Empirical Analysis of Field Data on HIV/AIDS Epidemic in Khartoum State, Sudan

Mohamed, Issam A.W.

Department of Economics, Al Neelain University, Khartoum, Sudan

2011

Empirical Analysis of Field Data on HIV/AIDS Epidemic in Khartoum State, Sudan

Professor Dr. Issam A.W. Mohamed[1]

[1] Professor of Economics, Alneelain University, Khartoum-Sudan. P.O. Box 12910-11111.
issamawmohamed@yahoo.com

1. Abstract

The study focuses on the HIV/AIDS in Khartoum state during the period (2003-2007). The main objectives are to study the situation of HIV/AIDS in Khartoum state through a sample of 1439 of volunteers for the three selected blood testing and counseling centers in Khartoum, Omdurman, and Khartoum North teaching hospitals. The data of the study were collected from secondary source namely the registered information about volunteers after testing blood for HIV. The data were analyzed by using descriptive statistics, chi-square test for dependency between demographic variables and HIV/AIDS incidence, and logistic regression model to discover the effect of predictors variables on the dependent HIV/AIDS incidence. The infected percent is found to be very high in Khartoum center (36.0), 33.5 for Omdurman center. The incidence rate of HIV/AIDS is found to 14.3 for Khartoum North center. The logistic regression model results have concluded social status, Occupation, and education level affect the HIV/AIDS incidence Khartoum state. The trends of the disease during 2007 have indicated that, the pandemic is at increasing rate for both sexes males and females positive although the positive cases for males were greater than females positive cases, except in one or two months. It is recommended that there must be assessment for the current situation of the HIV/AIDS so as to construct strategic plan to stop or eradicate the spread among the people mainly adults. The study has focused on the HIV/AIDS in Khartoum state during the period (2003-2007).Its main objective was to study the situation of HIV/AIDS in Khartoum state through a sample of 1439 of volunteers for the three selected blood testing and counseling centers in Khartoum, Omdurman, and Khartoum North teaching hospitals. The data of the study were collected from secondary source namely the registered information about volunteers after testing blood for HIV checking The data were analyzed by using descriptive statistics, chi-square test for dependency between demographic variables and HIV/AIDS incidence, and logistic regression model to discover the effect of predictors variables on the dependent HIV/AIDS incidence. The infected percent is found to be very high in Khartoum center (36.0), 33.5 for Omdurman center, and The incidence rate of HIV/AIDS is found to 14.3 for Khartoum North center. The logistic regression model results have concluded social status, occupation, and education level affect the HIV/AIDS incidence Khartoum state. The trends of the disease during 2007 have indicated that, the pandemic is at increasing rate for both sexes males and females positive although the positive cases for males were greater than females positive cases, except in one or two months. The study recommended that there must be assessment for the current situation of the HIV/AIDS so as to construct strategic plan to stop or eradicate the spread among the people mainly adults.

2. Introduction

In the last two decades since the first cases of AIDS were identified, HIV/AIDS has emerged as one of the leading challenges for global public health. Particularly in Sub-Saharan Africa, where the majority of HIV and AIDS cases are concentrated, the epidemic continuous to take an extraordinary human toll. To plan and evaluate control strategies effectively and to prepare for the vaccine efficiency trials, it is critical to estimate the magnitude and trajectory of the HIV/AIDS epidemic. Trade-off between alternative interventions and policies must be leased on the best possible information about current levels and trends in the epidemic. Unfortunately, population-based epidemiological data for Sub-Saharan Africa are extremely limited .Incidence data are rare because direct measurement is difficult, and because cohort studies are expensive and required long follow-up periods, AIDS notification data capture only fraction of new AIDS cases and are subject to reporting delays. Information on AIDS attributable mortality is also essential to assess the impact of the epidemic; about vital registration systems have extremely limited coverage in most of Sub-Saharan Africa. The most widely available epidemiological data on HIV/AIDS in Africa are seroprevelance data. Although population -based prevalence surveys would be the most useful, they have been undertaken in only a small number of locations. By contrast, sentinel surveillance systems which monitor the prevalence of HIV infections in specific sub populations have been established in countries throughout the region, and data are available for range of different population groups. Including Commercial Prostitutes, injecting drug users, blood donors, and pregnant women attending antenatal clinics. It is believed that the antenatal clinics data most closely approximate prevalence levels in the adult population, although the relationship between prevalence among clinic attendees and that of the general population remains uncertain. (WHO, 2002).

On the other side, the prevalence of Human Immune Deficiency Virus (HIV) among adults and mortality among under -5 years- olds have increased or stagnated in many countries. Although (HIV) and Acquired Immune Deficiency Syndrome (AIDS) originally emerged as adult health problems, they have become a major killer of under 5- years- old, specially in developing countries. Children of HIV- seropositive mothers can acquire the virus directly through vertical transmission, about (25-30)% of children born to infected mothers become infected with HIV and almost all of them die before 5 years of age in most developing countries with high HIV prevalence. Thus, under 5 mortality rates among children of HIV infected another is two to five times higher than those among children of HIV negative mothers. The joint United Nations program on HIV/AIDS (UNAIDS) estimated that by 1996 a cumulative total of 3 million children had been infected with HIV worldwide. Most of them had died. That number should have doubled by 2011. Since then, many move children are believed to have died as result of HIV infection and other AIDS- related problems. According to UNAIDS estimates, about 14 million of childbearing age currently live with HIV/AIDS in the world, giving birth to children with an elevated risk of HIV infection and death before 5 years .Even among children who are not infected, many will die because the resources needed to ensure their survival and health are used to care for HIV positive adults. Adult HIV prevalence and under 5 mortality rates therefore seem to have increased or stagnated in many countries .Recent World Bank report suggests that life expectancy at birth in countries with high HIV-prevalence, such as Botswana , Zambia and Zimbabwe is lower in 2000 than it was in 1975 . Although 14 million lives have already been lost to HIV/AIDS, it is estimated that only 10% of the total impact of illness and death from the pandemic has been so far. One of the concerns is whether the effects of the epidemic will, revere the substantial child survival gains that have resulted from huge international efforts over the past decades in several developing countries (WHO, 2000).

The importance of the study comes from the importance of the AIDS issue itself, because it destroys the human beings and threats them. So the world organizations, governments, researchers, and policy makers must give a high concern to this urgent issue. Our emphasis here is on the role of statisticians to express the dangerous levels of AIDS through statistical functions and models.

It is very clear that statistical models and techniques become useful in many types of life, and it is interesting because it is a means of using data to gain insight into real problems. As the continuing revolution in computing relieves the burden of calculating and graphing, on emphasis on statistical concepts and insight from data becomes both more important and more practical. Statistics in practice is concerned with gaining understanding from data; it is focused on problem solving rather than on methods that may be useful in specific settings. The aims of application of any methods focus on data and statistical reasoning rather than on either the presentation of as many methods as possible or the mathematical reasoning should be the most important objective of any researcher. The statistical analysis of lifetime or response time data has become a topic of considerable interest of statistical and workers in different areas such as engineering, medicine, biological, and social science. This field has expanded rapidly in recent years, and publications on the subject can be found in the literatures of several disciplines besides statistics. The aim is to deal with statistical methods and models so as to obtain a good measurement of AIDS in purpose to give a complete picture of this dangerous disease. The use of the models through the computer programs (statistical packages) like Statistical Analysis System (SAS) and Statistical Packages for Social Sciences (SPSS) to give abroad coverage of the variables denotes AIDS and its fields.

The models related to the study could be applied in the medical or biomedical sciences. These include the various parametric models and their association statistical methods, non-parametric and distribution –free models, and graphical procedures. The study attempts to illustrate the data of AIDS from the secondary data so as to develop theoretical frame to be ready for applications and modeling methodology. On the other side of the study, the models can be arranged through modern statistics for AIDS problems, that is necessity to give a good analysis and interpretation continuous models like the case of a single lifetime variables. On the other hand, our purpose is to apply and modify some important statistical models related to AIDS and its variables, numerous parametric models will be used in the analysis of lifetime data and in problems related to the modeling. Among univariate models, a few particular distributions occupy a central role because of their demonstrated usefulness in a wide range of situations. Foremost in this category are the Exponential, Weibul, Gamma, and Log-normal distributions. This section introduces models, as well as the Generalized Gamma Distribution, which can be considered as a generalization of each other. A few additional univariate models are also needed and multivariate regression models. The motivation for using particular model is a given situation is often mainly empirical, it has been found that the model satisfactorily describes the distribution of life times in the population under the study (Lawless, 1982). AIDS was recognized as an emerging disease only in the early 1980s. It has rapidly established itself throughout the world, and is likely to endure and persist well into the 21s century. AIDS has evolved from a mysterious illness to a global pandemic which has infected tens of millions less than 20 years. According to the recent estimates, the end of 1999 about 34.4 million people were living with HIV/AIDS. There were 5.4 million new infections and 2.8 million deaths during 1999, bringing the cumulative death toll to 18.8million Trends in AIDS incidence portray important difference between regions of the world .More than 95 percent new cases remain in developing countries. In Africa, HIV infection in women now out-numbers those in men. The number of AIDS deaths in industrial countries has recently been falling due to combine antiretroviral therapies introduced during the past few years .In developing countries where the vast majority of those infected with HIV live ,the number of new cases continue to increase. Beyond the death toll and human suffering, AIDS continues to roll back hard on development gains in many region of the world. The trend in HIV infection will have a profound impact on future infant, child and maternal, life expectancy and economic growth.

In South-East Asia region, the number of reported cases continues to increase and is likely to do so well into the early part of 21s century. The potential for continued spread of HIV/AIDS in Asia and Western Pacific is real requires determined and sustained prevention efforts. Several countries have already experienced intense epidemic in certain population or in the population at large .In these countries, including Cambodia, India, Myanmar, and Thailand, AIDS has imposed new demands on health care systems. In 1984, Thailand was the first country in the region to report a case of AIDS. In most other countries, HIV infection was not diagnosed till 1986 or later. Since then HIV infection has spread rapidly and WHO estimates that currently there are about 5 million HIV infected people in the region, an alarming 15 percent of the world's total .Besides persons with high risk behavior, HIV infection rates have now begun to increase in the general population as well as of January 2000, over 135000 cases of AIDS have been reported. India, Thailand and Myanmar reported the majority of cases with HIV/AIDS in the region. The analysis of the cases ensures that 91% of the AIDS cases are in the age group of (15-49) years and 4.6% are children. The male to female ratio is 4 to1.Heterosexual contact is the predominant mode of spread (85%), followed by injecting drug use 7%and mother to child transmission5% (Park, 2001). Most mathematical models involve unknown parameters that must be estimated from observed data. As new models are created, estimation methods appropriate to those models must be developed. Among the general approaches available for construction of such estimates are the methods of moments and the methods of likelihood. Methods of moments require the ability to compute population moments of the observation associate with the postulated models, whereas method of maximum likelihood requires the ability to compute the likelihood of the observed data under the given model (Lawless, 1982). When the model is complex such computations can prove challenging. In this study the natural history model of AIDS is the one of interest in its own right. On the other side, there is a main concept that found in the medical field there is no cure for AIDS but, there must be more efforts to eliminate this disease or to eradicate. In addition to that, people in many countries still lags behind in the treatment AIDS in terms of primary health care, medical laboratories, and health centers. That may effect on the high risk of exposure to HIV positive without any control or plans to stop the virus from dissemination over all the world among the early ages and teen-ager. Thus , our statistical model depend on the fact that, there is no definite equations or functions applied to AIDS, study will attempt to do that and find a relations between the variables through statistical models related to it. In purpose it is to provide statistical base and analysis of data to put them under-investigations and control.

1. What is the relation between HIV/AIDS and demographic variables?
2. What is the current level and incidence of HIV/AIDS among people in Khartoum State?
3. What is the effect of demographic variables on HIV/AIDS and what is the main factor that influences the HIV/AIDS situation?

It is hypnotized that

1. Non-linear models fit the HIV/AIDS data well, especially binary logistic regression.
2. The demographic variables affect the HIV/AIDS incidence in Khartoum state.
3. HIV/AIDS will be increasing among volunteers in the three centers (Khartoum, Khartoum North and Omdurman).

The main objective of the research is to study HIV/AIDS situation in Khartoum state through statistical models based on the facts known about the spread of the (AIDS).Virus and to predict its spread throughout the population that is statistically representative of the average population. The study selected as result of the fact AIDS is adversely impacting population throughout the world number of dangerous viruses have been eradicated in recent years via improved medical technology and treatment, however on cure has been yet discovered for AIDS. Because there is no positive treatment for the cure of AIDS, this situation needs in depth studies that will be required in both developed and undeveloped countries.

In Sudan, statistical models can be applied to evaluate multiple aspects involved in exposing a population to the AIDS virus and describe the level, trends of the disease and predicting the spread of the virus within population .The models will be applying for the historical data collected from the clinics, donors investigation centers, In purpose to obtain a good models of the AIDS virus.

Study has two types of objectives; they are main objective and sub-objectives. The main objective of the study is to apply statistical models for HIV/AIDS in Khartoum state so as to obtain a good analysis and know the situation of the disease:

1- To reveal the concept of AIDS through statistical models.

2- To discover the factors affecting the HIV/AIDS incidence in the three centers under the study.

3- To apply the logistic regression so as to discover the relation between variables denotes AIDS.

The research methodology is divided into two types, they are:

Study depend on secondary data was collected from the volunteers centers for blood testing inside Omdurman, Khartoum, and Khartoum North teaching hospitals. The data collect through questionnaires designed so as to provides all the information registered inside the centers. Study construct questionnaire filled from the records inside the three centers. It comprise of ten questions , question one is the age of respondents, question two sex of volunteers, question three investigate the place of residence, question four is about marital status, question five the occupation of individuals, question six is the education level , question seven religious status, question eight symptoms of the disease, question nine about the risk factors, and the last question is about the result of blood testing. The questionnaire filled direct from registered data in the counseling units without see the patients, because of the critical situation and security about the disease.

3. Analysis of Data

Study will stems the good methods and techniques either parametric or non-parametric models. The main idea of analysis is to identify statistical model that related to AIDS by using statistical packages or programs language to construct the models depend on these programs. Mainly the study depend on Statistical package for Social Science (SPSS) to generate the coefficient of the models and the other statistical tests.

Study will apply the statistical models related to the data of the AIDS, then definite models and choose the standard model. In the study plan, the focusing will be on binary logistic regression, because it's suitable to the data collected from the three centers inside the three hospitals in Khartoum state. Also estimate coefficients and statistical tests can be done to distinguish between the variables that related to HIV/AIDS incidence and spread through people in the three cities Khartoum, Khartoum North and Omdurman.

4. Previous Studies

Lippincott and Wilkins (2004) conducted a study about the knowledge, practices and behaviors related to HIV transmission among Brazilian population in the 15-4 years age group. The main objective of the study was to describe transmission vulnerability for acquiring HIV infection among the Brazilian aged (15-54) years. The study designed as a population based survey the sampling was stratified by geographical rejoin, total of 6006 interviews were conducted. Indicators of knowledge and sexual practices and the relative sizes of the vulnerable subgroups were estimated. Logistic regression analysis was used to determine the factors associated with safe sex practices. The results were as follows, regarding knowledge indicators in the age group (15-24) years, a high percentage 91 spontaneously cited sexual intercourse as a form of HIV transmission, and 62% had correct knowledge of the models of HIV transmission. The proportion of consistent condom use with causal partners was 52%, increasing to 59% in the youngest age group. Higher proportion of inconsistent condom use with any kind of partner were found among women and among the poorest. A multiplicity of sexual partners, low socio-economic status and cocaine use were important predictors of unprotected sex among men living without a companion. Among individuals age (15-49) years, 0.2% currently inject cocaine,4.6% of the men paid for sex at least once over he past year and 1% of the women were paid in exchange for sex. Among sexually active men of the same age group, 35% reported sexual relations with other men.

The study conclude that, besides the need to establish the role exercised by the vulnerable subgroups in the HIV transmission dynamics, results indicate that it is necessary to investigate unsafe sexual practices further among the poorer sectors of society. Johnson and Dorrington (2006) describes an approach to incorporating the impact of HIV/AIDS and the effect of HIV/AIDS prevention and treatment programs into a cohort component projection model of the south African population. The modeled HIV-positive population is divided into clinical and treatment stages, and it is demonstrated that the age profile and morbidity profile of the HIV-positive population is changing significantly over time. HIV/AIDS is projected to have a substantial demographic impact in South Africa. Prevention programs –social marketing, voluntary counseling and testing, prevention of mother to child transmission and improved treatment for sexually transmitted diseases are unlikely to reduce AIDS mortality significantly in the short term. However, more immediate reductions in mortality can be achieved when antiretroviral treatment is introduced.

Mahfouz (2007) conduct a research about modeling the impact of HIV/AIDS epidemic on mortality in sub-Sahara Africa with reference to Zimbabwe, Tanzania and Sudan (1985-2003),the main objective of the study was to improve the methodological basis for modeling HIV/AIDS epidemic in adults in sub-Sahara and to develop estimates of prevalence, incidence and mortality and to study the impact of HIV/AIDS on life expectancy and population growth with examples of Zimbabwe ,Tanzania and Sudan. Previous mathematical models were developed to estimate epidemic trends based on sentinel surveillance data from pregnant women. The study extended these models in order to take full advantage of the available data. The methodology involved Gauss Newton approach for the estimation of prevalence and incidence of HIV/AIDS in the studied countries using curve fitting procedure. Further back calculation method was used to estimate AIDS cases and mortality. Also life tables were constructed to asses the impact of HIV/AIDS on life expectancy. To trace the impact of HIV/AIDS on population growth study project the population growth with and without AIDS using demographic techniques for estimating population growth. For every country three models were estimated, double logistic curve was suitable for both Zimbabwe and Tanzania, the rational model was found to be suitable for Sudan. The application of models reveals the impact of HIV/AIDS in the target countries where adult HIV prevalence is significant. The epidemic has already had a number of serious consequences, including rise in the number of deaths and the reduction of life expectancy, particularly deaths of persons aged 15 to 45.Furthermore, some of the serious effects of the epidemic are expected to worsen in the future. By the year 2020 population of studied countries is expected to be lower than it would have been in the absence of AIDS. Increases of mortality have been particularly marked in Zimbabwe with the highest HIV prevalence. Also life expectancy at birth has already fallen dramatically, dropping within a decade or two levels last recorded before 1985, and in the case of Tanzania , the impact of HIV/AIDS on mortality and population growth is less of degree than Zimbabwe, this due to the fact that Tanzania epidemic is considered as moderate. For Sudan the impact of HIV/AIDS is less effective because the epidemic is of low level. Markus Haacker (2002) carried a paper about modeling the Macroeconomic of HIV/AIDS. The paper addresses the impact of HIV/AIDS on per capita output and income with particular emphasis on the role of labor mobility between the formal and informal sectors, and the impact of the epidemic on investment decision. The study found that HIV/AIDS affects both the supply of labor and the demand for labor in the formal sector. Only if there is a significant rise in the capital –labor ratio, will there be an increase in formal sector employment. However, this is associated with a decline in the rate of return to capital. To the extent that companies respond to this by reducing investment, conventional models underestimate the adverse impact on employment, per capita output, and income. The analysis of the impact of HIV/AIDS on output is complemented by an assessment of the impact on income.

Sharif (2003) conducted a study about clinical manifestations of AIDS and its correlation to the CD4-lymphocyte counts in adult Sudanese presenting to Khartoum Teaching Hospital in the period between October 2001 to June 2002, in an attempt to provide comprehensive description of the clinical manifestations of AIDS in adult Sudanese, and to correlate the manifestations to the CD4+ lymphocyte counts at presentation. Ninety individuals of ELISA confirmed, symptomatic HIV positive cases were enrolled. Various clinical and laboratory data were obtained and statistically analyzed. Anorexia with weight loss , chronic diarrhea , prolonged fever, dysphasia and chronic cough were the commonest presenting symptoms, followed by various skin and musculoskeletal complaints. Oral candidiasis cachexia, pallor and generalized lymphadenopathy were the most prevalent clinical signs, while molloscoum contagiousum was the predominant skin lesion detected in our cases. Anaemia and rapid ESR are useful indicators of the disease progression in AIDS patients, while prolonged diarrhea and weight loss were significant markers of low CD4+ counts. Most of the study subjects (70%) were either in stage III or IV of the disease with CD4+ counts < 200 cell/m1, 8.9% in early stage of the disease with CD4+ count > 500 cell/m1.The major symptoms and signs were significantly more relevant in patients with stage III and IV. It was concluded that most of Sudanese adult AIDS patients present late in the disease process and there is a significant correlation between the clinical manifestations of this disease and the CD4+ lymphocyte counts.

5. Literature Review

In the 1981, the first cases of the severe Immune System deterioration were recognized when several young men developed unusual infections and cancer. The new disease was later named AIDS. At that time, no one knew what was causing the disease. Since then, science has shown that HIV is the caused of AIDS. We know this because people who become infected with the virus can develop AIDS, and medicines that against the virus can prevent infected people from getting sick. As HIV infection progresses, it weakens a person ability to fight off diseases. By attacking the immune system, the virus leaves people susceptible to other disease. When a person with HIV contrasts one of several additional diseases, or when a person's immune system shows marked deterioration, that person is classified as having AIDS. Some individuals and groups have tied to push the false idea that AIDS is not caused by HIV. These claims can be really harmful. If people believe that they do not need to reduce their risk of getting or spreading HIV, they could become infected or pass the virus to others. Also, if people with HIV think they do not need to get medical care or consider treatment, they could become very sick and developed AIDS.

From time to time, the idea that HIV dose not caused AIDS resurfaces in the media. A march 2006 Harper's article out of control: AIDS and the corruption of medical science allege misconduct of certain AIDS research trials and dismiss the overwhelming scientific evidence that HIV causes AIDS. It is caused by the HIV by leading to destruction and functional impairment of cells of the immune system, notably CD4+ T cells; HIV progressively destroys the body's ability to fight infections and certain cancers. An HIV infected person is diagnosed with AIDS when his or her immune system is seriously comprised and manifestation of HIV infections are severe. The U.S centers for disease control and prevention (CDC) currently defines AIDS in adults or adolescent age 13 years or order as the presence of one of 26 conditions indicative of severe immune suppression associated with HIV infection, such as Pneumocystics Carini Pneumonia(PCP), a condition extraordinarily rare in people with HIV infection. Also most other AIDS defining conditions are also opportunistic infections which rarely cause harm in healthy individuals. A diagnosis of AIDS also is given to HIV infected individuals when their CD+4 T cell count falls below 200 cells cubic millimeter of blood. Healthy adults usually have CD+4 T cell count of 600-1500 mm of blood. In HIV infected children younger than 13 years, the CDC definition of AIDS is similar to that in adolescents and adults, except for the addition of certain infections commonly seen in pediatric patients with HIV *(CDC, 1992)*.

In many developing countries, where diagnostic facilities may be minimal, health care workers use of a World Health Organization (WHO). AIDS case definition based on the presence of clinical signs associated with immune deficiency and the exclusive of other known causes of immune suppression, such as cancer or malnutrition. As of the end of 2000, an estimated 36.1 million people world wide 34.7 million adults and 1.4 children younger than 15 year were living with HIV/AIDS. Through 2000, cumulative HIV/AIDS associated deaths worldwide numbered approximately 21.8 million 17.5 million adults and 4.3 million children younger than 15 years. In the United States, an estimated 800000 to 900000 people are living with HIV infection. As of December 1999 about 733,374 cases of AIDS and 430,441 AIDS related deaths had been reported to the CDC. AIDS is the fifth leading cause of death among all adults aged 25 to 44 in the United States. Among Africans – Americans in 25 to 44 age group, AIDS is leading cause of death for men and the second leading cause of death for women (*UNAIDS 1999).*

Among many criteria used over the years to prove the link between putative pathogenic (disease-causing) agents and disease, perhaps the most- sited are Koch's postulates, developed in the 19th century. Koch's postulates have been variously interpreted by scientists, and modifications have suggested accommodating new technologies, particularly with regard to viruses. However, the basic tenets remain the same, and for more than a century Koch's postulates, as listed below, have severed as the litmus test for determining the cause of any epidemic disease:

Epidemiological association: the suspected cause must be strongly associated with the disease.
Isolation: the suspected pathogen can be isolated and propagated outside the host.
Transmission pathogenesis: transfer of the suspected pathogen to an uninfected host, man or animal; produce the disease in the host. With regard to postulation 1, numerous studies from around the world show that virtually all AIDS patients are HIV seropositive; that is they carry antibodies that indicate HIV infection. With regard to postulate #2, modern culture techniques have allowed the isolation of HIV in virtually all AIDS patients, as well as almost all HIV seropositive individuals with both early –and late – stage disease. In addition, the polymerase chain (PCR) and other sophisticated molecular techniques have enabled researchers to document the presence of HIV genes in virtually all patients with AIDS, as well as in individuals in earlier stages of HIV disease.
Postulation 3, has been fulfilled in tragic incidents involving three laboratory workers with no other risk factors who have developed AIDS or severe immune suppression after accidental exposure to concentrated, cloned HIV in the laboratory. In all three cases, HIV was isolated from the infected individuals, sequenced and shown to be the infecting strain of virus. In another tragic incident, transmission of HIV from a Florida dentist to six patients has been documented by genetic analyses of virus isolated from both dentist and the patients. The dentist and three of the patients developed AIDS and died, and at least one of the other patients has developed AIDS. Five of the patients had no HIV risk factors rather than multiple visits to dentist for invasive procedures. (*Hardon and Goedret, 1996*). In addition, through 1999, the CDC had reports of 56 health care workers in the United States with documented, occupationally acquired HIV infection, of whom 25 have developed AIDS in absence of the risk factors. The development of AIDS known HIV seroconversion also has been repeatedly observed in pediatric and adult blood transfusion cases, in mother to child transmission, and in studies of hemophilia, injection, drug use and sexual transmission in which seroconversion can documented using serial blood samples. For examples, in a 10 years study in the Netherlands, researchers followed 12 children who had become infected with HIV as neonates by small aliquots of plasma from a single HIV infected donor. During the 10 years period, eight of the children died of AIDS. Five people worldwide die of AIDS every minute. HIV has hit every corner of the globe, infecting more than 42 million men, women and children; 5 million people were infected in 2002 alone. Started in 1988, World AIDS Day is not just about raising money, but also about raising awareness, education and fighting prejudice. World AIDS Day serves to strengthen global efforts to address the challenges of the AIDS pandemic, which continues to spread throughout every region of the world. This day is important in reminding people that HIV has not gone away, and that it should remain a major concern. The theme of the World AIDS Campaign 2002-2003 Live and let live, focuses on eliminating stigma and discrimination, the major obstacles to effective HIV/AIDS prevention and care. World AIDS Day is commemorated around the globe on 1 December. It celebrates progress made in the battle against the epidemic and brings into focus remaining challenges. On the occasion of World AIDS Day, UNAIDS launched the AIDS epidemic update 2003, on 25 November 2003. The annual report describes latest developments in the global HIV/AIDS epidemic and provides regional summaries.
In a global summary of the HIV/AIDS pandemic UNAIDS figures show that there are an estimated 37 million adults living with AIDS and a further 2.5 million children under 15 year's old living with the disease. They predict a total number of 5 million newly infected people during 2003. And a total number of 3 million AIDS deaths during 2003. HIV still remains a threat to people of all ages and nationalities. Stigma and Discrimination will be highlighted as a theme for the 2003 World AIDS Day. In many parts of the world, discrimination prevents people who are known to have HIV from securing a job or caring for their families. Discrimination can cause isolation and marginalizes people who have HIV and AIDS. This can prevent people from being offered or seeking the treatment which could save their lives. In order for HIV to be effectively tackled on an international level, efforts need to be made to end the discrimination against people with HIV/AIDS, educate them in safer sex and drug use, using appropriate media, providing condoms freely to people in the

developing world, providing financial and medical assistance so that people with HIV/AIDS can be treated. Negative attitudes about HIV can create a climate in which people become more afraid of the stigma and discrimination associated with the disease than of the disease itself. When fear and discrimination prevail, people may choose to ignore the possibility that they may be HIV-positive; even if they know they have taken risks. People may decide not to take measures to protect themselves in fear that in doing so they could be associating themselves with HIV. HIV/AIDS touches raw nerves in all our communities. The stigma of HIV/AIDS relates to deep taboos within society. For many the disease has a strong association with prolonged illness, death, sex and drug use; issues that many of us find difficult to talk about openly. Along with general discomfort about discussing these 'taboo' issues, many communities are also dealing with high levels of ignorance, denial, fear and intolerance about the disease itself. This potent combination can lead to rejection and even aggression against people living with HIV. All of this helps to create an environment in which the disease can more easily spread. This year's World AIDS Campaign encourages both individuals and institutions to reflect on how they respond to those living with HIV/AIDS. With challenging posters and television images the campaign clearly shows how the most painful symptoms of HIV/AIDS are often the reactions of others. When someone feels safe within their own community, they are more likely to take responsibility for their HIV status. This is why it is so important for all of us to examine our own attitudes. Only by confronting stigma and discrimination across the world will the fight against HIV/AIDS be won (*World AIDS campaign 2003*).

From the other side, Patients who are identified with HIV around the time they are first infected may benefit from potent antiviral therapy given at that time. The rationale for initiating this early treatment is primarily theoretical. Preliminary evidence, however, suggests that unique aspects of the body's immune response to the virus may be preserved by this strategy. It is thought that treatment during the primary infection may be an opportunity to help the body's natural defense system to work against HIV. Thus, patients may gain an improved control of their infection while on therapy, and perhaps even after therapy is stopped. At one time, the hope was that if therapy was started very early in the course of the infection, HIV could be eradicated. Most evidence today however suggests that this is not the case. Consequently, early treatment is not likely to result in a cure, although other benefits may still exist. Therefore, the current recommendation is that patients with primary infection should be referred to clinical studies where the potential role of therapy can be discussed and further explored. If emotional or social situations make adherence to such treatment questionable, however, the patients are clearly better off delaying therapy. After all, on the average, infected persons can expect to remain healthy for a prolonged period of time. Regardless, patients need to be aware that initiating treatment early puts them at risk for developing short and long-term side effects as well as resistance to the drugs.

6. Treatment for HIV during Pregnancy

One of the One of the greatest advances in the management of HIV infection has been in pregnant women. Prior to antiviral therapy, the risk of HIV transmission from an infected mother to her newborn was approximately 25-35%. The first major advance in this area came with studies giving ZDV after the first trimester of pregnancy, then intravenously during the delivery process, and then after delivery to the newborn for 6 weeks. This treatment showed a reduction in the risk of transmission to less than 10%. Although less data are available with more potent drug combinations, clinical experience suggests that the risk of transmission may be reduced to less than 5%. Current recommendations are to advise HIV-infected pregnant women regarding both the unknown side effects of antiviral therapy on the fetus, and the promising clinical experience with potent therapy in preventing transmission. In the final analysis, however, pregnant women with HIV should be treated essentially the same as non-pregnant women with HIV. Exceptions would be during the first trimester, where therapy remains controversial, and avoiding certain drugs that may cause greater

concern for fetal toxicity. All HIV infected pregnant women should be managed by an obstetrician with experience in dealing with HIV-infected women. Maximal obstetric precautions to minimize transmission of the HIV virus such as avoiding scalp monitors, and minimizing labor after rupture of the uterine membranes. In addition, the potential use of an elective Caesarean section. (C- section) should be discussed, particularly in those women without good viral control of their HIV infection where the risk of transmission may be increased. Breast feeding should be avoided if alternative nutrition for the infant is available. Despite the reduced risk of transmission associated with antiviral therapy, pregnant women with HIV need to be thoroughly counseled regarding all risks, as well as all options, including therapeutic abortions when appropriate.
Who grouped the HIV signs into two groups major and minor signs:
Major signs:

1. Prolonged fever (more than one month).
2. Prolonged and chronic diarrhea (usually over a month).
3. Significant weight loss (over a period of time and more than 10 percent of body weight).

The Minor Signs and symptoms:

1. Persistent cough for more than one month.
2. Persistent skin infection.
3. Aggressive skin cancer (Kaposi Sarcoma).
4. Oral thrush (canaidiasis).
5. Recurrent shingles (Ananse).
6. Enlargement of the lymph glands.

An individual with two of these major signs and symptoms and two of the minor signs and symptoms plus a positive HIV antibody test it said to have AIDS.

7. HIV/AIDS Statistics

As of the end of 2003, an estimated 37.8 million people worldwide 35.7 million adults and 2.1 million children younger than 15 years - were living with HIV/AIDS. Approximately two-thirds of these people (25.0 million) live in Sub-Saharan Africa; another 20 percent (7.4 million) live in Asia and the Pacific. Worldwide, approximately 11 of every 1000 adults aged 15 to 49 are HIV-infected. In Sub-Saharan Africa, about 7.5 percent of all adults in this age group are HIV-infected. Woman account for nearly half of all people worldwide living with HIV/AIDS. An estimated 4.8 million new HIV infections occurred worldwide during 2003; that is, about 14,000 infections each day. More than 95 percent of these new infections occurred in developing countries. In 2003, approximately 1,700 children under the age of 15 years, and 6,000 young people aged 15 to 24 years became infected with HIV every day. More than 20 million people with HIV/AIDS have died since the first AIDS cases were identified in 1981.In 2003 alone, HIV/AIDS-associated illnesses caused the deaths of approximately 2.9 million people worldwide, including an estimated 490,000 children younger than 15 years. The Centers for Disease Control and Prevention (CDC) estimate that 850,000 to 950,000 U.S residents are living with HIV infection, one-quarter of who are unaware of their infection. Approximately 40,000 new HIV infections occur each year in the United States, about 70 percent among men and 30 percent among women. Of these newly infected people, half are younger than 25 years of age (3-4) years. Of new infections among men in the United States, CDC estimates that approximately 60 percent of men were infected through homosexual sex, 25 percent through injection drug use, and 15 percent through heterosexual sex. Of newly infected men, approximately 50 percent are black, 30 percent are white, 20 percent are Hispanic, and small percentages are members of other racial/ethnic groups.
The estimated number of AIDS diagnoses through 2002 in the United States is 886,575. Adult and adolescent AIDS cases total 877,275 with 718,002 cases in males and 159,271 cases in females. Through the same time period, 9,300 AIDS cases were estimated in children under age 13.The

estimated number of new adult/adolescent AIDS diagnoses in the United States was 43,225 in 1998, 41,134 in 1999, 42,239 in 2000, 41,227 in 2001, and 42,136 in 2002.The estimated number of new pediatric AIDS cases (cases among individuals younger than age 13) in the United States fell from 952 in 1992 to 92 in 2002.The estimated rate of adult/adolescent AIDS diagnoses in the United States in 2002 (per 100,000 population) was 76.4 among blacks, 26.0 among Hispanics, 11.2 among American Indians/Alaska Natives, 7.0 among whites, and 4.9 among Asians/Pacific Islanders. From 1985 to 2002, the proportion of adult/adolescent AIDS cases in the United States reported in women increased from 7 percent to 26 percent. As of the end of 2002, an estimated 384,906 people in the United States were living with AIDS. As of December 31, 2002, an estimated 501,669 people with AIDS in the United States had died. The estimated annual number of AIDS-related deaths in the United States fell approximately 14 percent from 1998 to 2002, from 19,005 deaths in 1998 to 16,371 deaths in 2002. There is estimated 16,371 AIDS-related deaths in the United States in 2002, approximately 52 percent were among blacks, 28 percent among whites, 19 percent among Hispanics, and less than 1 percent among Asians/Pacific Islanders and American Indians/Alaska Natives. UNAIDS estimates that 40 million people are living with HIV/AIDS worldwide. Approximately 56,000 people are living with HIV/AIDS in Canada. Of those, 30% are unaware that they are infected. 2,482 Canadians were diagnosed with HIV in 2003. Health Canada estimates that over 4,000Canadians are actually infected with HIV each year.

Ontario has the highest number of positive HIV test reports in Canada, with a total of 24,408, or 44.2% of all positive reports from 1985 to 2003. Quebec has the second highest number, with 12,464 (22.6%), followed by British Columbia, with 11,552 (20.9%).From 1985 – 2002, Toronto accounted for 14,987 positive HIV test reports. That’s 67% of all positive test reports in Ontario and 28% of all positive test reports in Canada. Men who have sex with men (MSM) accounted for the majority of positive HIV test reports in Toronto, Ontario, and Canada between 1985 and 2003. During this period, MSM accounted for 60.7% of all positive HIV test reports in Canada. Women account for an increasing proportion of positive HIV test reports in Toronto, Ontario, and Canada. Women now account for about 25% of diagnoses each year in Canada. Canadians from countries with high rates of HIV (so-called ‘HIV-endemic countries’) account for an increasing proportion of positive test reports in Toronto, Ontario, and Canada. In 2003, they accounted for 10.2% of all positive test reports. From 1985 – 2003, youth aged 15 – 29 accounted for 27% of all positive HIV test reports in Canada. AIDS is still fatal. There is no cure and there is no vaccine. Prevention is still the best defense against HIV infection Sources. (*UNAIDS, 2004*).

Thousands of girls and boys snatched from their families remain unaccounted for. An estimated 20,000 children have been abducted since the conflict began – more than 5,000 in the last year alone. These children are often forced to commit atrocities against their own communities and to fight alongside their captors. They have been brutally used as human shields and sex slaves. Every resurgence in the fighting brings renewed expressions of concern from governments and a wide range of national and international organizations. But these concerns have yet to be transformed into a constructive peace process. This intensification of the armed conflict will make more vulnerable an already exhausted civilian population in an area plagued by conflict for the past 16 years. UNICEF fears that the increased fighting may further reduce access to the estimated 800,000 displaced people, causing a rapid and significant rise in child mortality and malnutrition. Furthermore, the broadening of military operations is a setback to hopes for peace at this crucial time when efforts were underway to establish a peace process.UNICEF urges both the Government of Uganda and the LRA to renew efforts for dialogue, agree on a cease-fire, and guarantee safe and unimpeded humanitarian access to all of northern Uganda.

According to table (1) world wide the joint United Nations Program on HIV/AIDS (UNIAIDS) estimates that more than 39.4 million people are currently living with HIV/AIDS, 37.2 million were adults, 17.6 million women and 2.20 million children. The newly infected during 2004 concentrated in adults (4.30 million), thus the deaths were (7.9 %) from the infected people.

Table (1) HIV/AIDS Situation in 2004

	Total	Adult	women	children
Living with HIV/AIDS	39.40	37.2	17.6	2.20
Newly infected in 2004	4.90	4.30	1.50	0.64
ADIS deaths in 2004	3.10	2.60	1.10	0.50

Source: UNIADIS 2004

8. Global Summary of AIDS Epidemic

According to table (1) bellow the estimated number of persons living with HIV worldwide in 2007 was 33.2 million, a reduction of 16% compared with the estimate published in 2006 (39.5 million). The reduction due to refinements in methodology, rather than trends in the pandemic itself .Also there is change in six countries Angola, India, Kenya, Mozambique, Nigeria, and Zimbabwe. In both Kenya and Zimbabwe, there is increasing evidence that a proportion of the declines is due to a reduction of the number of new infections which is in part due to a reduction in risky behaviors.

Table (2) Numbers of people living with HIV/AIDS in 2007

Total	33.2million{30.6-6.1million}
Adults	30.8million{28.2-33.6illion}
Women	15.4million{1.4-16.6million}
Children under 5 years	2.5million{2.2-2.6 million}

Source: UNIAIDS report 2007

According to table (2.2) bellow, the new infection of adults is high (2.1) million .The children was (420000) new case, which is critical situation for the disease.

Table (3) people newly infected with HIV/AIDS in 2007

Total	2.5 million{1.8-4.1 million}
Adults	2.1 million{1.4-3.6 million}
Children under 5 years	420000 million {350000-540000 }

Source: UNIAIDS report 2007

There is strong effect of HIV/AIDS on deaths cases in the community, it is very clear that (2.1) million was die during 2007 by last stage of HIV. Also the adults had a high level of mortality comparing with other age groups.

Table (4) AIDS Deaths in 2007

Total	2.1 million{1.9-2.4 million}
Adults	1.7 million{1.6-2.1 million}
Children under 5 years	330000{310000-380000 million}

Source: UNIAIDS report 2007

9. AIDS and Risk of Domestic Violence

Domestic violence against women or children is a theoretical threat to any other form of violence. In practice, risk of domestic violence may weak against the society and the community, including those victims. Women and children will acquire many rights and will be a fighting who improves the respect, dignity and honor. Women have a valuable, peace and tolerance attitudes, perfect condition, always in non-violence condition that they want to get rid of discrimination and violence against them. They have another strengthen that probably will always use, that one for an likely to fight for their violence. If we think one of these women can use this method, nation, governments or the organizations must give it to those victims at some mutually time. The aims are that women have lots of respect for their life. They are coming together for their voices, which gives them the rights to fight for violence. This is not the place to be in violence, when the women are so peaceful. If a

society respects women's prosperity, justice should bring social justice. In the African society, women always are doing fine in their jobs, home and with the family, but don't like discrimination or violence. Whether the women and the organizations are serving like the legislators remains to be effective. If women are empowered then economic progress for them is under the development. On the other hand, domestic violence against women has sharply increased in conflict torn areas. In the conflict women also always blindfolded, beaten and raped. Women face arrest specially in a society were they have to do odd jobs like moon-shining to raise their families where no jobs are available. Women face threats to their lives and security every day. Violence is endemic, in troubled African societies whether in the form of attacks by armed groups or criminal gangs. There is a continuum between the violence in conflict and violence in women's lives. Domestic violence against women takes place in all societies and cultures affecting women regardless of their race. Women are more vulnerable to trafficking, rape and other forms of sexual abuse. It admitted that violence against women is challenge, and is now being accompanied by the risk of HIV infection. We can not forget how young girls were having sex for food in Grenada, and armed hooligans stealing food in Haiti and raping women at shelters. Due to harmful practices like early marriage, wife inheritance and female genital mutilation and the use of rape as a tool of war, the spread of HIV/AIDS among females and violence against women are becoming two sides of the same coin in some areas.
A similar public health crisis could emerge in the war-torn region of Darfur in Sudan, where rape and sexual violence are being used as a weapon of war. The majority of women in Darfur have also undergone female genital mutilation, a factor putting them at increased risk of infection.In many parts of the world women often refrain from seeking medical treatment following rape because they fear being identified as rape victims and ostracized within their communities. In Colombia, Amnesty International has received testimonies about people from stigmatized groups, including those thought to have HIV/AIDS, who have disappeared or been persecuted or killed. Additionally, many women and girls lack awareness of how to protect themselves from HIV/AIDS. In Ethiopia, for example, some 80% of young married women have had no education and are unable to read. Ensuring access to education, including raising awareness about sex, health and HIV/AIDS, is fundamental to protecting the right of girls and women.
The international community has a responsibility to contribute material support. Women and girls in war zones suffer rape and violent abuse while offenders escape punishment. Because national authorities have failed to act to halt such abuses. Despite promises, treaties and legal mechanisms, governments are failed to protect women. While international laws recognized rape and sexual violence as war crimes, prosecutions were not being brought against offenders. Women and girls are not just killed; they are raped, sexually attacked. Women and girls are also affected by land mines. They are the ones that get blown up. In 1986, the UN Economic and Social Council declared domestic violence a serious violation of women's rights. In 1992, the Committee on the Elimination of Discrimination against Women adopted a recommendation urging governments to take steps to eradicate violence against women, seen as a form of gender discrimination. Latin America and the Caribbean was the first continent to secure a major legal instrument to protect women, namely the Convention on the Prevention, Eradication and Punishment of Violence against Women, which was adopted by the Organization of American States (OAS) in 1994 and ratified by 23 member countries. In 1995, the General Assembly urged member states to strengthen criminal, civil, labor-law and administrative penalties in national legislation in order to punish violence against women in both the private and the public spheres. Furthermore, it established that all forms of sexual abuse and trafficking in women and girls are a violation of human rights. This, the first resolution to be adopted on girls, was subsequently ratified in Binstock (1997).
Worldwide, 25% of women are sexually abused at some time in their lives. Close to 120 million women has suffered genital mutilation. Countless women, girls and whole families have been left devastated by rape during recent conflicts in Rwanda, Cambodia, Liberia, Peru, Somalia, Uganda and the former Yugoslavia. 47% of women in Bangladesh have been physically abused by their

husbands or partners at least once in their lives. In the United Kingdom, the police receive domestic violence-related calls at the rate of one a minute. About 87% of the victims are women. Every 23 seconds a woman is raped in South Africa. In America, every 15 seconds a woman is physically abused by her partner. Every 72 hours a woman is murdered by her husband or partner.
According to Inter-American Development Bank (IDB) in Nicaragua in 1997, the daughters of battered women are three times more likely to require medical care. These statistics also revealed that 63% of children who are exposed to domestic violence will have to repeat at least one grade in school. Even Non-violence Against Women on 25 November next, UNI Graphical calls on all member organizations and the international trade union movement as whole to implement all kinds of actions, campaigns, meetings, events, talks and conferences aimed at highlighting the persistence and indeed the worsening of violence against women and girls.
Millions of women and children are still being bought and sold. While domestic violence against women persisted in all countries, the international women's movement and human rights advocates are not doing enough to protect them. Harmful traditional or customary practices, including female genital mutilation and crimes committed in the name of honors, were forms of violence. Domestic violence, which encompassed marital rape as well as other forms of physical, psychological and sexual violence, was one of the most common, least visible forms of violence against women and girls. States had an obligation to eradicate all forms of violence against women. The absence of equal rights, gender-based discrimination and the denial of economic opportunities to women were factors that could lead to women's increased vulnerability to trafficking, which constituted the denial of the rights to liberty, and freedom from violence. On the other side a study involving HIV-positive women in Kenya suggests that providing antiretroviral therapy to individuals with advanced HIV disease may only have a limited impact on HIV transmission. The study, which is published in the April 15th edition of the Journal of Acquired Immune Deficiency Syndromes, found that women with a CD4 cell count below 200 cells/mm^3 had the highest genital shedding of HIV, but were the least sexually active. Women without any symptoms of HIV and a CD4 cell count above 350 cells/mm^3 shed a smaller, but still significant amount of HIV in their genitals, and were the most sexually active. It has been suggested that the widespread provision of antiretroviral therapy could decrease the number of new HIV infections by lowering the viral load and infectiousness of people with HIV. Current World Health Organization (WHO) guidelines recommend that antiretroviral therapy should be provided to individuals who are ill because of HIV or if they have a CD4 cell count below 200 cells/mm^3. However, the impact of antiretroviral therapy on HIV risk at a population level could be limited if infectivity or high-risk sexual behavior is high in individuals who do not qualify for treatment. Investigators from the United States and Kenya therefore looked at the genital shedding of HIV and sexual risk behavior of 650 HIV-positive women who had never taken antiretroviral therapy in Mombassa. They were stratified according to WHO criteria for initiating anti-HIV therapy. The women had a median age of 29, 50% were married and 6% reported selling sex. A total of 65% of women had symptomatic HIV infection. Genital shedding of HIV increased in a stepwise fashion with each advancement in the severity of HIV disease. Women with a CD4 cell count above 200 cells/mm^3 had the highest concentration of HIV in their genitals (3.8 $\log_{10}$ copies/swab) and women with a CD4 cell count above 350 copies/ml and no symptoms of HIV the lowest (2.4 $\log_{10}$ copies/swab). This difference was highly statistically significant ($p < 0.001$). Nevertheless, the concentrations of HIV observed in the genital swabs obtained from women with the least advanced HIV disease were still high and implied a significant risk of onward transmission of the virus. In addition, women with less advanced HIV disease reported higher levels of sexual activity. Of the 104 women with a CD4 cell count above 350 cells/mm^3 and no symptoms of HIV, 60 (58%) reported sexual intercourse in the last week. This compared to 35% of women with a CD4 cell count between 200 – 350 cells/mm^3 and 26% of women with a CD4 cell count below 200 cells/mm^3. Condom use was low in all groups of women, ranging from a low of 6% for women with the highest CD4 cell counts and no symptoms to a high of 13% for women with a CD4 cell count

below200cells/mm^3. Although the most important factor in deciding when to initiate antiretroviral therapy is the potential risk vs. benefit to the additional benefit for the individual patient, additional consideration of the public health implications of treatment guidelines is important because it may help maximize the benefit for communities where antiretroviral are being introduced, write the investigators. Although their study found that women with the most advanced HIV disease shed the most HIV in their genitals, they stress that significant levels of shedding and high-risk sexual activity was observed in patients in the best health. Earlier initiation of anti-HIV strategy together with risk reduction counseling and the promotion of condoms may help reduce HIV transmission in this population, the investigators conclude. *(McClelland*,*2006*).

10. Case of Kenya

From the other side a study involving HIV-Positive women in Kenya suggests that providing antiretroviral therapy to individuals with advanced HIV disease may only have a limited impact on HIV transmission. The study, which is published in the April 15th edition of the Journal of Acquired Immune Deficiency Syndromes, found that women with a CD4 cell count below 200 cells/mm^3 had the highest genital shedding of HIV, but were the least sexually active. Women without any symptoms of HIV and a CD4 cell count above 350 cells/mm^3 shed a smaller, but still significant amount of HIV in their genitals, and were the most sexually active. It has been suggested that the widespread provision of antiretroviral therapy could decrease the number of new HIV infections by lowering the viral load and infectiousness of people with HIV. Current World Health Organization (WHO) guidelines recommend that antiretroviral therapy should be provided to individuals who are ill because of HIV or if they have a CD4 cell count below 200 cells/mm^3. However, the impact of antiretroviral therapy on HIV risk at a population level could be limited if infectivity or high-risk sexual behavior is high in individuals who do not qualify for treatment.

Investigators from the United States and Kenya therefore looked at the genital shedding of HIV and sexual risk behavior of 650 HIV-positive women who had never taken antiretroviral therapy in Mombassa. They were stratified according to WHO criteria for initiating anti-HIV therapy.

The women had a median age of 29, 50% were married and 6% reported selling sex. A total of 65% of women had symptomatic HIV infection. Genital shedding of HIV increased in a stepwise fashion with each advancement in the severity of HIV disease. Women with a CD4 cell count above 200 cells/mm^3 had the highest concentration of HIV in their genitals (3.8 $\log_{10}$ copies/swab) and women with a CD4 cell count above 350 copies/ml and no symptoms of HIV the lowest (2.4 $\log_{10}$ copies/swab). This difference was highly statistically significant ($p < 0.001$). Nevertheless, the concentrations of HIV observed in the genital swabs obtained from women with the least advanced HIV disease were still high and implied a significant risk of onward transmission of the virus. In addition, women with less advanced HIV disease reported higher levels of sexual activity. Of the 104 women with a CD4 cell count above 350 cells/mm^3 and no symptoms of HIV, 60 (58%) reported sexual intercourse in the last week. This compared to 35% of women with a CD4 cell count between 200 – 350 cells/mm^3 and 26% of women with a CD4 cell count below 200 cells/mm^3. Condom use was low in all groups of women, ranging from a low of 6% for women with the highest CD4 cell counts and no symptoms to a high of 13% for women with a CD4 cell count below 200 cells/mm^3. Although the most important factor in deciding when to initiate antiretroviral therapy is the potential risk vs. benefit to the additional benefit for the individual patient, additional consideration of the public health implications of treatment guidelines is important because it may help maximize the benefit for communities where antiretroviral are being introduced. The study found that women with the most advanced HIV disease shed the most HIV in their genitals, they stress that significant levels of shedding and high-risk sexual activity was observed in patients in the best health. Earlier initiation of anti-HIV strategy together with risk reduction counseling and the promotion of condoms may help reduce HIV transmission in this population.

11. HIV/AIDS in Sudan

The first case of HIV/AIDS was diagnosed in Sudan in 1986. Since then the prevalence of the disease is on the increase. In 1987, Sudan National AIDS Control Program (SNAP) was established. During the period 1987 – 1998, two short term plans and two medium term plans were formulated. For the period 2003 – 2007, a more comprehensive strategic plan will be prepared. However, the data and information available about HIV/AIDS prevalence and the knowledge, attitude and behavior of the population with regards to HIV/AIDS was not sufficient. SNAP then decided to undertake a national survey that could assess the magnitude of HIV/AIDS in the country and provide information about the knowledge, attitude and the behavior of the different sectors of the community. The survey was also intended to assess the commitment of the different Government Ministries, National and International Organizations and other civil society organizations and whether they have specific plans and policies or activities addressing the HIV/AIDS problem. The main objective is to formulate a country wide national strategic plan to control HIV/AIDS based on the results of the behavioral, epidemiological and response analysis. As a collateral outcome of the survey, it is expected that an opportunity will be provided for building the capacity of the personnel in the HIV/AIDS control units in the states.(National survey, 2002).

A Task Force (TF) for strategic planning process was nominated and assigned to undertake the responsibility for planning and implementing the national survey and to submit a report. The TF was also to formulate, in collaboration with all concerned partners, a National Strategic Plan for HIV/AIDS prevention and control. For the assessment of HIV/AIDS prevalence, and to determine the knowledge, attitude, practice and behavior affecting AIDS transmission a stratified sampling method was adopted. The samples include eleven states out of sixteen states in the north and three states from the south. The targeted groups included women attending the antenatal clinics and this group is assumed to represent the general population. In addition, secondary data from blood donated and screened for HIV was also included to represent prevalence among the general population. The high risk groups included prostitutes, truck drivers, tea sellers, prisoners, university students, soldiers, street children, individuals in the displaced camps, patients attending the sexually transmitted disease clinics and tuberculosis patients in addition a group of refugees was also included. For the assessment of the HIV/AIDS prevalence, blood was collected from the target groups included for the survey and the suspected AIDS cases during the period when the survey was carried out and tested for HIV at the National Health Laboratory in Khartoum State.A questionnaire was designed and tested for the collection of the behavioral information from all those included in the survey. Structured close-ended questions were administered through direct interviews by trained interviewers. Focus group discussions were also carried out to provide more behavioral information.

The response analysis survey included representatives of different Government Ministries at state & federal levels, NGOs and wide spectrum of different civil society organizations. Meetings were held with concerned groups, ministers and directors of different departments.

The results of the behavioral survey showed that 31.9% of the respondents were males and 68.1% were females, 62% were married and 40% never married. The majority of the respondents (78.6%) heard about AIDS, the main channel of hearing about AIDS were the media i.e. Radio (40%), TV (34%), however, only 14% heard about AIDS from health workers. Less than 25% of the respondents know about the symptoms of AIDS, while less than 16.5% knew about the signs of the disease. Only about 20% of the respondents recognized HIV virus as the cause of AIDS. When asked about the mode of transmission 53.2% mentioned sexual intercourse, 29.9% blood transfusion, 26.7% skin penetration and 13.4% transmission from the mother to child. Less than 10% of the respondents mentioned use of condom as a mean of preventing AIDS transmission, practicing sex with one person was 27.3% and abandoning illegal sex was 41%. Transmission of AIDS through mosquito bite was mentioned by 27.6% of the respondents, while 24.4% believe that the transmission could happen by eating with people living with HIV/AIDS. More than one-third of the respondents think that teacher and school children should not be allowed to go to school. While 56.6% of the respondents would not buy food from food sellers with AIDS, 44.3% would not eat with individual infected with HIV/AIDS. More than two-third of the respondents never heard/saw about condom compared to only 21.1% who heard about it and 13.9% who heard/saw condom only .3.7% of the respondents ever used male condom. 31% of them would not nurse people living with HIV/AIDS and 30% would not allow teachers and school children to attend schools. More than two - third of the refugees 66.9% never heard or saw a condom, 27.5% heard about it, and only 4.7% heard/saw it. There were only 3.5% of the refugees who used condom. Only 2.9% of the refugees practiced sex outside marital relations. Urinary tract infection was prevalent among 19.5% of the refugees, secretion discharge among 6.0%, itching 4.1% and sore organ among 1.4%. Only 10.5% of the refugees had educational materials about HIV/AIDS. Radio was the main source of information, 73.5% of them mentioned radio and 48.7% from religious channels.

A total of 7385 blood samples were tested and out of those 118 were positive and the overall prevalence is 1.6%. Prevalence among women attending the antenatal clinics was 1.0%. Prevalence among the refugees was 4.0% while prevalence of the other high risk groups tested was 4.4% among prostitutes, 1.6% among TB patients, 2.5% among tea sellers, 25% of the suspected AIDS patients. A total of 470 blood samples were tested from refugees, and out of those 20 were positive HIV/AIDS, giving a prevalence of 4.3%. Only 9.1% of the respondents were practicing sex outside marital relations. More than half of the respondents 58.7% did not complain from health problems, however, 31.9% of them suffered from urinary tract infection, 19.0% from discharge and 9.3% from itching. The majority of the respondents75% did not have educational materials on AIDS. Out of 146 2.4% respondents who did HIV/AIDS testing, 22(0.4%) of them had the test as their wish. The majority of the refugees respondents 94.3% were females and 5.5% were males. The percentage of married among refugees was 85% and those who were never married `were 7.8%. About 20% of the refugees do not know how to read and write and the rest were literates. The results revealed that about 65% of the refugees heard about AIDS, while 35% never heard about it. About 47.4% of the refugees heard about AIDS through radio, 29.9% through health workers, 21.2% relatives and only 4.7% through newspapers. About 15% of the refugees know about the symptoms of the disease and about 5.7% the signs of the disease and the virus as its cause.

The modes of transmission were mentioned as follows: 32.4% sexual intercourse, 56.5% blood transfusion, 17.9% skin penetration and only 7.6% mentioned transmission from mother to child. What concerns prevention, 26.7% of them mentioned avoidance of illegal sex, 20.5% skin penetration, 14% practice sex with one person, only 7.8% mentioned use of condom. As regards wrong believes 27.7% of the refugees believe that HIV is transmitted by mosquito. The percentages of the refugees who would not buy from food seller or eat with a person infected with HIV/AIDS were 52.5% and 44.8% respectively (National survey 2002). The results of the response analysis revealed that invariably there were no strategic plans for controlling HIV/AIDS at government both at federal and state levels, civil society organizations and NGOs. However, there is strong political commitment expressed by all decision makers in the government departments, civil society organizations and NGOs. Though this commitment was not translated in terms of strategies, action plans and programs implementation to combat HIV/AIDS. The fragmented and sporadic activities related to HIV/AIDS are carried out as a part of the components of reproductive health activities. Initiatives to combat HIV/AIDS are usually proposed by external funding agencies i.e. WHO, UNAIDS, UNFPA, UNICEF and NGOs. There is lack of voluntary or involuntary testing of blood for HIV/AIDS, except in some blood banks and the National Health Laboratory for those who seek work in the Gulf. Epidemiological surveys were carried out only in Khartoum state, among the military and NGOs. The main activities conducted by stakeholders were confined to training workshop on how to control HIV/IDS and sending messages through media i.e. Radio, TV, newspapers and public lectures. Coordination is weak among stakeholders; however, it is better among NGOs. The main sources of finance to combat HIV/AIDS are the federal government, UN

Agencies and NGOs. The major lost opportunities are the lack of strategic plan, effective use of media and youth gatherings, blood safety and counseling of blood donors, and the use of religious forum. The overall attitude towards people living with HIV/AIDS is favorable; however, no clear policies were set by any of the stake holders in case of having an employee infected with HIV/AIDS.
It is concluded that the prevalence of HIV/AIDS according to WHO definition, Sudan is moving into the stage of a concentrated epidemic i.e. HIV prevalence is at 1% among Anti-Natal Care (ANC), and is close to 5% among refugees, prostitutes (SNAP, 2002).

12. HIV/AIDS in Khartoum State

A study was conducted by Agency for Co-Operation Research and Development (ACORD)and Ministry of Health- Khartoum State to analyze the situation of female Prostitutes in the context of Khartoum, the study attempted to draw a general profile on the socio economic, legal context in Khartoum in general, generating relevant information on illegal Prostitutes (ISW's) socio-cultural, economic conditions, and their perceptions and understanding about HIV/ AIDS. While sex work is an old phenomenon in Sudan, it is also illegal and, therefore, clandestine. This makes it difficult to determine the true extent of the sex work industry, although it is acknowledged to be substantial and has apparently been increasing in recent years. This increase has been attributed to various factors, including changes in political, civil and socio-economic conditions and increased population mobility. The study covered 65 individuals, who were approached through semi-structured interviews. The study used various qualitative methods, including literature review, oral testimonies and case studies, and interviews with female Prostitutes, mediators and key informants from the Community Security Department. Ages of more than half of the respondents was found to be less than 28 years. Mostly, the female Prostitutes (FSWs) started their profession when their ages under 22 years. Mostly FSWs are (single, divorced, or widows), most of them have economically dependent family members. Prostitutes in Khartoum take place in many kinds of ways and involves many different types of people, many of whom are in no way stereotypical. Women from various backgrounds and classes sell sex and they do so for many economic, social and cultural reasons. Most of those interviewed sell sex as a full-time occupation, while others sell sex only occasionally. Some female Prostitutes are willing to be whores and the majority claim that they have been pressured to sell sex. The major reason for those pressured to sell sex is poverty and limited economic opportunities, often because their education and marketable skills are limited.
Gender inequalities contribute to women entering into sex work. Although their rights and economic independence are often limited (in some instances, severely), women frequently bear the major burden of family obligations. Sex work is diversified. Diversification is observed in terms of work places, motivations of Prostitutes in terms of value and form (cash vs. goods). Stakeholders involved in the business, and professionalism. Most of Prostitutes tend to combine sex work with other sources of income. Professional sex worker who depend only on sex work for their livelihoods proved to be small percentage apart from students. Other sources of income are either salaries or income generating activities. Main income generating activity performed by Prostitutes is alcohol brewing (in the IDPs camps). Other businesses included tea selling, food processing, phone centers operators, handicrafts, as well as salary workers.
In Sudan, and according to a survey conducted by the Sudan National AIDS Program (SNAP) in 2002, the estimated overall HIV/AIDS prevalence among the surveyed samples at 1.6%. However, prevalence rate among Prostitutes, refugees and women tea sellers have recorded the highest rates, which were respectively 4.4%, 4% and 2.5%. Vulnerability to violence among female Prostitutes varies. For instance displaced Prostitutes and the less educated are particularly vulnerable to violence if compared with others. Forms of violence against them from their perception include physical violence and harassment by community security and clients who might some times refuse to pay their service incentives. However FSWs did not identified the psychological violence in considering

them as women who do not adhere to sexual and other behavioral norms, and excluded from mainstream society. Yet awareness on HIV modes of transmission proved to be very high among female Prostitutes in Khartoum, female Prostitutes are vulnerable for HIV and other sexual transmitted infections (STIs) because they are highly ignorant and have misconceptions to prevent themselves from these infections. Vulnerability to HIV infection can be illustrated as follows:

More than half of the respondents do not use any measures to protect themselves from getting STIs. However the main preventive method used by those revealed their adoption of measures, is reducing number of clients per day. Findings however have shown that most prevalent number of clients received per day is two, which is still high number. Reducing number of clients is followed by condom use. Ratio of female Prostitutes who use condom proved to be very low. Only few respondents negotiate condom use with clients. Most of those negotiate the use of condom perceive condom as a contraceptive measure rather than a tool for STIs prevention. *(ACORD and MOH, 2006)*. A burning question for almost everybody is: why do people continue to take risks even when they sometimes know the consequences? The horrific reality of HIV/AIDS in today's world does not give us the luxury of procuring actual or specific answers to this and other burning questions (including finding a medical cure). Realistically, then, the efforts exerted thus far concentrate on measures of preventing the pandemic. While policy makers, activists and those concerned about the HIV/AIDS pandemic need to discover whether people really understand the risks they face, whether they personally feel they are in danger, and how they calculate the balance of short-term gains against long-term hazards, it is imperative to stress the sad reality that HIV/AIDS is threatening the whole world. While some of the stricken nations managed to slow down the rates of infections (e.g. South Africa and Uganda), others (the Sudan) are considered high risk areas. The overall rate of the infection in Sudan in 2002 was found to be 1.6 %. But the result of this report, which is based on a study on one of the potential categories of the population in the Sudan, reveals that the rate might in fact be much higher than the official figure.

A report about Khartoum State prisoners' survey, it basically aims at aiding interventions that seek to raise the awareness of potential persons that are likely to contract the virus more than others. It is a part of a project undertaken by ACORD Sudan that seeks to illicit prevalence, knowledge, attitudes, behaviors and risk factors among prisoners, tea sellers, truckers and Prostitutes. Based on results from different parts of the world, these categories are believed to be potentially susceptible to contract the virus more than other categories in their societies. The report is based on a survey conducted in 5 prisons and 2 penitentiaries in Khartoum State, between November 2005 and January 2006. Both quantitative and qualitative methods were adopted in the study. A sample of 316 individuals was selected, from a population of 3766 prisoners. The following is a summary of the prisoners' survey results:

From a prevalence rate of 2% in 2002, the results of this survey indicated that the percentage of HIV/AIDS prevalence among prisoners jumped to 8.63% in 2006. This is an alarming increase that begs for immediate intervention. The prevalence rate among female prisoners is even more alarming. Out of the 48 blood samples, 13 (27.08 %) were positive. This supports the widely documented fact that females are more suspectable to contract HIV/AIDS than men. It also indicates that a sort of affirmative action is needed for women. This high prevalence rate among women prisoners should not come as a surprise; in fact it is consonant with the disadvantageous lot of women in the country.

Worrying is also the prisoner's knowledge about the different aspects related to HIV/AIDS. While the overwhelming majority of the prisoners heard about HIV/AIDS, from a variety of sources, their knowledge about the virus that causes the infection is little. Equally, the knowledge of prisoners in this study about both the virus that causes the infection and the symptoms are little. They cited a variety of sources from which they get information about HIV/AIDS. These sources include Radio, TV, newspapers, relatives, and friends, respectively. The knowledge about ways or means of protection is no better than knowledge about the virus and symptoms. Many prisoners would think that avoiding illegal sexual relations and skin piercing are the most effective ways of protection.

Still, some prisoners reported that mosquitoes bite causes infection. Inadequate knowledge results in stigmatizing those people living with HIV/AIDS. It also results in wrong or distorted attitudes. In this regard, sharing food with positive persons, buying stuff from them, and taking care of HIV/AIDS sick persons are some of the things that many prisoners would not do. This result emphasizes that still there is a long way to go with regard to fighting stigma.

The gap of the respondents' knowledge about condoms raises concerns about the effectiveness of the different educational and awareness raising campaigns carried out thus far. More than one third of the respondents reported that they have not heard about condoms. This is most likely the result of the fact that media outlets shy away from popularizing condoms. This tendency is not uncommon in a society that considers sex and sex-related matters as cultural taboos. But it must be stressed that individuals are also sensitive to talk about sex and answer sex-related questions. Cultural inhibitions are behind people's reluctance to indulge in a free conversation about things that are culturally dubbed taboos. It is interesting to note that 40.5% of prisoners in the study reported that pharmacies are the main sources of getting condoms. While relatives and families have been cited among the sources of knowledge about condom, none of the prisoners reported that families or relatives provide condoms. The use of condom as a preventive measure is also worrying so much so that only 15% of the prisoners said they used it before. More interesting is that the condom is used to avoid pregnancy. The prisoners' behaviors are found to be reckless. They do not seek proper medical assistance in cases of sexually transmitted infections. Furthermore, some use to have sex while they are having STIs. Worse still, those who have sex do not use condoms in the presence of sexually transmitted infections. Related to this, it was found that sex exists in prisons, and a little less than one third of the prisoners in this study explained that homosexuality is prevalent in prisons.

Apart from lack of knowledge, bad attitudes, and risk behaviors, there are other risk factors that contribute to the spread of the virus and other sexually transmitted infections. Prisons' setting is one of the factors that facilitate the spread of infections. Within prisons, the following factors were identified as contributing to the spread of the pandemic: overcrowding, unhealthy environment, poor medical facilities, unprotected sex, inadequate awareness-raising and educational programs within these facilities, and bullying (in the penitentiaries). These risk factors are present in all the prisons in Khartoum State. Prisons' authorities in all these facilities are conscious about the challenges of sexually transmitted infections and HIV/AIDS. They are also conscious about the importance of awareness-raising campaigns and educational programs, and are receptive to collaborate with NGOs and civil society organizations in the fight against HIV/AIDS (Manzoul, 2006)

In addition to that ,a report represents partial mapping research carried out by ACORD in coordination with the Ministry of Health under the global HIV/AIDS combat program of UNAIDS. This part of research was carried out in an attempt for social mapping of the truck drivers in Khartoum state with regard to prevalence, knowledge, attitudes and practices towards HIV/AIDS. Information generated is expected to feed into aids combat programs and interventions. The report covers qualitative information as well as quantitative information collected from truck drivers and assistants at the main five traffic entrance locations around Khartoum. Qualitative information was collected using an adapted family health questionnaire that was controlled by trained staff from Ministry of Health.

Total of 418 truck drivers and assistants (69.4% truck drivers, 29.4 % assistants and 1.2% other) were interviewed while 387 truck drivers and assistants voluntarily tested under supervision of Ministry of Heath for HIV prevalence.97.6% of interviewed are Sudanese drivers and assistants with 95% of them at the age group of 15-49 years old. 84% of the interviewed attended formal or informal education and 76.8% have the ability of comfortably reading and writing. Majority spent more than three years 69.6% as drivers and assistants. Majority spend 2-4 days on the highways 48.3% while only 24.4% spend more than a whole week traveling. Few drivers were found not practicing stop at roadside stations 2.2% others stop for periods ranging from few minutes to 24 hours. Few share serene or 3.3% or blades 10.8%. Most heard about HIV/AIDS 96.9% mainly through national radio 20.3%. 61.5% do not know HIV symptoms and 70.6% don't know the type of microbe that causes HIV infection. 54.2% think that HIV positive should look sick, 25.6% think he shouldn't necessarily look sick while 26.1% don't know what exactly happens upon getting HIV infection. About modes of transmission 90% mentioned different modes including mother to child, blood transfer, skin piercing or a combination of these.

On how to protect oneself from getting HIV infection majority 81.8% indicated avoidance of multi-sexual relationships 39%, avoiding both multi-sexual relationship and skin piercing 36.6% or having one partner free of HIV infection 6.2% as means of protection against HIV infection. 35.9% believe that mosquito bites can transfer HIV infection while 64.4% indicated that they are sure that infected mother transfers HIV infection to the child during pregnancy 21.5% or both pregnancy and delivery times 4.1%. Whether they would or wouldn't buy from food from an HIV positive person 74.2% said they wouldn't while 21.1% confirmed that they would buy from him/her. 64.4% said they wouldn't let HIV positive teacher teach at school and 63.6% said they wouldn't let an HIV positive student attend school. 50.2% mentioned that they would share food with a relative who is HIV positive and 62.7% said they would take of him. Only 14.8% have tested for HIV and non-mentioned that results were positive.25% have never heard or saw a condom, 47.1% heard about the condom but never seen one while 27.8% have heard about and seen a condom in their life. Only 8.9% have heard about, seen and used a condom. 69.4% of the drivers and assistants think that their fellow drivers and assistants involve in sexual practices during trips mainly 48.6% heterosexual or both 16.5% heterosexual and homosexual. 55.3% reported that they do not practice any sexual relation outside marriage relationship or before marriage for those single or divorced. 24.2% reported that they practice sex with three or more sex partners. 9.9% reported infection with STDs in the past or currently and 94.5% have taken drugs or currently taking drugs for healing.95.5% of possess radio, T.V. or both while national radio is the most listened to43.8%. 69.4% have never been exposed to HIV/AIDS educational material. 39% of the respondents think that radio and T.V. are the most appropriate for disseminating outreaching HIV related educational material.

A recent research about tea sellers conducted in Khartoum state, the objectives of the study were to understand and map the socio-economic conditions of women tea sellers in Khartoum, assess the risks surrounding their profession, especially that of infection of HIV/AIDS, and explore the different types of violence vis-à-vis vulnerability and risks of infection. The study relied on qualitative methods of data collection and on available secondary information. The researches available to date are quantitative and have its methodological limitations, and in some cases data were collected in a coercive and insensitive manner which may affect the accuracy and reliability of the information. Triangulation to identify and minimize biases is hardy used.

Structural poverty in the country is directly related to the individual decision of any given tea seller to take-up the profession and whether or not to get involved in risky behavior. Poverty represented by the Human Poverty Index (HPI) does not exhibit Violence Against Women (VAW) as a by-product of economic and social poverty. We suggest that there is a strong and direct link between the spread of HIV/AIDS and VAW, and if VAW is contained the risk of infection could be greatly reduced. Such a relationship is not easy to advocate given the power structure in Sudanese societies yet further research and advocacy are of high importance to explore the link and build alliances around it. Tea selling, although it is considered as a negative phenomenon by the authorities, the media and public at large, for most of the women interviewed it is the only source of income for families who are female headed. The majority of tea sellers have between 3 – 9 children and no source of income other than selling tea in a harsh antagonizing environment. Like any other profession, poverty creates vulnerability and hence increased risks of Gender-Based Violence and consequently a higher risk of infection.

There is no evidence that tea sellers are involved in sexual activities with their customers, it is rather the highly competitive business environment that makes it inevitable for any woman tea-seller to take care of her appearance in the way she does and flatter her customers in order to keep them and also to attract new ones. For those who combine the tea selling and sex selling there is a need for all of us to look beneath the surface to be able to distinguish the cycle of violence, vulnerability and risks. Although the level of awareness about HIV/AIDS is relatively high among women tea sellers, yet their awareness is mystified by misconception and high rate of stigma against those who are infected. The information needs to be taken to where people work and live, in a face-to-face communication. To date the official campaigns remain to be ambiguous and superficial.

A study was established to assess the knowledge, attitudes, practices, believes and behavior among secondary school students, that influence the spread and control of HIV/AIDS in Khartoum State (Omyma, 2006). Recent surveys have stressed that there was an increasing prevalence of HIV/AID in Sudan amongst the adolescent population. Sexual contact was found to be the major route of transmission. As adolescence is a period of sex awakening and increased sexual activity, and adolescents are more exposed to other risks of HIV transmission, they should be targeted regarding HIV/AIDS control. However, first policy makers and program mangers need to know about the current situation of their knowledge, behaviors, and practice that influence the spread of HIV so as to design and implement the prevention programs that fit the population and its behaviors.

A descriptive cross-sectional study was conducted among secondary school students in Khartoum State. The data was collected through a structured, pre-tested; close-ended questions were designed to expose depth of knowledge and attitude of respondents concerning HIV/AIDS. The results signify, the students had a deficient knowledge about the disease. They were not sure of the cause, nature, or modes of transmission and prevention, except that illicit sexual activity should be avoided. Their attitude and behavior were consequently inadequate. Their main source of information was the media, there is need to provide students with correct, detailed, and broad-based information on reproductive health as part of the school curriculum to help them acquire adequate knowledge and develop appropriate attitude and behavior towards HIV/AIDS and other STDs.(Abdelgadir, 2006).

13. The Statistical Models

The statistical methods considered the analysis of relationships between measurements made on groups of subjects or objects. For example, the measurements might be the heights or weights and the ages of boys and girls, or the yield of plants under various growing conditions. We use the terms response, outcome or dependent variable for measurements that are free to vary in response to other variables called explanatory variables or predictor variables or independent variables although this last term can sometimes be misleading. Responses are regarded as random variables. Explanatory variables are usually treated as though they are non- random measurements or observations; for example, they may be fixed by the experimental design. Responses and explanatory variables are measured on one of the following scales:

Nominal: why is classified as red, green, blue, yes, no, do not know, not applicable. In particular, for binary, dichotomous or binomial variables there are only two categories: male and female; dead and alive; smooth leaves and serrated leaves. If there are more than two categories the variable is called polychotomous, polychromous or multinomial.

Ordinal: the classifications in which there is the same as natural order or ranking between the categories: e.g., young, middle age, old; diastolic blood pressures grouped as ≤ 70, 71-90, 91-110, 111-130, ≥ 131 mm Hg.

Continuous: which are measurements where observations may, at last in theory, fall any where on continuum: e.g., weight, length, or time. This scale includes both interval scale and ratio scale measurements the latter have a well defined zero. A particular example of a continuous measurement is the time until specific event occurs, such as the failure on electronic component; the length of time from a known starting point is called the failure time.

Nominal and ordinal variables are sometimes called categorical or discrete variables and the numbers of observations, count or frequencies in each category are usually recorded. For continuous variable the individual measurement are recoded. The term quantitative is often used for a variable measured on a continuous scale and the term qualitative for nominal and sometimes for ordinal measurements. A qualitative, explanatory variable is called a factor and its categories are called the levels for the factor. A qualitative explanatory variable is sometimes called covariate. Methods of statistical analysis depend on the measurement scales of the response and explanatory variables.

The main idea of statistical modeling involves four steps:

1. Specifying models in two parts: equation linking the response and explanatory variables, and the probability distribution of response variable.
2. Estimating parameters used in the models.
3. Checking how well the models fit the actual data.
4. Making inferences; for example, calculating confidence intervals and testing hypotheses about the parameters.

Generalized linear model consider the outcome of the variables are measured on a binary scale. For example, the responses may be alive or dead; present or absent Success and failure are used as generic terms of the two categories.

First, we define the binary random variable

1 if the outcome is a success, z = 0 if the outcome is a faiture

With probabilities $P_r(z=1)=\Pi$ and $P_r(z=0)=1-\Pi$

If there are n such random variables $z_1, z_2 ....., z_n$ which are independent with Pr $(z_j=1)=\Pi_j$ then their joint probability is:

$$\prod_{j=1}^{n}\Pi_j^{z_j}\left(1-\Pi_j\right)^{1-z_j}=\exp\left[\sum_{j=1}^{n}z_j\log\left(\frac{\Pi_j}{1-\Pi_j}\right)+\sum_{j=1}^{n}\log\left(1-\Pi_j\right)\right] \quad (1)$$

Which is a member of the exponential family, for the case where the $\Pi_j's$ are all equal, we can define:

$$Y=\sum_{j=1}^{n}z_j \quad (2)$$

So that Y is the number of successes in n trials.

The random variable Y has the binomial distribution (n,Π)

$$P_r(Y=y)=\binom{n}{y}\Pi^y\left(1-\Pi\right)^{n-y} \quad y=0,1,.....n \ (3)$$

We consider the general case of N independent variables Y_1 , Y_2 , … , Y_N corresponding to the numbers of successes in N different sub- groups or strata.

If Y = binomial (n_i,Π_i) the log- likelihood function is:

$$L(y_1,y_2,y_3,y_N;\Pi_1,\Pi_2,....,\Pi_N)=\left[\sum_{i=1}^{N}y_i\log\left(\frac{\Pi_i}{1-\Pi_i}\right)+n_i\log(1-\Pi)+\log\binom{n_i}{y_i}\right] \quad (4)$$

Frequencies for N binomial distribution

Subgroups, 1, 2 ………… N

Success y_1 y_2 .... y_N

Failures n_1-y_1 n_2-y_2 …….. n_N-y_N

Total n_1 n_2 .... n_N

If we want to describe the proportion of success, $P_i = Y_i / n_i$, in each subgroup in terms of factor levels and other explanatory variables which characterized the sub group. As $E(Y_i)=n_i\Pi_i$ and so $E(p_i)=\Pi_i$, we modeled the probabilities Π_i as:

$$g(\Pi_i)=X_i^T\beta \quad (5)$$

Where x_i is a vector of explanatory variables (dummy variables for factor levels and measured values for covariates), ß is a vector of parameters and g is a link function.
The simplest case is the linear model:

$$\Pi = X^T \beta \qquad (6)$$

This is used in some practical application but it has the disadvantage that although Π is a probability, the fitted values may be less than zero or greater than one.
To ensure that Π is restricted to the interval [0,1] it is often modeled using a cumulative probability distribution.

$$\Pi = \int_{-\infty}^{t} f(s)ds \qquad (7)$$

Where $f(s) \geq 0$ and $\int_{-\infty}^{\infty} f(s)ds = 1$. The probability density function $f(s)$ is called the tolerance distribution. Historically, one of the uses of regression –like models for binomial data was for bioassay result. Responses were the proportions or percentages of successes; for example, the proportion of experimental animals killed by various dose levels of a toxic substance. Such data are sometimes called quintal responses. The aim is to describe the probability of success, Π, as a function of the dose, x; for example,

$$g(\Pi) = \beta_1 + \beta_2 x \qquad (8)$$

If the tolerance distribution f (s) is the uniform distribution on the interval [c_1 , c_2]
$f(s)$ and Π of the uniform distribution.

$$f(s) = \frac{1}{c_2 - c_1} \text{ if } c_1 < s < c_2 \qquad (9)$$

0 otherwise

This equation has the form $\Pi = \beta_1 + \beta_2 x$ (10)

$$\beta_1 = \frac{-c_1}{c_2 - c_1} \text{ and } \beta_2 = \frac{1}{c_2 - c_1} \qquad (11)$$

This linear model is equivalent to using the identity function g and imposing conditions on x, β_1 and β_2 corresponding to $c_1 < s < c_2$.These extra conditions mean that the standard methods for estimating β_1 and β_2 for generalized linear models cannot be directly applied. In practice, this model is not widely used. One of the original models used for bioassay data is called the probit model. The normal distribution is used as the tolerance distribution.

$$\Pi = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left[-\frac{1}{2}\left(\frac{s-\mu}{\sigma}\right)^2\right] ds \qquad (13)$$

$$= \Phi \left(\frac{x-\mu}{\sigma}\right)$$

Where Φ denotes the cumulative probability function for the standard Normal distribution N (0,1). Thus

$$\Phi^{-1}(\Pi) = \beta_1 + \beta_2 x \quad (4.14)$$

Where $\beta_1 = \frac{-\mu}{\sigma}$ and $\beta_2 = \frac{1}{\sigma}$ and the link function g is the inverse cumulative normal portability function Φ^{-1} probity models are used in. Another model that gives numerically results very much like those from the probit model, but which computationally is some what easier, is the logistic or logit model. The tolerance distribution is

$$f(s) = \frac{\beta_2 \exp(\beta_1 + \beta_2 s)}{[1 + \exp(\beta_1 + \beta_2 s)]^2} \qquad (15)$$

So :

$$\Pi = \int_{-\infty}^{x} f(s) ds = \frac{\exp(\beta_1 + \beta_2 x)}{1 + \exp(\beta_1 + \beta_2 x)]} \qquad (16)$$

This gives the link function:

$$\log\left(\frac{\Pi}{1-\Pi}\right) = \beta_1 + \beta_2 x \qquad (17)$$

The term $\log\left(\frac{\Pi}{1-\Pi}\right)$ is sometime called the logit function and it is a natural interpretation as the logarithm of odds. The logistic model is widely used for binomial data and is implanted in many statistical programs. The shapes of the functions f (s) and $\Pi(x)$ are similar to those for the probit model except in the tails of the distributions (Cox and Snell, 1989).

Several other models like extreme value distribution:

$$f(x) = \beta_2 \exp[(\beta_1 + \beta_2 s) - \exp(\beta_1 + \beta_2)] \qquad (18)$$

is used as the tolerance distribution then:

$$\Pi = 1 - \exp[-\exp(\beta_1 + \beta_2 x)] \qquad (19)$$

And so $\log[-\log(1-\Pi)] = \beta_1 + \beta_2 x$ this link , $\log[-\log(1-\Pi)]$ is called the complementary Log function. The model is similar to the logistic and probit models for values of Π near 0.5 .These models are illustrated in the following example.

Table (1) shows the proportion Pi = Y_i / n_i plotted against dose X_i (actually X_i is the logarithm of the quantity of carbon disulphide. We begin by fitting the logistic model.

$$\Pi_i = \frac{\exp(\beta_1 + \beta_2 x_i)}{1 + \exp(\beta_1 + \beta_2 x_i)} \qquad (20)$$

So $\log\left(\frac{\Pi}{1-\Pi}\right) = \beta_1 + \beta_2 x$ (4.21) and: $\log(1-\Pi_i) = -\log[1 + \exp(\beta_1 + \beta_2 x_i)]$ (4.22)

Therefore from the log- likelihood function is

$$l = \sum_{i=1}^{N} \left[y_i(\beta_1 + \beta_2 x_i) - n_i \log[1 + \exp(\beta_1 + \beta_2 x_i] + \log\binom{n_i}{y_i} \right] \qquad (23)$$

And the scores with respect to ß$_1$ and ß$_2$ are:

$$u_1 = \frac{\partial l}{\partial \beta_1} = \sum \left\{ y_i - n_i \left[\frac{\exp(\beta_1 + \beta_2 x_i}{1 + \exp(\beta_1 + \beta_2 x_i)} \right] \right\} = \sum (y_i - n_i \Pi_i) \quad (4.24)$$

$$u_2 = \frac{\partial l}{\partial \beta_2} = \sum \left\{ y_i x_i - n x_{ii} \left[\frac{\exp(\beta_1 + \beta_2 x_i}{1 + \exp(\beta_1 + \beta_2 x_i)} \right] \right\} = \sum x_i (y_i - n_i \Pi_i) \qquad (25)$$

Similarly the information matrix is:

$$J = \begin{matrix} \sum n_i \Pi_i (1-\Pi_i) & \sum n_i x_i \Pi_i (1-\Pi_i) \\ \sum n_i x_{ii} \Pi_i (1-\Pi_i) & \sum n x_i^{2}{}_i \Pi_i (1-\Pi_i) \end{matrix} \qquad (26)$$

Maximum likelihood estimates are obtained by solving the iterative equation:

$$J^{(m-1)}b^{m} = J^{(m-1)}b^{(m-1)} + u^{(m-1)} \quad (27)$$

Where the superscript (m) indicates the mth approximation and b is the vector of estimates. Starting with $b_1^{(0)} = 0$ and $b_2^{(0)} = 0$ successive approximation is done. The estimates converge by the sixth iteration. Also we can obtain the increase in values of the log- likelihood function, omitting the constant term $\log\binom{n_i}{y_i}$ the fitted values are $\hat{y}_i = n_i\hat{\Pi}_i$ calculated at each stage (initially $\hat{\Pi}_i = \frac{1}{2}$ for all i. For the final approximation, the estimated variance and covariance matrix for b, $[J\ (b)^{-1}]$, is shown at the bottom table (4.3) together with the deviance

$$D = 2\sum_{i=1}^{N}\left[y_i \log\left(\frac{y_i}{\hat{y}_i}\right) + (n_i - y_i)\log\left(\frac{n - y_i}{n - \hat{y}_i}\right)\right] \quad (28)$$

The estimates and their standard errors are:

$b_1 = -60.72 \quad SE = \sqrt{26.840} = 5.018$

$b_2 = 34.72 \quad SE = \sqrt{8.481} = 2.91$

If the model is a good fit of the data the deviance should approximately have the distribution $\chi^2(6)$because there are N = 8 covariate pattern {i-e, different values of X_i } and p = 2 parameters. The calculated value of D is almost twice the expected value of 6 and is almost as large as the upper 5%. Point of the $\chi^2(6)$distribution, which is 12.59. This suggests that the model dose not fit particularly well.

Table (2) Fitting a linear Model to the beetle mortality data:

	Initial estimate	First	Approximation Second	Six th
β_1	0	-37.856	-53.853	-60.717
β_2	0	21.337	30.384	34.270
Log Likelihood	333.404	-200.00	-187.274	-186.23

Several alternative models were fitted to the data. The results are shown in table (4.3). Among these models the extreme value model appears to fit the data best.(Annette and J.Dobson,2001).

Table (3) Observation and Fitted Values

Observations	Values		Fitted Values		
Y_1	6	29.5	8.505	4.543	3.458
Y_2	13	30.0	15.366	11.254	9.842
Y_3	18	31.0	24.808	23.058	22.451
Y_4	28	28.0	30.983	32.947	23.898
Y_5	52	31.5	43.362	48.197	50.096
Y_6	53	29.5	46.741	51.705	53.291
Y_7	61	31.0	53.595	58.061	59.222
Y_8	60	30.0	54.734	58.036	58.743

Logistic Regression Model:

Logistic regression provides approaches of handling by modeling the association between the response and a set of explanatory variables. Over the last decade the logistic regression model has become, in many fields the standard method of analysis of data concerned with describing the relationship between a binary variable and one or more explanatory variables. The goal of an analysis using logistic regression is the same as the usual linear regression model where the dependent variable is assumed to be continuous or discrete. The logistic regression model differs from the linear regression model is that the dependent variable in logistic is binary or dichotomous. This difference between logistic and linear regression is reflected both in the choice of a parametric model and in the assumptions. Thus the techniques used in linear regression analysis will motivate our approach to logistic regression. The first difference concerns the nature of the relationship between the dependent and independent variables. In many regression problem the key quantity is the mean value of the dependent variable, given the value of the independent variable. This quantity is called the conditional mean and expressed as $E(Y|x)$ where Y denotes the dependent variable and x denotes a value of the independent variable. In linear regression we assume that this mean may be expressed as an equation linear in x such as:

$$E(Y|x) = \beta_0 + \beta_1 x \qquad (4.29)$$

This expression implies that it is possible for $E(Y|x)$ to take on any values as x ranges between $-\infty$ and $+\infty$. With dichotomous data the conditional mean must be greater than or equal to zero and less than or equal to one, i. e., $0 \le E(Y|x) \le 1$ the change in the $E(Y|x)$ per-unit change in x becomes progressively smaller as the conditional mean gets closer to zero or one. The curve is said to be S-shaped. It resembles a plot of a cumulative distribution of random variable.

Many distribution functions have been proposed for use in the analysis of a dichotomous dependent variable Cox and Snell (1989) discuss some of these. There are two primary reasons for choosing the logistic distribution. First, from a mathematical point of view, it is an extremely flexible and easily used function, and second, it lends itself to a clinically and biologically meaningful interpretation.

In order to simplify notation, we will use the quantity $P(x) = E(Y|x)$ to represent the conditional mean of Y given x when the logistic distribution is used. The specific form of the logistic regression model we will use is as follows:

$$P(x) = \frac{\exp(\beta_0 + \beta_1 x)}{1 + \exp(\beta_0 + \beta_1 x)} = \frac{1}{1 + \{\exp(\beta_0 + \beta_1 x)\}^{-1}} \qquad (4.30)$$

A transformation of $P(x)$ that is central to our study of logistic regression is the logit transformation. This transformation is defined, in terms of $P(x)$, as follows:

$$g(x) = \log it(p(x)) = \ln \frac{p(x)}{1 - p(x)} = \beta_0 + \beta_1 x \qquad (4.31)$$

Where, β_o *and* β_1 are the logistic intercept and coefficient, respectively. The importance of this transformation is that $g(x)$ has many of the desirable properties of a linear regression model. The logit ($P(x)$) is linear in its parameters, may be continuous and may range from $-\infty$ to $+\infty$ depending on the range of x. The second important difference between the linear and logistic regression model concerns the conditional distribution of the dependent variable. In the linear regression model we assume that an observation of the dependent variable may be expressed as $Y = E(Y \mid x) + \varepsilon$ the quantity ε is called the error and expresses an observation's deviation from the conditional mean. The common assumption is that ε follows a normal distribution with mean zero and some variance that is constant across levels of the independent variable. It follows that the conditional distribution of the dependent variable given x will be normal with mean $E(Y \mid x)$, and variance that is constant. This is not the case with dichotomous dependent variables. In this situation we may express the value of the dependent variable given x as $y = P(x) + \varepsilon$. Here the quantity ε may assume one of the two possible values, if $y = 1$ then $\varepsilon = 1 - P(x)$ with probability $P(x)$, and if $y = 0$ then $\varepsilon = -P(x)$ with probability $1 - P(x)$. Thus ε has a distribution with mean zero and variance equal to $P(x)(1 - P(x))$. That is, the conditional distribution of the outcome variable follows a binomial distribution with probability given by the conditional mean, $P(x)$.

In summary, we have seen that in a regression analysis when the outcome variable is dichotomous:

a. The conditional mean of the regression equation must be formulated to be bounded between zero and one. We have stated that the logistic regression model, $P(x)$ given in equation (4.30) satisfies this constraint.

b. The binomial, not the normal, distribution describes the distribution of the errors and will be the statistical distribution upon which the analysis is based.

c. The principles that guide an analysis using linear regression will also guide using logistic regression.

The logistic regression model contains parameters which may be estimated from the data. 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 a dichotomous (outcome) dependent variable coded as zero or one, and x_i is the value of the independent variable for the ith subject. To fit the logistic regression model to a set of data requires that we estimate the values of the unknown parameters β_0 and β_1.

In linear regression the method used most often for estimating unknown parameters is least squares. In that method we choose those values of β_o and β_1 which minimize the sum of squared errors (deviation) of the observed values of Y from the predicted values based upon the model. Under the usual assumptions for linear regression 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 dependent variable the estimators no longer have the same properties.

There are three methods for estimating the unknown parameters. These methods are: maximum likelihood method, non-iterative weighted least squares, and discriminant function analysis.

14. Maximum Likelihood Method

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. This method will provide the foundation for our approach to estimation with the logistic regression model. In a very general sense the method of maximum likelihood yields values of 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 the likelihood function. This function expresses the probability of the observed data as a function of the unknown parameters. 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 observe data. We now describe how to find these values from the logistic regression model.

If Y is coded as zero or one then the expression for P(x) given in equation (4.30) provides the conditional probability that Y is equal to zero 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 is xi. A convenient way to express the contribution to the likelihood function for the pair (x_i, y_i) is through the expression:

$$[p(x_i)]^{y_i}[1-p(x_i)]^{1-y_i} \quad (4.32$$

Since the observations are assumed to be independent, the likelihood function is obtained as the product of the terms given in expression (4.32) as follows:

$$L(\beta)=\prod_{i=1}^{n} P_i^{y_i}(1-P_i)^{1-y_i} \quad (4.33)$$

Where P_i = P (x_i)

The principle of maximum likelihood states that we use as our estimate of β = (β_o, β_1) the value which maximize the expression in equation (4.33). However, it is easier mathematically to work with the log of equation (4.33). This expression, the log likelihood, is defined as:

$$L(\beta)=\ln(l(\beta))=\sum_{i=1}^{n}[y_i \ln P_i+(1-y_i)\ln(1-P_i)] \quad (4.34)$$

to find the value of β that maximizes L(β) we differentiate L(β) with respect to βo and β1 set the resulting expressions equal to zero. These equations are as follows:

$$\sum_{i=1}^{n} [y_i - P_i] = 0 \qquad (4.35)$$

$$\sum_{i=1}^{n} x_i(y_i - P_i) = 0 \qquad (4.36)$$

and called the likelihood equations.

In linear regression the likelihood equations, obtained by differentiating the sum of squared errors function with respect to β, are linear in the unknown parameters and thus are easily solved. For logistic regression the expressions in equations (4.35) and (4.36) are nonlinear in β_o and β_1, and thus require special methods for their solution. These methods are iterative in nature and have been programmed into available logistic regression software.

The value of β given by the solution to equation (4.34) and (4.35) is called the maximum likelihood estimator and will be denotes as $\hat{\beta}$ an interesting consequence of equation (4.35) is that:

$$\sum_{i=1}^{n} y_i = \sum_{i=1}^{n} \hat{P}_i \qquad (4.37)$$

Where $\hat{p}_i$ is the maximum likelihood estimate of pi that is, the sum of the observed values of Y is equal to the sum of the predicted (expected) values. This property will be especially useful in the discussion on assess in the fit of the model.

15. Non-Iterative Weighted Least Squares Method

The method of maximum likelihood describe in previous section is the estimation method used in the logistic regression routines of the major Software Packages. However, the two other methods used for estimating the coefficients of logistic regression are non-iterative weighted least squares and discriminant function analysis. A linear models approach to the analysis of categorical data was proposed by weighted least squares. They demonstrate that the logistic regression model is an example of a very general class of models that can be handled with their methods. The approach suggested by Grizzle et al. (1969), uses only, one iteration in the process.

A major limitation of this method is that we must have an estimate of p(x) which is not zero or one for most values of x. After estimating the coefficients, our first look at the fitted model commonly concerns an assessment of the significance of the variable in the model. This usually involves formulation and testing of a statistical hypothesis to determine whether the independent variables in the model are significantly related to the dependent variable. The method for performing this test is quite general and differs from one type of model to the next only in the specific details. We begin by discussing the general approach for a single independent variable.

The method for testing the significance of the coefficient of a variable in logistic regression follow a similar approach to that used in linear regression but uses the likely- hood function for a dichotomous dependent variable. In linear regression the assessment of the significance of the slope coefficient is approached by first forming the analysis of variance table. This table partitions the total sum of square deviations of observations about their mean into tow components: the sum of squared deviations of observations about the regression line SSE, (or error sum-of-square), and the sum of squares of predicted values, based on the regression model, about the mean of the dependent variable SSR, (or regression sum-of-squares). If y_i denotes the observed value and $\hat{y}_i$ denotes the predicted value for the ith individual under the model, then the statistic used to evaluate this comparison is

$$SSE = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \qquad (4.38)$$

Under the model not containing the independent variable the only parameter is β_o, and $\hat{\beta}_o = \bar{y}$ the mean of the response variable. In this case $\hat{y}_i = \bar{y}$ and SSE equal the total variance. When we include the independent variable in the model any increase in SSE will be due to the fact that the slope coefficient for the independent variable is not zero. The change in the value of SSE is the due to the regression source of variability, denoted SSR that is,

$$SSR = \left[\sum_{i=1}^{n}(y_i - \bar{y})^2\right] - \left[\sum_{i=1}^{n}(y_i - \hat{y}_i)^2\right] \quad (4.39)$$

In linear regression, interest focuses on the size of SSR. A large value suggests that the independent variable is important, whereas a small value suggests that the independent variable is not helpful in predicting the response. The guiding principle with logistic regression is the same: compare observed values of the response variable to predicted values obtained from models with and without the independent variable. In logistic regression comparison of observed to predicted values is based on the log likelihood function defined in equation (4.34). To better understand this comparison, it is helpful conceptually to think of an observed value of the response variable as also being a predicted value resulting from a saturated model. A saturated is one that contains as many parameters as there are data points. The comparison of observed to predict values using the likelihood function are based on the following expression:

$$D = -2\ln\left\{\frac{likelihood\ \ of\ \ the\ \ fitted\ \mathrm{mod}\ el}{likelihood\ \ of\ \ saturated\ \mathrm{mod}\ el}\right\} \quad (4.40)$$

The quantity inside the large brackets in the expression above is called the likelihood ratio. Using minus twice its log is necessary to obtain a quantity whose distribution is known and can therefore be used for hypothesis testing purposes. Such a test is called the likelihood ratio test. Using equation (4.34) and (4.39) becomes:

$$D = -2\sum_{i=1}^{n}\left[y_i \ln\left(\frac{\hat{p}_i}{y_i}\right) + (1-y_i)\ln\left(\frac{1-\hat{p}_i}{1-y_i}\right)\right] \quad (4.41)$$

And since the value of the outcome variable is binary, the likelihood of the saturated model is 1. Specifically, it follows from the definition of a saturated model that $\hat{p}_i = y_i$ and the likelihood is $\prod_{i=1}^{n} y_i^{y_i}(1-y_i)^{1-y_i} = 1$, thus it follows from equation (4.39) that the deviance is: $D = -2\ln($ likelihood of the fitted model). The statistic, D in equation (4.40) is called the Deviance by some authors [see, for example Mc Cullagh & Nelder (1983)] and plays a central role in some approaches to assessing goodness of fit. The deviance for logistic regression plays the same role that the residual sum of squares plays in linear regression. In fact, the deviance as shown in equation (4.40), when computed for linear regression, is identically equal to SSE.

For purposes of assessing the significance of an independent variable we compare the value of D with and without the independent variable in the equation. The change in D due to including the independent variable in the model is obtained as follows:

G = D(model without the variable) – D(model with the variable)

This statistic plays the same role in logistic regression as the numerator of the partial F test does in linear regression. Because the likelihood of the saturated model is common to both values of D being differenced to compute G, it can be expressed as:

$$G = -2\ln\left[\frac{likelihood\ without\ the\ \mathrm{var}\,iable}{likelihood\ with\ the\ \mathrm{var}\,iable}\right] \quad (4.41)$$

For the specific case of a single independent variable, it is easy to show that when the variable is not in the model, the maximum likelihood estimate of β_o is $\ln(n_1/n)$ where $n1 = \sum y_i$ and $n = \sum(1 - y_i)$ and the predicted value is constant, n1/n . In this case, the value of G is as follows:

$$G = -2 \ln \left[\frac{\left(\frac{n_1}{n}\right)^{n_1} \left(\frac{n_o}{n}\right)^{n_o}}{\prod_{i=1}^{n} \hat{p}_i^{\,y_i} (1 - \hat{p}_i)^{1-y_i}} \right] \tag{4.42}$$

or

$$G = 2\left\{ \sum_{i=1}^{n} \left[y_i \ln(\hat{p}_i) + (1-y_i)\ln(1-\hat{p}_i) \right] - \left[n_1 \ln(n_1) + n_o \ln(n_o) - n\ln(n) \right] \right\} \tag{4.43}$$

Under the hypothesis that β_1, is equal to zero, the statistic G follows, a chi-square distribution with one degree of freedom.

The calculation of the likelihood and the likelihood ratio test are standard features of any good logistic regression software. This makes it easy to check for the significance of the addition of new terms to the model. In the simple case of a single independent variable, we first fit a model containing only the constant term. We then fit a model containing the independent variable along with the constant. This gives rise to new log likelihood. The likelihood ratio test is obtained by multiplying the difference between these two values by (– 2). Two other similar, statistically equivalent tests have been suggested. These are the Wald test and Score test. The assumptions needed for these tests are the same as those of the likelihood ratio test in equation (4.42).

The Wald test is obtained by comparing the maximum likelihood estimate of the slope parameter, $\hat{\beta}_1$, to an estimate of the standard error as follows:

$$W = \frac{\hat{\beta}_1}{SE(\hat{\beta}_1)} \tag{4.44}$$

The resulting ratio, under the hypothesis that $\beta_1 = 0$, will follow a standard normal distribution.

Haulk & Donner (1977) examined the performance of the Wald test and found that it behaved in an aberrant manner, often failing to reject the null hypothesis when the coefficient was significant. They recommended that the likelihood ratio test be used.

Jennings (1986) has also looked at the adequacy of inferences in logistic regression based on Wald statistic. His conclusions are similar to those of Haulk & Donner. , the likelihood ratio test, G, and the Wald test, W, require the computation of the maximum likelihood estimate for β_1 for a single variable this is not a difficult or costly computational task. However, for large data sets with many variables, the iterative computation needed to obtain the maximum likelihood estimates can be considerable. A test of the significance of a variable which does not require these computations is the Score test. Use of the test is limited by the fact that it can not be obtained easily from some software packages. The Score test is based on the distribution theory of the derivatives of log likelihood. In the univariate case this is based on the conditional distribution of the derivative in equation (4.36), given the derivative in equation (4.35). In this case, we can write down an expression (4.36), computed using $\beta_o = \ln\left(\frac{n_1}{n_o}\right)$ and $\beta_1 = 0$. As noted easier, under these parameter values, $\hat{P} = \frac{n_1}{n} = \bar{y}$ thus, the left hand side of equation (4.36) becomes $\sum x_i\left(y_i - \bar{y}\right)$. It may be shown that the estimated variance is $\bar{y}(1-\bar{y})\sum(x_i - \bar{x})^2$ .The test statistic for Score test (ST) is:

$$ST = \frac{\sum_{i=1}^{n} x_i (y_i - \bar{y})}{\sqrt{\bar{y}(1-\bar{y}) \sum_{i=1}^{n} (x_i - \bar{x})^2}} \quad (4.45)$$

An important adjunct to testing for significance of the model is calculation and interpretation of confidence intervals for parameters of interest. As is the case in linear regression we can obtain these for the slope, intercept and the line , (i.e., the logit) in some settings it may be of interest to provide interval estimates for the fitted values (the predicted probabilities). The basis for construction of the interval estimators is the same statistical theory we used to formulate the tests for significance of the model. In particular, the confidence interval estimators for the slope and intercept are based on their respective Wald tests. The endpoints of a 100(1–α)% confidence interval for the slope coefficient are:

$$\hat{\beta}_1 \pm z_{1-\alpha/2} SE(\hat{\beta}_1) \quad (4.46)$$

And for the intercept they are

$$\hat{\beta}_O \pm z_{1-\alpha/2} SE(\hat{\beta}_0) \quad (4.47)$$

Where $z_{1-\alpha/2}$ is the upper 100 (1 - α/2) % point from the standard normal distribution and SE denote a model-based estimator of the standard error of the respective parameter estimator.

As in the case with any regression model, the constant term provides an estimate of the response in the absence of x unless the independent variable has been centered at some clinically meaningful value. The constant is important when considering point and interval estimators of the logit. The logit is the linear part of the logistic regression model and, as such, is most like the fitted line in a linear regression model. The estimator of the logit is

$$\hat{g}(x) = \hat{\beta}_0 + \hat{\beta}_1 x \quad (4.48)$$

The estimator of the variance of the estimator of the logit requires obtaining the variance of a sum as follows:

$$\operatorname{var}[\hat{g}(x)] = \operatorname{var}(\hat{\beta}_o) + x^2 \operatorname{var}(\hat{\beta}_1) + 2x \operatorname{cov}(\hat{\beta}_o, \hat{\beta}_1) \quad (4.49)$$

In general the variance of a sum is equal to the sum of the variance of each term and twice the covariance of each possible pair of terms formed from the components of sum. The endpoints of a 100(1 – α) % Wald based confidence interval for the logit are

$$\hat{g}(x) \pm z_{1-\alpha/2} SE[\hat{g}(x)] \quad (4.50)$$

Where $SE[\hat{g}(x)]$ is the positive square root of the variance estimator in (4.49), the endpoint of the 100 (1 – α) % Wald based confidence interval for the fitted value are:

$$\frac{\exp\left[\hat{g}(x) \pm z_{1-\frac{\alpha}{2}} SE[\hat{g}(x)]\right]}{1 + \exp\left[\hat{g}(x) \pm z_{1-\frac{\alpha}{2}} SE[\hat{g}(x)]\right]} \quad (4.51)$$

One application of fitted logistic regression models that has received a lot of attention in the subject matter literature is the use of model based fitted values to predict the value of a binary dependent value in individual subjects (Sawsan, 2008).

4-6 General Logistic Regression Model:

The simple linear logistic model $\log\left[\frac{\Pi_i}{1-\Pi_i}\right] = \beta_1 + \beta_2 x_i$ used in the Previous example is a special case of the general logistic regression model.

$$\log it\Pi_i = \log\left(\frac{\Pi_i}{1-\Pi}\right) = x_i^T\beta \quad (4.52)$$

Where xi is a vector continuous measurement corresponding to covariates and dummy variables corresponding to factor levels and β is the parameter vector. This model is very widely used for analyzing data involving binary or binomial responses and several explanatory variables it provides a powerful technique analogous to multiple regression and ANOVA for continuous responses.

Maximum likelihood estimates of the Parameters β, and consequently of the probabilities $\Pi_i = g(x_i^T, \beta)$ are obtained by maximizing the log- likelihood function table (4.3) comparison of observed numbers killed with fitted values obtained from various dose- response models for the beetle mortality data. Deviance statistics are also given.

Table (3): Comparison of observed number killed with fitted values obtained from various dose-response models for the beetle mortality data. Deviance statistics are also given:

observed value of Y	Logistic model	Probit model	Extreme value model
6	3.46	3.36	5.59
13	9.84	10.72	11.28
18	22.45	23.48	20.95
28	33.90	33.82	30.37
52	50.10	49.62	47.78
53	53.29	53.32	54.14
61	59.22	59.66	61.11
60	58.74	59.23	59.95
D	11.23	10.12	3.45

$$l(\Pi, y) = \sum_{i=1}^{N}\left[y_i \log\Pi_i + (n_i - y_i)\log(1-\Pi_i) + \log\binom{n_i}{y_i}\right] \quad (4.53)$$

The estimation process is essentially the same whether the data are grouped as frequencies for each covariate pattern (I.e., observation with the same values of all explanatory variables) or each observation is coded 0 or 1 and its covariate pattern is listed separately. If the data can be grouped, the response Y_i , number of success for covariate pattern i , may be modeled by the binomial distribution. If each observation has a different covariate pattern, then $n_i = 1$ and the response y_1 is binary.

The deviance:

$$D = 2\sum_{i=1}^{N}\left[y_i \log\left(\frac{y_i}{\hat{y}_i}\right) + (n_i - y_i)\log\left(\frac{n_i - y_i}{n_i - \hat{y}_i}\right)\right] \quad (4.54)$$

Thus D has the form:

$$D = 2\sum \log\frac{o}{e} \quad (4.55)$$

Where O denotes the observed frequencies y_i and $(n_i - y_i)$ from the cells of table (4.4) and e denotes the corresponding estimated expected frequencies or fitted values $\hat{y}_i = n_i\hat{\Pi}_i$ and $(n_i - \hat{y}_i) = (n_i - n_i\hat{\Pi}_i)$

summation is over all 2 × N cells of the table.

Notice that D dose not involve any nuisance parameters (like σ^2 normal response data), so goodness of fit can be assessed and hypotheses can be tested directly using the approximation where p is the number of parameters estimated and N the numbers of covariate patterns.

The estimation methods and sampling distributions used for inference depend on asymptotic results. For small studies or situations where there are few observations for each covariate pattern, the asymptotic results may be poor approximations. However software, such as stat-Xact and log-Xact, has been developed using exact methods that can be used even when sample sizes are small.((Annette J.Dobson,2001)

4-7 Goodness of fit Statistics:

Instead of using maximum likelihood estimation we could estimate the parameters by minimizing the weighted sum of squares

$$S_w = \sum_{i=1}^{N} \frac{(y_i - n_i\Pi_i)^2}{n_i\Pi_i(1-\Pi_i)} \quad (4.56)$$

Since $E(y_i) = n_i\Pi_i$ (4.57)

$$\chi 2 = \sum_{i=1}^{N} \left(\frac{(y_i - n_i\Pi_i)^2}{n_i\Pi_i(1-\Pi_i)} \right)(1 - \Pi_i + \Pi_i) = S_w \quad (4.58)$$

When χ^2 is evaluated at the estimated expected frequencies, the statistic is

$$\chi^2 = \sum_{i=1}^{N} \left(\frac{(y_i - n_i\hat{\Pi}_i)^2}{n_i\hat{\Pi}_i(1-\hat{\Pi}_i)} \right) \quad (4.59)$$

Which is asymptotically equivalent to

$$D = 2\sum_{i=1}^{N} \left[y_i \log\left(\frac{y_i}{n_i\Pi_i} \right) + (n_i - y_i) + \log\left(\frac{n_i - y_i}{n_i - n_i\Pi_i} \right) \right] \quad (4.60)$$

The asymptotic distribution of D, under the hypothesis that the model is correct, is $D \approx \chi^2(N-p)$

The choice between D and χ^2depends on the adequacy of the approximation to the χ^2 (N – P) distribution. There is some evidence to suggest that χ^2 is often better than D because D is unduly influenced by very small frequencies.

Both the approximations are likely to be poor, however, if the expected frequencies are too small (e.g., less than 1).

In particular, if each observation has a different covariate pattern so Y_i is zero or one, then neither D nor χ^2 provides a useful measure of fit. This can happen if the explanatory variables are continuous, for example. The most commonly used approach in this situation is due to Hosmer and lemeshow (1980). Their idea was to group observation into categories on the basis of their predicted probabilities. Typically about 10 groups are used with approximately equal numbers of observations in each group.

Sometimes the log- likelihood function for a minimal model, in which the values Π are all equal (in contrast to the saturated model which is used to define the deviance).

Under the minimal model $\tilde{\Pi} = \frac{\sum y_i}{\sum n_i}$ let $\hat{\Pi}_i$ denote the estimated probability for Y_i under the model of interest (so the fitted value is $\hat{y}_i = n_i\hat{\Pi}_i$)

The statistic is defined by:

$$C = 2\sum \left[y_i \log\left(\frac{\hat{y}_i}{n_i\hat{\Pi}_i} \right) + (n_i - y_i)\log\left(\frac{n_i - \hat{y}_i}{n_i - n_i\Pi_i} \right) \right] \quad (4.61)$$

In the beetle mortality example ; c = 27297 with df = 1, indicating that slope parameter is definitely needed.

If the response variable is categorical, with more than two categories, then there are two options for generalized linear models. One relies on generalization of logistic from dichotomous responses to nominal or ordinal responses with more than two categories. For nominal or ordinal logistic regression one of the measured or observed categorical variables is regarded as response, and all other variables are explanatory variables. For log- linear models, all the variables are treated alike. The choice of which approach to use in a particular situation depends on whether one variable is clearly a response or several variables have the same status (as may be the situation in a cross-sectional study). Additionally, the choice may depend on how the results are to be presented and interpreted. Nominal and ordinal logistic regression yield odds ratio estimates which are relatively easy to interpret if there are no interactions (or only fairly simple interactions). Log- linear models are good for testing hypotheses about complex interactions, but the parameter estimates are less easily interpreted. Multinomial distribution provides the basis for modeling categorical data with more than two categories. Then the various formulations for nominal and ordinal logistic regression models are discussed, including the interpretation of parameter estimates and methods for checking the adequacy of a model. A numerical example is used to illustrate the methods.

4-9 Nominal Logistic Regression:

Nominal Logistic Regression models are used when there is no natural order among the response categories. One category is arbitrarily chosen as the reference category. Suppose this is the first category. Then the logit for the other categories are defined by

$$\log it\ (\Pi_j) = \log\left(\frac{\Pi_j}{\Pi_1}\right) = X_j^T\ B_j\ \ ,\ for\ j = 2,\dots,J\ \ (4.62)$$

The (J-1) logit equations are used simultaneously to estimate the parameters ß$_j$ once the parameter estimates b_j have been obtained, the linear predictors can be calculated. $X_j^T b_j$, $j = 2,\dots,J$ From the previous equation:

$$\hat{\Pi}_j = \hat{\Pi}_1\ \exp\ (X_j^T\ b_j)\ \ for\ J = 2,\dots,J\ \ (4.63)$$

$$But\ \hat{\Pi}_1 + \hat{\Pi}_2 + \dots + \hat{\Pi}_j = 1\quad So\ (4.64)$$

$$\hat{\Pi}_1 = \frac{1}{1+\sum_{j=2}^{J}\ \exp\ (X_j^T\ b_j)}\ (4.65)$$

and $\hat{\Pi}_j = \dfrac{\exp\ (X_j^T\ b_j)}{1+\sum_{j=2}^{J}\ \exp\ (X_j^T\ b_j)}$ for j = 2,…, J (4.66)

Summary statistics for goodness of fit are analogous to those for binomial logistic regression:

(i) chi- square statistic: $\chi^2 = \sum_{i=1}^{N} r_i^2$ (4.67)

(ii) Deviance, defined in terms of the maximum values of the log- likelihood function for the fitted model, L(b), and for the maximal model, L(max), $D = 2\left[L\ (b\ \max\ ) - (b\ )\right]$ (4.68)

(iii) Likelihood ratio chi- squared statistic, defined in terms of the maximum value of the log-likelihood function for the minimal model, L (b max), and L (b), C=z$\left[L(b) - L(b\min)\right]$; (4.69)

$$pseudo\ R^2 = \frac{L(b\min) - L(b)}{L(b\min)}\ (4.70)$$

If the model fits well then both χ^2 and D have, asymptotically, the distribution $\chi^2_{(N-P)}$ where P is the number of parameters estimated. C has the asymptotic distribution $\chi^2\left[P - (J-1)\right]$ because the minimal model will have one parameter for each logit.

Often it is easier to interpret the effects of explanatory factors in terms of odds ratios than the parameters B. for simplicity, consider a response variable with J categories and a binary explanatory variable x which denotes whether an 'exposure' factor is present (x=1) or absent (x= 0).
The odds ratio for exposure for response j = (j = 2,…, j) relative to the reference category j=1 is $OR_j = \frac{\Pi_j p}{\Pi_j a} \Big/ \frac{\Pi_i p}{\Pi_i a}$ (4.71)

Where $\Pi_j p$ denote the probabilities of response category j (j = 1,…,J) according to whether exposure is present or absent, respectively. For the model $\log\left(\frac{\Pi_i}{\Pi_i}\right) = \beta_0 j + \beta_1 x$, j = 2,…, J (4.72)

The log odds are $\log\left(\frac{\Pi_j a}{\Pi_i a}\right) = \beta_0 j$ when x = 1 , indicating the exposure is present.

Therefore the logarithm of the odds ratio can be written as:

$$\log OR_j = \log\left(\frac{\Pi_j p}{\Pi_i p}\right) - \log\left(\frac{\Pi_j a}{\Pi_i a}\right) = \beta_1 j \quad (4.73)$$

Hence $OR_j = \exp(\beta_{ij})$ which is estimated by exp (b_{ij}). If $\beta_{1j} = 0$ then $OR_j = 1$ which corresponds to the exposure factor having no effect. Also, for example, 95% confidence limits for OR_j are given by exp [V_1 j] , where s.e. (b_1 j) denotes the standard error of b_1 j . confidence intervals which do not include unity correspond to ß values significantly different from zero. For nominal logistic regression , the explanatory variables may be categorical or continuous. The choice of the difference category for the response will affect the parameter estimates b but not the estimated probabilities $\hat{\Pi}$ or the fitted values (Dobson, 2001)

16. Ordinal Logistic Regression

If there is an obvious natural order among the response categories then this can be taken into account in the model specification. Example of ordinal regression on car preference provides an illustration as the study participants rated the importance of air conditioning and power steering in four categories from 'not important' to 'very important'. Ordinal response like this are common in areas such as market research, opinion polls and fields like psychiatry where 'soft' measures are common.
In some situations there may, conceptually, be a continuous variable z which is difficult to measure, such as severity of disease. It is assessed by some crude method that a mounts to identifying 'cut points', C_i , for the latent variable so that, for example, patients with small values are classified as having 'no disease', those with larger values of z are classified as having 'mid disease' or 'moderate disease' and those with high values are classified as having 'severe disease'. The cut points C_1,…, C_{J-1} define j ordinal categories with associated probabilities $\Pi_1, \ldots, \Pi J \left(with \ \sum_{j=1}^{j} \Pi_j = 1\right)$ not all ordinal variables can be thought of in this may, because the underlying process may have many components.
Nevertheless, the idea is helpful for interpreting the results from statistical models. For ordinal categories, there are several different commonly used models which are described later.
4-11 Cumulative Logit Model:
The cumulative odds for the j th category is

$$\frac{P(Z \le C_j)}{P(Z * C_j)} = \frac{\Pi_1 + \Pi_2 + \ldots + \Pi j}{\Pi_{j+1} + \ldots + \Pi_j} \quad (4.74)$$

The cumulative logit model is:

$$\log \frac{\Pi_1 + \ldots + \Pi_j}{\Pi_j + 1 + \ldots + \Pi_j} = C_j^T \beta_j \quad (4.75)$$

Proportional Odds Model:

In the linear predictor $X_j^T \beta_j$ has an intercept $\beta_0 j$ which depend on the category j, but the other explanatory variable do not depend on j, then the model is

$$\log \frac{\Pi_1 + \ldots + \Pi_j}{\Pi_j + 1 + \ldots + \Pi_j} = \beta_0 j + \beta_1 X_1 + \ldots + \beta_{P-1} . \quad (4.76)$$

4-12 Wald Statistics:

To test the significance of each coefficient (β_i) is measured by the wald statistics, given by:

$$\text{Wald} = \left(\frac{\beta}{S.E(\beta)}\right)^2 \quad (4.77)$$

Which is distributed as chi-square with one degree of freedom. In addition to that 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 particular logit(effect)coefficient is zero. For dichotomous independents, the wald statistics is the squared ratio of unstandardized logit coefficient to its standard error.

Menard warns that for large logit coefficients standard error is inflated, lowering the Wald statistic leading to type two error. That is, there is flaw in the wald statistic such that very large effects many lead to large standard errors and small wald chi-square values. For models with large logit coefficients or when dummy variables are involved, it is better to test the difference using the likelihood ratio test of the difference of models with and without the parameter. Also wald statistics is sensitive to violations of the large –sample assumption of logistics regression. For these reasons, the likelihood ratio test of individual model parameter is generally preferred. (Statistics Solution,)

4-13 The Likelihood Ratio:

It is a function of log likelihood, because -2LL has approximately a chi-square distribution, -2LL can be used for assessing the significance of logistic regression, analogous to the use of the sum of squared errors in ordinary least square (OLS) regression. The -2LL statistic is the likelihood ratio. It is also called goodness of fit, deviance chi-square, scaled deviance, deviation chi-square, DM, or L-square. It reflects the significance of the unexplained variance in the dependent. In SPSS out put, this statistic is found in the -2log likelihood column of the iteration history table or the likelihood ratio test table. The likelihood ratio is not used directly in significance, but it is the basis for the likelihood ratio test, which is the test of difference between two likelihood test, which is the test of the difference between two likelihood ratios (two-2LL's). The likelihood ratio test is based on -2LL(deviance).The likelihood ratio test is a test of the significance of the difference between the likelihood ratio(-2LL) for the researcher's model minus the likelihood ratio for a reduced model. It is an alternative to the wald statistic and is also called the log-likelihood test. There are three main forms of the likelihood ratio test. The following formula of likelihood Ratio test is :

$-2\log(L_0/L_1) = -2\{\log(L_0)-\log(L_1)\} = -2(L_0 - L_1)$ (4.78)

4-14 Test of the Overall Model

When the reduced model is the baseline model with the constant only, the likelihood ratio tests the significance of the researcher's model as a whole. A well-fitting model is significance at the 0.05 level or better, meaning the researcher model is significantly different from the one with constant only. The likelihood ratio test appears in the model fitting information table in SPSS output. Thus, the likelihood ratio test of a model tests the difference between -2LL for the full model and -2LL for initial chi-square in the null model. This is called the model chi-square test. The null model, also called the initial model, is logit(p) = the constant. That is, initial chi-square is -2LL for the model which accepts the null hypothesis that all the b coefficients are zero. This implies that none of the independents are linearly related to the log odds of the dependent. Model chi-square thus tests the

null hypothesis that all population logistic coefficients except the constant are zero. It is an overall model test which dose not assure that every independent is significance. Degrees of freedom in this test equal the number of terms between the two models minus one (for the constant).This is the same as difference in the number of terms between the two modes, since the null model has only one term. Model chi-square measures the improvement in fitting that the explanatory variable make compared to null hypothesis. Model chi-square is a likelihood ratio test which reflects the difference between error not knowing the independents (initial chi-square) and error when the independents included in the model (deviance). When probability (model chi-square) <= 0.05, we reject the null hypothesis that knowing the independents make no difference in predicting the dependent in logistic regression.
$\chi^2 = -2LL_R - (-2LL_F) = -2\ln(\text{likelihood}_R / \text{likelihood}_F)$ (4.79)

17. Individual Model Parameters

The likelihood ratio test assesses to overall logistic model but dose not tell us if particular independents are more important than others. This can be done, however, by comparing the difference in -2LL for the overall model with a nested model which drops one of the independents. We can use the likelihood ratio test to drops one variable from the model to create a nested reduced model. In this situation, the likelihood ratio test tests if the logistic regression coefficient for the dropped variable can be treated as zero, thereby, justifying drooping the variable from the model to create a nested. A none significance likelihood ratio test indicates no difference between the full and the reduce model, hence justifying drooping the given variables so as to have a more parsimonious model that works just as well. Note that the likelihood ratio test of individual's parameters is a better criterion than the alternative wald statistic when considering which variables to drop from the logistic regression model.
In general, the likelihood ratio test can be used to test the difference between a given model and any nested model which is a subset of the given model. It cannot be used to compare two non-nested models. Chi-square is the difference in likelihood ratios (-2LL) for the two models, and degrees of freedom for the two models. If the computed chi-square is equal or greater than critical value of chi-square (in chi-square table) for the given degree of freedom, then the models are significantly different. If the difference is significant, then the researcher concludes that the variable dropped in the nested model to matter significance in predicting the dependent. If the difference is below the critical value, there is a finding of non-significance and the researcher concludes that dropping the variables makes no difference in prediction and for reasons of parsimony the variables are dropped from the model. That is, chi-square difference can be used to help decide which variables to drop from or add to the model.
Hosmer and Lemeshow's goodness of fit tes*t*, not to be confused with a similarly named, obsolete goodness of fit test discussed below, is another name for a chi-square goodness of fit test. It is available under the Options button in the SPSS binary logistic regression dialog. The test divides subjects into deciles based on predicted probabilities, then computes a chi-square from observed and expected frequencies. Then a probability (p) value is computed from the chi-square distribution with 8 degrees of freedom to test the fit of the logistic model. If the H-L goodness-of-fit test statistic is greater than 0.05, as we want for well-fitting models, we fail to reject the null hypothesis that there is no difference between observed and model-predicted values, implying that the model's estimates fit the data at an acceptable level. That is, well-fitting models show non significance on the H-L goodness-of-fit test, indicating model prediction is not significantly different from observed values. This does not mean that the model necessarily explains much of the variance in the dependent, only that however much or little it does explain is significant. As the sample size gets large, the H-L statistic can find smaller and smaller differences between observed and model-predicted values to be significant. On the other hand, the H-L statistic assumes sampling adequacy, with a rule of thumb being enough cases so that no group has an expected value < 1 and 95% of cells (typically, 10

deciles groups times 2 outcome categories = 20 cells) have an expected frequency > 5. Collapsing groups may not solve a sampling adequacy problem since when the number of groups is small, the H-L test will be biased toward non significance will overestimate model fit. There are many important research topics for which the dependent variable is limited for example voting, morbidity or mortality and participation data is not continuous or distributed normally. Binary logistic regression is a type of regression analysis where the dependent variable is a dummy variable: coded 0(did not vote) or 1 (did vote).(John Whitehead)

18. Data Analysis and Application

Khartoum state is the capital of Sudan contain three areas they are , Khartoum, Khartoum North, and Omdurman. It is smallest state with an area of 20,140 km with the largest population 5.6 million divided into seven localities urban, suburban, and rural areas. There were counseling services for AIDS patients through 80 counselors inside hospitals, health centers, and higher education institutes. There are 20 volunteers centers for testing blood. The Data of the study were collected from Omdurman, Khartoum, and Khartoum North Teaching Hospitals denoted by the records of Testing and Counseling centers. The centers were ranked as the old and best centers in the state. The centers cover the majority of Khartoum state and represent all the area of it. During that period (939) persons were subjected to HIV test in Omdurman center. Some of them (550) volunteers and the other (389) were suspected cases. So the (939) members represent the sample of the Omdurman center. Also (200) cases collected from Khartoum hospital and (300) from Khartoum North hospital. So the total sample size of the study is (1439) cases represent Khartoum state for the period under the study. The questionnaire of the study depends on the socio-economic variables and demographic variables existing in the records during the period of the study.

Descriptive statistics is done to explain the main characteristics of the study population through the sample of Omdurman (939) cases. The result is as shown below:

Table (5) Frequency distribution of tested people according to age:

	Volunteers		Suspected cases	
Age group	Frequency	Percent	Frequency	Percent
0 – 10	5	0.9	4	1
11 - 20	39	7.1	23	5.9
21 – 30	158	28.7	7	1.8
31 – 40	176	32.0	22	5.7
41 – 50	90	16.4	12	3.1
51 +	82	14.9	321	82.5
Total	550	100.0	389	100

Source: SPSS software outputs

Table (5) presents the age of investigators through volunteers , the two middle age groups (31-40) and (21- 30) years had a high percentages 32% and 28.7%, respectively, 16.4% for the group (41-50), 14.9% for the elder age greater than 50. While the age group (11- 20) took the percent 7.1%, and the children less than ten years show the small one 0.9%. The percentages of age among suspected investigators reveal that the two middle age group (21-30) and (31-40) year had a high percentages 42.4% and 25.2% respectively. 9.5% For the age (41-50), 3.9% for the elder age group 51+, the teenagers (11-20) had 13.9% and only 5.1% for the children less than ten years. The two middle age groups for both volunteers and suspected cases had the high numbers; because the volunteers inside this age they face the problem of youth.

Table (6) shows the sex of volunteers and suspected cases, the majority of volunteers were males 60.5% and 39.5% females. While more than have were males 59.4% while 40.6% females among suspected cases. This reflects the high percent for males who attend the centre.

Table (6) Frequency distribution of tested people according to sex

Sex	Volunteers		Suspected cases	
	Frequency	Percent	Frequency	Percent
Male	333	60.5	231	59.4
Female	217	39.5	158	40.6
Total	550	100.0	389	100

Table (7) illustrates the place of residence for the volunteers , the people who investigate were from Omdurman, Khartoum, and Khartoum North. Their percentages were 71.1%, 6% and 22.9%, respectively. Also the place of residence for the suspected cases, the percentages are 84.3% came from Omdurman, 7.2% from Khartoum North and only 8.5% from Khartoum. Thus the high percent from Omdurman because the center is inside the city, although its well known to the other cities of the state.

Table (7) Frequency distribution of tested people according to residence

Residence	Volunteers		Suspected cases	
	Frequency	Percent	Frequency	Percent
Khartoum	25	4.5	204	52.4
Khartoum North	33	6.0	160	41.1
Omdurman	391	71.1	17	4.4
Rural	101	18.4	8	2.1
Total	550	100.0	389	100

Table (8) Show the social status of the volunteers , 30.5% of them were single, 55.5% married 9.5% divorced, and 4.5% widowed. Beside the results of the Social status of suspected cases , 52.4% single, 41.1% married, 4.4% divorced and only 2.1% widowed. The married people gain the high 55.5%.That is due to their desire to check their health status.

Table (8) Frequency distribution of tested people according to social status

Social Status	volunteer		Suspected cases	
	Frequency	Percent	Frequency	Frequency
Single	168	30.5	204	52.4
Married	305	55.5	160	41.1
Divorced	52	9.5	17	4.4
Widowed	25	4.5	8	2.1
Total	550	100.0	389	100

Table (9) reports the job of volunteers , their job classified as, 34.2% do not work, 17.3% were labor, 8% employee, 28.4% had free job, 4.9% student 3.1% tee sellers, and only 4.2% drivers. The classification of job among suspected cases were , 33.7% for the people who not work, 16.7% free job, 14.7% employee, 8.5% lab our, 5.7% tee seller and 3.1 drivers. People without work represent considerable percent inside the two categories.

Table (9) Frequency distribution of the occupation tested people

Occupation	volunteers		Suspected cases	
	Frequency	Percent	Frequency	Percent
Not work	188	34.2	131	33.7
Labour	95	17.3	33	8.5
Employee	44	8.0	57	14.7
Free job	156	28.4	65	16.7
Student	27	4.9	69	17.7
Tee seller	17	3.1	22	5.7
driver	23	4.2	12	3.1
Total	550	100.0	389	100

Table (10) gives the education level for volunteers, 48.2% of them were illiterate, 27.5% primary education, 13.8%secondary education, 8.2% for the university / postgraduate education. Illiterate investigators represent near half of the sample. The percentages of education levels into suspected cases were 25.7% secondary, 24.9% primary / basic, 22.4% illiterate, 22.1% university / postgraduate and 4.9% for the intermediate education. The results reflects the considerable effect of the illiteracy level on the HIV prevalence inside Khartoum state.

Table (10) Frequency distribution of tested people according to education level

Education level	volunteers		Suspected cases	
	Frequency	Percent	Frequency	Percent
Illiterate	265	48.2	87	22.4
Primary / basic	151	27.5	97	24.9
Intermediate	45	8.2	19	4.9
Secondary	76	13.8	100	25.7
University / postgraduate	13	2.4	86	22.1
Total	550	100.0	389	100

Table (11) reports religious status through volunteers, 80.9% were Muslim and 19.1% Christian. This reflect that, the people who live in Khartoum most of them Muslims. The percentages of religious status among suspected cases, 82.8% were Muslim and 17.2% Christian. The majority of volunteers were Muslims.

Table (11) Frequency distribution of tested people according to religion

Religion	Volunteers		Suspected cases	
	Frequency	Percent	Frequency	Percent
Muslim	445	80.9	322	82.8
Christian	105	19.1	67	17.2
Total	550	100.0	389	100

Table (12) Signify the symptoms of HIV/AIDS 26.5% lost of weight, 23.8% diarrhea, 21.1% fever, 14% cough, 4% skin rash, 3.3% (fever, lost of Weight, and diarrhea), and 0.2% had no symptoms. The percentages of symptoms for suspected cases of HIV/AIDS, 82.5% indicates no clear symptoms, 5.7% for the diarreoha, 5.9% lost of weight, 3.1% skin rash and 1.8%, 1.0% for the cough and fever respectively.

Table (12) Frequency distribution of tested people according to symptoms

Symptoms	volunteers		Suspected cases	
	Frequency	Percent	Frequency	Percent
Fever	120	21.8	4	1
Lost of weight	146	26.5	23	5.9
Cough	77	14.0	7	1.8
Diarrhea	131	23.8	22	5.7
Skin rash	22	4.0	12	3.1
Fever +lost of weight	23	4.2	321	82.5
Fever+ diarrhea	4	0.7	-	-
Fever +loss +Weight +diarrhea	18	3.3		
No symptoms	1	0.2		
TB	8	1.5		
Total	550	100.0	389	100

Table (13) display the risk factors, 57.5% were worried, 39.8% sexual, 1.6% blood transmission, 0.7% husband / wife, and 0.4 %from parents to their children. It is very clear that people worried about their health status. Also, the percentages of risk factors for suspected cases , 44.5% worried, 23.7% sexual, 13.6% parents, 12.3% has band / wife, 4.1% for the purpose of procedure, and 1.8% blood transmission.

Table (13) Frequency distribution of tested people according to risk factors

Risk Factors	volunteers		Suspected cases	
	Frequency	Percent	Frequency	Percent
Sexual	219	39.8	92	23.7
Worried	316	57.5	173	48.6
Blood transmission	9	1.6	7	1.8
Parents	2	0.4	53	13.6
Husband / wife	4	0.7	48	12.3
Total	550	100.0	389	100

Source: SPSS software outputs

Table (14) shows the result of blood investigator, 66.5% were negative HIV/AIDS and 33.5% Positive. Suspected cases reports the HIV test result, 77.4% were negative while 22.6% were positive. The results of volunteers and suspected cases is reflecting the incidence rate of HIV / AIDS indicate the terrible rate need more integrated efforts to stop increasing. Also the high present of HIV positive among individuals reflect terrible situation need policy makers to stop on it, and immediate intervention from the government and Ministry of Health.

Table (14) Frequency distribution of tested people according to result

Result	Volunteers		Suspected cases	
	Frequency	Percent	Frequency	Percent
Negative	20	66.5	301	77.4
Positive	54	33.5	88	22.6
Total	550	100.0	389	100

19. Dependency Analysis

The purpose of this section to investigate the relation (dependency) between HIV incidence and some socioeconomic variables, results are reported using (χ^2) test for dependency. Table (15) reports the result of χ^2 is 11.428, df = 4 and p-value (0.022), that indicate the dependency of the two variable at 5% level of significance (result of HIV investigation and education level). As a result of the test the null hypothesis of the independent between the two variables is rejected. Also table (5.11) express the high number of illiterate persons are more infected with HIV. This result confirms the need of awareness regarding HIV infection, particularly among the illiterate.

Table (15) chi- square test for Dependency between HIV incidence and education level

HIV Test	Education Level				Tota l	χ^2	p-value
	illiterate	Primary basic	Intermediate	University/ Postgraduate			
Negative	61	77	11	75	301	11.721	0.02
Positive	26	20	8	11	88		
Total	87	97	19	100	389		

Table (16) reports the result of the χ^2 test for dependency. The value of chi- square is calculated at 16.29 with df =6 and P-value = 0.012, which is significant at 5% level of significance. So the result support the idea of association between HIV infection and employment status of individuals. People who not work they have a high positive number (41), so this result ensures the fact that leisure time of youth affects the personal behavior.

Table (16) chi- square test for Dependency between HIV incidence and occupation:

HIV Test	Occupation							Total	χ^2	p-value
	Not work	Labor	Employee	Free job	student	Tee selle	drive r		16. 9	0.01

						r				
Negative	90	26	47	49	64	16	9	301		
Positive	41	7	10	16	5	6	3	88		
Total	131	33	57	65	69	22	12	389		

Table (17) reports the result of the χ^2 test for dependency between HIV infection and social status. The value of the χ^2 is calculated at 24.9 with df = 3, P-value 0.0001, which is significant at 1% , this indicate the significance result for the test. Thus, the hypothesis of dependency between the two variables is accepted. We may note that although single individuals are assumed to have more positive HIV cases, the high number of positive HIV test is found among married persons. This result may be attributed to the fact that the infection of one married partner increases the probability of infection of the other partner, giving rise to higher incidence of the disease.

Table (17) Chi- square test for dependency between HIV incidence and social status:

HIV Test	social status				Total	χ^2	p-value
	Single	married	Divorced	Widowed			
Negative	178	107	12	4	301	25.326	0.0001
Positive	26	53	5	4	88		
Total	204	160	17	8	389		

Table (18) illustrates the χ^2 test between HIV incidence and religion χ^2 = 12.11 with df = 1, and p-value = 0.001 which is significant at 1% level, this indicate the high significance. We notice that, Christians (26) are relatively more infected with HIV than Muslim, although Muslims gain (62) infected person.

20. Volunteers cases

Table (18) show the Chi- square test for dependency between incidence and education level χ^2 calculated = 5.00 with p-value (0.08), the result indicate independency of the two variables. So education didn't influence HIV/AIDS incidence.

Table (18) Chi-square test for Dependency between HIV incidence and education level

HIV Test	Education Level				Total	χ^2	p-value
	illiterate	Primary basic	intermediate	University/Post			
Negative	185	99	46	6	366	5.000	0.08
Positive	80	52	30	7	184		
Total	265	151	76	13	550		

Table (19) reports the χ^2 test for dependency between HIV incidence and social status. χ^2 calculated = 16.104 with p-value (0.001), which indicate a dependency between the two variable at 1% level of significance. Thus, positive HIV test of the married persons is very high (55.5%) which suggest that, there must be awareness of the individuals then all the society.

Table (19): chi-square test for dependency between HIV incidence and social status:

HIV Test	social status				Total	χ^2	p-value
	single	Married	divorced	Widowed			
Negative	119	211	25	11	366	16.014	0.001
Positive	49	94	27	14	184		
Total	168	305	52	25	55		

Table (20) illustrate χ^2 test for dependency between HIV incidence and occupation.. χ^2 calculated is 12.59 with p-value (0.05)so there is dependency between the job of individuals and HIV infection (i.e. the job affect the HIV/AIDS incidence).

Table (20): chi- square test for dependency between HIV incidence and occupation:

HIV Test	Occupation	Total	χ^2	p-value

	Not work	Labour	employee	Free job	student	Tee seller	driver		12.59	0.05
Negative	137	62	23	95	22	9	15	366		
Positive	51	33	21	58	5	8	8	184		
Total	188	95	44	156	27	17	23	550		

Table (21) reports frequency distribution of volunteer according to age, the middle age groups (21-30), (31-40) had 48.5%, 28% respectively, 1.5% for the age group (0-10), 8.5% for the age group (11-20) year, 9 % for (41-50), and 4.5% for the elder group (51+) people have 51 years and more.

Table (20) Frequency distribution of volunteer age

	Frequency	Percent
0-10	3	1.5
11-20	17	8.5
21-30	97	48.5
31-40	56	28
41-50	18	9
51+	9	4.5
Total	200	100

We note that, volunteer are concentrated in the two middle age group (20-30) (31-40), that is due to concerned of these age groups about their health situation and they seek about security health.

Table (21) reports frequency distribution of volunteer according to sex; about quarter percent (74%) of volunteer were males and 26% females.

Table (21): Frequency distribution according to Sex.

	Frequency	Percent
Male	148	74.0
Female	52	26.0
Total	200	100.0

Table (22) illustrate frequency distribution of volunteer according to education level, illiterate people (12%) of the volunteers, (31%) primary/Basic education, 4% were intermediate education, while (26.5%) for both secondary and postgraduate levels. Also we can say there is a high percent of tested people among the last two classification of education level.

Table (22):Frequency distribution according to education level.

	Frequency	Percent
Illiterate	24	12
Primary/basic	62	31
Intermediate	8	4
Secondary	53	26.5
University/postgraduate	53	26.5
Total	300	100

Table (23) reports Frequency distribution of tested people 21% for employee and free job, 19.5% for labor and people without work and 19% for the student.

(23): Frequency distribution according to occupation.

	Frequency	Percent
Not work	39	19.5
Labour	39	19.5
employee	42	21
Free job	42	21
student	38	19

Total	200	100

Table (24) shows the frequency distribution of volunteer according to religion, most of them 86% were Muslims and 14% Christian.

Table (24) Frequency distribution according to religion

	Frequency	Percent
Muslim	172	86
Christian	28	14
Total	200	100

Table (25) reports frequency distribution of volunteer according to social status, about 54% were single, 37% married, 5% divorced, and only 4% widowed. The percent of single people indicate they worried about the health.

Table (25) Frequency distribution according to social status.

	Frequency	Percent
Single	108	54
Married	74	37
Divorced	10	5
Widowed	8	4
Total	200	100

These results indicate the high percent of single volunteer of Khartoum, which may influence the HIV/AIDS situation. There were 61.5% of volunteer attend to the center because of their sexual behavior, 16.5 of them were worried, 5% for blood transmission, 4.5 for the infected parent, and 12.5% for the husband/wife reasons. People doing sex always, they live in dangerous situation. The high percent of the sexual behavior reflect that people didn't care about the claim of sexual as one of the major method to be infected with HIV/AIDS.

Table (26) Frequency distribution according to the risk factors

	Frequency	Percent
Sexual	123	61.5
Worried	33	16.5
Blood transmission	10	5
Parent	9	4.5
Husband/wife	25	12.5
Total	200	100

Table (27) reports the HIV/AIDS incidence, 64% were negative HIV/AIDS, while 36% were positive HIV/AIDS, and this percent of positive people indicates high prevalence of HIV/AIDS among Khartoum hospital volunteers and suspected cases.

Table (27) Frequency distribution of HIV/AIDS result

	Frequency	Percent
Negative	128	64
Positive	72	36
Total	200	100

To test the dependency between the HIV/AIDS incidence and some socio-economics variables, study used chi-square test for this purpose. Table (28) reports the dependency test between HIV/AIDS incidence and education level. χ^2 value = 41.85 df = 4, P- value = 0.00001, P-value is less than the significance level 1% so we accept the alternative hypothesis there is association between the HIV/AIDS indicate and education level. It is evident that HIV/AIDS relate to education level of people, and it is the crucial factor of determination the incidence of HIV/AIDS among citizens.

Table (28) Chi-square test for dependency between HIV/AIDS incidence and education level

HIV Test	Education Level	Total	χ^2	P-value

	illiterate	Primary /basic	Intermediate	Secondary	University/post			
Negative	8	26	7	40	47	8	41.85	0.00001
Positive	16	36	1	13	6	16		
Total	24	62	8	53	53	24		

Table (29): illustrate dependency between the HIV/AIDS and occupation, the $\chi^2 = 21.27$, df = 4, P-value = 0.0001 which is less than the level of significance 1% so we fail to reject the alternative hypothesis that, the two variables are dependent. This result ensures the fact that the single people tend to have positive HIV/AIDS because of their behavior according to their situation in the community, they behave badly.

Table (29) Chi-square test for dependency between HIV/AIDS and occupation:

HIV Test	Education Level					Total	χ^2	P-value
	Not work	Labour	Employee	Free job	student			
Negative	18	18	31	28	33	128	21.271	0.001
Positive	21	21	11	14	5	72		
Total	39	39	42	42	38	200		

Table (30): reports the test of independency between HIV/AIDS and social status. $\chi^2 = 42.299$, df = 3, P-value = 0.0001. This result indicates the relation between HIV/AIDS incidence and social status. Always single people their sexual behavior is not legal, so this may lead to illegal relation between males and females. Which help the dissemination of the disease a mange youth.

Table (30): chi-square test for dependency between HIV/AIDS and Social Status

HIV Test	Social Status				Total	X^2	P-value
	Single	Married	Divorced	widowed			
Negative	91	29	5	3	128	42.29	0.00001
Positive	17	45	5	5	72		
Total	108	74	10	8	200		

Table (31) reports frequency distribution of tested people according to sex, there were 64.7% males and 35.3% Females.

Table (31): Frequency distribution of sex

	Frequency	Percent
Male	194	64.7
Female	106	35.3
Total	300	100.0

Source: spss software outputs

Table (32) shows the frequency distribution of tested people according to age, the two middle age groups (21-30) (31-40) had 33.3% , 29.3%respectively, people age (11-20) years had 12.7% , persons less than ten years score 2% only, (41-50) age group had 8% and the last age group 51+ gain 14.7%.

Table (31): Frequency distribution of volunteer age

	Frequency	Percent
0-10	6	2
11-20	38	12.7
21-30	100	33.3

31-40	88	29.3
41-50	24	8
51+	44	14.7
Total	300	100

Table (31) illustrates frequency education level of study population, illiterate people 30% primary/Basic education 27.3%, 22.3% for the secondary and 20.7% for the university/postgraduate volunteers. We note the high percent of illiterate people who attend the centre, which indicate the concerning oft AIDS among this group.

Table (31): Frequency distribution of education level.

	Frequency	Percent
Illiterate	88	30
Primary/basic	82	27.3
Secondary	67	22.3
University/postgraduate	62	20.3
Total	300	100

Table (32) reports frequency distribution of volunteer according to religion, most of them were Muslims (85.7%) and 14.3% for the Christian.

Table (32): Frequency distribution of religion

	Frequency	Percent
Muslim	257	85.7
Christian	43	14.3
Total	300	100

Table (33) shows that, the high percent of volunteers (46.3) were single, 42.7% married, and (6% , 5%) for the divorced and widowed respectively. it is very clear that the social status of volunteers concentrated in single and married group.

Table (33): Frequency distribution of social status

	Frequency	Percent
Single	137	46.7
Married	126	42.7
Divorced	18	6
Widowed	14	4.7
Total	300	100

Table (34) reports Frequency distribution of tested people according to occupation, 26.3% without work, 23.3% labours, 22% employee, 15.3% free job 11.7% student, 1.3% tee sellers .

Table (34): Frequency distribution of occupation

	Frequency	Percent
Not work	75	25
Labor	69	23
Employed	64	21.3
Free job	43	14.3
Student	35	11.7
Total	300	100

Table (35) display the distribution of place of residence, 83.3% from Behree, 11% for rural areas, 3% from Khartoum and only 2.7% from Omdurman. We note that the high percent 83.3% for Khartoum North because the center inside it and there were considerable percent 2.7% from Omdurman, it may due to stigma of AIDS patient. People tend to check their blood away from residence, because of stigma

Table (35): Frequency distribution of place of residence

	Frequency	Percent
Khartoum	9	3
Khartoum North	250	83.3
Omdurman	8	2.7
Rural	33	11
Total	300	100

Table (36) shows the risk factors, 28% sexual, 22.3% worried, 4% blood transmission, 1% parent, 25.3% symptoms, 19.3% procedures. The results explain the fact that sexual behavior and symptoms that appear on the volunteers became reasons for attending the centre and check HIV/AIDS. There is a rational assumption that the two percentages are the same (i-e the symptoms of patients due to their sexual practice).

Table (36): Frequency distribution of risk factors

	Frequency	Percent
Sexual	84	28
Worried	67	22.3
Blood transmission	12	4
Parent	3	1.5
Symptoms	76	25.3
Procedures	58	19.3
Total	300	100

Table (37) reports the HIV/AIDS incidence, through the result of test, 85.7% were negative HIV/AIDS, while (14.3%) were positive.

Table (37): Frequency distribution of HIV/AIDS result

	Frequency	Percent
Negative	257	85.7
Positive	43	14.3
Total	300	100

Table (38) reports the dependency between HIV/AIDS and education level. χ^2 value = 10.07, df = 3, P- value = 0.018, According to P-value, we reject null hypothesis that the two variable are independent and accept the alternative one they are dependent. So education affect HIV/AIDS incidence in Khartoum North area.

Table (38) Chi-square test for dependency between HIV incidence and education level

HIV Test	Education Level				Total	χ^2	P-value
	Illiterate	Primary /basic	intermediate	Secondary			
Negative	78	63	57	59	257	10.07	0.018
Positive	11	19	10	3	43		
Total	89	82	67	62	300		

Table (39): reports chi-square test for dependency between HIV/AIDS incidence and occupation. χ^2 = 6.964, df = 7 , P-value = 0.433 the P-value is greater than (0.055) the level of significance, so we can‘t reject the null hypothesis of the test there is no relation between the occupation of the tested people and the HIV/AIDS result .

Table (39) Chi-square test for dependency between HIV/AIDS and occupation:

HIV Test	Education Level					Total	χ^2	P-value
	Unemploy	Labour	employee	Free job	Student			

	ed							
Negative	18	18	31	28	33	128	21.271	0.0001
Positive	21	21	11	14	5	72		
Total	39	39	42	42	38	200		

Table (40) reports the dependency analysis between HIV/AIDS incidence and the social status of volunteers who enter the center. The χ^2 = 25.14, df = 3, P-value = 0.0001. This suggests the acceptance of the alternative hypothesis there is relation between the two variables. This result coincides with the result of Khartoum result.

Table (40):chi-square test for dependency between HIV/AIDS and social status:

HIV Test	Social Stat				Total	χ^2	P-value
	Single	married	Divorced	widowed			
Negative	130	107	10	10	257	25.14	0.0001
Positive	9	21	8	5	43		
Total	139	128	18	15	300		

21. Modeling HIV/AIDS in Khartoum State Data

Next we deal with the collected data (Khartoum state data) the sample was 1439 members of volunteers and suspected cases . The model tries to find the relation between dependent variables (HIV/AIDS) incidence and socio-economic and demographic variables (independent variables) through the application of binary logistic regression. The value labels of the dependent variable (HIV/AIDS) incidence (0 denote negative test and 1 is positive test of result). Step one of the model create a reduce model (i.e. model without predictor) or the model with the constant only, which is gives the results, Wald statistic = 282.935 with P-value (0.0001). So the model represented by the constant is significance at 1% level of significance. And this is the null model (model without predictor). In purpose of product full model or final model (with predictor) binary logistic regression is done to obtain the model with education level, the model results as follows:

Table (41) The estimated model coefficients of education level

Effect	B	S.E	Wald	Sig	OR	95%CL of OR
Education	0.173-	0.042	17.189	0.0001	0.841	0.76,0.94

Source: SPSS output

Table(5.41):Test of the model

Model	-2log likelihood	Chi-square	Sig
Null	1677.21		
Final	1657.989	17.669	0.00001

Table (42) Goodness of fit test

	Chi-square	Sig
Hosmer-lemeshow	18.145	0.00000
Deviance	18.6	0.000033

Table (40) reports Wald statistic = 17.189, P-value (0.0003). which support the rejection of the null hypothesis that (the coefficient is zero). So the education level is statically significance at 1% level of significance. It is very clear through this result, the education level of volunteer affect the HIV/AIDS incidence (i-e the education level of a person play an important role of getting negative or positive HIV/AIDS). The odd ratio (0.84) indicates that (illiterate people) have HIV/AIDS positive 0.84 times more than the other types of education level. Or we can say, people without education have HIV/AIDS 0.84 times more than educated people. To test the overall model, the chi-square value (17.667) with P-value (0.00016),which indicate the statistically significance at 1%level of significance. So the model can be put into consideration and its result is acceptable.

hosmer and lemeshow test, the L-H value = 18.146, P-value (0.002), we can judge that the test is significant at 1% level of significance, so there is difference between observed and predicted value, the model is not fitting the data well. Also deviance statistic (18.6) confirm the result of fitting.

Table (43): Estimated coefficient for binary logistic regression(occupation)

Effect	B	S.E	Wald	Sig	OR	95%CL of OR
Occupation	0.071-	0.037	3.741	0.053	0.932	0.87,1.01

Table (44) Test of the model

Model	-2log likelihood	Chi-square	Sig
Null	1675.567		
Final	1671.781	3.786	0.052

Table (45) Goodness of fit test

	Chi-square	Sig
Hosmer-lemeshow	18.145	0.00000
[Deviance	18.6	0.000033

Table (43) as followed by table (44) reports that Wald statistic = 3.741, P-value (0.053) which indicate insignificance of the occupation at 1% Level of significance. The result of insignificance explains , there is no direct effect of the predictor variable occupation on the HIV/AIDS incidence. Odd ratio of the variable is (0.932), so the people without job have positive HIV/AIDS 0. 932 time more than people without work. The overall model is statistically significance according to χ^2 value = 3.786 with P-value (0.05). So the occupation has no significant affect on HIV/AIDS of volunteers who attend the centers. Hosmor-lemeshow test value = 31.69, P-value = 0.0001, so the model is not well fitted. According to table (46) and table (47) bellow Wald statistic = 70.398, P-value (0.0001) which indicate the statistically significance of the coefficient, the independent variable (Social Status) influence the HIV/AIDS incidence positively. The odds ratio EXP (B) = 1.94 which means the single people of Khartoum area have positive HIV/AIDS 1.94 times more than people (married, divorced, widowed).This can be support by the test of the over all significance of the model. χ^2 = 73.449whith P-value (0.000012), so the model is statistically significance at 1% level of significance. H-L value = 3.196, P-value (0.074), so the model is well-fitting the data of Khartoum state.

Table (5.46) Estimated coefficient for binary logistic regression (social status)

Effect	B	S.E	Wald	Sig	OR	95% CL of OR
Social	0.661	0.079	70.39	0.00	1.937	1.66,2.26

Table (47) Fitting information

Model	-2log likelihood	Chi-square	Sig	Hosmer-lemeshow
Null	1675.21			
Final	1602.118	73.449	0.072	0.074

In purpose of constructed binary logistic regression depend on the two predictors variables (social and education level) to see the effect of them on dependent variable (HIV/AIDS incidence) the results are, Wald statistic for the social status = 70.398, p-value = 0.011and Wald statistic for education = 6.493 , p-value =o.ooo1, the result of wald test explain the significance of the two important variables. Also odd ratio of social and education were 1.937, 0.895 respectively. Thus the change in the log odd ratio of HIV/AIDS per social and education is 1.937, 0.895.

In order to test the overall significance of the χ^2 =80.019, p-value = 0.00001,which insures the significance of the model. Hosmer-Lemeshow test done to test the goodness of fitting H-L = 31.45, p-value = 0.0001, so the model didn't fit the data.

Table (48): Estimated coefficient for binary logistic regression (social status, education)

Variable	B	S.E	Wald	Sig	OR	95%CL of OR
Constant	-1.944	0.182	114.519	0.000	0.143	
Social	0.623	0.079	70.398	0.011	1.937	1.66,2.26
Education	-0.11	0.043	6.493	0.000	0.895	0.823, 0.975

Table (49) Test of the model

Model	-2log likelihood	Chi-square	Sig	Hosmer-lemeshow
Null	1675.21			
Final	1595.548	80.019	0.000	0.074

According to table (49) the model include two important variable (social status and Occupation).The result signify that Wald statistics = 60.385, with p-value = 0.011 , odds ratio 1.865, which suggests that the change in the log odds of the HIV/AIDS incidence per social status illustrates the association between HIV/AIDS and social status , which is statistically significance .Beside that, occupation variable influence the dependent variable (HIV/AIDS) incidence significantly at 5% level of significance wald = 6.493, p-value = 0.00012. The odd ratio (0.895) suggest that, changing in the log odds ratio of HIV/AIDS incidence per occupation between 0.823 and 0.975.

Table (50) Estimated coefficient for Binary logistic regression Social Status and Job

Variable	B	S.E	Wald	Sig	OR	95%CL of OR
Constant	-2.143	0.183	137.599	0.000	0.117	
Social	0.655	0.079	67.235	0.000	1.937	1.66,2.26
Occupation	-0.016	0.038	0.176	0.675	0.984	0.914, 1.060

Table (51) Testing the model

Model	-2log likelihood	Chi-square	Sig
Null	1675.21		
Final	1602.118	73.449	0.072

Table (52) reports Wald statistic for the education level = 1.246with P-value = 0.264, and for the second variable occupation is 14.784with P-value = 0.0001. The occupation is not significant and education statistically significance at 1% level of significance,. This result reflects the influence of education level to HIV/AIDS among volunteers. Also χ^2 test is done to test the overall mode, χ^2 = 18.923, P-value = 0.00001, so the overall model is statistically significance at 1% level of significance. Hosmer and lemeshow test explain the result as H-L = 31.589 with P-value = 0.000, which suggests that , the model fail to fill the data well).

Table (52) illustrate the results of modeling HIV/AIDS depend on two predictors variable occupation and education level , wald statistics for occupation 1.246 with p-value 0.264 which is not significant , While education level is appear to be statistically significant wald = 6.493 , p-value = 0.000.So the model influence by the education level. Also odds ratio of occupation (0.96) indicate that people without job have positive HIV/AIDS 0.96 times more than working people. Beside that odds ratio of education level(0.85).Thus people without education they have a chance to infected by the disease 0.85 times more than educated people. The over all model is significance at 5% level of significance according to table (5.52) χ^2 = 18.9, P-value = 0.00012 .The goodness of fitting is tested through hosmer-lemeshow test and deviance statistics , H-L =31.6 , p-value = 0.0003 , D = 85.4, p-value = 0.00001.Therefore , the model fail to fit the data well.

Table (52): Estimated coefficient for binary logistic regression(occupation, education)

variable	B	S.E	Wald	Sig	OR	95%CL of OR

Constant	-.671	0.101	43.909	0.000	0.511	
Occupation	-0.042	0.037	1.246	0.264	0.959	0.892,1.032
Education	-0.11	0.043	6.493	0.000	0.849	0.781, 0.923

Table (53): Test of the model

Model	-2log likelihood	Chi-square	Sig
Null	1675.568		
Final	1656.65	18.923	0.000

Table (54) bellow illustrate the full model with the three predictors variables(social status, occupation, education).According to the model estimation , it is very clear that , social status and education level both of them gains statistically significance , wald statistics for social status = 59.26, p-value = 0.0001 and Wald statistics for education 6.321pvalue = 0.012 .The overall model significance is tested through $\chi^2 = 114.845$, p-value = 0.000 , the model is acceptable at 1% level of significance. Hosmer-lemeshow = 34.15, p-value = 0.000, .This result suggests that ,the model is not well fitted. The results insures the validity of the model , these three factors (social, education and occupation) influence HIV/AIDS in spite of the insignificance of occupation variable. The result verify the hypothesis of the study that, the demographic variables affect HIV/AIDS incidence in Khartoum state.

Table (54) Estimated Coefficients of binary logistic regression (social status, occupation, education)

Variable	B	S.E	Wald	Sig	OR	95%CL of OR
Constant	-1.94	0.198	96.726	0.000	0.143	
Social	0.623	0.081	59.26	0.000	1.865	1.591,2.186
Occupation	-0.0016	0.038	0.0018	0.997	1.00	0.928, 1.078
Education	-0.11	0.044	6.321	0.012	0.895	0.822, 0.976

Table (55) Test of the model

Model	-2log likelihood	Chi-square	Sig
Null	1675.567		
Final	1595.548	80.02	0.000

Table (56) Estimated coefficients for a multiple logistic regression model

variable	B	S.E.	Wald	Sig.	$\hat{OR}$
Social status			49.25	0.0000	
Single	-1.519	0.303	25.16	0.0000	0.219
Married	-0.805	0.288	7.817	0.005	0.447
Divorced	-0.23	0.35	0.432	0.511	0.795
Occupation			10.805	0.147	
Labor	4.85	13.5	0.129	0.719	127.8
Employed	5.022	13.501	0.138	0.710	151.7
Free Job	4.859	13.502	0.130	0.719	123.9
Student	5.019	13.501	0.138	0.710	151.2
Tee seller	4.61	13.504	0.095	0.758	64.1
Driver	5.27	13.504	0.153	0.696	195.1
Solider	5.04	13.506	0.139	0.709	154.6
Education level			17.995	0.001	
Primary/basic	0.56	0.248	5.108	0.024	1.75
Intermediate	1.006	0.252	15.995	0.000	2.74
Secondary	0.631	0.287	4.855	0.028	1.9
University	0.682	0.254	7.203	0.007	1.98

22. HIV/AIDS Trends of Three Centers

Table (57) presents the trend of positive HIV/AIDS for Omdurman center for both males and females. It is very clear that males positive cases is large than female positive cases except in the two months may and august.

Table (57) Trends of HIV/AIDS in Omdurman 2007

months	male positive	female positive
January	1	2
February	6	6
March	7	7
April	7	1
May	6	3
June	8	7
July	14	3
August	4	3
September	4	7
October	4	2
November	4	1
December	5	5

Figure (1): Trends of HIV /AIDS Positive Cases in Omdurman center 2007

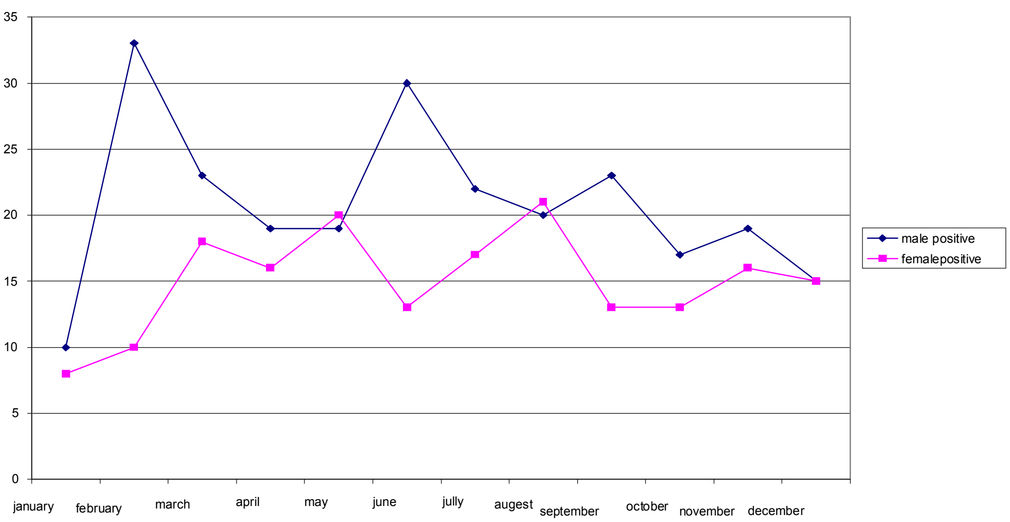


Table (58) presents HIV/AIDS situation which is increased through months in January there (1) case of positive male and (2) cases of positive female. The number grow over the time until it reached its maximum point in July (14) males and (7) for females in September.

The trend of female positive cases is less than the trend of male's positive cases except in September female positive cases more than males positive cases. In the last month of the year the cases of positive HIV/AIDS are equal for both males and females.

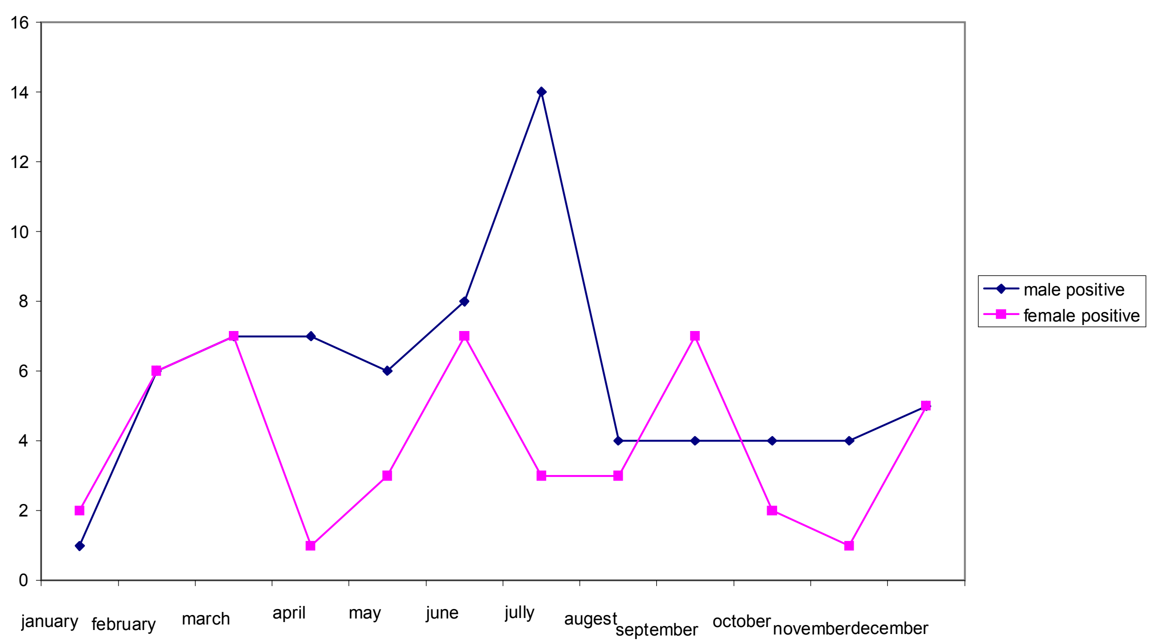
Table (58): Trend of HIV/AIDS in Khartoum 2007

Months	male positive	female positive
January	1	2
February	6	6
March	7	7
April	7	1

May	6	3
June	8	7
July	14	3
August	4	3
September	4	7
October	4	2
November	4	1
December	5	5

Source : Ministry of Health 2007

Figure(5.2): Trends of HIV/AIDS Incidence of Khartoum



source: Minestry of health(Khartoum state)

Table (59) figures trends of HIV/AIDS incidence of Khartoum North center, the graph indicate increasing in the incidence over time, in January was(10) positive of males while (8) positive of females. The numbers of positive males and females rise up gradually, and then fell down into low level (2) cases for both males and females in December.

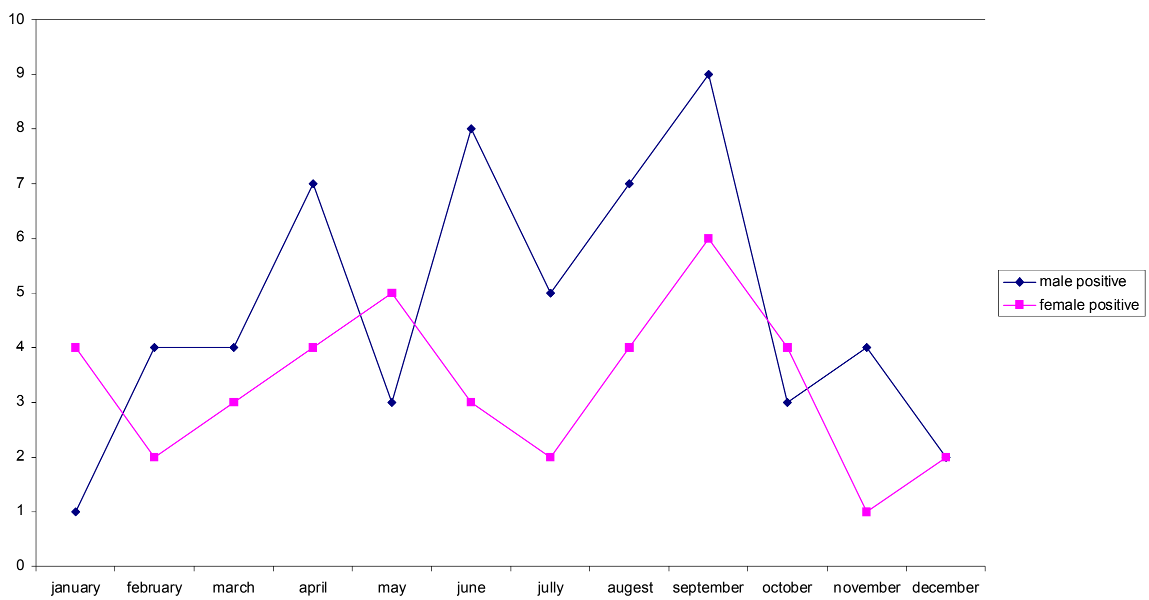
Table (59): Trend of HIV/AIDS in Khartoum North 2007

Months	male positive	female positive
January	1	4
February	4	2
March	4	3
April	7	4
May	3	5
June	8	3
July	5	2
August	7	4
September	9	6
October	3	4
November	4	1

December	2	2

Source : Ministry of Health 2007

Figure(3) Trends of HIV /AIDS Incidence of Bahree



source:Minestry of Health (Khartoum State)

In general, HIV/AIDS positive cases among volunteers and suspected cases of the three centers under the study are in critical situation of the disease. There must be more effort need to be exerted to stop the spread of the pandemic into Khartoum state and other states of the century.
According to the data Omdurman center has the high positive cases either for males or females. It seems that there is considerable number of people attend the center more than the other two centers Khartoum and Khartoum North, so the high incidence is acceptable.

23. Conclusions

The study was conducted in Khartoum state for the period (2003-2007) aims to apply statistical models for the HIV/AIDS data in Khartoum state centers of testing blood and counseling. The study contains six chapters , chapter one is an introductory chapter , chapter two illustrates the important and recent literature about HIV/AIDS, chapter three focus on the HIV/AIDS situation in Khartoum state during the last ten year. While chapter four introduce the statistical models to estimate the models for the Khartoum state data and statistical tests related to them. Chapter five signifies the analysis of data and modeling , so as to obtain the results that help to verify the hypothesis and conduct a relation between HIV/AIDS incidence and demographic variables. Finally chapter six illustrates the main findings through conclusion and followed by selected recommendation derived from the findings. Recognized as an emerging disease only in the early 1980s, AIDS has rapidly established itself throughout the world, and is likely to endure and persist well into the 21s century. AIDS has evolved from a mysterious illness to a global pandemic which has infected tens of millions less than 20 years.
The study construct three main hypotheses they are first, Non-linear models fit the HIV/AIDS data well especially binary logistic regression. Second, demographic variables affect the HIV/AIDS incidence in Khartoum state. Third, HIV/AIDS will be increasing among volunteers in the three centers Khartoum, Khartoum North, and Omdurman. The main objective of the study is to apply statistical models for HIV/AIDS in Khartoum state so as to obtain a good analysis, beside other sub

objectives. The Study Methodology was divided into two parts. Source of Data, Study depend on secondary data collected from the volunteers centers for blood testing and counseling inside Omdurman, Khartoum, and Khartoum North teaching hospitals. The data collected through questionnaires was designed so as to get all the information registered inside the three centers. Study stems the good models and techniques.The main idea of analysis is to apply and identify statistical model that related to AIDS by using statistical packages to construct the models depends on the collected data about HIV/AIDS of Khartoum state. Mainly the study depends on Statistical package for Social Science (SPSS) and Log Xact to estimate the coefficients of the models and the other statistical test related to them. The importance of the study comes from the importance of the AIDS issue itself, because it is destroys the human beings and threats them. Study applies the statistical models related to the data of the AIDS, then define models and chooses the standard model. Our focus here is on binary logistic regression, because it's suitable to the data collected from the three centers inside the three hospitals in Khartoum. Also estimates coefficient and statistical tests done to distinguish between the variables that related to HIV/AIDS incidence and spread through people in the three cities Khartoum, Khartoum North, and Omdurman.Through analysis of data we obtain the main findings that are: the incidence rate of HIV/AIDS among volunteers and suspected cases Omdurman center is 33.5, Khartoum state is 36, and 14.3 for Khartoum North center. So the rate of Khartoum data is the highest one. In order to test the dependency between HIV/AIDS incidence and some socio-economic variables, the study obtains that there is no effect of education level on n HIV/AIDS infection for the data collected from Omdurman. But there is dependency between HIV/AIDS incidence and occupation of volunteers. So the job of individuals affects the HIV/AIDS incidence inside Omdurman area. Also there is a association between HIV/AIDS incidence and social status of individuals; we notice that the high number of positive HIV/AIDS in Omdurman center is among married people.

In the suspected cases of Omdurman center, there is no effect of education level on HIV/AIDS incidence, So education is not influence AIDS Situation for the suspected cases. For Khartoum data the study obtains the dependency between HIV/AIDS and some socio-economic variables, there is dependency between HIV/AIDS incidence and education level of volunteers, it confirms that AIDS related to education status of individuals. Also social status is related to HIV/AIDS among volunteers according to chi-square test which insures the dependency. For Khartoum North data dependency analysis reports that, there is dependent between HIV/AIDS incidence and education level, so education influence HIV/AIDS incidence among people inside Khartoum North area. Also the same result was found through the test between HIV/AIDS incidence and occupation, which reflect the fact that the job of the individuals affects their HIV/AIDS infection. The study construct binary logistic model for Khartoum state , the predictors variables of the models were(social status, education level ,and occupation) .The variables social, education found to be significance where Occupation is insignificant .We conclude that, the three variables are very important to affect the dependent variable HIV/AIDS incidence inside Khartoum state during the period of the study. The overall model test gain statistically significance at 1% level of significance. However, there are insignificance associations between HIV/AIDS incidence and Occupation. Also study conducted diagram for the HIV/AIDS positive cases during 2007 for the three centers. The figure of Omdurman centers reports increasing of positive cases for males and female .It seems that, males positive cases were greater than females positive cases except in may and Augest.For Khartoum hospital cases the incidence is increasing over the time for both sexes. But in September , females positive cases is higher than males positive cases. In Khartoum North center, trends of HIV/AIDS incidence during 2007 is at increasing rate for the positive cases of both males and females rising gradually then fell down into low level in the last month of the year (December).

We conclude that, education level and occupation influence HIV/AIDS incidence into Khartoum data only and social status is found to be the strong factor affect HIV/AIDS incidence for the three centers under the study. For more description of the HIV/AIDS in Khartoum state represented by the testing and Counseling centers, data about positive HIV/AIDS for both males and females during 2007.The trend of HIV/AIDS incidence is increased among both sexes, but males were infected more than females.

The study has the following recommendation

1. Providing a system for HIV/AIDS data so as to follow up the prevalence of the disease annually.
2. Assessing the current situation of the pandemic and try to establish strategic plan to stop or eradicate the disease among people mainly adults.
3. Studying the interrelation between HIV/AIDS incidence and socio-economic variables need to be evaluated by the other specific studies.
4. Spreading awareness about the dangerous of the disease still very weak so community, agencies, government, civil society, and non-governmental organization they should play an important role to raise the awareness.
5. There must be follow up for the volunteers in terms of their Place of residence and reasons for blood testing to know the contribution of them in the HIV/AIDS incidence through logistic regression model.

24. References

1. Center for Diseases Control and Prevention (CDC).(2001). HIV and AIDS in United States. .
2. Centers for Diseases Control and Prevention (CDC) 2002 HIV/AIDS Surveillance Report.
3. Jacob Adetunji (2000) Trends in under-five mortality rates and the HIV/AIDS Epidemic.
4. Jerald F. Lawless (1982) Statistical models and methods for lifetime data.
5. John Whitehead, An introduction to logistic regression
6. Joshua, N. and Christopher, S. (2001), Modeling HIV/AIDS in Sub-Saharan Africa using sero-prevalence data from antenatal clinics.
7. Omyma Sayeed Ahmed and Ahmed Mukhtar (2006) Women, tea-sellers in Khartoum and HIV/AIDS surviving against the odds. Nile basin research Journel, 8. Khartoum, Sudan.
8. Manzoul.A.M. Assal.(2006) HIV Prevalence, Knowledge, Attitude, practices, and risk factors among prisoners in Khartoum state.
9. Mohomed Salih and Salah Hussien (2006) HIV Prevalence, Knowledge, Attitude, practices, and risk factors among truck drivers in Khartoum state.
10. Mohomed Salih Mahfouz (2007) Modeling the impact of HIV/AIDS Epidemic Mortality in Sub-Saharan Africa: with reference to Zimbabwe, Tanzania and Sudan (1985-2003),.
11. MOH(2005) Baseline study on knowledge, attitudes and practices on sexual Behaviors and HIV/AIDS prevention among young people in selected states in Sudan.
12. McClell R. (2006) A comparison of Genital shedding of HIV-1 and sexual risk behavior among Kenyan women based on eligibility for initiation of HAART according to WHO guidelines Acquire Immune Syndrome.
13. Ropert Remis (2003) Report on HIV/AIDS in Ontario, , MD, University of Toronto.
14. Surveillance Report (2003), HIV and AIDS in Canada.
15. SNAP (2002) Situation Analysis: Behavioral and Epidemiological Surveys response analysis.
16. UNIADS. (2002) Report on the global HIV/AIDS epidemic.
17. UNIADS. (2004) Report on the global AIDS epidemic, Geneva.
18. UNAIDS (2004) HIV/AIDS statistics updates.
19. UNAIDS/WHO (2004) HIV/AIDS EPI updates.
20. UNAIDS and WHO (2007) AIDS epidemic update. Global summary of the AIDS epidemic.
21. WHO (1994) The HIV/AIDS epidemic, overview. Available at www.who.org.
22. Park. K. (2000) Text Book of Preventive and Social Medicine, (India). Attitude, practices, and risk factors among prisoners in Khartoum state.

23. Sharif Abu Fatima (2003) Clinical manifestations of AIDS & its correlation to the CD4-lymphocyte count in adults Sudanese presenting to Khartoum Teaching Hospital.