	.043
	ALCOH(2)	-.673	.894	.566	1	.452	.510
	Constant	-.803	1.054	.580	1	.446	.448
Step 6(f)	NO_ROOMS(1)	1.177	.597	3.883	1	.049	3.245
	PART_HIV			14.475	2	.001	
	PART_HIV(1)	-.669	.971	.474	1	.491	.512
	PART_HIV(2)	2.386	.673	12.559	1	.000	10.873
	BASE_WEIG(1)	1.389	.545	6.502	1	.011	4.012
	BASE_CD4(1)	1.863	.621	8.987	1	.003	6.442
	CONDOM(1)	-3.778	.779	23.521	1	.000	.023
	ALCOH			18.657	2	.000	
	ALCOH(1)	-3.304	.995	11.026	1	.001	.037

	ALCOH(2)	-.890	.912	.952	1	.329	.411
	Constant	-1.298	1.118	1.350	1	.245	.273
Step 7(g)	NO_ROOMS(1)	1.458	.649	5.053	1	.025	4.297
	PART_HIV			14.689	2	.001	
	PART_HIV(1)	-.581	.967	.360	1	.548	.560
	PART_HIV(2)	2.475	.690	12.863	1	.000	11.881
	BASE_WEIG(1)	1.268	.555	5.219	1	.022	3.555
	BASE_CD4(1)	1.752	.636	7.586	1	.006	5.769
	CONDOM(1)	-3.603	.787	20.932	1	.000	.027
	ALCOH			14.668	2	.001	
	ALCOH(1)	-3.447	.998	11.938	1	.001	.032
	ALCOH(2)	-1.531	.947	2.615	1	.106	.216
	DRUG(1)	1.404	.619	5.142	1	.023	4.073
	Constant	-1.611	1.139	2.001	1	.157	.200

a Variable(s) entered on step 1: CONDOM.
b Variable(s) entered on step 2: ALCOH.
c Variable(s) entered on step 3: PART_HIV.
d Variable(s) entered on step 4: BASE_CD4.
e Variable(s) entered on step 5: BASE_WEIG.
f Variable(s) entered on step 6: NO_ROOMS.
g Variable(s) entered on step 7: DRUG.

Model if Term Removed

Variable		Model Log Likelihood	Change in -2 Log Likelihood	df	Sig. of the Change
Step 1	CONDOM	-100.700	31.924	1	.000
Step 2	CONDOM	-92.946	38.671	1	.000
	ALCOH	-84.738	22.254	2	.000
Step 3	PART_HIV	-73.611	21.642	2	.000
	CONDOM	-84.018	42.457	1	.000
	ALCOH	-74.979	24.378	2	.000
Step 4	PART_HIV	-65.942	19.870	2	.000
	BASE_CD4	-62.790	13.565	1	.000
	CONDOM	-80.379	48.743	1	.000
	ALCOH	-68.309	24.604	2	.000
Step 5	PART_HIV	-59.711	15.874	2	.000
	BASE_WEIG	-56.007	8.467	1	.004
	BASE_CD4	-58.758	13.968	1	.000
	CONDOM	-72.827	42.106	1	.000
	ALCOH	-64.656	25.763	2	.000
Step 6	NO_ROOMS	-51.774	4.201	1	.040
	PART_HIV	-57.698	16.049	2	.000
	BASE_WEIG	-53.081	6.816	1	.009
	BASE_CD4	-55.005	10.664	1	.001
	CONDOM	-70.411	41.476	1	.000
	ALCOH	-61.823	24.300	2	.000

Step 7	NO_ROOMS	-49.937	5.718	1	.017
	PART_HIV	-55.313	16.469	2	.000
	BASE_WEIG	-49.773	5.390	1	.020
	BASE_CD4	-51.473	8.788	1	.003
	CONDOM	-64.209	34.261	1	.000
	ALCOH	-56.103	18.050	2	.000
	DRUG	-49.673	5.190	1	.023

Variables not in the Equation

			Score	df	Sig.
Step 1	Variables	SEX(1)	1.149	1	.284
		MARIT_STAT	3.042	2	.219
		MARIT_STAT(1)	2.853	1	.091
		MARIT_STAT(2)	1.131	1	.288
		EDUC_LEV	4.299	2	.117
		EDUC_LEV(1)	.275	1	.600
		EDUC_LEV(2)	2.555	1	.110
		COUNS(1)	1.491	1	.222
		EMP_STAT	7.647	2	.022
		EMP_STAT(1)	7.295	1	.007
		EMP_STAT(2)	.269	1	.604
		NO_ROOMS(1)	7.568	1	.006
		PART_HIV	21.312	2	.000
		PART_HIV(1)	6.070	1	.014
		PART_HIV(2)	18.654	1	.000
		BASE_WEIG(1)	10.736	1	.001
		BASE_CD4(1)	12.322	1	.000
		STAGE	6.243	3	.100
		STAGE(1)	1.870	1	.171
		STAGE(2)	2.274	1	.132
		STAGE(3)	.064	1	.800
		REGIMEN	8.552	2	.014
		REGIMEN(1)	2.471	1	.116
		REGIMEN(2)	7.635	1	.006
		TOBAC	12.263	2	.002
		TOBAC(1)	11.752	1	.001
		TOBAC(2)	5.633	1	.018
		ALCOH	23.381	2	.000
		ALCOH(1)	21.143	1	.000
		ALCOH(2)	11.431	1	.001
		DRUG(1)	17.127	1	.000
		EDUC_LEV * EMP_STAT	5.006	4	.287

		EDUC_LEV(1) by EMP_STAT(1)	.257	1	.612
		EDUC_LEV(1) by EMP_STAT(2)	1.061	1	.303
		EDUC_LEV(2) by EMP_STAT(1)	2.465	1	.116
		EDUC_LEV(2) by EMP_STAT(2)	.126	1	.722
		BASE_WEIG * STAGE	3.334	3	.343
		BASE_WEIG(1) by STAGE(1)	.009	1	.923
		BASE_WEIG(1) by STAGE(2)	.092	1	.762
		BASE_WEIG(1) by STAGE(3)	3.259	1	.071
	Overall Statistics		83.301	30	.000
Step 2	Variables	SEX(1)	.141	1	.707
		MARIT_STAT	1.471	2	.479
		MARIT_STAT(1)	.912	1	.340
		MARIT_STAT(2)	1.165	1	.280
		EDUC_LEV	3.452	2	.178
		EDUC_LEV(1)	.015	1	.903
		EDUC_LEV(2)	2.979	1	.084
		COUNS(1)	1.807	1	.179
		EMP_STAT	8.038	2	.018
		EMP_STAT(1)	7.771	1	.005
		EMP_STAT(2)	.379	1	.538
		NO_ROOMS(1)	9.021	1	.003
		PART_HIV	23.342	2	.000
		PART_HIV(1)	6.621	1	.010
		PART_HIV(2)	20.433	1	.000
		BASE_WEIG(1)	12.036	1	.001
		BASE_CD4(1)	14.270	1	.000
		STAGE	4.100	3	.251
		STAGE(1)	2.098	1	.147
		STAGE(2)	1.453	1	.228
		STAGE(3)	.662	1	.416
		REGIMEN	7.803	2	.020
		REGIMEN(1)	2.266	1	.132
		REGIMEN(2)	6.949	1	.008
		TOBAC	.274	2	.872
		TOBAC(1)	.260	1	.610
		TOBAC(2)	.159	1	.690
		DRUG(1)	10.884	1	.001
		EDUC_LEV * EMP_STAT	5.032	4	.284
		EDUC_LEV(1) by EMP_STAT(1)	.489	1	.484

		EDUC_LEV(1) by EMP_STAT(2)	.495	1	.482
		EDUC_LEV(2) by EMP_STAT(1)	2.982	1	.084
		EDUC_LEV(2) by EMP_STAT(2)	.038	1	.846
		BASE_WEIG * STAGE	6.844	3	.077
		BASE_WEIG(1) by STAGE(1)	.031	1	.860
		BASE_WEIG(1) by STAGE(2)	.036	1	.850
		BASE_WEIG(1) by STAGE(3)	6.478	1	.011
	Overall Statistics		65.912	28	.000
Step 3	Variables	SEX(1)	.803	1	.370
		MARIT_STAT	1.921	2	.383
		MARIT_STAT(1)	1.581	1	.209
		MARIT_STAT(2)	1.073	1	.300
		EDUC_LEV	4.338	2	.114
		EDUC_LEV(1)	.081	1	.776
		EDUC_LEV(2)	3.857	1	.050
		COUNS(1)	.916	1	.338
		EMP_STAT	6.519	2	.038
		EMP_STAT(1)	5.623	1	.018
		EMP_STAT(2)	.001	1	.974
		NO_ROOMS(1)	9.165	1	.002
		BASE_WEIG(1)	8.262	1	.004
		BASE_CD4(1)	12.713	1	.000
		STAGE	3.150	3	.369
		STAGE(1)	1.625	1	.202
		STAGE(2)	1.165	1	.281
		STAGE(3)	.935	1	.334
		REGIMEN	4.603	2	.100
		REGIMEN(1)	1.908	1	.167
		REGIMEN(2)	4.293	1	.038
		TOBAC	.279	2	.870
		TOBAC(1)	.005	1	.943
		TOBAC(2)	.100	1	.751
		DRUG(1)	7.471	1	.006
		EDUC_LEV * EMP_STAT	5.700	4	.223
		EDUC_LEV(1) by EMP_STAT(1)	.004	1	.950
		EDUC_LEV(1) by EMP_STAT(2)	.478	1	.489
		EDUC_LEV(2) by EMP_STAT(1)	3.391	1	.066
		EDUC_LEV(2) by EMP_STAT(2)	.714	1	.398
		BASE_WEIG * STAGE	4.016	3	.260

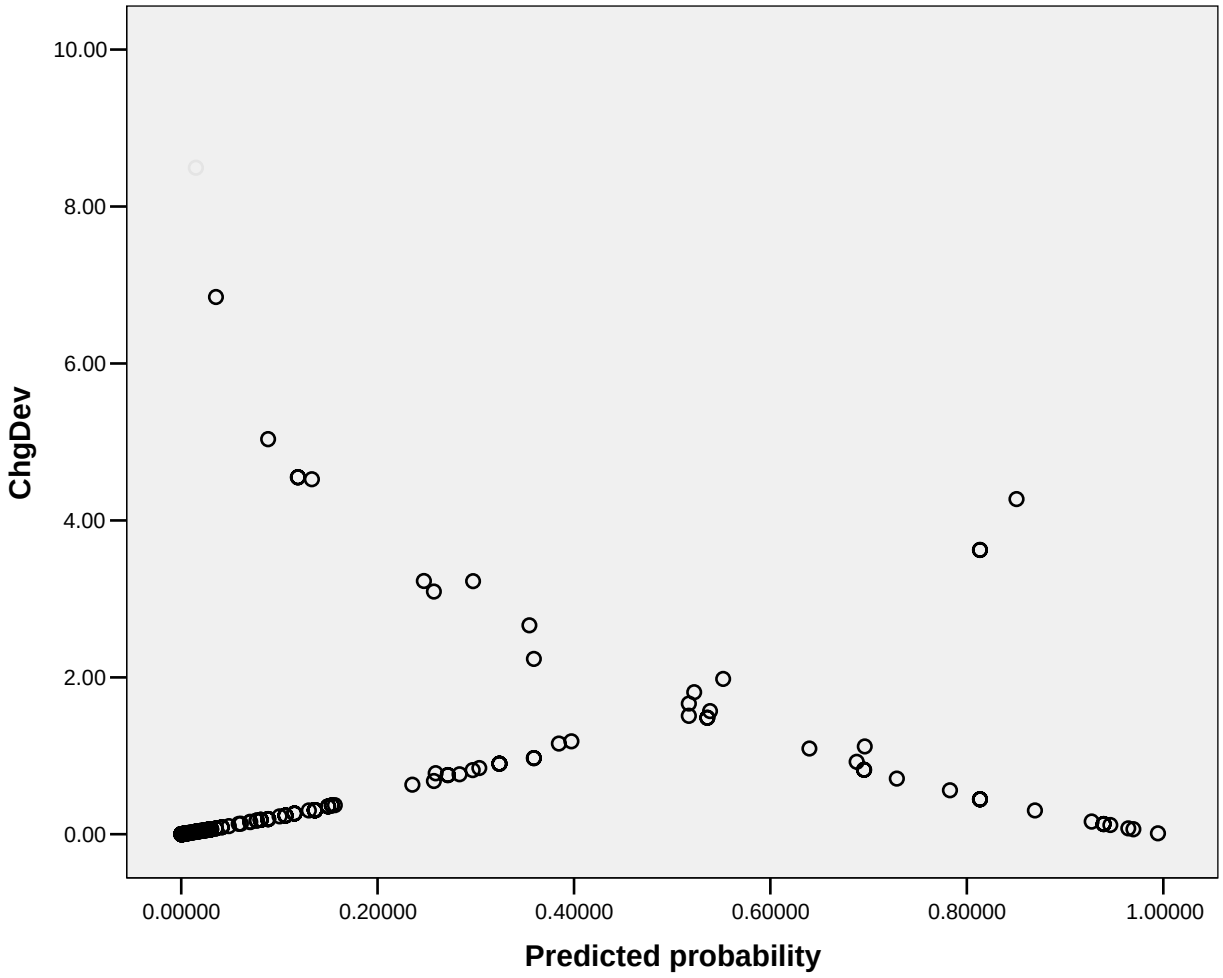
		BASE_WEIG(1) by STAGE(1)	.034	1	.853
		BASE_WEIG(1) by STAGE(2)	.002	1	.965
		BASE_WEIG(1) by STAGE(3)	3.814	1	.051
	Overall Statistics		45.710	26	.010
Step 4	Variables	SEX(1)	.211	1	.646
		MARIT_STAT	2.703	2	.259
		MARIT_STAT(1)	1.109	1	.292
		MARIT_STAT(2)	2.546	1	.111
		EDUC_LEV	1.653	2	.438
		EDUC_LEV(1)	.153	1	.696
		EDUC_LEV(2)	1.617	1	.204
		COUNS(1)	.455	1	.500
		EMP_STAT	4.150	2	.126
		EMP_STAT(1)	4.091	1	.043
		EMP_STAT(2)	.266	1	.606
		NO_ROOMS(1)	5.614	1	.018
		BASE_WEIG(1)	8.619	1	.003
		STAGE	2.941	3	.401
		STAGE(1)	2.057	1	.151
		STAGE(2)	.041	1	.840
		STAGE(3)	1.763	1	.184
		REGIMEN	5.689	2	.058
		REGIMEN(1)	2.174	1	.140
		REGIMEN(2)	5.186	1	.023
		TOBAC	.261	2	.878
		TOBAC(1)	.155	1	.694
		TOBAC(2)	.058	1	.810
		DRUG(1)	5.229	1	.022
		EDUC_LEV * EMP_STAT	2.806	4	.591
		EDUC_LEV(1) by EMP_STAT(1)	.008	1	.928
		EDUC_LEV(1) by EMP_STAT(2)	.101	1	.751
		EDUC_LEV(2) by EMP_STAT(1)	2.482	1	.115
		EDUC_LEV(2) by EMP_STAT(2)	.017	1	.897
		BASE_WEIG * STAGE	5.785	3	.123
		BASE_WEIG(1) by STAGE(1)	.294	1	.587
		BASE_WEIG(1) by STAGE(2)	.005	1	.943
		BASE_WEIG(1) by STAGE(3)	4.946	1	.026
	Overall Statistics		38.212	25	.044
Step 5	Variables	SEX(1)	.946	1	.331

		MARIT_STAT	2.435	2	.296
		MARIT_STAT(1)	1.868	1	.172
		MARIT_STAT(2)	1.628	1	.202
		EDUC_LEV	1.144	2	.564
		EDUC_LEV(1)	.019	1	.890
		EDUC_LEV(2)	1.024	1	.312
		COUNS(1)	.075	1	.784
		EMP_STAT	3.480	2	.175
		EMP_STAT(1)	3.238	1	.072
		EMP_STAT(2)	.038	1	.845
		NO_ROOMS(1)	4.089	1	.043
		STAGE	2.520	3	.472
		STAGE(1)	2.168	1	.141
		STAGE(2)	.079	1	.779
		STAGE(3)	1.106	1	.293
		REGIMEN	4.199	2	.123
		REGIMEN(1)	1.811	1	.178
		REGIMEN(2)	3.836	1	.050
		TOBAC	.498	2	.780
		TOBAC(1)	.182	1	.670
		TOBAC(2)	.028	1	.868
		DRUG(1)	3.834	1	.050
		EDUC_LEV * EMP_STAT	1.774	4	.777
		EDUC_LEV(1) by EMP_STAT(1)	.069	1	.793
		EDUC_LEV(1) by EMP_STAT(2)	.241	1	.623
		EDUC_LEV(2) by EMP_STAT(1)	1.051	1	.305
		EDUC_LEV(2) by EMP_STAT(2)	.092	1	.762
		BASE_WEIG * STAGE	.815	3	.846
		BASE_WEIG(1) by STAGE(1)	.015	1	.902
		BASE_WEIG(1) by STAGE(2)	.412	1	.521
		BASE_WEIG(1) by STAGE(3)	.010	1	.919
	Overall Statistics		36.467	24	.049
Step 6	Variables	SEX(1)	.773	1	.379
		MARIT_STAT	1.988	2	.370
		MARIT_STAT(1)	1.416	1	.234
		MARIT_STAT(2)	1.420	1	.233
		EDUC_LEV	1.292	2	.524
		EDUC_LEV(1)	.037	1	.848
		EDUC_LEV(2)	.836	1	.361
		COUNS(1)	.016	1	.899
		EMP_STAT	3.165	2	.205

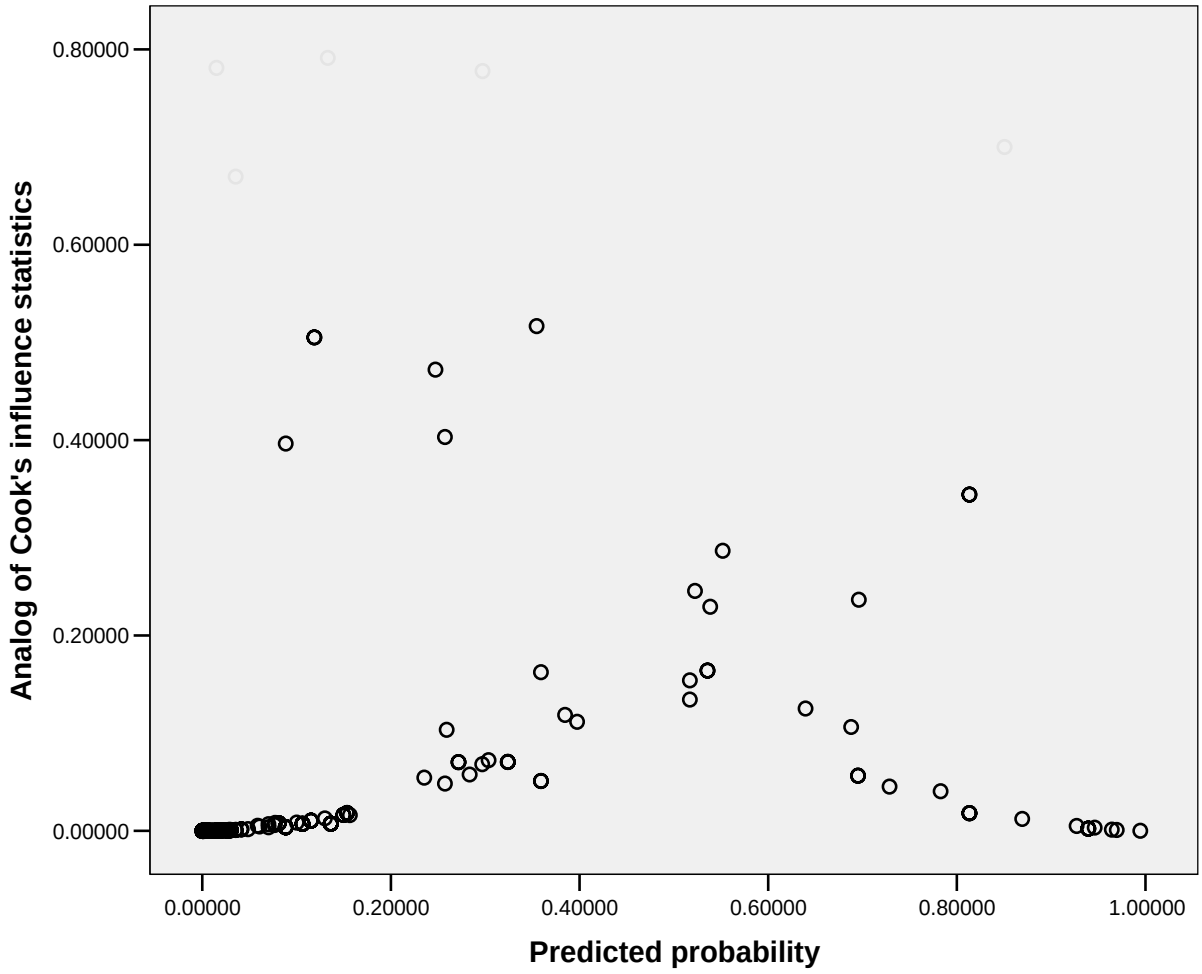
		EMP_STAT(1)	2.824	1	.093
		EMP_STAT(2)	.000	1	.985
		STAGE	2.649	3	.449
		STAGE(1)	2.405	1	.121
		STAGE(2)	.001	1	.978
		STAGE(3)	.952	1	.329
		REGIMEN	3.823	2	.148
		REGIMEN(1)	1.360	1	.244
		REGIMEN(2)	3.280	1	.070
		TOBAC	.332	2	.847
		TOBAC(1)	.155	1	.694
		TOBAC(2)	.042	1	.838
		DRUG(1)	5.436	1	.020
		EDUC_LEV * EMP_STAT	2.679	4	.613
		EDUC_LEV(1) by EMP_STAT(1)	.021	1	.884
		EDUC_LEV(1) by EMP_STAT(2)	.693	1	.405
		EDUC_LEV(2) by EMP_STAT(1)	1.628	1	.202
		EDUC_LEV(2) by EMP_STAT(2)	.001	1	.971
		BASE_WEIG * STAGE	.699	3	.873
		BASE_WEIG(1) by STAGE(1)	.001	1	.978
		BASE_WEIG(1) by STAGE(2)	.235	1	.628
		BASE_WEIG(1) by STAGE(3)	.031	1	.861
	Overall Statistics		31.263	23	.116
Step 7	Variables	SEX(1)	.450	1	.502
		MARIT_STAT	1.022	2	.600
		MARIT_STAT(1)	.850	1	.357
		MARIT_STAT(2)	.535	1	.465
		EDUC_LEV	1.444	2	.486
		EDUC_LEV(1)	.213	1	.644
		EDUC_LEV(2)	.622	1	.430
		COUNS(1)	.007	1	.934
		EMP_STAT	3.568	2	.168
		EMP_STAT(1)	2.636	1	.104
		EMP_STAT(2)	.071	1	.790
		STAGE	2.787	3	.426
		STAGE(1)	2.113	1	.146
		STAGE(2)	.044	1	.834
		STAGE(3)	1.290	1	.256
		REGIMEN	2.440	2	.295
		REGIMEN(1)	.739	1	.390
		REGIMEN(2)	1.949	1	.163

	TOBAC	3.183	2	.204
	TOBAC(1)	3.104	1	.078
	TOBAC(2)	2.094	1	.148
	EDUC_LEV * EMP_STAT	3.012	4	.556
	EDUC_LEV(1) by EMP_STAT(1)	.068	1	.794
	EDUC_LEV(1) by EMP_STAT(2)	1.019	1	.313
	EDUC_LEV(2) by EMP_STAT(1)	1.000	1	.317
	EDUC_LEV(2) by EMP_STAT(2)	.197	1	.657
	BASE_WEIG * STAGE	.581	3	.901
	BASE_WEIG(1) by STAGE(1)	.018	1	.895
	BASE_WEIG(1) by STAGE(2)	.036	1	.849
	BASE_WEIG(1) by STAGE(3)	.188	1	.664
Overall Statistics		24.885	22	.303

Change in Deviance against Predicted probabilities



Cook’s Distance against Predicted probabilities



—

DECLARATION

I, the undersigned, declare that the thesis is my original work, has not been presented for degrees in any other University and all sources of material used for the thesis have been duly acknowledged.

Name: Nuredin Ibrahim
Signature: .....
Place: Faculty of Science, Addis Ababa University
Date: August, 2007

This thesis has been submitted for examination with my approval as a University advisor.

……...
Professor Eshetu Wencheko